



<i>Kartverk</i>	<i>Rapport</i>	<i>1</i>
Dokument tittel: GeoJSON i Kartverket	Side 1 av 11	
Dokument nr.:	Versjon: 1	
Filident: GeoJSON i Kartverket	Dato: 2020-05-20	



GeoJSON i Kartverket

En rapport som utreder hvilke type geodata som egner seg for å gjøres tilgjengelig for brukerne som GeoJSON.

Innholdsfortegnelse

Innhold

1. Introduksjon	2
2. Hva er GeoJSON	3
2.1 Kort beskrivelse	3
2.2 GeoJSON og Geonorge	3
2.3 GeoJSON validering	3
3. Brukergruppen og brukerdata angående GeoJSON.....	4
4. GeoJSON og krav til koding i Geodataloven	4
5. Teknologiske vurderinger – hvilke data egner GeoJSON seg for	5
5.1 Generelle betraktninger	5
5.2 Skjema konverteringsregler (Schema conversion rules).....	5
5.3 Kodingsregler for instanser (Instance Encoding rules).....	7
5.4 Tematiske kodingsregler (Theme encoding rules).....	7
6. Konklusjoner:	8
6.1 Anbefalinger	8
6.2 Frarådinge	9
6.3 Anbefalinger om framtidige aktiviteter	9

1. Introduksjon

Kartverket får flere henvendelser vedrørende leveranser av data som GeoJSON. Det har imidlertid vært en diskusjon om hvilke data som egner seg for å bli publisert som GeoJSON.

Som et ledd i denne avklaringen ble det nedsatt en arbeidsgruppe som skulle lage er rapport om bruk eller mulig bruk av GeoJSON i Kartverket.

Denne rapporten bygger i stor grad på konklusjonene i «INSPIRE action 2017.2 – on alternative encodings», se <https://github.com/INSPIRE-MIF/2017.2>.

Selv om denne utredningen er gjort med tanke på Kartverket vil mange av vurderingen også være av interesse for andre parter, av denne grunn gjøres rapporten tilgjengelig på GeoNorge.

2. Hva er GeoJSON

2.1 Kort beskrivelse

JSON (JavaScript Object Notation) er utvidet med en "geo" modul for å håndtere stedfestet informasjon. Spesifikasjon finnes her: <https://geojson.org/>

GeoJSON kan lett tas i bruk av ikke-geospesifikke applikasjoner. Det har mye støtte bl.a. i JavaScript og Python, som bidrar til popularitet sammenlignet med GML. Den har stor utbredelse i web og desktop klient programvare, slik som ArcMap, QGIS, LeafletJS og OpenLayers.

Testing av INSPIRE data i ulike klienter er tilgjengelig i form av en rapport fra JRC, se <https://inspire-mif.github.io/caniuse/generator/out.html>. Siste del av rapporten omhandler GeoJSON. Forklaring til selve testen av GeoJSON og testdatasett er tilgjengelig på <https://inspire-mif.github.io/caniuse/docs/geoJSON.html>

Mye av poenget med GeoJSON er at det skal være "schema"-løst. Data skal være så enkelt som mulig for brukere å ta tak i. Derfor har GeoJSON spesifikasjonen (IETF RFC 7946) noen få, ganske enkle, men bestemte regler. Den mest omdiskuterte regelen internt i kartverket er restriksjonen på bruk av referansesystem, hvor standarden angir WGS 84 med lengde og bredde i desimalgrader som default CRS. <https://tools.ietf.org/html/rfc7946#section-4>. Dette tilsvarer (OGC) URN `urn:ogc:def:crs:OGC::CRS84`.

[Det er også vært å merke seg at bruk av alternative koordinatreferansesystemer var spesifisert i en tidligere versjon av GeoJSON fra 2016. Dette har man i denne versjonen gått bort fra med utgangspunkt i at en erfarte interoperabilitetsproblemer. GeoJSON implementasjoner har i liten grad tilgang til registre over koordinatreferansesystemer og forventet WGS 84. Dette er imidlertid ikke til hinder for å bruke andre CRS i form av spesielle avtaler/leveranser.]

2.2 GeoJSON og Geonorge

Geonorge leverer i dag data på mer enn 20 forskjellige formater, deriblant GeoJSON. En årsak til at GeoJSON ikke benyttes i større grad er relatert til modell-kompleksitet. GeoJSON tillater, kort fortalt, bare flate modeller. i.e. modeller hvor dataene kan bli representert i sin helhet av en enkel tabell (dette er en forenklet beskrivelse). Standardleveranser fra Geonorges produksjonsløyper er GML og SOSI, som er formater som kan realiseres i henhold til SOSI produktspesifikasjoner

Geonorge leverer allerede i dag noe data på GeoJSON format, men disse er ikke nødvendigvis implementert i henhold til GeoJSON spesifikasjonen. Spesifikt er det mange filer som ikke holder seg til WGS-84 koordinater.

2.3 GeoJSON validering

SOSI og GML realiseringene støttes blant annet via validering. SOSI formatet bruker SOSI-kontroll og GML formatet bruker XSD applikasjonskjema, men vi har ikke noen måte å gjøre dette for GeoJSON per i dag. Samtidig er det usikkert om den samme type validering er nødvendig for GeoJSON. Realiseringsregler er nyttig for å definere en datastruktur, men er det nødvendig å validere filens datastruktur eller datainnhold i etterkant av fil-produksjonen, så lenge bruker har blitt informert at filen ikke representerer datasettets fulle struktur i henhold til produktets spesifikasjon?. Er det nødvendig for mottaker av et datasett å kjenne strukturen på et datasett og

eventuelt validere om det er i henhold til en angitt datastruktur ved mottak og eventuell prosessering av et datasett

Det har tidligere vært diskutert å lage realiseringsregler for GeoJSON utfra en SOSI-produktspesifikasjon, og det har vært foreslått å lage dette i form av SOSI del 1 – Realisering i GeoJSON. Det bør også tydeliggjøres for bruker at GeoJSON nedlastingen ikke nødvendigvis følger den opprinnelige produktspesifikasjonen, men en delmengde av dette. Et slikt forenklet UML applikasjonsskjema som utgangspunkt for realisering i GeoJSON må dokumenteres.

3. Brukergruppen og brukerdata angående GeoJSON

Hvem er brukere av geoJSON? Hendelser håndtert av CRM systemet, og det faktum at GeoJSON formatet er støttet av mange ikke-geo domenespesifikke klienter og programpakker, gir en indikasjon på at GeoJSON brukere ikke er domene eksperter. Som hintet til i seksjonen GeoJSON og Geonorge, kan det være lite motivasjon fra Geonorge eller dataeier sin side å implementere en delmengde i henhold til produktspesifikasjonen. Derfor er det naturlig at forespørselen om data på GeoJSON formatet også kommer fra et annet sted i kartverket, for eksempel Formidlingstjenesten, som har mulighet til å aggregere tilbakemeldinger fra brukerne, og prioritere forespørslene til Geonorge/IT-utvikling ang. hvilke datasett som burde produseres/endres.

Om brukergruppen faller utenfor geomatikkfag domenet, kan det spekuleres i at behovet er en forenklet dataleveranse, hvor GeoJSON stiller som leveringsformat. I disse tilfellene kan en bruker trolig klare seg med en dataleveranse basert på en forenklet datamodell, og kan leveres i tillegg til formater som tilbyr data i henhold til den fulle datamodellen

4. GeoJSON og krav til koding i Geodataloven

KOMMISJONSFORORDNING (EU) nr. 1089/2010 av 23. november 2010 om gjennomføring av europaparlaments- og rådsdirektiv 2007/2/EF med hensyn til samvirkingsevnen til geodatasett og –tjenester, Artikkel 7 gir følgende regler for koding av geodata:

1. Alle kodingsregler som anvendes for å kode geodata, skal være i samsvar med EN ISO 19118 (Geografisk informasjon - koderegler). Reglene skal særlig omfatte regler for skjemakonvertering for alle geografiske objekttyper og alle attributter og assosiasjonsroller samt den strukturen for utdata som anvendes.
2. Alle kodingsregler som anvendes for å kode geodata, skal gjøres tilgjengelige.

Selv om GML ikke er beskrevet som et krav i Geodataloven er dette et format som kan benyttes for utveksling av alle data i henhold til de harmoniserte modellene, og oppfattes som default koding.

I regi av JRC (Joint Research Centre of the European Commission) er det vurdert ulike typer koding, men med fokus på GeoJSON.

Her skiller en mellom alternativ koding og ytterligere (additional) koding.

Alternativ koding

Tilfredsstill alle krav i modellen, i vårt tilfelle produktspesifikasjonen

Ytterligere koding	Trenger ikke å tilfredsstille alle krav i modellen, kan innføre forflatning eller utelate de deler av et produkt som det ikke finnes kodingsregler for. Her bør det være dokumentasjonsplikt på hva en faktisk får av data.
--------------------	---

INSPIRE modellene er basert på internasjonale standarder. De samme standarder ligger til grunn for SOSI del 1 – Regler for UML modellering.

Følgelig vil de teknologiske utfordringer, begrensninger og koderegler spesifisert i INSPIRE prosjektet også være et godt utgangspunkt for oss.

[Andre alternative formater som det også jobbes med er «simplified GML», database formater (geopackage, PostGIS, ESRI geodatabase) og Linked data.]

5. Teknologiske vurderinger – hvilke data egner GeoJSON seg for

Dette kapitlet er i hovedsak basert på JRC's vurdering av alternativ koding. En utfyllende rapport finnes på <https://github.com/INSPIRE-MIF/2017.2/blob/master/GeoJSON/geojson-encoding-rule.md>. Det som gjentas her er et utplukk. En analyse av bruken av INSPIRE data er tilgjengelig i [2017.3 D2.2 Usability of INSPIRE.pdf](#). Kapittel 2.2 fokuserer på GeoJSON som alternativ koding til INSPIRE data og utfordringer med dette.

5.1 Generelle betraktninger

Kodingsregler (eller realiseringsregler) for GeoJSON må ta i betraktning noen generelle utfordringer, slik som:

- De fleste GIS applikasjoner kan ikke fullt ut handtere komplekse (non-simple) egenskaper og nøstede strukturer.
- Multiple verdier for egenskaper kan ikke benyttes direkte i de fleste vanlige GIS-verktøy.
- Referanser til andre objekttyper håndteres ofte ikke i GIS programvare.
- Abstrakte geometrier (eks GM_Object) innebærer at en rekke segmenttyper og interpolasjoner kan brukes for de respektive objekttypene.
- Flere geometrityper i en "FeatureCollection" er vanligvis ikke støttet.

Kodingsregler kan inndeles i følgende nivåer:

- Skjema konverteringsregler (Schema conversion rules)
- Kodingsregler for instanser (Instance Encoding rules)
- Tematiske kodingsregler (Theme encoding rules)

5.2 Skjema konverteringsregler (Schema conversion rules)

På samme måte som vi for GML konverterer fra UML til XML skjema, kan vi for GeoJSON konvertere til JSON skjema. Men JSON skjema er lite brukt og GeoJSON oppfattes som skjema-løst.

Det forutsettes at dataene er i overensstemmelse med en UML modell i henhold til SOSI del 1 – Regler for UML modellering.

Objekttyper

Alle typer som har stereotypen <featureType> konverteres til GeoJSON objekt, navn beholdes.

Egenskapstyper

Alle egenskapstyper konverteres i henhold til følgende tabell:

UML Model property type	JSON datatype	Merknad
CharacterString	string	
LocalisedCharacterString	string	LanguageCode legges til som en egen egenskap.
Boolean	boolean	
Integer	integer	
Real	number	
Decimal	number	
DateTime	string	Datatypesen må følge ISO 8601 format, inkludert tidssone informasjon (e.g. 2008-10-31T15:07:38-05:00).
Date	string	Datatypesen må følge formatet yyyy-mm-dd.
Length	double	uom legges til som en egen egenskap.
Measure	double	uom legges til som en egen egenskap.
URI	string	

Andre UML egenskaps typer konverteres til <string>, med spesielle regler definert case by case.

Geometrityper:

ISO 19107 typer	GeoJSON datatype	Merknad
GM_Aggregate	GeometryCollection	Begrensninger på hvilke typer som kan inngå i samlingen.
GM_Curve	LineString	I GML kan kurver også representeres som ikke-lineære segmenter eller buer. GeoJSON støtter bare lineære segmenter.
GM_MultiCurve	MultiLineString	I GML kan kurver også representeres som ikke-lineære segmenter eller buer. GeoJSON støtter bare lineære segmenter.
GM_MultiPoint	MultiPoint	
GM_MultiPrimitive	Ikke støttet	GM_MultiPrimitive er en abstrakt type.
GM_MultiSurface	MultiPolygon	
GM_Object	Any GeoJSON Geometry type	GM_Object er en abstrakt type, enhver GeoJSON geometri kan benyttes.
GM_Point	Point	
GM_PolyhedralSurface	Ikke støttet	3D meshes støttes ikke
GM_Primitive	Any GeoJSON Geometry type	GM_Object er en abstrakt type, enhver GeoJSON geometri kan benyttes.
GM_Surface	Polygon	En GM_Surface kan ha mange SurfacePatches, mappes til MultiPolygon.
GM_Tin	Ikke støttet	En TIN uten trinagulering kan konverteres til et MultiPoint object.
GM_Triangle	Polygon	

Temporale typer

I Kartverket bruker vi kun tid som en tematisk (ikke temporal) egenskap. Følgelig ingen utfordring.

Metadataelementer

Vi bruker i liten grad (kanskje overhodet ikke) metadata fra ISO 19115-1 i våre applikasjonskjemaer. Men må i så fall løses "case by case".

Union

"Case by case"

Kodelister og enumererte typer

Ingen mapping påkrevd. Verdier identifiseres gjennom (resolvable) http URI's.

Arrays

Egenskapstyper for egenskaper med kardinalitet større enn 1 kan benytte arrays.

Assosiasjonsroller

Assosiasjonsroller kan beholdes slik de foreligger.

5.3 Kodingsregler for instanser (Instance Encoding rules)

Ytterligere regler for avledning fra den konverterte konseptuelle modellen.

Karaktersett

Karaktersett for all koding i GeoJSON skal være UTF-8.

Koordinatreferansesystem (CRS)

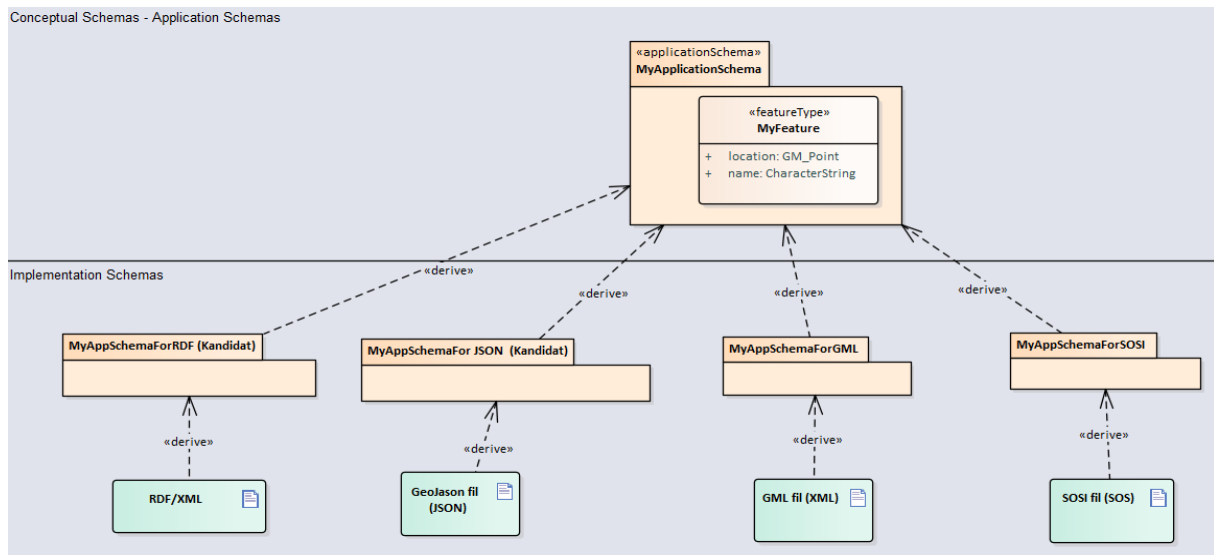
Koordinatreferansesystem for GeoJSON er WGS84 med lengde og bredde i desimalgrader. Dette oppfattes som både en stor hindring med tanke på mer presise koordinatsystemer men er også en stor fordel med tanke på å holde det enkelt, som igjen er ett av grunnprinsippene i GeoJSON. Utfordringene rundt dette er godt beskrevet i introduksjonen tidligere i dokumentet.

Identifikatorer

Default Encoding property	Property Type	GeoJSON Property	Property Type	Conversion Notes
<code>gml:id</code>	ID	<code>id</code>	string	
<code>gml:identifier</code>	CodeWithAuthorityType	<code>properties.identifier</code>	string	Can be omitted if empty in the source object.
<code>base:inspireId</code>	IdentifierPropertyType	<code>properties.inspireId_localId</code>	string	SOSI: LokalId
		<code>properties.inspireId_namespace</code>	string	SOSI: Navnerom
		<code>properties.inspireId_versionId</code>	string	SOSI: VersjonId (valgfritt)

5.4 Tematiske kodingsregler (Theme encoding rules)

De tematiske kodingsreglene kan lages for hvert applikasjonsskjema der det ikke er mulig å representere hele modellen i GeoJSON. I utgangspunktet går disse på å tilpasse UML applikasjonsskjema til det som kan representeres i GeoJSON.



Denne tilpasningen er ikke spesiell for GeoJSON. Med unntak av GML (som inneholder regler for mapping fra ethvert applikasjonskjema til GML skjema), vil vi også ha samme utfordring med f.eks SOSI formatet. UML applikasjonskjemaet må da tilpasses det som kan realiseres i GeoJSON. Denne tilpasningen bør dokumenteres.

Generelle regler for forenkling av modeller slik at de kan realiseres i GeoJSON

Disse forenklingsreglene er fremkommet gjennom JRC's arbeid med å gjøre data tilgjengelige i form av alternativ koding, slik som GeoJSON.

Her presenteres bare konseptene med en referanse til ytterligere beskrivelse.

- Forflattning av komplekse strukturer – se [General Flattening of nested structures](#)
- Bruk av "arrays" – se [Extract Primitive Array](#)
- Assosiasjon/aggregering til komposisjon med "Hard Typing" – se [Association/Aggregation to Composition with Hard Typing](#)
- Assosiasjon/aggregering til komposisjon med "Soft Typing" – se [Association/Aggregation to Composition with Hard Typing](#)
- Restriksjon av egenskapskoding til "only by – reference", se [Property Composition to Association](#)
- Referanse (Citation), se [Simple Citation](#)
- Kodelister – alternativ struktur for spesielle typer eller typer hierarkier, se [Simple Codelist Reference](#)
- Forenklet beskrivelse av periode, se [Simple Period](#)

6. Konklusjoner:

6.1 Anbefalinger

GeoJSON er et godt alternativ til GML og SOSI for enkle modeller/data for brukere av web og desktop klient programvare, slik som ArcMap, QGIS, Leaflet og OpenLayers.

6.2 Fraråding

Generelt frarådes å gjøre data generelt nedlastbart som GeoJSON med andre CRS enn WGS84 og tro at brukerne er innforstått med dette. For noen typer data, hvor forskjellen mellom f.eks EUREF 89 og WGS 84 ikke har noen praktisk betydning kan dette likevel vurderes, dette er en vurdering som de respektive dataeiere må ta.

Det var ulike oppfatninger om denne frarådingen. Det ble foreslått at der det fremkommer i nedlastingsløsningen og/eller i GeoJSON fila at koordinatsystemet er annerledes enn WGS84 er dette greit. Men det ble også fremhevet at når GeoJSON benyttes utenfor vår Geomatikk bransje (f.eks i eGovernment løsninger) vil dette kunne gi interoperabilitetsutfordringer.

Selv om GeoJSON har en enkel struktur og er å betrakte som skjematøst har ulike web og desktop klient programvare ulike støtte av GeoJSON funksjonalitet. Foreløpige tester viser at f.eks håndtering av flere objekttyper i samme GeoJSON fil ikke støttes i ArcGIS Pro v2.3, ArcMap v10.7 og QGIS v3.4. Egenskaper fra ulike objekttyper samles, dvs at alle objekttyper får samme struktur. Tilsvarende gjelder også OpenLayers og Leaflet. Noen klientverktøy har begrensninger på filstørrelse, andre har utfordringer med å takle "Array's".

6.3 Anbefalinger om framtidige aktiviteter

For å sikre en omforent måte å bruke GeoJSON i den nasjonale geografiske infrastrukturen kan det formaliseres egne mapping- regler fra UML modeller til GeoJSON. Disse mappingreglene kan også utvides med forflatningsregler for UML applikasjonsskjema spesielt tilpasses GeoJSON. En slik formalisert beskrivelse foreslås utgitt som en ny standard, SOSI del 1 – Realisering i GeoJSON.

Det anbefales også at det tas en vurdering av GeoJSON i det nye QMS-API'et ut fra de retningslinjer som ligger i dette dokumentet.

Tilleggspunkt etter siste møte.

Før vi går videre med en eventuell implementasjon av dataleveranser på GeoJSON (f.eks. via Geonorge), bør vi gjøre en utvidet kartlegging av brukergruppen, for å inkludere brukergruppens behov i våre beslutninger. Vi bør identifisere hvilke brukergrupper vi har, som faller utenfor det tradisjonelle GIS/geomatikk miljøet. Eksempler på disse kan være web-utviklere som ønsker små filer for rask opplasting/ lett håndtering i nett-klient, studenter som ønsker enkle kartvisualiseringer eller journaliser som skal vise en hendelse i et kart. Vi kan begynne med å se på data samlet inn fra Hack4no, og igjennom CRM systemet. I tillegg bør statistikk over bruk av formater og tjenester tas til etterretning.

VEDLEGG

Skjema for vurdering av om datasett er egnet for distribusjon som GeoJSON

Prosjekt: Utredning av GeoJSON i
Kartverket - GeoJSON - IETF RFC 7946

Hensikten med dette regnearket er å se hvilke typer data som er tenkt representert i UML og i hvilke sammenheng, samt om det er kompleksitet i dataene som vil være en utfordring for GeoJSON

Vurderingskriterie	Datasett
Formål med GeoJSON	
Alternativt format (skal kunne representere hele produktet)	
Tilleggsformat (representere deler av et produkt)	
Dersom tillegg, er det greit om deler av produktet ikke lar seg realisere	
Desktop/klient som ligger til grunn for ønske om GeoJSON (om noen)	
Leaflet	
OpenLayers	
ArcMap	
QGIS	
Phyton	
Andre (forklar)	
Brukergruppen	
Profesjonelle (eks: Norge digitalt parter)	
Sluttbrukere utenfor geomatikk miljøet	
Internt i produksjon og forvaltningssammenheng	
Andre (forklar)	
Utvekslingsbehov	
Bare ut fra distribusjonsløsningen	
Tilbake til forvaltningsløsningen	
Internt i forvaltningssammenheng	
Er dataene modellert (Ja / Nei)	
KODEREGLER FRA UML TIL GEOJSON (bare dersom data er modellert)	
Hva slags geometrityper benyttes	

GM_Aggregate	
GM_Curve (SOSI:LINJE,KURVE)	
GM_MultiCurve	
GM_MultiPoint (SOSI:SVERM)	
GM_MultiSurface	
GM_Point (SOSI: PUNKT)	
GM_PolyhedralSurface	
GM_Surface (SOSI:FLATE)	
GM_Tin	
GM_Triangle	
Benyttes modellelementet UNION(JA/NEI)	
Har modellen assosiasjoner (Ja/Nei)	
Har modellen egenskaper med multiplisitet > 1 (Ja/Nei)	
Koordinatreferansesystem	
Er forskjellen mellom WGS84 og EUREF89 signifikant for produktet	
Målestokk	
Brukergruppen	
Profesjonelle (eks: Norge digitalt parter)	
Sluttbrukere utenfor geomatikk miljøet	
Internt i produksjon og forvaltningssammenheng	
Andre (forklar)	
Utvekslingsbehov	
Bare ut fra distribusjonsløsningen	
Tilbake til forvaltningsløsningen	