

Kommenterings-skema

15. januar 2018
Sekretariatet for
Initiativ 8.1.

BEMÆRK: Alle indsendte kommentarer offentliggøres (på arkitektur.digst.dk). Såfremt du *ikke* ønsker en kommentar offentliggjort, bedes du gøre opmærksom på det i mailen.

Den udfyldte skabelon sendes til arkitektur@digst.dk.

Dokumentet/dokumenter der kommenteres på: Retningslinjer for stabile http-URIer

Organisationen der kommenterer: Sundhedsdatastyrelsen

Kontaktperson for evt. uddybelse: Thor Schliemann

Kommentarer:

Oversigt

- Indledning til kommentarer
- Kommentarer til Retningslinjer for stabile http-URIer
- URI anvendt af FHIR
- Konklusion

Indledning til kommentarer

I Indledningens afsnit 3, skriver man at ”... *formål er at formidle 'best practice' for design og publicering af stabile URIer.* Dog synes der mangler en egentlig beskrivelse af formål og anvendelse for *Retningslinjer for stabile http-URIer.* Dog skriver man noget om anvendelsesområder uden at konkretisere det: *Retningslinjernes fokus er på data, datasæt, web-APIer, datamodeller og modelementer, mens de ikke anses for nødvendige for adresser for websider i almindelighed.* Så, hvad er det URI kan understøtte dansk offentlige forvaltning med? Der har på tidligere møder i arbejdsgruppen ytret ønske om konkretisering og gerne understøtte af konkrete eksempler i offentlig dansk forvaltning.

SDS kan nu godt følge den overordnede idéen om stabile idéer, og det hjælper med forståelse når man henter yderligere råd fra ISAs rapport ”*D7.1.3 - Study on persistent URIs, with identification of best practices and recommendations on the topic for the MSs and the EC*”. En underfundig bemærkning: kontaktpersonen kun under reviewet ikke hente ISAs rapport fra det angivne link. Når man har læst retningslinjerne fra ISA får man

dog et bedre indblik i anvendelse, dog kan vi kun opfordre vi jer til at indlede bedre om formål og anvendelse i en dansk kontekst.

Vi bidrage med få konkrete eksempler HL7 FHIR, hvor URI anvendes, og hvor der også er behov for at URI'er er stabile. Da FHIR standard endnu ikke er normativ skal disse eksempler betragtes som udkast. Dog forventes FHIR som overordnet standardisering ramme at blive normativt i år (2018), men flere del-elementer vil fortsat være under udvikling efter en normativ udgivelse, dvs. disse ikke alle har fuld modenhed, og ændringer kan forekomme. Om der vil forekomme ændringer af anvendelsen for URI'er, ved vi for nuværende ikke.

Én af grundforudsætningerne i sundhedsvæsnet måder for at øge den semantiske interoperabilitet er ved hjælp af *Object Identifiers* (OID), som telecom industrien med mange flere har anvendt siden 1980^{erne}. Det er en metode til opmærkning af objekter herunder dataobjekter, klassifikationer, datasæt m.m., anvendt på tværs af brancher internationalt. Selvom OID ikke har helt det samme formål, er der dog ligheder fx at man kan dokumentere sine objekter.

Når der indledes med at gøre opmærksom på OID som standard, er det for, at sikre at eksisterende metodik ikke overses, og at sikre forståelse for transformationsprocesser og ved eventuel målarkitekturer hvor stabile-URI indgår. I gennemgangen af FHIR's anvendelse af URI, bliver dette belyst en smule, og her er muligheder for at bære modne opmærkningsmetodik/-standard videre. For non-FHIR systemer, dvs. tidligere version af HL7 V3 (version 3) og i særdeleshed Clinical Document Architecture (CDA) anvendes alene OID.

Kommentarer til Retningslinjer for stabile http-URL'er

ISAs 10 regler er i dansk regi blevet til 14 retningslinjer.

Hvad skyldes denne udvidelse af retningslinjerne? Er det en præcisering eller udvidelse ift. de oprindelige retningslinjer? De fire yderligere retningslinjer har man tilføjet for at præcisere hvad der menes med at "overholde det vedtagne mønster". Dette læser vi som en profilering og indsnævring af retningslinjerne, hvilket kan være en sund disciplin.

En URI skal defineres som en HTTPS-URI.

Mht. at *opslog krypteres kommunikationen, og dermed forøges sikkerheden*. Hvad er baggrunden for denne præcisering og strammere profilering? Er det for, at sikre integritet mellem anvender og URI-udstilleren? Skyldes det at data er følsomme? Hverken ISA eller FHIR's anvendelse af URI kræver HTTPS, hvilket kan udfordre vores standarder på sundhedsområdet. En grund begrundelse for HTTPS kan anvendes til at sikre internationale standarder modnes, ved at vi bidrager med disse argumenter i internationale arbejdsgruppe der videre udvikler standarder, regler og "retningslinjer".

Hvis dette bunder i fornuftige begrundelse, er det argumenter man kan bringe til arbejdsgrupper, der standardisere stabile-URI, eksempelvis ISA og HL7. Hvis det ikke er tilfældet, vil vi opfordre til at ændre ”kravet” fra ”obligatorisk” til at HTTPS er anbefalet, når det er muligt. Og det vil ofte være for stabile-URI'er der er danske. Se endvidere *Genbrug i videst muligt omfang eksisterende identifikatorer*.

Lad alle http-URIer tilhøre domænet <https://data.gov.dk/> eller et af dets underdomæner.

Kan dette betragtes som styringsredskab/governance af stabile URI'er i fælles offentligt regi? Kan private anvende dette domæne eller skal de have deres eget?

Er der tænkt værktøjsunderstøttelse af URI'er? Man kan eksempelvis lade sig inspirere af OID-INFO.com, som er et repository for OID. Her har man et værktøj hvor man kan forespørge om et Objekt (OID) findes, inden man opretter nye. Noget lignende bør tænkes ind, så redundans har en realistisk mulighed for at blive elimineret.

At kun anvende et dansk domæne synes kan have fordele når genbruge af URI'er skal sikres. Dog se bemærkninger under *Genbrug i videst muligt omfang eksisterende identifikatorer*.

URIer skal i videst muligt omfang være menneskeligt læselige ord.

En interessant pointe. Dog måske udfordret af *Termer anvendt i URIer skal være engelsksprogede*. Oversættelser af danske begreber til engelsk kan til tider være en krævende opgave, og man kan eventuelt opfordre til dokumentation for valg af engelsk sproget term/begreb.

Det synes lidt underligt at termer ikke skal være meningsbærende, og dog bruges der meget spaltepads på at gøre dem 'forståelige'. Hvor vigtigt er det reelt for URI'er der er overladt til maskiner, at være nemme at læse? Er det til fejlsøgning for udviklere, så URI'er er nemme at læse og huske. Altså en hjælp for mennesker? Er det strengt nødvendig, som en retningslinje?

Genbrug i videst muligt omfang eksisterende identifikatorer.

Den danske oversættelse overholder ikke gængse regler for profilering, da den er bredere end den oprindelige på engelske: *Re-use existing identifiers*. Hvorfor har man ikke valgt en mere direkte oversættelse, her et fattigt forsøg: *Genbrug eksisterende identifikatorer*.

Hvad er baggrunden for denne mangel på præcisering? Eksempelvis, hvis man anvender en internationale URI i det offentlige. Så skal denne http-URI tilhøre domænet <https://data.gov.dk/> og dernæst kan den re-direktes til dens internationale URI. Dette er ikke fornuftig genbrug. Genbrug en internationale URI direkte, og undgå re-direkte. Det skaber mindre administrativ overhead og gavner transparens.

Http-URler anvendes til entydig identifikation af data, datasæt, datamodeller og modelementer.

Man skriver: *En http URI identificerer entydigt en ressource. En http URI kan også fungere som en adresse, der kan åbnes for at hente yderligere information om ressourcen.*

OID gør altid begge dele, identificere og beskriver et objekt, som her er synonym med en ressource. Hvorfor anvender man ikke noget der allerede er anvendt internationalt. Hvad er baggrunden for at lave noget nyt, når det findes? OID er velbeskrevet og har værktøjer til sikre entydighed af en ressource – ét unikt objekt, har ét unikt OID-nummer, og en beskrivelse. Der kan være gode grunde til stabile URl'er.

OID er gode til semantisk interoperabilitet, men indeholder ingen måde for teknisk opslagsmuligheder i eksempelvis online datasæt. Måske er det her man finder en symbiose mellem en gammelkendt (OID), og den nye i klassen – den stabile URI. Hvor URI kan give opslagsmekanismen, der øger den tekniske og computerprocesserbare interoperabilitet.

URI anvendt af FHIR

FHIR anvender URI fx ift. RFC, eksempelvis 3986:

<https://tools.ietf.org/html/rfc3986>

Det betyder bl.a. at den anvender URI bredere end blot http og HTTPS. Og derfor er de fleste FHIR anvendelser af URI udenfor anvendelse ift. *Retningslinjer for stabile http-URl'er*.

http URI fra FHIR - CodeSystem.url er af type uri

Definition: An absolute URI that is used to identify this code system when it is referenced in a specification, model, design or an instance. This SHALL be a URL, SHOULD be globally unique, and SHOULD be an address at which this code system is (or will be) published. The URL SHOULD include the major version of the code system. For more information see Technical and Business Versions. This is used in [Coding]{datatypes.html#Coding}.system¹.

...the canonical URL that never changes for this code system - it is the same in every copy. Ideally, the URL should also be the location of the master version of the code system, though this is not always possible².

Dette eksempel synes at på områder kunne falde indenfor *Retningslinjer for stabile http-URl'er*. Det fortæller om data – et kodesystem. Bemærk endvidere at man her beder om kodesystemets version, som går imod retningslinjerne. Pointen her er at

¹ <https://www.hl7.org/fhir/codesystem-definitions.html>

² <https://www.hl7.org/fhir/codesystem.html>

der skal beskrives en klar afgrænsning af *Retningslinjer for stabile http-URIer*, så retningslinjer bliver anvendt og forstået optimalt.

Eventuelle eksempler på URI anvendt i andre standarder:

HL7 FHIR: <https://www.hl7.org/fhir/datatypes.html>

Konklusion

Retningslinjerne er udarbejdet med udgangspunkt i tilsvarende internationale retningslinjer, og de skal sikre, at http-URIer holdes unikke og vedvarende, så de, når de anvendes til opslag, altid peger på samme ressource. Meget enig i dette, dog synes det at fordrer understøttelse af værktøjer der kan sikre netop dette, dvs. at URI er stabile, unikke og entydige. Derfor kan der opfordres til at forstå OID der løser tilsvarende krav til 'objekter' der skal være stabile, unikke og entydige, og disse har værktøjsunderstøttelse. OID synes her at være semantisk interoperabilitet for mennesker og computere, dvs. de kan processeres, men mangler et automatiseret (maskine) opslagsmulighed. Det er bl.a. en af styrkerne i stabile-URI'erne. Jeg tror vi skal tænke dette bedre sammen, så 'vi' bruger det bedste af begge verdenerne.

Hvis disse retningslinjer ikke for nuværende løser konkrete og efterspurgte behov og ikke er tilstrækkeligt markedsanvendt, kan man lade disse retningslinjer ligge til egentlige behov for stabile URI trænger sig på. Vi tror det kan finde anvendelse og når et mere klart formål og en anvendelse forligger. Da kan man gøre det til en profil, der er et regelsæt man anbefaler at overholder, og ikke blot er 'retningslinjer', som man kan frygte ingen vil efterleve.