

Strukturella Metadata - Referensgrupp 6 den 24 februari 2025

Närvarande

Andrew Mercer (RAA)
Benny Lund (Bolagsverket)
Caspar Almalander (SKR)
Cilla Öhnfeldt (Naturvårdsverket)
Erik Mossing (Bolagsverket)
Fredrik Emanuelsson (Riksarkivet)
Fredrik Persäter (Lantmäteriet)
Johan Oelrich (Arbetsförmedlingen)
Marcus Smith (RK)
Martin Brandhagen (Vetenskapsrådet)
Matthias Palmér (Digg)
Mattias Ekhem (Digg)
Michalis Vassilas (Digg)
Lars Nässlund (Trafikverket)
Olov Johansson (SGU)
Olof Olsson (SND)
Patrick Wahlgren (SCB)
Peter Israelsson (Digg)
Stefan Jakobsson (Svensk nationell datatjänst)
Tomas Lindberg (SGU)

Bakgrund och syfte

Arbetet sker inom byggblock metadata, en del av Ena, Sveriges digitala infrastruktur och drivs av Myndigheten för digital förvaltning. Syftet med referensgruppen är att arbeta fram en profil för att främja interoperabilitet vid datadelning genom återanvändning av klasser, egenskaper och koncept.

Allt material kommer att publiceras till Github: specifications

<https://github.com/diggsverige/interoperable>

Materialet som tagits fram är på engelska är för att underlätta dialog med bland annat SEMIC (Europeiska kommissionens initiativ för att förbättra den semantiska interoperabiliteten) på EU nivå. Referensgruppsdeltagarnas namn kommer att synas på Digg:s GitHub. Om ni inte vill att era namn ska synas på GitHub hör av er till Michalis Vassilas.

Tidigare förstudie finns även den på Github: <https://github.com/diggsverige/information-models-investigation/tree/main>

Syftet med applikationsprofilen är att tydligare specificera hur klasser, egenskaper och koncept ska återanvändas i nya sammanhang. Behovet kan till stor del täckas av SHACL Shapes Constraint Language, ett språk för att validera RDF-grafer mot ett antal villkor. Dock är SHACL för flexibelt för användningsfallet med applikationsprofiler, därför definierar vi SHACL-INSPEC för att fånga de specifika krav/begränsningar som behöver uppfyllas vid användning av SHACL för att uttrycka applikationsprofiler.

Genomgång av den interoperabla specifikationen (INSPEC) v 0.9

Mathias presenterade den senaste beskrivningen av INSPEC baserat på feedback från Semic, referensgruppen och egna anpassningar: <https://github.com/diggsverige/interoperable-specifications?tab=readme-ov-file>

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

Sedan sist har ordningen förändrats och ett nytt avsnitt har tillkommit i form av - [Using the same property for different purposes](#) samt nya regler i regelavsnittet som ett sätt att inkorporera tankar från SEMIC kring hur olika varianter av DCAT-AP ska hanteras. Överlag kan sägas att profilen mottogs väl av SEMIC med detaljerade och engagerade diskussioner. Dialogen med SEMIC har lett till två justeringar av INSPEC kopplat till hur interoperabla specifikationer ska förhålla sig till varandra.

1. **Införa inspec: refine och inspec: variant för att hantera förfiningar inom interoperabla specifikationer.** Tidigare hade dterms: isVersionOf och sh:and använts, men den nya lösningen inkluderar en dubbel konstruktion med sh:and för property shapes för att bibehålla SHACL-kompatibilitet. På applikationsprofilnivå föreslås att prof: isProfileOf kombineras med inspec: variant, vilket ska möjliggöra en mer exakt beskrivning av hur specifikationer relaterar till varandra. Regler har definierats för att säkerställa enhetlighet och förutsägbarhet vid implementering.
2. **Den andra justeringen handlade om dubbelanvändning av properties** och i vilka fall det är möjligt att återanvända en egenskap istället för att skapa en ny. Motivationen bakom detta är att minska kostnaden för att etablera och förvalta olika profilvarianter. Återanvändning av en property kan ske när den används med olika nodtyper, olika datatyper, separerbara URI-mönster, ett begränsat antal URI:er eller blanka noder med unika egenskaper. Ett exempel på dubbelanvändning av en property är dct: subject, där en datamängd kan klassificeras både med en språktaggad text och en URI från en kontrollerad vokabulär, vilket illustrerar hur en property kan användas flexibelt men fortfarande hanteras inom tydliga ramar.

Synpunkter och funderingar

Gruppen var överens om att det är många termer och ganska abstrakta begrepp som behövs för att förstå de olika avsnitten kopplat till INSPEC. Begreppen behöver förklaras så att en relativt bred målgrupp förstår. Ett eget avsnitt i form av en ordlista kommer att läggas till i samband med att Matthias bearbetar texten och svarar på inkomna synpunkter.

Frågan huruvida Standard för svensk yrkesklassificering (SSYK) skulle kunna vara en kandidat för kommande INSPEC-pilot lyftes och kommer att övervägas.

Frågan om en formell remissrunda till fler myndigheter lyftes och Digg:s ståndpunkt är att en sådan mer formell och bredare remiss inte är aktuell i detta skede.

Inför fortsättningen

Digg kommer att fånga och hantera de synpunkter och förbättringsförslag som inkommit samt uppdatera diagram och göra vissa redaktionella förändringar i början av mars. När dessa justeringar är gjorda och eventuella ytterligare synpunkter av redaktionell karaktär hanterats anser Digg att referensgruppens arbete är klart för denna gång. Under våren kommer Digg att föra en dialog med SEMIC för att förankra de senaste justeringarna kopplat till interoperabilitet mellan olika profiler på EU-nivå och fånga eventuella justeringar i en officiell version som förhoppningsvis kan lyftas till EU-nivå under SEMIC:s paraply. I anslutning till detta kommer en första version INSPEC publiceras på dataportalen och utgöra grunden för det fortsatta arbetet.

I nästa skede kommer fokus att flyttas till att pröva tillämpningen av INSPEC i praktiken genom fortsatt arbete med piloten vilket innebär att ett webbgränssnitt kommer att tas fram för att utforska strukturella metadata som följer de rekommendationer som tagits fram i förstudien om "systemstöd för informationsmodeller". Även ett koncept för uppackning av specifikationer som

uppfyller INSPEC profilen kommer tas fram vilket också innebär funktionalitet för uppackning av format 2 (UML) enligt tidigare plan.

Tanken är att utgå från de nationella grunddataområdenas informationsmodeller för att testa den praktiska tillämpningen av INSPEC då den här datan i grunden behöver vara interoperabel med tanke på dess värde för offentlig sektor och samhället i stort. Digg har inte landat exakt hur arbetet ska bedrivas och hur aktuella myndigheter ska involveras, men förhoppningen är att piloten ska bygga på riktiga specifikationsdelar såsom en informationsmodell för att testa Oslo-ramverket.

Digg kommer att kontakta aktuella myndigheter för att diskutera deras möjlighet att bidra till arbetet med piloten, antingen med exempeldata och/eller mer aktivt engagemang. I takt med att piloten växer fram kan det bli aktuellt med demotillfällen där berörda myndigheter och intresserade kan få en inblick i hur INSPEC kan realiseras i praktiken och förhoppningsvis få en bättre bild av nytta och vad som kommer krävas för att göra en specifikation interoperabel och maskinläsbar. Piloten beräknas vara klar oktober 2025.