

Interoperabla specifikationer

En utredning inom byggblock metadata

Matthias Palmér
Michalis Vassilas

(metadataexpert för Digg)
(byggblocksansvarig)

Byggblock metadata

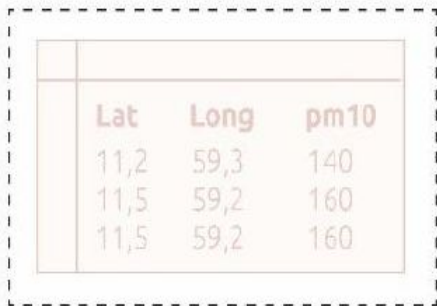
Myndigheten för digital förvaltning, Digg leder arbetet med att etablera en förvaltningsgemensam digital infrastruktur (Ena) för att information ska kunna utbytas på ett säkert och effektivt sätt.

- Samverkan
- Nationella grunddata (ramverk)
- Kompetensområden
- Förutsättningsskapande byggblock

<https://www.digg.se/ledning-och-samordning/ena---sveriges-digitala-infrastruktur>

Beskrivande metadata (Fokus på datamängden)

Titel: Luftkvalitet
Utgivare: Naturvårdsverket
Uppdat: Dagligen

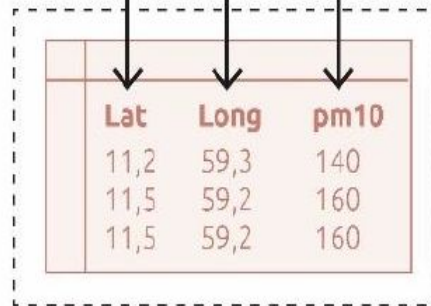


A dashed box containing a table with three columns: Lat, Long, and pm10. The table contains three rows of data. An arrow points from the descriptive metadata box above to the top of this table.

Lat	Long	pm10
11,2	59,3	140
11,5	59,2	160
11,5	59,2	160

Strukturella metadata (Fokus på datauttrycket)

Kolumn: lat
long
pm10
Label: latitud
longitud
partiklar
Datatyp: decimaltal
decimaltal
heltal



A dashed box containing a table with three columns: Lat, Long, and pm10. The table contains three rows of data. Three arrows point from the structural metadata box above to the column headers of this table: 'lat' to 'Lat', 'long' to 'Long', and 'pm10' to 'pm10'.

Lat	Long	pm10
11,2	59,3	140
11,5	59,2	160
11,5	59,2	160

Lite förenklat motsvarar beskrivande metadata FAIR principerna Findable och Accessible medan strukturella metadata motsvarar Interoperable och Reusable.

Syfte med referensgruppen

Syftet med referensgruppen är att arbeta fram en profil för att främja interoperabilitet och återanvändning av delar i en specifikation.

En specifikation kan innehålla en blandning av bakgrund, motivering och mer formella beskrivningar. Det är ett paket med olika delar/resurser, vissa riktade mot mänsklig konsumtion, andra maskinläsbara.

Genom att ge användare möjlighet att söka djupare i en datamängd, på den bakomliggande olika delarna möjliggörs interoperabilitet och återanvändning av dem.

Tidigare förstudie: <https://github.com/diggsweden/information-models-investigation>

Referensgruppsarbete

Material:

- Allt läggs up på Github
<https://github.com/diggsweden/interoperable-specifications>
- Bakgrundsmaterial, specifikation, exempel etc.
- Material på engelska då vi förutser Nordisk / Europeisk samverkan

Process:

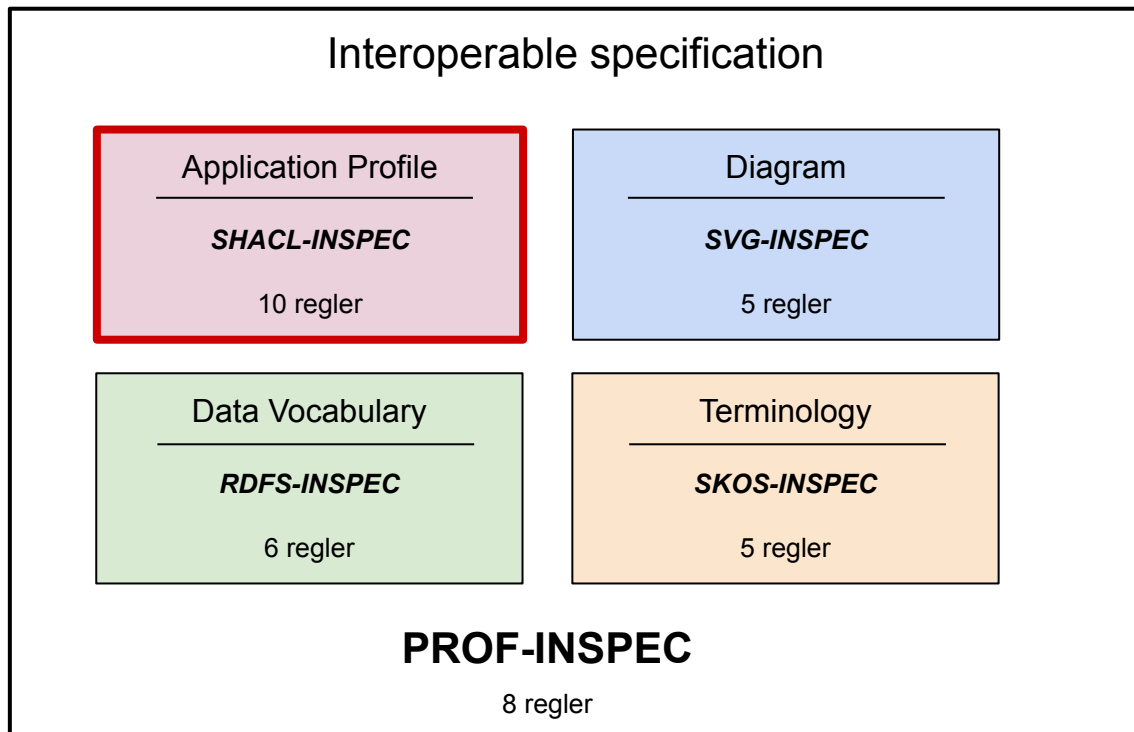
- Ärendehantering på Github - mallar för styra kommunikationen
- Ändringsförslag (PR = Pull Requests)
- Referensgrupp cirka 6 möten - deltagares namn visas på github
- Referensgruppsarbete November 24 - Februari 25

Interoperable specificationer har mognat

Profile parts

1. [Background and motivation](#)
2. [Conceptual design and definitions](#)
3. [Application profiles and subclassing](#)
4. [Rules for interoperable specifications](#) ← **(the formal specification)**
5. [Application Profiles as SHACL-INSPEC](#)
6. [Bootstrapping specifications](#)
7. [Harvesting interoperable specifications](#)
8. [Example specification according to the profile](#)

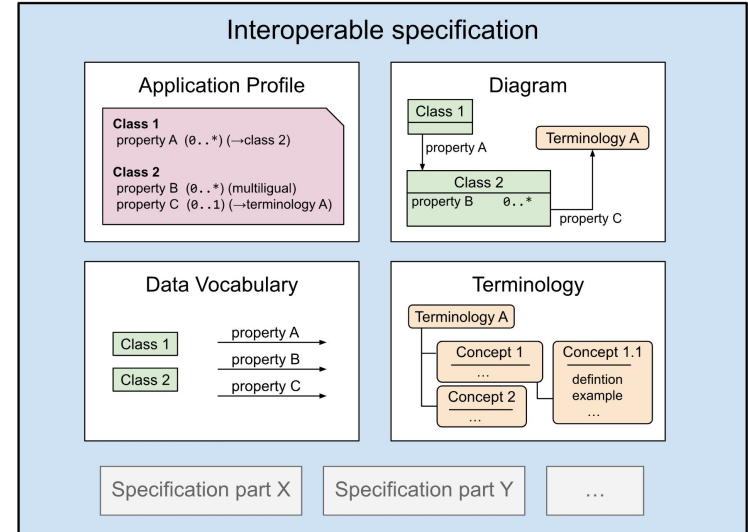
Lyfta fram "Application Profile" mer i INSPEC



Application Profile
och Diagram nu på
toppen

Omstrukturerad text "Conceptual design"

- Introducera delarna redan från början med illustrationen för att ge översikt
- Dela upp kraven:
 - Machine readability requirements
 - Human readability requirements
 - Practical requirements
- Byta till "foundational specifications"
 - För att fånga både data vocabularies och terminologies
 - Är "foundational" ett bra ord här?



Separerat ut diskussionen om subklassning

Diskussionen blev för lång och passade in i "design och motivation" längre.

Nu en egen sida som pratar om varför applikations profiler behöver en egen lösning som inte kan lösas med RDFS eller OWL.

Avslutas med jämförelse med hur Semantic Style Guide talar om återanvändning.



Regler för diagram SVG-INSPEC

- Hur representera INSPEC entiteter i diagram
(class, property, node-shape, property-shape, concept, terminology, concept collection, INSPEC, foundational specification)
- Modelleringsspråk är valfritt, men UML klass diagram är bra

SVG-1: Hyperlänka INSPEC entiteter till deras URI:er.

SVG-2: Attribut: typ av INSPEC entitet, t.ex. `data-inspec-type="class"`

SVG-3: id med md5 summa av entitetens URI (möjliggör CSS)

SVG-4: Attribut: reusable markering, t.ex. `data-inspec-reusable="true"`

SVG-5: Attribut: "main" eller "supportive", t.ex. `data-inspec-weight="main"`

Bootstrap specifications

Dessa specifikationer pekas ut via `dcterms:conformsTo` för att markera att en:

- specifikation är en interoperabel specifikation
 - specifikationsdel är en data vokabulär
 - specifikationsdel är en terminologi
 - specifikationsdel är en applikationsprofil
 - Specifikationsdel är ett diagram
- `inspec:PROF`
 - `inspec:RDFS`
 - `inspec:SKOS`
 - `inspec:SHACL`
 - `inspec:SVG`

och därmed följer motsvarande regler

Skördningen dokumenterad (vanilla INSPEC)

- Detektion via bootstrap specifikationerna
- Hur man laddar specifikationsdelen
- Att man validerar motsvarande regler
- Att man lägger till (vid behov) `dct:requires` och `dct:hasPart`
- Att vissa delar måste finnas i triplestoren andra inte

Följande behöver redas ut

- Skördning av UML enligt `inspec:UML-OSLO`
- Skördning av CSV enligt `inspec:CSV`

UML och INSPEC

- OSLO ramverket - initiativ från Flamländska myndigheten "Digital Flanders" ("Open Standards for Linking Organisations")
 - Tekniskt ramverk för transformation:
<https://github.com/Informatievlaanderen/OSLO-UML-Transformer>
- Delvis tagits över av Semic, framför allt:
 - Core Vocabularies använder detta
 - Semantic Style Guide, mer högnivå
 - Tekniskt ramverk för transformation:
<https://github.com/SEMICEU/toolchain-manual>
- Bygger vidare på ISO-TC211 - Geographic information/Geomatics (The Harmonized Model Maintenance Group)

UML och INSPEC

Data vocabulary

- Klassdiagram
- Klasser, attribut och associationer
- **Ingen** kardinalitet
- Arv
- **Introducera** nya klasser & properties
- Referera terminologier som enumerations
via ap-codelist

Applikationsprofiler

- Klassdiagram
- Klasser, attribut och associationer
- Kardinalitet
- **Inga** arv
- Alltid **referera** klasser & properties
- Referera terminologier som enumerations
via ap-codelist

Terminologier - NJA

- Kan beskrivas men kräver egen
UML-modell, bättre tillhandahålla i SKOS

UML taggade värden

Tag	Explanation
label	Human readable name of the concept, e.g. Public Organization if the identifier of the concept in the diagram is PublicOrganization. This name is used in the documentation of the diagram.
definition	Summary description of the concept.
usageNote	Additional description, e.g. with examples.
uri	URI of the concept.

A language code is added to the tag name to indicate the language in which the tag value is set, e.g. definition-nl.

The tag name is preceded by the prefix ap if the tag refers to the AP instead of the VOC, e.g. ap-definition-nl.

Från SEMIC Style Guide

Examples:

The class Address in Core Location has several properties. All the properties have a displayed name (e.g. "addressArea") with the UML class diagram and associated UML Tags, such as a label for the HTML specification ("address area"), definition, usage note and the related URI.

