Stedsnavn og Administrative enheter som linked open data.

# Linked open data

Linked open data er en måte å levere data på som bygger på standard web-teknologier, men som i motsetning til vanlig websider, gjør det mulig for datasystemer å lese informasjonen og forstå dens mening. Informasjon som leveres på den måten kan videre knyttes opp mot andre data som levers på den måten, og på den måten vise relasjonene dem imellom.

Hver enkelt bit av informasjon identifiseres av hver sin unike URI[[1]](#footnote-1). RDF-formatet brukes så til å beskrive dataene, og publiseres gjennom HTTP. Samlet, har får man informasjon som kan brukes og tolkes av både programvare og mennesker.

For å gi dataene mening, beskrives de ved hjelp av vokabularer eller ontologier. Innholdet i dataene knyttes da opp mot begreper som defineres gjennom den aktuelle ontologien. For mer informasjon, se W3C - Vocabularies[[2]](#footnote-2).

Relasjonene mellom data er kanskje det viktigste aspektet ved RDF-data. For å bruke et eksempel fra data vi leverer til Hack4no, Administrative enheter, så består det datasettet f.eks. av flere kommuner. Hver enkelt kommune identifiseres med en URI. Hver enkelt kommune har også flere egenskaper, som navn og geometri. I linked open data bruker man begrepene subjekt, predikat og objekt for å beskrive denne informasjonen. I dette eksempelet har kommunen (subjektet) en egenskap (predikatet), f.eks. navn, hvor selve egenskapens innhold (objektet) er «Ringerike». Selv om dette ikke kommer fram i datasettet vi leverer, kan en også tenke seg at kommunen har den egenskap at den «ligger i fylke» (predikat) «Buskerud» (objekt), hvor «Buskerud» er et subjekt som igjen har sine egenskaper. På den måten kan subjekter også være objekter, og man kan presentere et komplekst nett av data. Dette nettet av data kalles i linked open data-verdenen grafer.

For å gjøre spørringer i en graf, bruker man SPARQL[[3]](#footnote-3). Dette er et SQL-lignende spørrespråk, men det at det forholder seg til triples gjør det likevel ganske annerledes. Man kan gjøre spørringer etter alle subjekter, predikater og objekter, og kombinasjoner av disse. Man kan f.eks. finne ut hvilke subjekter som har egenskaper navn, og hva navnene er, eller finne subjekter med navnet «Ringerike».

Eksempel på det siste:

*PREFIX na: <http://www.link\_til\_ontologi/navn/>*

*Select ?subjekt WHERE ?subjekt na:navn ”Ringerike”*

For å forstå dette bedre anbefaler vi:

* “Brief introduction to linked data” av Lukas Koster.[[4]](#footnote-4)
* “Introducing Linked Data and the Semantic Web” fra Linked Data Tools.[[5]](#footnote-5)

# Data for hack4no

Vi har i forbindelse med Hack4no levert 2 datasett i RDF-format: Stedsnavn fra Sentralt stedsnavnsregister (SSR) og Administrative enheter. Realiseringen av dataene som RDF bygger på datamodeller fra INSPIRE[[6]](#footnote-6), henholdsvis Geographical names[[7]](#footnote-7) og Administrative units[[8]](#footnote-8). Vi har også opprettet ontologier basert på disse datamodellene, som inneholder klassene/begrepene og deres definisjoner[[9]](#footnote-9). For å realisere den geografiske delen av dataene, har vi benyttet GeoSPARQL-standarden[[10]](#footnote-10) for geometriene, noe som muliggjør romlige spørringer gjennom et SPARQL endpoint som støtter standarden.

Data er lagt ut i Parliament[[11]](#footnote-11), som er en triple store/SPARQL endpoint med støtte for GeoSPARQL. På grunn av ytelsesproblematikk, er det kun stedsnavn innenfor Hole kommune som er lagt ut, men administrative enheter-datasettet er fullstendig. Stedsnavn for hele Norge er imidlertid lagt ut på data.kartverket.no, og det samme er administrative enheter. Derfor er det en mulighet å laste ned data for å jobbe lokalt, f.eks. ved å sette opp sin egen triple store.

Datamodellene fra INSPIRE kan være vanskelig å forstå, og det samme kan måten SPARQL-spørringer settes opp være, for de som ikke har kjennskap til det. Derfor har vi satt om noen ferdige spørringer som fokuserer på å gjøre romlige søk med geoSPARQL. Disse kan man så forsøke å endre på, eller lage egne spørringer.

**Merk:** Dette er en demo, der det kan være problemer med ytelsen. Tjenesten kan også gå ned. Hvis det skulle skje, anbefaler vi å laste ned RDF-data og arbeide med det lokalt.

## Tilgang til data

* For å prøve SPARQL-spørringer, gå til: <http://rdf.kartverket.no/>.
* For å laste ned alle stedsnavn, gå til: <http://data.kartverket.no/data/stedsnavn/landsdekkende/inspire_gn.rdf.zip>
* For å laste ned administrative enheter, gå til: <http://data.kartverket.no/data/grensedata/landsdekkende/inspire_au.zip>
* REST API: <http://rdf.kartverket.no/parliament/sparql>
* Tilgjengelige grafer:
	+ urn:INSPIRE:AU
	+ urn:INSPIRE:GN:Hole
* Ontologier ligger her:
	+ [http://rdf.kartverket.no/ont/inspire/au#](http://rdf.kartverket.no/ont/inspire/au)
	+ [http://rdf.kartverket.no/ont/inspire/gn#](http://rdf.kartverket.no/ont/inspire/gn)
1. Uniform Resource Identifier: http://en.wikipedia.org/wiki/Uniform\_resource\_identifier [↑](#footnote-ref-1)
2. http://www.w3.org/standards/semanticweb/ontology [↑](#footnote-ref-2)
3. http://www.w3.org/TR/sparql11-query/ [↑](#footnote-ref-3)
4. https://docs.google.com/document/d/1W6UOCLgxTyM0BlPfd5hs58dh4k6CUdLW354AjjtnJfk/edit?hl=en\_US# [↑](#footnote-ref-4)
5. http://www.linkeddatatools.com/semantic-web-basics [↑](#footnote-ref-5)
6. <http://inspire.ec.europa.eu/> [↑](#footnote-ref-6)
7. http://inspire.ec.europa.eu/documents/Data\_Specifications/INSPIRE\_DataSpecification\_GN\_v3.1.pdf [↑](#footnote-ref-7)
8. http://inspire.ec.europa.eu/documents/Data\_Specifications/INSPIRE\_DataSpecification\_AU\_v3.1.pdf [↑](#footnote-ref-8)
9. [http://rdf.kartverket.no/ont/inspire/au#](http://rdf.kartverket.no/ont/inspire/au) [http://rdf.kartverket.no/ont/inspire/gn#](http://rdf.kartverket.no/ont/inspire/gn) [↑](#footnote-ref-9)
10. http://www.opengeospatial.org/standards/geosparql [↑](#footnote-ref-10)
11. http://parliament.semwebcentral.org/ [↑](#footnote-ref-11)