Datafordeler

Dataleverancespecifikation

Version 1.7

2015 09 16



# Versions historik

|  |  |  |
| --- | --- | --- |
| **Dato** | **Version** | **Bemærkning** |
| 19-08-2014 | 1.0 | Document was created |
| 20-08-2014 | 1.1 | Struktur og overskrifter /CPR |
| 22-08-2014 | 1.2 | First draft indhold /CPR |
| 19-09-2014 | 1.3 | Uge 39 version, rettet bl.a. p.b.a. GST/DIGST review |
| 29-09-2014 | 1.4 | Uge 40 version tilrettet udvikling i løsningsbeskrivelsen /JAF |
| 07-11-2014  13-11-2014 | 1.5  0.6 | Tilrettet efter internt review /JAF  Tilrettet af GST/DIGST |
| 28-11-2014 | 0.7 | Tilrettet af KMD efter møde med GST/DIGST |
| 04-12-2014 | 0.8 | Tilrettet af KMD efter afklaringer med GST/DIGST |
| 17-12-2014 | 0.9 | Tilrettet af KMD efter møde mellem GST og KMD |
| 16-01-2015 | 0.95 | Udvidet med Geo specifikke informationer og navn ændret til dataleverancespecifikation |
| 02-02-2015 | 0.96 | Udvidet med hændelses specifik information samt sikkerhedsbeskrivelse og metadata |
| 26-02-2015 | 1.0 | Kosmetiske rettelser og ophøjet til version 1.0 |
| 06-03-2015 | 1.1 | Mulighed for indberetning af hændelses formater m. mapning, samt tilpasning af bilags nummereringer. |
| 09-03-2015 | 1.2 | Fjernet mulighed for nær-realtids pull jvf. løsningsbeskrivelsen. |
| 01-07-2015 | 1.3 | Tilføjet beskrivelse af at der kan medsendes yderlige oplysninger i ajourføringer af Grunddata, som kan benyttes i dannelse af hændelsesbeskeder. |
| 17-07-2015 | 1.4 | Splittet XSD i to dele. Datafordeler specifik del og dataleverandør specifik del. |
| 10-08-2015 | 1.6 | Tilpasning af versionering på tværs af DLS dokumenterne. |
| 16-09-2015 | 1.7 | Indføjelse af navne konvention for DLS-dokumenter samt nulstillet ændrings markering.  Tilpasning af zipnings krav ved forskellige datasæt typer.  Metadata afsnit opdateret.  Afsnit om ’ SSH nøgler til SFTP’, ’Kald af tjenester med relation til replikeringkanalerne’ samt ’DLS ændringslog’ tilføjet.  Mindre rettelser og trykfejl. |
|  |  |  |

## DLS ændringslog

Siden version 1.6 er der fortaget følgende ændringer:

1. Dataleverancespecifikation - skabelon.docx er opdateret. Bl.a. er afsnit om metadata forbedret og der er indført navnekonvention for DLS-dokumenter.
2. Bilag 2 - Tjeneste skabelon.xlsx er opdateret med én fane for metadata. REST skabelonen er medtaget som et faneblad i dette dokument. REST skabelonen er blevet forbedret, således at der nu er en systematisk måde at angive parametre.
3. Bilag 2b - REST tjeneste skabelon.xlsx udgår. I stedet benyttes REST type i Bilag 2.
4. Bilag 4 - Sikkerhedsbeskrivelse.docx information om krigsreglen fjernet samt mulighed for at benytte ’ALLE’.
5. Bilag 5 - Metadata for udstillingsmodel.xlsx afløser det tidligere ’Bilag 5 - Metadata.docx’
6. Bilag 19 - Tjenesteoversigt.xlsx erstatter det tidligere ’Bilag 19 - Tjeneste oversigt.docx’

Indholdsfortegnelse

[Versions historik 2](#_Toc430780776)

[DLS ændringslog 3](#_Toc430780777)

[Indledning 6](#_Toc430780778)

[Om dette dokument 6](#_Toc430780779)

[Aktører 6](#_Toc430780780)

[Dokument versionering 6](#_Toc430780781)

[Afsnit 1 – Process for ny leverance 8](#_Toc430780782)

[Hvad skal der ske? 8](#_Toc430780783)

[Afsnit 2 – Replikeringskanalen 9](#_Toc430780784)

[Replikeringsmønster 9](#_Toc430780785)

[Batch-Port 11](#_Toc430780786)

[SOA-Port 11](#_Toc430780787)

[Dataudtræk 12](#_Toc430780788)

[Driftsforstyrrelser 12](#_Toc430780789)

[Databeskrivelse 12](#_Toc430780790)

[Afsnit 3 – Dataleverancen 14](#_Toc430780791)

[Pakker og Sekvensmekanisme 14](#_Toc430780792)

[Filopbygning og Sanity Check 15](#_Toc430780793)

[XML leverancer 15](#_Toc430780794)

[Rasterdata 18](#_Toc430780795)

[Punktsky-data 22](#_Toc430780796)

[Kvittering for dataleverancer 23](#_Toc430780797)

[Afsnit 4 – Synkronisering 25](#_Toc430780798)

[Om Synkronisering på denne Replikeringskanal 25](#_Toc430780799)

[Proaktiv Synkronisering 25](#_Toc430780800)

[Ekstraordinær Synkronisering 26](#_Toc430780801)

[Afsnit 5 – Tjenester 27](#_Toc430780802)

[Generelt 27](#_Toc430780803)

[Navne konvention i DLS dokumenter 27](#_Toc430780804)

[Afsnit 6 – Metadata 28](#_Toc430780805)

[Afsnit 7 – Sikkerhedsbeskrivelse 30](#_Toc430780806)

[Sikkerhedsbeskrivelse 30](#_Toc430780807)

[Afsnit 8 – Hændelsesbeskeder 31](#_Toc430780808)

[Datanære hændelsesbeskeder 31](#_Toc430780809)

[Forretningsmæssige hændelsesbeskeder 31](#_Toc430780810)

[Afsnit 9 – Diverse oplysninger 33](#_Toc430780811)

[Kontaktoplysninger 33](#_Toc430780812)

[Incidents 33](#_Toc430780813)

[End-points, FTP-adresser mv. 33](#_Toc430780814)

[SSH nøgler til SFTP 34](#_Toc430780815)

[Kald af tjenester med relation til replikeringkanalerne 34](#_Toc430780816)

[Oversigt over Bilag til denne Dataleverancespecifikation 34](#_Toc430780817)

# Indledning

### Om dette dokument

Dette dokument fastlægger en specificering af data leverance fra en Dataleverandør til Datafordeleren.

Denne specificeringsaftale omhandler leverance af: DAR  
Udfyldt af: Netcompany dato: 2016-06-22

Aftalen omfatter de forhold, der skal afklares for at kunne etablere integration mellem Dataleverandørens register og Datafordeleren og implementere udstilling af dataleverandørens data på Datafordeleren.

Dette omfatter blandt andet valg om leveringsmønstre (pull/push), angivelse af hvilke typer data der indgår i leverancen, hvilke mekanismer der ønskes anvendt til sikring af dataintegritet i overførslen og hvorvidt der ønskes at anvende både filleverancer og webservices til leverancer eller kun filleverancer.

Endelig omfatter specifikationens bilag specifikationer af de data leverancen indeholder og de tjenester, som skal udstilles på Datafordeleren.

### Aktører

Denne specifikation er til brug for de tre parter:

* **Dataleverandøren** er den organisation, der leverer data til udstilling på Datafordeleren.
* **Leverandøren** er KMD, som er ansvarlige for udvikling og drift af Datafordeleren. Leverandøren er også Single Point of Contact (SPoC) for Datafordelerens dataanvendere.
* **Operatøren** er den enhed i Geodatastyrelsen, som varetager kunderollen overfor Leverandøren. Operatøren har den daglige kontakt med Leverandøren og alle bestillinger af ændret eller ny funktionalitet som skal implementeres af Leverandøren skal gå igennem Operatøren.

Alle 3 aktører kan anfordre total-udtræk og ekstraordinære synkroniseringer når de konstatere et behov, men det er udelukkende dataleverandøren der kan tage beslutningen om det i praksis skal gennemføres, og i så fald indenfor hvilken tids-horisont, da sådanne total-udtræk kan have indflydelse på dataleverandørens produktions systemer.

### Dokument versionering

Af hensyn til klarhed over oprindelse af dokumentation, samt versioner ved gentagne leverancer, skal følgende navngivning anvendes når de udfyldte skabeloner indsendes:

”REGISTER NAVN”\_ VERSION\_DATO\_ ”DOKUMENTNAVN”

Eksempel: DAGI\_v1.0\_2015.02.15\_Bilag 2\_Tjeneste skabelon.doc

Hvad "Register Navn" angår, så vær opmærksom på at det i datafordeler sammenhænge er navnet for hvert datasæt der indsendes på egen replikeringskanal. Dvs. at der kan forekomme adskille forskellige datafordeler "registre" som hos dataleverandøren tilhører samme register. Dette forekommer når dataleverandøren af praktiske grunde ønsker at opdele leverancerne til datafordeleren i forskellige delmængder.

I dokumentationsnavngivningen er det derfor væsentligt at det er delmængdens navn der anvendes, og ikke dataleverandørens interne betegnelse.

# Afsnit 1 – Process for ny leverance

### Hvad skal der ske?

1. Opstart  
   Leverandøren fremsender dataleverancespecifikations-skabelonen til Dataleverandøren.
2. Udfyld skabelon

Dataleverandøren udfylder de relevante dele af dataleverancespecifikationen. Bemærk, at der er information, der efterfølgende vil blive udfyldt af Leverandøren eller Operatøren.

1. Bilag 1-5 udarbejdes.

Dataleverandøren udarbejder de bilag, der omhandler datamodeller, tjenester, hændelser, sikkerhed og metadata.

1. Materialet sendes til Operatøren til QA.  
   Den udfyldte dataleverancespecifikation samt udfyldte bilag sendes til info@datafordeler.dk.
2. Aftalen registreres.  
   Aftalen bliver efter QA videresendt af Operatøren til KMD. Under denne proces vil de udestående informationer blive indsat i dataleverancespecifikationen hvorefter den vil blive returneret til Dataleverandøren. Dataleverancespecifikationen vil blive returneret til Dataleverandøren med de af leverandøren udfyldte felter inden den egentlige behandling i testmiljøet startes, således at dataleverandøren har relevante oplysninger om replikerings kanals identifikation, ftp-sites og kontakt informationer så tidligt som muligt i forløbet.
3. Opsætning i testmiljø.  
   På baggrund af dataleverancespecifikationen vil indlæsningsdatamodellen, udstillingsdatamodellerne, tjenester og hændelser blive implementeret af Leverandøren og herefter stillet til rådighed for test på Datafordeleren.
4. Testperiode  
   I en periode vil Leverandøren afprøve de forskellige elementer i henhold til kontrakten om Datafordeleren. Sideløbende med og uafhængigt af dette kan Dataleverandøren foretage sine egne afprøvninger af eksempelvis integrationer til andre systemer.
5. Idriftsættelse.  
   Når Leverandørens afprøvninger er afsluttet succesfuldt, sættes replikeringskanalen og de implementerede udstillingstjenester i drift. Produktionssystemet opsættes, og den initiale indlæsning af data foretages. I denne forbindelse vil Leverandøren kontakte Dataleverandøren for udveksling af login informationer til ftp sites.

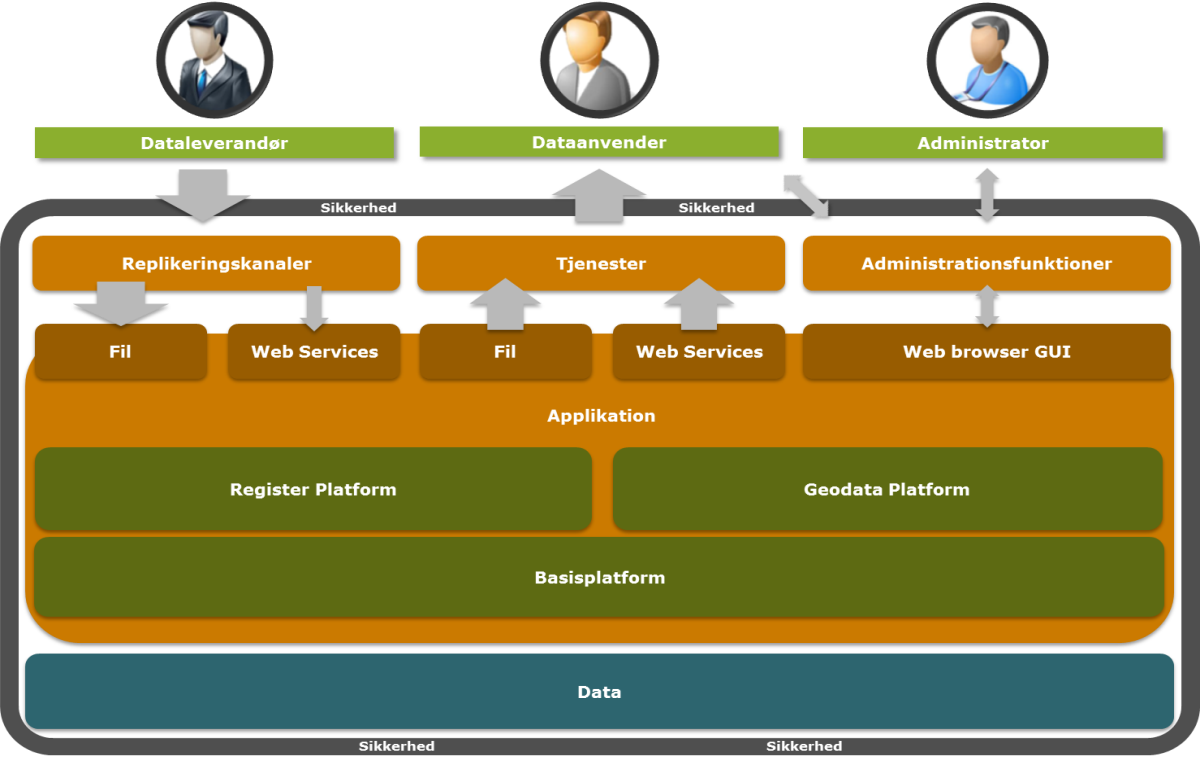
# Afsnit 2 – Replikeringskanalen

Den tekniske hjørnesten under dataleverancen er i Datafordelerens begrebsverden en *replikeringskanal*.

Replikeringskanalen er en samlet betegnelse for de tekniske snitflader mellem Dataleverandørens register og Datafordeleren. Dette omfatter blandt andet mulighed for overførsel af filer, en tjeneste til opslag af, at overførsler er gået godt samt evt. web services til nærrealtids-opdatering af enkelte forretningsobjekter fra registeret.

Replikeringskanalen er specifik for den enkelte dataleverandør, således, at den kan indeholde kontroller for integriteten af det pågældende registers konkrete data.

Figur 1: Replikeringskanalernes placering i Datafordelerens arkitektur



For at gøre det muligt at opsætte kommunikationen imellem Dataleverandøren og Datafordeleren, skal der træffes en række beslutninger om replikeringskanalen, idet Datafordeleren som en fleksibel platform stiller flere såkaldte replikeringsmønstre til rådighed.

Replikeringskanalen har det unikke ID: 22

Replikeringskanalens ID angives ved alle opdateringer og synkroniseringsaktiviteter.

### Replikeringsmønster

En replikeringskanal kan sættes op til at bruge to forskellige replikeringsmønstre for opdateringer: Push og Pull.

Push betyder at Dataleverandøren uploader dataleverancer til Datafordeleren, mens Pull betyder, at Datafordeleren trækker data fra registeret.

Indlæsning af data påbegyndes i begge tilfælde så snart de er fuldt modtaget i Datafordeleren og alle tidligere opdateringer fra Dataleverandøren er behandlet.

Konkret er Leverandøren forpligtet til at påbegynde indlæsning af batch-opdateringer senest 5 minutter efter, at opdateringsfilen er fuldt modtaget af Datafordeleren.

For nærrealtids-opdateringer skal Leverandøren sikre, at de opdaterede data er udstillet på Datafordeleren senest 8 sekunder efter modtagelsen.

#### Push-mønsteret

Med Push-mønsteret er det Dataleverandøren, der tager initiativ til opdatering af data på Datafordeleren. Det er op til Dataleverandøren, om man vil gøre dette periodisk – eksempelvis dagligt eller en gang i timen – eller ad hoc når der er sket ændringer i registeret. Dermed giver Push-mønsteret mulighed for at ændringer i data hos Dataleverandøren kan fremsendes til og udstilles fra Datafordeleren i nærrealtid.

Hvis Push mønstervælges, skal Dataleverandøren etablere funktionalitet i sine systemer, som opdaterer data på Datafordeleren via de servere og/eller tjenester, der stilles til rådighed af Datafordeleren.

#### Pull-mønsteret

Pull betyder at Dataleverandøren udstiller sine dataleverancer på sine egne servere hvorfra Datafordeleren så henter dem.

Med Pull-mønsteret er der ikke mulighed for ad hoc-opdateringer, da Datafordeleren ikke kan vide, hvornår der sker ændringer i registeret, men intervallet mellem Datafordelerens forespørgsler kan sættes så lavt, at effekten kommer til at minde om – eksempelvis ved opslag hvert minut.

Hvis Pull mønster er valgt skal der tages stilling til, med hvilken frekvens Datafordeleren skal hente data fra dataleverandøren.

Desuden kan der angives et tidsinterval for hentning og eksekvering. Det vil sige, at Dataleverandøren kan specificere et tidsrum, hvor Datafordeleren må hente dataleverancer fra registeret. Datafordeleren vil så ikke tilgå registeret udenfor dette interval. Tilsvarende kan der angives et tidspunkt for gennemførelse af opdatering af Datafordeleren. Dette giver dataleverandøren mulighed for at bestemme hvornår Datafordeleren skal påbegynde indlæsning af de hentede opdateringer.

Angives intet tidsinterval vil Datafordeleren med det aftalte mellemrum hente opdateringer fra dataleverandøren og så snart en leverance er fuldt modtaget og alle tidligere opdateringer er behandlet, indlæses den modtagne leverance ligegyldigt hvad tid på døgnet den hentes på.

#### Anvendt replikeringsmønster

Dataleverandøren har valgt at anvende det angivne mønster: Push [ X ] Pull [ ]

Nedenfor angives parametrene for Pull-opdatering, hvis dette mønster er valgt:

|  |  |
| --- | --- |
| Der skal forespørges efter opdateringer for hver: | \_\_\_\_ timer. |
| Evt. interval for hentning af opdateringer:  Der hentes data på de afkrydsede ugedage i det angivne interval.  Hvis der ikke afkrydses nogen ugedage, opdateres på alle dage i det angivne tidsinterval.  Hvis der ikke angives noget tidsinterval, opdateres døgnet rundt på de valgte ugedage. | Mandag: \_\_ Tirsdag: \_\_ Onsdag: \_\_  Torsdag: \_\_ Fredag: \_\_  Lørdag: \_\_ Søndag: \_\_  Fra kl.: \_\_:\_\_ Til kl.: \_\_:\_\_ |
| Evt. indlæsningstidspunkt: | Indlæsning påbegyndes kl. \_\_:\_\_ |

### Batch-Port

Der er altid en batch-port tilknyttet til en dataleverancespecifikation. Batch-porten anvendes til fremsendelse af filleverancer.

Alle tabulære datasæt der sendes via batch-porten, sendes som filer, der af hensyn til overførselstid og pladsforbrug, skal være zip'et, dvs. leveret i et .ZIP archive format der overholder "PKWARE ZIP Application Note" som anbefalet af ISO (ISO/IEC JTC 1/SC 34 N 1621). Af disse vil ZIP64 være at foretrække da dette format understøtter zip filer der er større end 4GB.

Rasterdatasæt skal fremsendes uden at være zippet som en liste af tiff filer og afslutningsvis en vrt fil.

Se ”Afsnit 9 – Diverse oplysninger” for information om end points og adgangsoplysninger for denne dataleverance.  
Hvis Dataleverandøren har valgt at bruge Pull-replikeringsmønsteret skal Dataleverandøren i ”Afsnit 9 – Diverse oplysninger” angive oplysninger om Dataleverandørens end point så Datafordeleren kan konfigureres til at hente filerne, I forbindelse med opsætning af produktions miljøet vil Leverandøren kontakte Dataleverandøren for udveksling af adgangsoplysninger.

Datafordeleren stiller en standard SFTP-klientadgang til rådighed for Dataleverandøren i relevante miljøer \*).

### SOA-Port

SOA-porten består af en web service, der kan anvendes når Dataleverandøren ønsker at foretage nærrealtids-opdateringer, det vil sige opdateringer, der skal udstilles straks og ikke afvente næste batch-opdatering.

Anvendelse af SOA porten er valgfri.

SOA porten er beregnet til enkeltopdateringer, og for at undgå overbelastning af Datafordeleren skal større datamængder forsendes som filleverancer via batch-porten.

I kontekst af nærværende dataleverancespecifikation betragtes det som større datamængder når [udfyldes af Leverandøren efter aftale med Dataleverandøren og på baggrund af specifikationerne i bilag 1-5].

Dataleverandøren skal beslutte om replikeringskanalen skal indeholde en SOA-port: Anvendes

Datafordeleren udstiller en web service (WSDL inkluderes som bilag WS-01) på relevante miljøer \*) der skal anvendes ved SOA-port-leverancer.

Såfremt Dataleverandøren ønsker mulighed for at indlevere data på SOA-porten, skal Dataleverandøren selv implementere kald af denne Web Service.

### Dataudtræk

Dataleverandøren skal udtrække, pakke, og enten fremsende data til Datafordelerens SFTP-Server eller fremsende data gennem datafordelerens udstillede webservice RegisterData ved push mønstret.

Ved pull mønstret udstilles data på dataleverandørens egen SFTP-Server.

Reglerne for datastrukturen i leverancerne er fastlagt i Dataleverancespecifikationens afsnit 3 og den konkrete datastruktur for opdateringer under nærværende aftale er specificeret i bilag 1.

### Driftsforstyrrelser

Datafordeleren udstiller webservicen ”DeliveryDisturbance” [WSDL inkluderet som bilag WS-02] på relevante miljøer. \*)

I tilfælde af planlagte eller spontane driftsforstyrrelser skal dataleverandøren vurdere dataanvendernes behov for at vide, at datareplikeringen er på pause. Såfremt dataleverandøren skønner et sådant behov, skal dataleverandøren kalde ”DeliveryDisturbance”, med oplysninger om den forventede periode for driftsforstyrrelsen. Der vil være mulighed for at angive en kategori af omstændighederne, som Datafordeleren kan videregive til dataanvenderne via supporten. Med gentagne kald kan dataleverandøren opdatere den forventede periode for driftsforstyrrelsen såvel som kategorien.

### Databeskrivelse

Dataleverandør skal levere en *databeskrivelse* som bilag 1 til nærværende aftale. Databeskrivelsen skal indeholde:

* XMI-Datamodel for de data der skal udstilles på Datafordeleren. XMI-modellen skal indeholde klasser, attributter, typer, identifikatorer, relationer og evt. andre relevante UML-data. Hver klasse og hver attribut skal have angivet en UUID i modellen.
* Datamodel skal leveres i format XMI 2..4.1 med UML 2.X.
* Rasterdatamodeller beskrives i VRT-format.

Desuden oplyses

* Anslået initiale antal objekter af hver klasse.
  + NavngivenVej: 130.000
  + Adresse: 3.800.000
  + Adressepunkt: 18.000.000
  + Husnummer: 6.800.000
  + NavngivenVejKommunedel: 140.000
  + NavngivenVejPostnummer: 250.000
  + NavngivenVejSupplerendeBynavn: 250.000
  + Postnummer: 1.000
  + SupplerendeBynavn: 10.000
* Forventet udvikling i antal objekter af hver klasse.
  + Dette er ukendt på nuværende tidspunkt, 22.06.2016
* Overslag over hvor hyppigt, objekterne i hver klasse forventes at blive opdateret.
  + Dette er ukendt på nuværende tidspunkt, 22.06.2016

Dataleverandør skal desuden levere en *leverancebeskrivelse* som bilag 3 til nærværende aftale. Leverancebeskrivelsen skal indeholde:

* XSD for de data, der skal leveres fra registeret til Datafordeleren via replikeringskanalen. Her skal kun angives den del af forsendelsen der er specifik for leverancen, se bilag 14 for skabelonen. De generelle informationer der er gældende for alle dataleverancer er beskrevet i bilag 18.
* Hvert element i leverancebeskrivelsen skal indeholde henvisning til UUID’en for det modsvarende element i databeskrivelsen, så Datafordeleren kan behandle opdateringer mv. korrekt. Disse UUID referencer angives i XSD'et som <xs:annotation><xs:appinfo>, Se bilag 6 for eksempel på udfyldt XSD med UUID referencer.

*\*) ”Relevante miljøer” inkluderer som minimum miljøerne Produktion og kundetestområdet.*

# Afsnit 3 – Dataleverancen

## Pakker og Sekvensmekanisme

Dataleverandøren skal aflevere data i afgrænsede enheder kaldet ”pakker” til Datafordeleren. En pakke skal enten afleveres på replikeringskanalens batch port eller via replikeringskanalens SOA port.

For at sikre, at data indlæses i den rigtige kronologiske rækkefølge på datafordeleren, dvs. at en oprettelse af en entitet kommer før ændringer til samme entitet, skal dataleverandøren tildele pakkerne et fortløbende sekvensnummer der er gældende for begge af replikeringskanalens porte. Leverandøren tildeler hver Replikeringskanal et startnummer.

Denne Replikeringskanal starter med sekvensnummer 100000001.

Pakkens ID i Datafordeleren vil være [ReplikeringskanalID]-[Sekvensnummer]. Første pakke på denne Replikeringskanal skal således have pakkeID = ”RC22 -100000001”.

Pakker til Batch-porten afleveres i filer der navngives [pakkeID].zip eller [pakkeID].vrt. Som eksempel kan en Replikeringskanal starte med at aflevere filen ”RC1-100000001.zip”. Næste fil skal så hedde ”RC1-100000002.zip” osv.

Såfremt data afleveres på SOA-porten medsendes pakkeID og sekvensnummeret i inputfelter til servicen.

Et sekvensnummer må aldrig anvendes 2 gange på samme Replikeringskanal, medmindre en pakke fejlede valideringen, og der må ikke være huller i sekvensen. Dette gælder også såfremt replikeringskanalen anvender begge porte til dataleverancer: Sekvensnummereringen skal være fortløbende og entydig på tværs af portene.   
Pakker må kun fremsendes med tidligere anvendte sekvensnummer som korrektion på fejlede valideringer. Fejler en validering af en pakke skal dataleverandøren korrigere fejlen og sende pakken igen med samme sekvensnummer som første gang. Modtages pakker med tidligere anvendte sekvensnumre der oprindeligt ikke fejlede i valideringen vil disse pakker blive kasseret og altså ikke blive processet af replikeringskanalen. Status på de indsendte pakker, inklusivt hvorvidt disse har fejlet valideringen, kan ses ved kald af DeliveryReceipt servicen (WSDL i bilag WS03).

Datafordeleren bruger blandt andet sekvensmekanismen til at sikre korrekt opdateringsrækkefølge, således at filer der eventuelt overhaler hinanden i transportlaget ikke medfører fejl i opdateringerne.

Dataleverandøren har ansvaret for, at data sekvensnummereres i forretningsmæssigt korrekt rækkefølge (inserts-kommer-før-updates osv.).

Dataleverandøren skal være klar til at genfremsende en datapakke indtil Dataleverandøren via anvendelse af Datafordelerens kvitterings-tjeneste har fået bekræftet, at Datafordeleren har gennemført opdateringen.

En pakke består for registerdata af en mængde forretningsobjekter i xml-format, alle individuelt påhæftet en ajourføringskode, der angiver om der er tale om en oprettelse, en opdatering eller en sletning. Geodata, der skal pålægges datafordelerens database fremsendes i XML på samme måde som almindelige registerdata, dog er information om geometrien indlejret som et XML element i WKT formatet.  
En pakke består for raster eller punktskyer af nedenstående:

* Et antal filer – selve dataleverancen
* En indeksfil der katalogiserer filerne

Raster-filer skal leveres i GEO-TIFF format som beskrevet i afsnit 3 "Rasterdata ".  
Punkt-skyer skal leveres i LAZ format som beskrevet i sidst i afsnit 3 "Punktsky-data"  
Hvis det er nødvendigt at anvende andre formater, skal det aftales med og bestilles igennem Operatøren.

Alle forsendelser bør være scannet for virus inden indsendelse til datafordeleren.

## Filopbygning og Sanity Check

### XML leverancer

Såfremt pakker afleveres på batch-porten skal dataleverandøren muliggøre et ”sanity check” i Datafordeleren.

Sanity-checket kan bestå af forskellige mekanismer udvalgt hensigtsmæssigt på baggrund af den konkrete replikeringskanal og afvejet med hensyn til performance.

Datafordelerens replikeringskanaler understøtter følgende muligheder for konfiguration af Sanity-check-mekanismer på dataleverancerne:

* Skema-validering (XML-format) er obligatorisk for alle XML forsendelser.
* Checksum på dataindholdet kan tilvælges.
* Medfølgende kontrolfil eller indlejrede kontroloplysninger der indeholder data for første og sidste record, alternativt checksum på første og sidste record. Kan tilvælges.

Her skal angives hvilke, om nogen, af de to valgfrie kontrolmekanismer der ønskes anvendt: N/A

Leverance XML'et skal basere sig på datafordelerens grundlæggende skema (se bilag 18) samt det for dataleverancen specifikke skema (se bilag 14).

Vær opmærksom på at det specifikke skema for dataleverancen skal have påført UUID referencer til de modsvarende elementer i databeskrivelsen (bilag 1). Se bilag 6 OpdateringsEksempel.xsd for eksempel på udfyldte UUID referencer.

Leverance XML'et opbygges så det har følgende udseende:

<?xml version='1.0' encoding='UTF-8'?>  
< cmn:DatafordelerDataDelivery   
 xmlns:cmn="http://data.gov.dk/rest\_of\_namespace\_to\_be\_decided"  
 xmlns:abbr="http://data.gov.dk/abbr/rep"  
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"   
 xsi:schemaLocation="http://data.gov.dk/rest\_of\_namespace\_to\_be\_decided replication\_channel\_common.xsd http://data.gov.dk/abbr/rep replication\_channel\_specific.xsd">   
 <cmn:Header>  
 <cmn:DataProviderName>[Name]</cmn:DataProviderName>  
 <cmn:ReplicationChannelID>[RCID]</cmn:ReplicationChannelID>  
 <cmn:SequenceNumber>[Number]</cmn:SequenceNumber>  
 <cmn:TimeOfCreation>[CreationTS]</cmn:TimeOfCreation>  
 <cmn:CoversPeriodFrom>[StartPeriodTS]</cmn:CoversPeriodFrom>  
 <cmn:CoversPeriodTo>[EndPeriodTS]</cmn:CoversPeriodTo>  
 <cmn:DeliveryType>[Total/Delta/Sync]</cmn:DeliveryType>  
 </ cmn:Header>

<cmn:Data>  
 <cmn:DataEntity>  
 <cmn:Action>[Create/Update/Delete]</cmn:Action>  
 <cmn:CreateEvent>[NoEvent/CreateEvent/UpdateEvent/DeleteEvent]</cmn:CreateEvent>   
 <!-- CreateEvent er optional og kan undlades helt -->  
 <abbr:[EntityName]”>  
 <abbr:[DataField]>[Data]</abbr:[DataField]>  
 <abbr:[DataField]>[Data]</abbr:[DataField]>  
 ...  
</abbr:[EntityName]>  
</cmn:DataEntity>  
  
<cmn:DataEntity>  
 <cmn:Action>[Create/Update/Delete]</cmn:Action>  
 <cmn:CreateEvent>[NoEvent/CreateEvent/UpdateEvent/DeleteEvent]</cmn:CreateEvent>   
<abbr:[EntityName]>  
 ...  
</abbr:[EntityName]>  
</cmn:DataEntity>  
...  
 </cmn:Data>  
</cmn:DatafordelerDataDelivery>

I Header sektionen angives oplysninger om den specifikke leverance.

DataProviderName: Her angives dataleverandørens navn i almindelig klar tekst.  
 ReplicationChannelID: Indeholder den af Datafordeleren tildelte unikke id for leverancens replikerings kanal.   
SequenceNumber: Det relevante sekvensnummer der unikt identificerer den enkelte forsendelse fra dataleverandøren.  
TimeOfCreation: Timestamp for genereringen af forsendelsen.  
CoversPeriodFrom: Ved deltaleverancer angives timestamp for starten på perioden som udtrækket dækker over. Ved totalleverancer sættes dato'en blot til null for at angive at start datoen ikke er relevant.  
CoversPeriodTo: Ved deltaleverancer angives timestamp for slutningen på perioden som udtrækket dækker over. Ved totalleverancer sættes denne timestamp blot til null.  
DeliveryType: Her angives hvad det er for en slags forsendelse. Total = Totaludtræk af hele leverancen, Delta = Ændrings udtræk. Indeholder alle entiteter der er blevet ændret siden sidste leverance, Sync = Indeholder et totaludtræk til synkroniserings formål.

I Data sektionen angives de egentlige forsendelsesdata i en flad tabel form. Dvs. at der laves en xml række for hver rækker i en tabel der ønskes ajourført.

<Action>: Feltet skal udfyldes med "Create", "Update" eller "Delete" og skal være det første felt i alle data rækker. Feltet fortæller Datafordeleren hvad den skal gøre med den enkelte række i dataleverancen.  
<CreateEvent> Feltet kan undlades hvis der ikke ønskes skabt datanær hændelse i datafordelren for denne specifikke række. Anvendes feltet, kan følgende værdier bruges "NoEvent", "CreateEvent", "UpdateEvent" og "DeleteEvent"   
<[EntityName]>: Navnet på den dataentitet, der skal ajourføres, f.eks. "<Bygning>".   
<[DataField] >[Data]: Angives for hvert enkelt felt i en dataentitet, der leveres. [DataField] erstattes med feltets navn. Ved Creates skal alle felter medtages, men [DataField] kan undlades hvis feltet må være NULL. Ved Update skal nøglefelter og de datafelter der ønskes ændret være udfyldt. Ved Delete skal nøglefelter være udfyldt. Et enkelt felt kan f.eks. se sådan ud <abbr:Boligareal>85</abbr:Boligareal>

Se bilag 6 for eksempel på et skema for en deltaleverance til Datafordeleren.

#### Vektordata

For felter med vektorgeometri, der som ovenfor nævnt leveres i WKT, findes i Datafordeleren geometrityperne, jf. OGC 06-103r4, som beskrevet i tabellen nedenfor. For hvert felt vælges een - og kun een - af disse typer, således leverancen er typestærk også på geometrien.

Desuden vælges for hvert geometrifelt een - og kun een - af de af Datafordeleren understøttede projektioner, som feltets geometrier leveres i.

|  |
| --- |
| Type |
| POINT |
| LINESTRING |
| POLYGON |
| MULTIPOINT |
| MULTILINESTRING |
| MULTIPOLYGON |

I nedenstående tabel beskrives for hvert felt geometritypen og projektionen

|  |  |  |  |
| --- | --- | --- | --- |
| Klassenavn | Attributnavn | Type | EPSG-kode for projektion |
| Adressepunkt | position | POINT | 25832 |
| Husnummer | husnummerretning | POINT | 25832 |
| NavngivenVej | vejnavnelinje | MULTILINESTRING | 25832 |
| NavngivenVej | vejnavneområde | POLYGON | 25832 |
| NavngivenVej | vejtilslutningspunkter | MULTIPOINT | 25832 |

### Rasterdata

Herunder beskrives rasterdata ifbm Datafordeleren. Nedenstående beskrivelse gælder medmindre andet specifikt aftales.

Datafordeleren modtager rasterdata i GeoTIFF-format. Afhængigt af data anvendes forskellige profiler af GeoTIFF.

#### Arketyper for raster-data

Arketyperne nedenfor er navngivet, så de matcher GSTs dokument "Regelsæt for lagring af data i Geodatabanken".

* Digitalt Producerede og Scannede Billeder, der er tre- eller firekanals-data, kendetegnet ved mange unikke farver, og som kan komprimeres med tabsgivende kompression. Udstilles som RGB.
* Digitalt Producerede Kortværker, der er trekanals-data, f.eks. baseret på vektordata. Kendetegnet ved relativt få unikke farver. Udstilles som RGB
* Højdemodeller, der er 32 bits Enkeltkanals-data med "måledata", f.eks. højdedata eller dybdedata, der kan komprimeres med en ikke-tabsgivende kompression. Udstilles som fortolket resultat i RGB på afbildningsservices som rådata i andre services og downloades som rådata

GeoTIFF-opsætningerne for hver arketype, der skal følges medmindre andet specifikt aftales, er:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Arketype | Antal bånd (type) | Bånd datatype | Tiled | Kompression | Andet |
| Digitalt Producerede og Scannede Billeder | 3 (R,G,B) | Byte | Ja (256x256) | JPEG, JPEG\_QUALITY=95 | PHOTOMETRIC=YCBCR, BIGTIFF=YES |
| Digitalt Producerede og Scannede Billeder | 4 (R,G,B,A) | Byte | Ja (256x256) | JPEG, JPEG\_QUALITY=95 | BIGTIFF=YES |
| Digitalt Producerede og Scannede Billeder | 4 (R,G,B,NIR) | Byte | Ja (256x256) | JPEG, JPEG\_QUALITY=95 | BIGTIFF=YES |
| Digitalt Producerede Kortværker | 3 (R,G,B) | Byte | Ja (256x256) | DEFLATE, PREDICTOR=2, ZLEVEL=6 | BIGTIFF=YES |
| Højdemodeller | 1 (Værdi) | 32bit Float | Ja (256x256) | DEFLATE, PREDICTOR=2, ZLEVEL=6 | BIGTIFF=YES, NODATA=-999 |

Datasættets arketype og GeoTIFF-opsætning er: N/A

Datas projektion, een (1) projektion, aftales for hele datasættet i den enkelte DataLeveranceSpecifikation. Projektionen vælges blandt de af datafordeleren understøttede.

Den valgte projektion er: N/A

Den enkelte fils projektion, affin transformation og nodata-værdi mv *skal* angives i GeoTIFF-headeren for hver enkelt fil af hensyn til sikkerhed for korrekt placering ved upload. Filens filnavn, headerindhold med videre vil blive sammenlignet med de i DataLeveranceSpecifikationen definerede værdier, og filen blive afvist, hvis der er uoverenstemmelser. Se desuden eksempel på påkrævet header-indhold i rasterfiler hos GSTs Geodatabank.

De leverede GeoTIFF-fil må IKKE indeholde overviews, da disse så fejlagtigt kan blive brugt ved datafordelerens videre arbejde og forringe datakvaliteten. GeoTIFF-filerne må heller ikke indeholde andre georefereringsmetoder end affin transformation (Feks må der ikke anvendes transformationsmatrix eller Ground Control Points).

Den affine transformation må ikke indeholde rotationsparametre forskellig fra nul. Akserne i rasteren skal således være parallelle med akserne i den anvendte projektion.

Påkrævede headers i raster-filer (eksemplificeret ved GDAL-udskrift af TIFF), jvf. GSTs "Regelsæt for lagring af data i Geodatabanken" kan ses i Bilag 9.

#### Aftaleindgåelse og initial-upload

Ved indgåelse af aftale om raster-register udarbejdes en VRT-fil, der beskriver det komplette datasæt. VRT-filen skal overholde det XSD-schema som Datafordeleren udstiller på N/A, som kun tillader de korrekte felter og kræver de felter, som beskrevet i dette dokument.

VRT-filen benyttes her til at beskrive, hvorledes en række homogene rasterfiler kan samles til een raster. Ved homogene rasterfiler forstås filer, der alle har samme egenskaber bortset fra translationsparametrene i den affine transformation og dataindholdet af de enkelte pixels. Populært sagt: Hvordan en række tiles tilsammen udgør en raster.

VRTen skal beskrive: - Projektionen for datasættet (angivet i <SRS>-elementet). - Datasættets affine transformation (angivet i <GeoTransform>-elementet). - Et antal VRTRasterBand svarende til antal bånd. - For hvert bånd angives en eller flere <SimpleSource>-elementer (een pr fil), der må ikke angives extra attribute tags på elementet (f.eks. resampling) - Filnavne skal være angivet ved relative stier (relativeToVRT="1" i <SourceFilename>), og der må ikke anvendes underbiblioteker. Dvs filerne skal ligge i samme bibliotek som VRT-filen - Der skal angives et fuldt udfyldt SourceProperties-element for hver fil. - SrcRect skal *altid* give tilefilens fulde udstrækning. Dvs xOff="0", yOff="0" og xSize og ySize i <SrcRect> skal være identiske med RasterXSize og RasterYSize fra SourceProperties. - DstRect skal give den tilsvarende rektangel i destinationsdatasættet. Dvs xSize og ySize skal være identiske med xSize og ySize i <SrcRect>. - Inden for hvert bånd, må de forskellige DstRect ikke overlappe hinanden. - For billedbærende data (R,G,B,NIR, R,G,B,A eller R,G,B), skal R,G,B altid ligge i henholdsvis bånd 1, 2 og 3. Alle bånd tilknyttes et ColorInterp-element.

Den affine transformation må ikke indeholde rotationsparametre forskellig fra nul. Akserne i rasteren skal således være parallelle med akserne i den anvendte projektion.

I <MetaData>-elementet angives meta-data om datasættet på følgende form (feltnavngivning følger XML for andre register-opdateringer):

<Metadata>

<MDI key="DataProviderName">[NAME]</MDI>

<MDI key="ReplicationChannelID">[RCID]</MDI>

<MDI key="SequenceNumber">[Number]</MDI>

<MDI key="TimeOfCreation">[Creation Timestamp]</MDI>

<MDI key="CoversPeriodFrom">[Start Period Timestamp]</MDI>

<MDI key="CoversPeriodTo">[End Period Timestamp]</MDI>

<MDI key="DeliveryType">[Total/Delta/Sync]</MDI>

</Metadata>

Initalupload vil have sekvensløbenummer N/A og DeliveryType have værdien ”Total”. Øvrige oplysninger udfyldes.

Det er vigtigt at bemærke, at Datafordeleren ikke nødvendigvis benytter VRT-filen, når rasterdata skal anvendes. Derfor skal VRT-filen og de levererede GeoTIFF-filer være udformet, så der opnås samme resultat ved at læse VRT-filen, som der opnås ved at indlæse GeoTIFF-filerne direkte. VRT-filen skal således blot beskrive de leverede GeoTIFF-filer.

Læsning af data gennem VRT'en må således ikke modificere data. Derfor kan en række VRT-funktionaliterer *ikke* anvendes. Dette gælder bl. a. følgende tags og atributter:

* ComplexSource
* GCPList
* MaskBand
* HideNoDataValue
* ColorTable
* Description
* Offset
* Scale
* Overview
* CategoryNames
* AveragedSource
* ComplexSource
* KernelFilteredSource

Der uploades ved initialupload een fil pr tile, således at datasættet fra starten består af samtlige beskrevne GeoTIFF-Filer. Disse filer kan være ensfarvede (f.eks. sorte) eller komplet fyldte med NODATA-værdier. Filerne skal i øvrigt overholde specifikationer under "opdatering" og som angivet ovenfor.

Alle efterfølgende uploads af rastertiles til det givne rasterregister på Datafordeleren er således opdateringer til det eksisterende datasæt.

Et eksempel på en VRT-fil kan ses i Bilag 10.

#### Opdatering

Ved opdatering leveres en pakke bestående af:

* X GeoTIFF-filer (hver med egenskaber som defineret i specifikationen for replikationskanalen)
* 1 VRT-fil (navngivet jf. navngivning for replikationskanalen - "[PakkeId].vrt", f.eks. "RC0001-100000001.vrt")

VRT-filen udpeger hver enkelt "tile"/fil i leverancen og skal overholde præcis samme krav som for initialupload - bemærk at <Metadata>-tagget skal udfyldes med løbenummer mv. Filerne (og derfor også deres beskrivelser i VRT-filen) skal være en ægte delmængde af det oprindelige datasæt.

En Raster leverance anses for fuldt modtaget når VRT-filen modtages, det er derfor væsentligt at alle GeoTIFF filerne er afleveret inden VRT-filen afleveres.

#### Transport af visse formater direkte til download

Der er mulighed for at transportere uploads direkte til download-tjeneste - eksempelvis JP2, CADRG mv (for visse formater er dette eneste mulighed, idet datafordeleren kun skal levere disse til download).

Datafordelerens datastore, for Ortofotos f.eks. JPEG-komprimerede GeoTIFF, vil blive benyttet til at generere aflevering af alt andet end de originale filer, ligesom download med areal-begræsninger eller lignende også vil blive håndteret denne vej.

#### SAnity checks af rasterdata

Ved upload testes det at den medfølgende VRT-fil opfylder kravene fra Datafordeleren, samt at hver enkelt rasterfil i en opdatering er (har en header) som beskrevet i den oprindelige konfiguration af datasættet. Hvis en opdaterings filer ikke udgør en ægte delmængde af den oprindelige konfiguration, afvises opdateringen. Der må således ikke være filer i en opdatering, der ikke fandtes i den oprindelige beskrivelse, ligesom filer der findes i opdateringen skal være præcis som oprindeligt beskrevet.

### Punktsky-data

Punktsky-data leveres som LAZ (se http://www.laszip.org/), der skal kunne læses af LasZIP version 2.2.0. LAZ er en industristandard for komprimering af filer i ASPRS LAS-standard (http://www.asprs.org/Committee-General/LASer-LAS-File-Format-Exchange-Activities.html) og der understøttes en underliggende LAS-standard op til og med version 1.3.

Datasættet skal være opdelt i tiles, der berører *men ikke overlapper hinanden*, som navngives standardiseret og ensartet.

#### Arketyper for punktskydata i Datafordeleren

Arketypen er et sæt af punkter (typisk mange) med varierende tæthed. De enkelte punkter opnås typisk fra laser scanning af et objekts overflade. Punktskyer er grundlag for højdemodellerne.

Headeren for LAZ-filen skal indeholde de i LAS-standarden påkrævede headerinformationer, der skal udfyldes med meningsfyldte værdier. Især skal det iagtages, at følgende felter udfyldes korrekt (navngivning følger ASPRS' standard):

* X scale factor
* Y scale factor
* Z scale factor
* X offset
* Y offset
* Z offset
* Max X
* Min X
* Max Y
* Min Y
* Max Z
* Min Z

Som beskrevet ovenfor skal datasættet være opdelt i tiles (se også "Regelsæt for lagring af data i Geodatabanken" fra GST) der grænser op til hinanden men ikke overlapper. Således skal den udstrækning, som (Min X, Min Y) og (Max X, Max Y) beskriver, svare til tileopdelingen.

Filens filnavn, headerindhold med videre vil blive sammenlignet med de ved initialupload definerede værdier, og filen blive afvist, hvis der er uoverenstemmelser. Se desuden eksempel på påkrævet header-indhold i rasterfiler hos GSTs Geodatabank.

Eksempel på udskrift fra lasinfo kan ses i Bilag 11.

#### Aftaleindgåelse og initial-upload

Ved indgåelse af aftale om punktsky-register udarbejdes en XML-fil, der beskriver det komplette datasæt. XML-filen indeholder en komplet liste af filnavne og for hver fil min X, min Y, max X, max Y (svarende til header-informationen for hver fil). XML-filen skal overholde det XSD-schema som Datafordeleren udstiller på N/A, som kun tillader de korrekte felter og kræver felter, som beskrevet i dette dokument.

Der uploades initielt et komplet datasæt (inkl evt. tomme LAZ-filer, hvor der endnu ikke er data) og alle følgende uploads er derfor en opdatering af de eksisterende filer. Filerne leveres i een og kun een af Datafordeleren understøttet projektion. Initialupload vil have sekvensløbenummer N/A og DeliveryType have værdien ”Total”. Øvrige oplysninger udfyldes.

For dette datasæt er valgt: N/A.

#### Opdatering

Ved opdatering af punktskydata leveres ligeledes en XML-fil, der beskriver opdateringen. XML-filen indeholder en header og et XML-element for hver fil, der opdateres med: filnavn, min X, min Y, max X, max Y. Hvert element i opdateringsfilen skal svare til et element i det ved initialupload definerede datasæt. Ved opdatering er der således altid tale om at overskrive en eksisterende fil, hvorfor DeliveryType ved opdatering altid er ”Delta”.

#### Eksempel på XML-fil

Nedenstående er et eksempel på XML, til opdatering af den fiktive ”1”-replikeringskanal. Opdateringen indeholder en enkelt opdateret LAZ-file:

<?xml version="1.0" encoding="utf-8"?>

<DatafordelerPointcloudDelivery xmlns="pointcloud.xsd">

<Header>

<DataProviderName>Geodatastyrelsen</DataProviderName>

<ReplicationChannelID>1</ReplicationChannelID>

<SequenceNumber>100000001</SequenceNumber>

<TimeOfCreation>2015-01-20T10:07:00</TimeOfCreation>

<CoversPeriodFrom>2014-01-01T00:00:00</CoversPeriodFrom>

<CoversPeriodTo>2014-12-31T23:59:59</CoversPeriodTo>

<DeliveryType>Delta</DeliveryType>

</Header>

<Data>

<File FileName="punktsky\_6371\_572.laz" MinX="572000" MinY="6371000" MaxX="573000" MaxY="6372000" />

</Data>

</DatafordelerPointcloudDelivery>

## Kvittering for dataleverancer

Datafordeleren udstiller en webservice hvor status på de modtagne leverancer kan hentes, herunder det senest afsluttede sekvensnummer. Dataleverandøren skal være forberedt på at genfremsende leverancer som Datafordeleren ikke har markeret som værende færdigbehandlet, såfremt Leverandøren anmoder om dette.

WSDL for kvitterings-webservicen fremgår af bilag WS-03.

# Afsnit 4 – Synkronisering

### Om Synkronisering på denne Replikeringskanal

Datafordeleren indeholder processer allokeret til at detektere og udbedre uoverensstemmelser mellem dataleverandørens registre og de data der er replikeret til Datafordeleren. Disse processer kaldes under et for Synkronisering.

Datafordelerens Synkronisering består af 2 hovedspor

• En proaktiv Synkronisering, der kører i en løbende proces under normal drift. Det er valgfrit hvorvidt proaktiv synkronisering ønskes anvendt.

• En ekstraordinær synkronisering, der afvikles såfremt konkrete uoverensstemmelser er konstateret. De nødvendige mekanismer for ekstraordinær synkronisering skal være til stede.

### Proaktiv Synkronisering

Den proaktive synkroniserings primære opgave er at kontrollere, om Datafordelerens kopi af Dataleverandørens data er korrekte.

Kontrollen udføres ved at tilfældigt udvalgte dataentiteter fra Datafordeleren sammenlignes med de samme dataentiteter hos dataleverandøren. Findes der forskelle, logges fejlen og korrigeres i Datafordelerens kopi.

Ved Raster data (filsæt) kontrollerer den proaktive synkronisering dog udelukkende metadata for dataentiteten, filstørrelse og oprettelses tidspunkt, og altså ikke indholdet af filen. Findes forskelle i metadata hentes den berørte Raster fil ved seperat kald og filen udskiftes Datafordeleren.

For at sikre størst mulig vished for korrektheden af data på Datafordeleren uden at belaste registeret eller Datafordeleren unødigt, er det i nærværende aftale besluttet at gennemføre synkroniseringskontrol af 5000 dataentiteter i døgnet.

Proaktiv synkronisering udføres for hverdage kun før kl. 08 og efter kl. 17; herudover må proaktiv synkronisering ske hele døgnet rundt for weekender.

Proaktiv synkronisering kræver, at Dataleverandøren implementerer og udstiller web servicen "getData" i overensstemmelse med WSDL vedlagt som bilag WS-04.

End points til synkroniseringstjenester i de relevante miljøer hos Dataleverandøren anføres nedenfor:

|  |  |  |
| --- | --- | --- |
|  | Testmiljø | Produktionsmiljø |
| URL | https://pp-sg.danmarksadresseregister.dk/  getDataService.svc | https://sg.danmarksadresseregister.dk/  getDataService.svc |
| IP | 5.44.137.144 | 257.257.257.257 |

Autentifikation ved service-kaldet foretages ved hjælp af et funktionscertifikat, der tilvejebringes af Leverandøren.

Det er Dataleverandørens ansvar at sikre servicen tilstrækkeligt.

Skal der være proaktiv synkronisering for denne leverance: **Ja**

### Ekstraordinær Synkronisering

Den ekstraordinære synkronisering anvendes når der er konstateret større forskelle mellem et Register og Systemet, og består af et dataudtræk fra Dataleverandøren der er markeret som værende et totalt synkroniseringsudtræk. Dette udtræk skal have et sekvensnummer så det indgår i den almindelige opdaterings forsendelse.

Denne ekstraordinære synkronisering sker via en manuel proces.

Dataleverandøren har i afsnit 8 anført, hvem der skal kontaktes for at få udført den ekstraordinære synkronisering.

# Afsnit 5 – Tjenester

### Generelt

Tjenester, der skal udstilles på Datafordeleren, skal specificeres af Dataleverandøren.

*Tjenestebeskrivelsen*, der vedlægges denne aftale som bilag 2, indeholder oversigt og detailbeskrivelser af de tjenester, der skal udstille Dataleverandørens data.

Detailbeskrivelserne skal være så grundige, at de kan danne grundlag for tjenesteudstilling på Datafordeleren. Detailbeskrivelserne skal blandt andet indeholde stillingtagen til hvorvidt svardata skal arkiveres med henblik på revisionsspor, og i givet fald i hvor lang tidsperiode de fremsøgte datasæt skal opbevares.

For hver tjeneste skal mindst angives:

* Specifikation af dataindhold og snitflader til tjenesten.

1. For SOAP tjenester leveres WSDL
2. For REST tjenester leveres JSON schema samt tilhørende beskrivelse af operationer udfyldt i bilag 2.
3. For WMS tjenster leveres Map fil
4. For WFS tjenester leveres GPP fil
5. For WMTS, TMS og WCS tjenester udfyldes tabel i bilag 2 med specifikationerne
6. For Fildownload udfyldes tabel i bilag 2 med specifikationerne

* Adgangskontrolkrav – om tjenesten er frit tilgængelig eller kræver tildelt adgang
* For tjenester, hvortil adgang skal tildeles anføres navn og e-mail-adresse på den person, der skal orienteres om anmodninger om adgang.
* Om brug af tjenesten skal logges, og i givet fald med hvilke oplysninger. Det anføres også, hvor længe, de loggede data skal bevares.
* For filudtræk angives også periode for arkivering af gamle udtræk.
* Om der skal registreres et faktureringsgrundlag for tjenesten, i form af logning af hvilke brugere der tilgår tjenesten samt hvor ofte de gør det og hvor mange data, de trækker ud.

### Navne konvention i DLS dokumenter

Når Bilag 2 udfyldes for hver tjeneste, skal tjeneste navnet indgå i dokument navnet på følgende vis:

”REGISTER NAVN”\_ VERSION\_DATO\_ ”DOKUMENTNAVN”\_"TJENESTENAVN"  
Eksempel: DAGI\_v1.0\_2015.02.15\_Bilag 2\_Tjeneste skabelon\_TjenesteNavn.doc

Hermed tilsikres at den enkelte tjenestes definition er entydigt identificeret, både i første leverance men især også når der opstår ændrings ønsker til tjenesten i fremtiden.

# Afsnit 6 – Metadata

For at gøre data og tjenester på Datafordeleren lette at finde og anvende, udstiller Datafordeleren metadata.

Metadata for datamodeller

Alle metadata der er beskrivende for datamodellen indberettes direkte i XMI'et på leverance, klasse og attribut niveau. (Dette skal være en del af dokumentationen som etableres af dataleverandøren og indgår i den færdige DLS som Bilag 1). Bilag 5 indeholder oversigt over hvilke metadatafelter der gældende for den pågældende datamodel. Bilaget er forud udfyldt med de metadatafelter der er angivet i modelreglerne for grunddata. Såfremt man har yderligere metadatafelter på datamodellen, som ønskes udstillet på Datafordeleren, kan de angives nederst i dette bilag.

Metadata for tjenester

Alle metadata, der er beskrivende for tjenester, indberettes i tjenestespecifikationen under fanebladet ”Metadata for tjenester” (en del af Bilag 2). Bemærk at for geografiske tjenester er metadataskabelonen baseret på ISO19139 og INSPIRE, og der er tale om metadata, som i dag kendes fra geodata-info.dk. I fanebladet ”Metadata for tjenester” er der angivet eksempler på hvordan de enkelte felter skal udfyldes. Der er desuden 3 kolonner, som angiver, om et felt skal benyttes til en given type af tjeneste:

INSPIRE: Angiver om feltet skal benyttes for at være leve op til INSPIRE direktivet

Geodata: Angiver om feltet er påkrævet for geotjenester, der ikke er omfattet af INSPIRE

Øvrige: Angiver om feltet er påkrævet for ikke-geotjenester

Om et felt skal benyttes, er angivet med følgende:

Auto – feltet er udfyldt på forhånd eller udfyldes automatisk.

Obl. – feltet er obligatorisk og skal udfyldes.

Betinget – for INSPIRE-metadata skal feltet udfyldes, når visse betingelser er til stede. Denne angivelse er kun benyttet for INSPIRE-metadata. Udfyldelse af INSPIRE-metadata er nærmere beskrevet i metadatavejledningen, som findes på  
<http://inspire-danmark.dk/media/gst/64390/INSPIRE%20metadatavejledning%20v1_0.pdf> eller via link på geodata-info.dk. Kolonnen med beskrivelse er baseret på metadatavejledningen.

Valgfri – feltet er valgfrit.

N/A – feltet er ikke relevant for den pågældende tjenestetype.

Man kan have brug for at angive flere værdier for nogle felter og grupper af felter. Der er angivet et R ud for de felter og grupper, hvor det kan være relevant. Når man udfylder regnearket kan man indsætte nye rækker for at angive flere værdier af felter eller grupper af felter. Nøgleord er et eksempel på metadata, hvor det kan være relevant at angive flere værdier. For at angive en liste af nøgleord, indsætter man nye rækker for flere nøgleord (med eller uden kontrolleret ordliste).

Nogle metadatafelter hører sammen i en gruppe. Det er angivet i en kolonne med a, b, c og d, således at a angiver tilhørsforhold til én gruppe, b angiver tilhørsforhold til en anden gruppe osv. Når der er angivet grupper med a, b, c og d, vælger man typisk at udfylde én af grupperne. Det er især tilfældet for INSPIRE-metadata, hvor der henvises til metadatavejledningen på <http://www.inspire-danmark.dk> eller <http://www.geodata-info.dk> for nærmere detaljer om INSPIRE’s krav til udfyldelse af metadatafelterne.

Bemærk at nogle af metadatafelterne er beregnet til links til tjenesten og derfor først kan udfyldes, når det konkrete endpoint for tjenesten er kendt.

Service Metadata

For geografiske tjenester der specificeres via .MAP og .GPP filer kan der angives service metadata, dvs. de metadata, som optræder i svar på operationen GetCapabilities, i de respektive .MAP og .GPP filer. Derudover kan der vedlægges capabilities filer med GetCapabilities-svar.

Metadata for Datasæt

For nuværende er det ikke muligt at angive dette i DLS-skabelonen

# Afsnit 7 – Sikkerhedsbeskrivelse

### Sikkerhedsbeskrivelse

Dataleverandøren skal levere en *Sikkerhedsbeskrivelse* af de leverede data.

Beskrivelsen skal angive, hvilke af de leverede data, der er fortrolige og/eller følsomme samt hvilke tjenester, der anvender fortrolige og/eller følsomme data. I hvert tilfælde anføres, hvilke data der betragtes som følsomme eller fortrolige – eksempelvis, at de er personhenførbare og dermed omfattet af Persondataloven.

Sikkerhedsbeskrivelsen vedlægges denne aftale som bilag 4.

Der udstilles ikke personhenførbare data.

# Afsnit 8 – Hændelsesbeskeder

### Datanære hændelsesbeskeder

Dataleverandøren har mulighed for at specificere om der skal dannes datanære hændelsesbeskeder når Grunddata bliver ajourført i Datafordeleren. De datanære hændelsesbeskeder har en unik beskedtype, der en sammenstilling af objekttype/entitetsnavn ([EntityName]) og objekthandling/CreateEvent (Action: CreateEvent, UpdateEvent eller DeleteEvent), f.eks. ’BygningUpdateEvent’. Der dannes ikke datanære hændelsesbeskeder for raster data opdateringer og andre filopdateringer.

Datafordeleren skaber kun datanære hændelsesbeskeder når datasektionen for den enkelte ajourføring indeholder et ’CreateEvent’-element med en af følgende værdier ’ CreateEvent', 'UpdateEvent' eller 'DeleteEvent’, ’<CreateEvent>UpdateEvent</CreateEvent>’ ’CreateEvent’-elementet er ikke krævet og indeholder elementet andre værdier end de ovenfor listede, skabes der ikke en datanær hændelsebesked. Det er således op til Dataleverandøren at styre hvornår der skal skabes datanære hændelsesbeskeder.

Dataleverandøren skal for hver af de beskedtyper, de ønsker at kunne danne datanære hændelsesbeskeder for, specificere indholdet til de elementer i beskedformatet der ønskes udfyldt. Dette inkluderer både konstanter og variable data.

Det er muligt at medsende oplysninger udelukkende til brug i datanære hændelsesbeskeder, når Grunddata bliver ajourført. Dette kræver at oplysningerne medtages (evt. som valgfrie felter) i den XSD der beskriver dataleverancen til Datafordeleren. Anvendelsen af data fra opdateringen specificeres i bilag 2 med brug af Load([DataField])-syntaksen.

En hændelsesbesked kan indeholde Beskeddata. Beskeddata er enten en objektreference eller en repræsentation af objektet som hændelsesbesked vedrører og en reference til et tilhørende dataskema. Vælger Leverandøren objektrepræsentationen skal det specificeres hvilket skema der skal benyttes til objektet repræsentation og hvorledes der skal mappes (angives som i leverancebeskrivelsen med en henvisning til UUID’en i udstillingsmodellen). I bilag 15 og 16 er der eksempler på hvordan mapningen med UUID’er skal foretages for henholdsvis XML og JSON. Der skal leveres skemaer for både XML og JSON.

Tjenestespecifikationen – bilag 2 udfyldes for hver beskedtype der skal være mulighed for at danne hændelsesbeskeder for.

### Forretningsmæssige hændelsesbeskeder

Dataleverandøren har mulighed for at aflevere (push) forretningsmæssige hændelsesbeskeder på en separat replikeringskanal. De forretningsmæssige hændelsesbeskeder er hændelsesbeskeder der er generet af Dataleverandøren og som overholder det aftalte beskedformat. Datafordeleren tilføjer Leverancerute information til beskeden, men distribuerer den ellers uberørt.

Vær opmærksom på at filer der indeholder tabulære data og som sendes via batch kanalen skal være zippet. Dette gælder også for forretningsmæssige hændelser. Rasterdata forsendes uzippet.

Replikeringskanalen har det unikke ID: 22

Denne Replikeringskanal starter med sekvensnummer 100000001

Tjenestespecifikationen – bilag 2 udfyldes med specifikation af de forretningsmæssige hændelsesbeskeder.

# Afsnit 9 – Diverse oplysninger

### Kontaktoplysninger

Kontaktoplysninger for Leverandøren:

|  |  |  |  |
| --- | --- | --- | --- |
| Rolle | Navn | Telefonnummer (direkte) | E-mail-adresse |
| Driftsansvarlig | Udfyldes af Leverandøren | Udfyldes af Leverandøren | Udfyldes af Leverandøren |

Kontaktoplysninger for Operatøren:

|  |  |  |  |
| --- | --- | --- | --- |
| Rolle | Navn | Telefonnummer (direkte) | E-mail-adresse |
| Driftsansvarlig | Udfyldes af Leverandøren | Udfyldes af Leverandøren | Udfyldes af Leverandøren |

Kontaktoplysninger for Dataleverandøren:

|  |  |  |  |
| --- | --- | --- | --- |
| Rolle | Navn | Telefonnummer (direkte) | E-mail-adresse |
| Dataansvarlig | Netcompany Service Desk | +45 33 34 89 25 | servicedesk@netcompany.com |
| Driftsansvarlig | Netcompany Service Desk | +45 33 34 89 25 | servicedesk@netcompany.com |
| Kontaktperson ved incidents | Netcompany Service Desk | +45 33 34 89 25 | servicedesk@netcompany.com |

### Incidents

### End-points, FTP-adresser mv.

Dette afsnit indeholder alle relevante tekniske domæneadresser på services og snitflade-servere relevante for denne Replikeringskanal. Vær opmærksom på at det er dataleverandøren der skal udstille WS-04 og WS-06.

|  |  |  |
| --- | --- | --- |
| **Servicenavn** | **Integrationstest** | **Produktion** |
| SFTP Server (Push) | Udfyldes af Leverandøren | Udfyldes af Leverandøren |
| RegisterData (WS-01) | Udfyldes af Leverandøren | Udfyldes af Leverandøren |
| DeliveryDisturbance(WS-02) | Udfyldes af Leverandøren | Udfyldes af Leverandøren |
| DeliveryReceipt (WS-03) | Udfyldes af Leverandøren | Udfyldes af Leverandøren |
| getData (WS-04) | https://pp-sg.danmarksadresseregister.dk/getDataService.svc | https://sg.danmarksadresseregister.dk/getDataService.svc |
| SynchronisationService (WS-06) | N/A | N/A |
| SFTP Server (Pull) | N/A | N/A |

### SSH nøgler til SFTP

Til udveksling af data via SFTP skal den offentlige del af en SSH-2 RSA nøgle udveksles.

Vejledning til hvorledes en sådan nøgle kan genereres på hhv. Windows eller Mac/Linux findes f.eks. her:

<https://support.google.com/youtube/answer/3071034?hl=en>

Bemærk, at hvis der benyttes PuttyGen, så skal det være SSH-2 RSA optionen der skal vælges.

Den offentlige del af en SSH-2 RSA nøglen sendes til: [DatafordelerSwSupport@kmd.dk](mailto:DatafordelerSwSupport@kmd.dk)

### Kald af tjenester med relation til replikeringkanalerne

De tjenester der udstilles af Datafordeleren med relation til replikeringskanalerne, skal kaldes med en tjenestebruger der har et tilknyttet FOCES eller VOCES certifikat. Tjenestebrugeren skal autoriseres inden det er muligt at kalde. Send en anmodning til [DatafordelerSwSupport@kmd.dk](mailto:DatafordelerSwSupport@kmd.dk) med brugernavnet på tjenestebrugeren og hvilket register der ønskes adgang til.

### Oversigt over Bilag til denne Dataleverancespecifikation

Bilag 1: ”Datamodel.xmi”. *Databeskrivelsen*. Beskrivelse af de data der skal udstilles på Datafordeleren (se nærmere detaljer i afsnit 1, kapitel ”Databeskrivelse”). **Udarbejdes af Dataleverandøren** med udgangspunkt i bilag 12 "Datamodel XMI skabelon"

Bilag 1a: ”DFSample.XML”. Eksempel på XML opdaterings leverance til Datafordeleren. XML'et skal overholde skemaet der er defineret i bilag 18, samt det dataleverandør specifikke skema.

Bilag 2: ”Tjeneste skabelon.xlsx”. *Tjenestebeskrivelsen*. Indeholder dokumentation på tjenester knyttet til de leverede data. Tjenestebeskrivelsen indeholder oversigt og detailbeskrivelser af kendte tjenester der udstiller data fra Replikeringskanalens datadomæne. Detailbeskrivelserne skal være så grundige, at de kan danne grundlag for tjenesteudstilling på Datafordeleren. Detailbeskrivelserne skal blandt andet indeholde stillingtagen til hvorvidt svardata skal arkiveres med henblik på revisionsspor, og i givet fald i hvor lang tidsperiode de fremsøgte datasæt skal opbevares. **Tjenestebeskrivelsen udarbejdes af Dataleverandøren.**

Bilag 3: "Opdatering.xsd". Specifikation af det tekniske format på de data, der leveres fra Dataleverandøren til Datafordeleren. Specifikationen er angivet i XML Schema (xsd) og anvendes til validering af de data, som modtages på Datafordeleren. **Udarbejdes af Dataleverandøren** med udgangspunkt i bilag 14 "Opdatering skabelon.xsd".

Bilag 4: ”Sikkerhedsbeskrivelse.doc”. Sikkerhedsbeskrivelsen (se afsnit 5). **Sikkerhedsbeskrivelsen udarbejdes af Dataleverandøren.**

Bilag 5: ”Metadata for udstillingsmodel.xlsx” Beskriver metadata for udstillingsmodellen. **Udarbejdes af Dataleverandøren, hvis der ønskes udvidelser til metadata.**

Bilag 6: "OpdateringEksempel.xsd" Eksempel på udfyldt XML skema (xsd), der definerer den dataleverandør specifikke del af et opdaterings sæt.

Bilag 7: Databehandleraftale mellem Dataleverandøren og Leverandøren. **Udelades, hvis der ikke skal udstilles persondata.**

Bilag 9: ”RasterHeaders.docx” Påkrævede headers i raster-filer (eksemplificeret ved GDAL-udskrift af TIFF), jvf. GSTs "Regelsæt for lagring af data i Geodatabanken"

Bilag 10: ”VRTsample.docx” Eksempel på VRT-fil

Bilag 11: ”lasinfoSample.docx” Eksempel på udskrift af LAZ-fil fra lasinfo-værktøjet

Bilag 12 - "Datamodel XMI skabelon.docx”. Skabelon til udfyldelse af XMI der definerer en udstillingsmodel i Datafordeleren.

Bilag 13 - "XMISample.xmi". Et eksempel på en udfyldt datamodel der er udfyldt på baggrund af Bilag 12.

Bilag 14: ”Opdatering skabelon.xsd”. Skabelon til udfyldelse af XSD der definerer den del af dataleverancerne til Datafordeleren der er dataleverandør specifik.

Bilag 15: ”Bygning.xsd”. Eksempel på et XML skema for et dataobjekt til benyttelse i datanære hændelsesbeskeder.

Bilag 16: ”Bygning.json.txt”. Eksempel på et JSON skema for et dataobjekt til benyttelse i datanære hændelsesbeskeder.

Bilag 17: "IP adresser Whitelist.docx". Dataleverandørens liste over IP adresser der skal tilgå datafordeleren. **Whitelisten udarbejdes af Dataleverandøren.**

Bilag 18: ”DataDelivery.xsd”. Grund skemaet til forsendelser til datafordeleren. Dette skema indeholder de dele der er fælles for alle leverancer til datafordeleren.

Bilag 19: ” Bilag 19 - Tjenesteoversigt.xlsx”. Dataleverandørens liste over tjenester i Datafordeleren.

Bilag WS-01: ” RegisterDataService.wsdl” udarbejdes p.b.a. databeskrivelsen. Bilag WS-01 udarbejdes kun såfremt Replikeringskanalen indeholder en SOA-Port. **Udstilles af Datafordeleren.**

Bilag WS-02: ”DeliveryDisturbanceService.wsdl”. **Udstilles af Datafordeleren.**

Bilag WS-03: ”DeliveryReceiptService.wsdl”. **Udstilles af Datafordeleren.**

Bilag WS-04: ”GetDataService.wsdl”. **Udstilles af Dataleverandør.**

Bilag WS-06: ”SynchronisationService.wsdl”. Bilag WS-06 udarbejdes kun såfremt dataleverandøren vil stille muligheden for automatisk bestilling af synkronisering til rådighed. **Udstilles af Dataleverandør.**

Bilag WS-07: "GetRasterService.wsdl". **Udstilles af Dataleverandør.**