

Notat om metadata om grunddata

6. december 2013
SAR & PLACE

Indledning

Metadata – data om data – betegner ikke en entydig klasse af data. Anvendelsen af betegnelsen metadata knytter sig typisk til et specifikt perspektiv på data: De data, som er forretningsrelevante i en given sammenhæng, kaldes data, mens de data, som giver kontekst til de forretningsrelevante data, kaldes metadata.

I det følgende beskrives en række typer af metadata, som man ofte opererer med inden for den offentlige dataforvaltning.

For hver type angives:

Beskrivelse	Beskrivelse af typen af metadata.
Eksempler	Eksempler på denne type af metadata og de informationer, denne type metadata kunne indeholde.
Grunddata	Beskrivelse af hvordan denne type metadata håndteres i grunddata, herunder om der er fastsat fælles regler for indhold eller formater.
Service	Overvejelser og anbefalinger om hvordan denne type metadata bedst håndteres i forhold til services.

Datasæt-metadata

Beskrivelse	Data, som beskriver en samling af dataobjekter, som findes i et register, og/eller som er etableret ud fra samme specifikation.
Eksempler	Et register kan operere med datasæt, indeholdende dataobjekter, som er indsamlet med fælles tilblivelseshistorie, kvalitetsbeskrivelse, omfang, lovhjemmel etc. Eksempelvis kan et helt geodata-datasæt (som Danmarks Højdemodel) have metadata, som gælder for hele bestanden af dataobjekter, fx koordinatsystem, publikationsdato, præcision.
Grunddata	Der er pt. ingen fælles regler eller standarder gældende for denne type metadata inden for grunddataprogrammet. Udstillingen specificeres af den pågældende registermyndighed.

Det kan anbefales, at data modelleres på en måde, så der skabes relation mellem de enkelte data-objekter og deres generelle metadata; eksempelvis ved at de generelle metadata repliceres som attributter på hvert data-objekt, eller ved at data-objektet har en relation til et data-objekt, der repræsenterer datasættet og dets metadata.

Service	Hvis data er modelleret, som beskrevet ovenfor, kan registermyndigheden specificere services, som leverer datasæt-metadata i tilknytning til data.
---------	--

Dataobjekt-metadata

Beskrivelse	Data, som beskriver dataobjektet, men ikke er centrale forretningsdata.
Eksempler	<p>Bitemporale egenskaber, oprindelsesoplysninger, kvalitetsbeskrivelse, aktører. Disse metadata er - alt efter synsvinkel - enten centrale data eller data, som beskriver det enkelte data-objekts historik og organisatoriske tilknytning.</p> <p>Hvis data-objektet eksempelvis omfatter Ejendomsnummer og Ejerforholdskode og Landsejerlavskode/tekst, så kan data som Revisionsdato, Teknisk standard, Nøjagtighedsklasse og Kilde siges at være dataobjekt-metadata.</p>
Grunddata	<p>“Modelregler for grunddata” definerer en række generelle egenskaber (kapitel 6), som kunne betegnes som metadata, bl.a. bitemporale egenskaber samt oplysninger om aktør og status. Der er i modelreglerne fastsat krav til formater og indhold af de generelle egenskaber.</p> <p>På Datafordeleren vil dataobjektet også indeholde oplysninger om tidspunkt for Datafordelerens kopi.</p> <p>Med hensyn til oplysninger om oprindelse, kvalitet og lignende har datamodel-projektet ikke fundet det hensigtsmæssigt at fastsætte en generel standard herfor, da oplysningerne vil have vidt forskellig karakter fra datadomæne til datadomæne. Det er derfor op til den enkelte registermyndighed at fastsætte regler herfor efter behov.</p>
Service	Dataobjekt-metadata udstilles sammen med data på linje med

andre attributter. Registermyndigheden kan vælge at specificere services, som udelukkende leverer dataobjekt-metadata.

Definitioner og beskrivelser

Beskrivelse	Beskrivelser af de forretningsobjekter, som data omhandler.
Eksempler	<p>Definitioner af forvaltningsobjektet, definitioner af attributter, indhold af værdilister.</p> <p>Hvis datamodellen for eksempel indeholder et objekt “Byggesag”, vil det være relevant, at der findes en definition af objektet, som beskriver, hvad der forstås ved en byggesag. Sådanne beskrivelser er vigtige, dels for at sikre datakvaliteten ved at udbrede en fælles forståelse af, hvad dataobjektet beskriver, dels som udgangspunkt for applikationsudvikling, hvor beskrivelserne kan indgå som forklaringer for de fremviste eller behandlede data.</p>
Grunddata	<p>“Modelregler for grunddata” stiller krav om dokumentation af datamodellen (se regel 5.9) i form af beskrivelser af alle model-elementer (klasser, attributter, relationer) inkl. bl.a. navn og definition. Beskrivelserne er indlejret i og vedligeholdes sammen med UML-modellen. Der er fastsat krav til formater og indhold. Værdilister skal i grunddatamodellen modelleres og fremgå eksplicit af datamodellen og modelreglerne anviser forskellige formater herfor (se regel 5.6).</p> <p>Beskrivelser og definitioner, som er korrekt placeret i UML-modellen kan outputtes i XMI-versionen af modellen, og kan derfra transformeres til det ønskede snitfladeformat. Fx kan definitionerne indlejres som kommentarer i XML schema eller placeres direkte i XML dokumentet.</p>
Service	<p>Registermyndighederne kan vælge at konstruere services således, at det er muligt, at vælge om objektdefinitioner skal leveres indlejret i data - evt. også om det skal være i skemaet eller i dokumentet/instansen.</p> <p>Datamodellen udstyres med tjenester, som kan generere samlinger af definitioner (feature catalogs) i et antal (maskin/menneskelæselige) formater.</p>

Transaktions-metadata

Beskrivelse	Data, som beskriver de data, som leveres i en service.
Eksempler	<p>Transaktionstidspunkt, overført datamængde, paging-information, aktualitet ift. kilderegister. Evt. også relevante data-resuméer eller statistikker, fx min/max værdi samt sum for given attribut, kvalitetsbeskrivelser etc.</p> <p>Så snart et service-kald har returneret data, kan disse data beskrives med metadata, som knytter sig til netop denne data-bestand. Disse metadata kan være nyttige ift. applikationsdesign. De kan bruges til at karakterisere et søgeresultat uden nødvendigvis at vise alle poster, eller applikationen kan foretage relevante handlinger baseret på datamængde, aktualitet eller måske endda konfidentialitet.</p>
Grunddata	Der er ingen regler gældende specielt for grunddata, omhandlende denne type metadata.
Service	Det er god skik at specificere services - både i REST og SOAP - som leverer transaktionsmetadata som udgangspunkt. For eksempel i headere eller indkapslende elementer. Prøv fx http://api.digitaliser.dk/rest/groups .

Service-metadata

Beskrivelse	Data, som beskriver den service, der leverede data.
Eksempler	<p>Version, ejerskab, data-skema.</p> <p>Disse er statiske oplysninger, som beskriver den service, som leverede data. Sådanne data kan være interessante især i forbindelse med udvikling af de applikationer, som binder sig til servicen.</p>
Grunddata	Der er ingen regler gældende specielt for grunddata, omhandlende denne type metadata.
Service	Det anbefales at specificere services - både i REST og SOAP - som leverer service-metadata.

Sammenfatning

Registermyndighederne har ansvar for at specificere services for deres grunddata og på den måde sikre, at relevante metadata leveres i tilknytning til data. Oversigt over data såvel som metadata vil på den måde blive tilgængelige på Datafordelerens metadatakatalog. Datamodeller vil ligeledes blive tilgængelige her.

Referencer

Modelregler for grunddata:

<http://www.digst.dk/Loesninger-og-infrastruktur/Grunddata/Projekterne/Faelles-datamodel>

Krav til Datafordelerens metadatakatalog findes i udbudsdokumenter version 2.0, “Bilag 3 Leverancebeskrivelse” (se afsnit 4.4.4):

<http://www.digst.dk/Loesninger-og-infrastruktur/Grunddata/Projekterne/Faelles-Datafordeler/~//media/Files/L%C3%B8sninger%20og%20infrastruktur/grunddata/Udbudsmateriale%20til%20Datafordeleren%20november%202013/Bilag%2003%20%20Leverancebeskrivelsev20.pdf>