

Moderniseringsbehov til arbejdsgrupper

Drift og modernisering af Datafordeleren

August 2023

Version [11-08-2023](#)

Version	Beskrivelse	Ansvarlig	Dato
1.0	Oplæg til arbejdsgrupperne	SDFI	11/08-2023

Indholdsfortegnelse

1. Introduktion til bilagets udformning og indhold	1
1.1 Kapitel 2 "Udvikling af den moderniserede Datafordeler"	1
1.2 Kapitel 2.1 "Generelle behov i relation til den moderniserede Datafordeler"	1
1.3 Kapitel 3 "Funktionelle behov og særlige egenskaber i den moderniserede Datafordeler"	1
2. Udvikling af den moderniserede Datafordeler	2
2.1 Generelle behov i relation til den moderniserede Datafordeler	3
3. Funktionelle behov og særlige egenskaber i den moderniserede Datafordeler	5
3.1 MB1: Etablering af en modulær it-arkitektur	6
3.2 MB2: Udvidet understøttelse af versionering	7
3.3 MB3: Etablering af funktionalitet ved brug af standardprogrammel	9
3.4 MB4: Øget brug af dynamisk skalering af kapacitet til drift	10
3.5 MB5: Øget automatisering af testmiljøer	11
3.6 MB6: Øget brug af automatiseret test	12
3.7 MB7: Optimering af hændelser samt hændelsesgenerering	13
3.8 MB8: Optimering af logning og overvågning via logning	15
3.9 MB9: Optimering af distribution af geodata	17
3.10 MB10: Optimering af indlæsning af data	18
3.11 MB11: Optimering af forespørgsels- og udstillingsværktøjer	25
3.12 MB12: Optimering af understøttelse af udstilling af data i nær realtid	26
3.13 MB13: Etablering af automatisering af konfiguration af tjenester	28
3.14 MB14: Optimering af filudtræk	30
3.15 MB15: Optimering af funktionalitet til selvbetjening	32
3.16 MB16: Optimering af understøttelse af kopiregistre	34
3.17 MB17: Optimering af håndtering af metadata	37

1. Introduktion til bilagets udformning og indhold

Nærværende bilag er tiltænkt brugt i forbindelse med de arbejdsgrupper, der i projektets udviklingsfase har brug for indsigt i de behov, der blev stillet fra SDFIs side til en kommende leverandør. Bilaget indeholder SDFIs funktionelle behov ("moderniseringstiltag") for modernisering af Datafordeleren og er angivet i overordnede behov samt user stories.

Dette bilag eksisterer i 5 versioner, en for hvert af de 5 temaer/arbejdsgrupper, hvor de alene er indledningen i kapitel 3 der adskiller versionerne.

Bilaget er målrettet arbejdsgrupperne i det agile samarbejde for modernisering af Datafordeleren og kan med fordel læses i sammenhæng med Netcompanys bilag vedrørende produktvisionen for det tilsvarende tema/arbejdsgruppe. Bilaget må ikke videredistribueres til andre end deltagerne i arbejdsgrupperne.

Nedenfor gennemgås bilagets indholdet samt opbygningen af kapitlet omkring moderniseringsbehovene.

1.1 Kapitel 2 "Udvikling af den moderniserede Datafordeler"

Forløbet omkring udvikling af den moderniserede Datafordeler er beskrevet i kapitel 2, herunder gennemførelse af udviklingsfasen.

1.2 Kapitel 2.1 "Generelle behov i relation til den moderniserede Datafordeler"

De generelle behov i relation til den moderniserede Datafordeler er beskrevet i kapitel 2.1.

1.3 Kapitel 3 "Funktionelle behov og særlige egenskaber i den moderniserede Datafordeler"

Der er i kapitel 3 angivet en række moderniseringsbehov for den moderniserede Datafordeler samt dertil hørende funktionelle behov og behov for særlige egenskaber, der samles i en backlog.

1.3.1 Funktionelle behov

De funktionelle behov angivet i nærværende bilag er forretningsorienterede og giver værdi for en bestemt gruppe af brugere. Disse behov er i dette bilag udtrykt som user stories med angivelse af "brugeren", den *funktionalitet*, som stilles til brugerens rådighed, og hvilken *fordel*, der opnås med funktionaliteten. Formatet er det meget udbredte user story format¹, der ofte omtales som "Som en..., vil jeg..., således at...", jf.:

- Som en ; **hvem**, der ønsker noget opnået ("**brugeren**")
- Vil jeg ; **hvad** de vil opnå (**Handlingen**)
- Således at ; **hvorfor** de vil opnå den ting (**Fordelen**)



Formatet følger således følgende skabelon:

Som "brugeren", [vil/skal/ønsker] [jeg/man] <ønsket funktionalitet>, således at ["brugeren"/jeg/man] <opnået fordel>

¹ Fra <https://www.agilealliance.org/glossary/user-story-template/>

Som det fremgår af ovenstående skabelon, kan fordelene både være for brugeren, der starter historien, men fordelene kan også være for andre brugere. "Brugeren" kan i nogle tilfælde dække over forskellige brugere, der ønsker samme funktionalitet for at opnå den samme fordel.

Typiske eksempler på behov beskrevet som user stories med dette format er:

- Som **bankkunde**, vil jeg **trække penge** fra en pengeautomat, således at **jeg ikke er begrænset af bankens åbningstider**
- Som **bruger** af systemet, vil jeg **sende feedback**, således at **systemejerne kan lære af mine oplevelser**

1.3.2 Særlige egenskaber

Behov for særlige egenskaber ved den moderniserede Datafordeler er ikke på samme måde forretningsorienterede, men er udtrykt som et behov til en særlig teknisk egenskab relateret til den moderniserede Datafordeler.

Formatet for beskrivelsen af disse egenskaber følger følgende skabelon:

Den moderniserede Datafordeler skal *<besidde en egenskab>*

De særlige egenskaber udtrykker ikke-funktionelle egenskaber for den moderniserede Datafordeler, men adskiller sig fra generelle ikke-funktionelle krav, ved at være knyttet op på et specifikt moderniseringstiltag.

1.3.3 Backlog

Hvert moderniseringsbehov afsluttes med en tabel indeholdende en "backlog", der samler user stories og/eller særlige egenskaber. For hvert element i backloggen er der i separate kolonner angivet:

- 1) En entydig identifikation, der følger formatet "mm.nn", hvor
 - mm: er nummeret på moderniseringsbehovet
 - nn: er et unikt nummer for behovet i det pågældende moderniseringsbehov
- 2) En beskrivelse formuleret som user story, jf. ovenstående om funktionelle behov, eller et behov om en egenskab (jf. ovenstående om særlige egenskaber)

2. Udvikling af den moderniserede Datafordeler

Den moderniserede Datafordeler vil blive udviklet gradvist gennem udviklingsfasen i en række udviklingsleverancer. Hver udviklingsleverance vil dække over udvikling af en eller flere dele af den moderniserede Datafordeler, f.eks. ét eller flere moderniseringsbehov. Hver udviklingsleverance vil være opdelt i et eller flere Sprints.

De konkrete forretningsmæssige ønsker til den moderniserede Datafordeler er formuleret i en række moderniseringsbehov, jf. kapitel 3. Netcompanys produktvision for den moderniserede Datafordeler, herunder løsning af de enkelte moderniseringsbehov, følger af tilsvarende bilag for Netcompanys produktvision.

Netcompany skal sikre, at den moderniserede Datafordeler efter afslutningen af udviklingsfasen tilbyder samme services som den nuværende Datafordeler inklusive en række forbedringer (moderniseringstiltag).

2.1 Generelle behov i relation til den moderniserede Datafordeler

Udvikling af den moderniserede Datafordeler forventes grundlæggende at skulle gennemføres som en re-implementering af den nuværende Datafordelers funktionalitet i en ny modulær it-infrastruktur og med forbedring af en række funktionaliteter.

Det er afgørende for SDFI, at den moderniserede Datafordeler ved endt udvikling kan levere de samme services som i den nuværende Datafordeler, men med en række arkitekturmæssige og funktionelle forbedringer. Hermed menes:

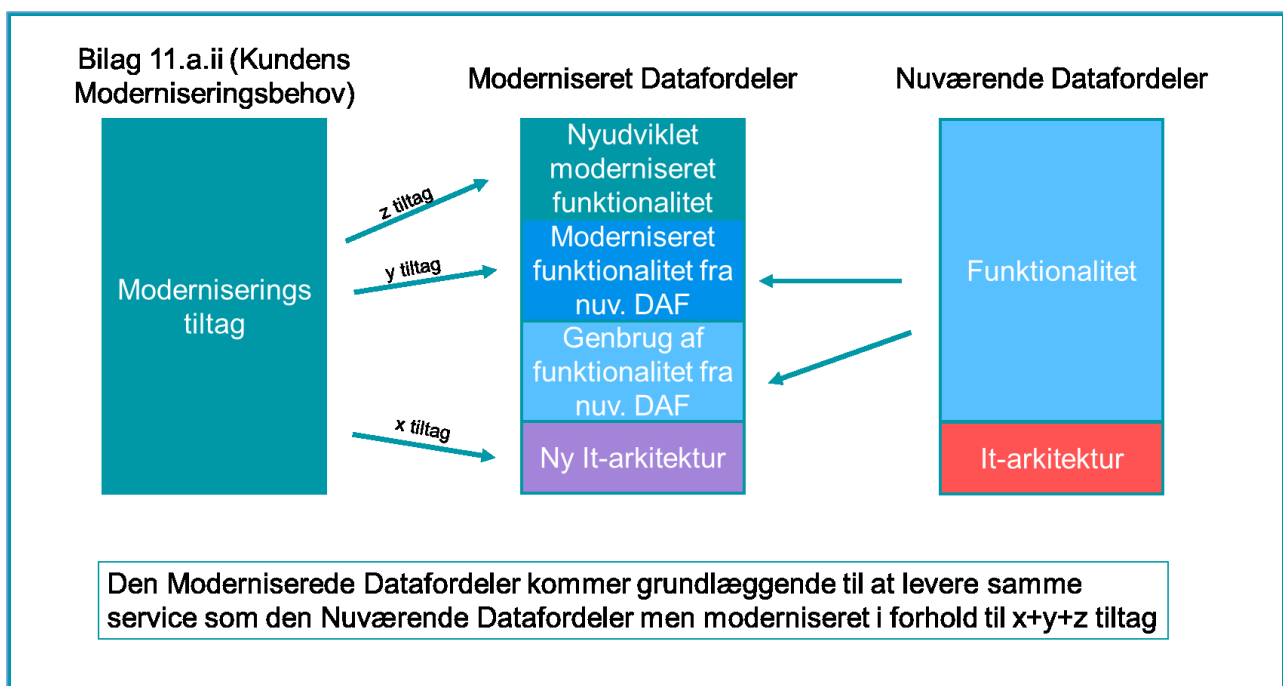
- 1) At den moderniserede Datafordeler skal indeholde de samme tekniske foranstaltninger, funktionaliteter og grænseflader, som forefindes i den nuværende Datafordeler, således at brugerne påvirkes mindst muligt ved overgangen til den moderniserede Datafordeler, og
- 2) At den moderniserede Datafordeler forbedres i forhold til de funktionelle og non-funktionelle områder, som er dækket af de moderniseringsbehov, som indgår i udviklingsfasen.

2.1.1 Opbygning af den moderniserede Datafordeler

Det er ikke alle funktionaliteter i den moderniserede Datafordeler, som nødvendigvis skal udvikles fra bunden for at opfylde SDFIs forretningsmæssige mål og behov. Der vil forventeligt være dele af den nuværende Datafordeler, som hensigtsmæssigt kan genbruges i den moderniserede Datafordeler som grundlag for en modernisering eller som ren videreførelse.

De dele af Datafordeleren, der anbefales moderniseret, de dele, der anbefales nyudviklet, og de dele, der anbefales genbrugt, fremgår af Netcompanys produkt vision.

En principtegning af opbygning af den moderniserede Datafordeler fremgår af Billede 2.1.



Billede 2.1: Opbygning af den moderniserede Datafordeler

2.1.2 Overordnede behov

Den moderniserede Datafordeler skal fortsat understøtte en effektivisering af offentlig datadistribution ved at sikre, at der fremadrettet eksisterer en fælles platform, som offentlige data kan distribueres fra.

Den moderniserede Datafordeler skal derfor tilbyde følgende overordnede funktioner og services:

For anvenderorganisationerne:

- Levere tjenester i relevante formater med tilgængelige grunddata
- Levere filudtræk i forskellige formater med tilgængelige grunddata
- Levere beskeder omkring ændringer på Dataentiteter, som man har opsat abonnement på
- Levere adgang til opsætning af abonnement (hændelser og filudtræk)
- Levere adgang til at opsætte og konfigurere dataudtræk
- Levere metadatakatalog, hvor beskrivelser og dokumentation af tjenester kan tilgås
- Levere adgang, hvor man kan registrere sig
- Levere adgang til information om status som helhed og de enkelte tjenester, f.eks. om de er tilgængelige

For registermyndighederne:

- Modtage fuld kopi af data fra et register
- Modtage enkelt dataopdatering
- Modtage ændringer (eksempelvis deltafiler) fra registre
- Behandle og opdatere grunddata i Datafordeleren på baggrund af alle ovenstående dataleverancer fra registre
- Tilgå logudtræk, der pr. tjeneste giver information om enkelt brugers forbrug
- Tilgå logudtræk, der over en periode viser tilgængelighed af de enkelte tjenester
- Levere adgang til konfiguration af og ændringer til egne tjenester
- Levere adgang, hvor brugeres rettigheder kan vedligeholdes

For SDFI:

- Tilgå logudtræk, der over en periode viser tilgængelighed og anvendelse (antal tilgange og overførte datamængder) af enkelte og grupper af tjenester
- Tilgå adgang til information – både på forretnings- og teknikniveau, der understøtter monitorering
- Levere adgang til vedligeholdelse af brugere og rettigheder

3. Funktionelle behov og særlige egenskaber i den moderniserede Datafordeler

I nærværende kapitel er SDFIs funktionelle behov til modernisering af den nuværende Datafordeler beskrevet.

Punktet er struktureret i følgende underpunkter:

1. [MB1: Etablering af den moderniserede Datafordeler i en modulær it-arkitektur](#)
2. [MB2: Udvidet understøttelse af versionering](#)
3. [MB3: Etablering af funktionalitet ved brug af standardprogrammel](#)
4. [MB4: Øget brug af dynamisk skalering af kapacitet til drift](#)
5. [MB5: Øget automatisering af testmiljøer](#)
6. [MB6: Øget brug af automatiseret test](#)
7. [MB7: Optimering af hændelser samt hændelsesgenerering](#)
8. [MB8: Optimering af logning og overvågning via logning](#)
9. [MB9: Optimering af distribution af geodata](#)
10. [MB10: Optimering af indlæsning af data](#)
11. [MB11: Optimering af forespørgsels- og udstillingsværktøjer](#)
12. [MB12: Optimering af understøttelse af udstilling af data i nær realtid](#)
13. [MB13: Etablering af automatisering af konfiguration af tjenester](#)
14. [MB14: Optimering af filudtræk](#)
15. [MB15: Optimering af funktionalitet til selvbetjening](#)
16. [MB16: Optimering af understøttelse af kopiregistre](#)
17. [MB17: Optimering af håndtering af metadata](#)

De moderniseringsbehov, der vedrører arbejdsgruppen for "Geodata" fremgår af nedenstående figur, idet de moderniseringsbehov, der dækkes fuld ud af et moderniseringstiltag er markeret med "x", mens dem, der er delvis dækket er markeret med et "(x)".

	MT8: Optimering af distribution af geodata
Udviklingsleverance	U3
MB-03 – Standardprogrammel	(x)
MB-09 – Geodata	x
MB-17 – Metadata	(x)

Figur 3.1 Relevante moderniseringsbehov vedrørende arbejdsgruppen for "Geodata"

3.1 MB1: Etablering af en modulær it-arkitektur

3.1.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Den nuværende Datafordeler er designet og oprindeligt udviklet med en lagdelt software arkitektur med veldefinerede løst koblede komponenter. Hver komponent havde en klar funktionalitetsafgrænsning med veldefinerede snitflader til at tilgå funktionaliteten.

Videreudvikling af den nuværende Datafordeler har over tid medført, at den oprindelige softwarearkitektur ikke er bevaret, og har introduceret teknisk gæld, således at der nu er en større kobling mellem en række komponenter. Samtidig er den designede lagdelte arkitektur ikke blevet overholdt stringent.

Disse forhold har resulteret i, at det er svært at udskifte/erstatte komponenter i arkitekturen, dels grundet den tætte kobling og dels de mange stærke afhængigheder, der er opbygget.

Udskiftning/erstatning af komponenter med f.eks. standardkomponenter er reelt gået fra at være en mindre vedligeholdelsesopgave med tilhørende regressionstest til at være et større udviklingsprojekt med risiko for at introducere Fejl eller fejlsituationer i andre komponenter.

Den moderniserede Datafordeler

Moderniseringsbehovet ”modulær it-arkitektur” dækker over et ønske om, at den moderniserede Datafordeler opbygges af fysiske og logiske moduler med klare grænseflader. Denne modulære arkitektur ses som grundlaget for at implementere løsninger til en række af nedenstående moderniseringsbehov. Den modulære arkitektur skal derfor indtænkes i løsningsdesign og realisering af moderniseringsbehovene, således at den modulære arkitektur understøtter moderniseringen og fremtidige behov.

Modulerne skal opbygges, således at det er muligt at udvikle og/eller udskifte moduler uafhængigt af andre moduler i den moderniserede Datafordeler. Grænsefladerne mellem moduler skal være velbeskrevet og bl.a. kunne håndtere implementering af et nyt modul, uden at dette påvirker de eksisterende moduler. Modulerne ønskes udviklet og implementeret med moderne teknologier, herunder brugen af moderne container- og orkestreringsteknologier, således at driftsomkostninger minimeres.

3.1.2 Backlog

ID	Beskrivelse
01.01	Den moderniserede Datafordeler skal opbygges af fleksible og løst koblede moduler.
01.02	Den moderniserede Datafordelers moduler skal kunne genbruges.
01.03	Den moderniserede Datafordelers moduler skal kunne sættes i produktion enkeltvist.
01.04	Den moderniserede Datafordeler skal have en høj grad af indlejret robusthed og redundans, der sørger for, at selv hvis dele af den moderniserede Datafordeler bliver utilgængelige eller overbelastede, kan den fortsætte normal drift.

[Tilbage til oversigt](#)

3.2 MB2: Udvidet understøttelse af versionering

3.2.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Den nuværende Datafordeler understøtter versionering af tjenester, men det er kun taget i anvendelse i meget begrænset omfang. Således er alle udstillede tjenester fortsat i version 1. Dette er bl.a. understøttet gennem versionsangivelse i URL. Eksempel: EBR-tjenesten Ejendomsbeliggenhed har følgende URL: <https://services.datafordeler.dk/EBR/Ejendomsbeliggenhed/1/REST/BFEnrAdresse>, hvor 1-tallet i URL angiver version for tjenesten.

Ved udvikling af nye tjenester i den nuværende Datafordeler angives versionen for den enkelte tjeneste i dataleverancespecifikationen (herefter benævnt "DLS" nedenfor). Hvis en tjeneste ændres, og der ikke angives en ny version, vil den ændrede tjeneste overskrive den eksisterende (f.eks. hvis den ændrede version stadig er angivet som version 1). Den nye version af tjenesten med uændret versionsnummer betyder, at anvenderorganisationer ikke har mulighed for at tilgå den gamle version i f.eks. en overgangsperiode, samt at der er stor sandsynlighed for, at applikationer, som anvender den gamle version, ikke længere vil fungere, da den nye tjeneste er forskellig fra den gamle.

Ved fejlrettelser til en tjeneste sker der i den nuværende Datafordeler ingen ændringer i versionen, også selvom om fejlrettelsen medfører, at tjenesten har en anden logik, eller at dataresultatet fra tjenesten er ændret, f.eks. at antallet af dataobjekter, som tjenesten returnerer, er ændret afledt af fejlrettelsen.

For versionering af tjenester er der i dag følgende udfordringer:

- Alle tjenester har version 1.0.
- Der benyttes ikke faste processer eller værktøjer, der understøtter versionering, ud over den manuelle proces for håndtering af DLS.
- Der mangler stringente regler for versionsangivelse, og som klart beskriver, hvornår en tjeneste skal skifte version.
- Versionering af fejlrettelser til tjenester afleder ingen ny version, selvom rettelse set fra anvenderorganisationernes perspektiv er en breaking change.

For versionering af datamodeller for grunddata er der i dag følgende udfordringer:

Der arbejdes kun med en gældende version af en datamodel, hvilket er den version, der er repræsenteret i grunddatamodellen, jf. datafordeler.dk. Eventuelle nye versioner af datamodeller håndteres og beslutes i Arkitekturforummet.

Registermyndighedernes systemer understøtter som udgangspunkt også kun en version af en datamodel, hvilket er den model der anvendes ved myndighedsudøvelsen. Det vil også være denne version, der er udstillet på den nuværende Datafordeler.

Vælger en registermyndighed at have forskellige versioner af datamodeller på den nuværende Datafordeler, kan dette gøres ved at have dem som separate registre på Datafordeleren med tilhørende separate tjenester. Der er ikke tale om egentlig versionering, da der ikke er nogen sammenhæng mellem de forskellige datamodeller og datasæt, som hver skal opdateres af registermyndigheden.

Den moderniserede Datafordeler

Den moderniserede Datafordeler skal tilbyde versionering, således at anvenderorganisationer kan tilgå flere versioner af den samme tjeneste. Samtidig skal der sikres rettidig og fyldestgørende dokumentation af nye versioner af tjenester, sådan at overgang til nye versioner kan planlægges og gennemføres inden for en rimelig tid. Anvenderorganisationer skal ikke ændre i deres egen opsætning, medmindre de skal anvende en nyere version.

Dokumentation for versioner af tjenester (version, release, patch), der er afledt af konfigurationer, skal automatisk vedligeholdes med visning og beskrivelser (jf. MB15), således at brugere kan få et retvisende overblik over tjenesterne.

3.2.2 Backlog

ID	Beskrivelse
02.01	Den moderniserede Datafordeler skal understøtte parallel afvikling af tjenester i forskellige versioner.
02.02	Som anvenderorganisation skal man have mulighed for at benytte ældre versioner af en tjeneste, således at man har tid til at tilpasse egne it-systemer til tjenester med ny eller opdateret funktionalitet.
02.03	Som datafordelermyndighed skal man, hvor der er behov, kunne udvikle og implementere nye versioner af tjenester uden hensyntagen til det eksisterende, således at brugere kan fortsætte med at benytte de eksisterende versioner af tjenesten.
02.04	Som registermyndighed skal man kunne udvikle og implementere nye Versioner af tjenester uden hensyntagen til de eksisterende, således at anvenderorganisationer kan fortsætte med at benytte de eksisterende versioner af tjenesten.
02.05	Som bruger af en tjeneste skal man i god tid varsles og informeres om en ny version af tjenesten, således at brugere med denne information kan udvikle og planlægge deres løsninger mod den nye version.

[Tilbage til oversigt](#)

3.3 MB3: Etablering af funktionalitet ved brug af standardprogrammel

3.3.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Den nuværende Datafordeler er udviklet som en .Net-applikation med en underliggende platform bestående af standardteknologier, f.eks. Microsoft SQL og PostgreSQL databaser.

Komponenter for den nuværende Datafordeler, der indeholder forretningsfunktionalitet, er i vidt omfang udviklet specifikt til den nuværende Datafordeler. Den nuværende Datafordeler indeholder således megen kundespecifikt programmel. F.eks. er indlæsningsfunktionalitet og udstilling af tjenester i høj grad baseret på kundespecifikt programmel, specielt for tabulære data. Tjenester til udstilling af geodata anvender dog en række typer af standardprogrammel, f.eks. MapServer, men der indgår altid kundespecifikt programmel eksempelvis switchboardet.

Den moderniserede Datafordeler

Den moderniserede Datafordeler skal for de komponenter for den nuværende Datafordeler, der er udviklet specifikt til den nuværende Datafordeler, og i det omfang det understøtter SDFIs forretningsmæssige behov samt giver den mest optimale økonomiske og teknologiske løsning, baseres på standardprogrammel, herunder open source, der kan opsættes gennem konfiguration.

Der er behov for, at det anvendte programmel, bl.a. medfører lettere adgang til skalering og implementering af tjenester.

3.3.2 Backlog

ID	Beskrivelse
03.01	Den moderniserede Datafordeler skal i relevant omfang anvende markedsudbredt, veletableret, supporteret og gennemprøvet standardprogrammel, herunder open source.
03.02	Anvendt programmel skal være konfigurerbart.
03.03	Som datafordelermyndighed ønskes brug af standardprogrammel, herunder open source, i det omfang det medfører færre ressourcer til at udvikle og vedligeholde den moderniserede Datafordeler samt understøtte de Forretningsmæssige mål og behov.

[Tilbage til oversigt](#)

3.4 MB4: Øget brug af dynamisk skalering af kapacitet til drift

3.4.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Den nuværende Datafordeler er oprindeligt designet og udviklet til at blive driftet på en cloud-platform, mere specifikt Microsoft Azure opsat som private cloud i den afgivende leverandørs eget miljø. Denne Microsoft-Azure-baserede private cloud var ikke dedikeret til Datafordeleren, men var en delt ressource blandt den afgivende leverandørs kunder. Den nuværende Datafordelers arkitektur for udstilling af data er derfor designet til horisontal skalering for at udnytte de cloud-kapabiliteter, som den private cloud tilbyder.

Grundet tekniske udfordringer med performance og stabilitet med den etablerede private cloud-infrastruktur, er der som erstatning herfor etableret en infrastruktur med dedikerede HW-ressourcer til Den nuværende Datafordeler, primært baseret på virtuelle servere, men hvor udvalgte servere kører på dedikeret hardware uden virtualisering, eksempelvis en række databaseservere.

Den fornødne kapacitet i infrastrukturen til den nuværende Datafordeler er sikret ved, at der i produktionsmiljøet er etableret ekstra servere, der kan tages i anvendelse ved behov for skalering. Dette betyder også, at der i det nuværende setup for den nuværende Datafordeler er en betragtelig overkapacitet, som ikke anvendes i den daglige drift. Hvis der f.eks. generelt er behov for at have mulighed for at skalere belastning på switchboard-servere med 25% for at håndtere normale spidsbelastninger, og der til normal drift er behov for otte servere, vil der være yderligere to ekstra servere, der ikke anvendes, men alene afventer at en belastningssituation, der afleder et skaleringsbehov.

Er der behov for yderligere skalering, skal der manuelt tilføres HW-ressourcer, hvilket alt andet lige er en langsommelig proces. Dette betyder også, at driftskapaciteten ikke hurtigt kan skaleres op eller ned i forhold til de ændringer i belastningen, som en given anvendelse afspejler.

Den moderniserede Datafordeler

Forbrugsmønstret på den nuværende Datafordeler viser en stadig større anvendelse af tjenester samt store og stigende datamængder. Der er i den moderniserede Datafordeler behov for en driftsplatform med høj performance og stabilitet, der samtidig tilbyder hurtig tilpasning af kapacitet, og som kan skaleres i forhold til både ændringer i anvendelse og datamængder samt ændrede anvendelsesmønstre.

3.4.2 Backlog

ID	Beskrivelse
04.01	Den moderniserede Datafordeler skal have den rette grad af automatisk skalering til at være tilgængelig i en belastet driftssituation.

[Tilbage til oversigt](#)

3.5 MB5: Øget automatisering af testmiljøer

3.5.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Opsætning og konfiguration af den nuværende Datafordeler er i stor udstrækning baseret på manuelle processer. Det har den konsekvens, at opsætning af testmiljøer til brug for specifikke formål er en langvarig opgave. Den manuelle proces giver også større risici for Fejl i opsætning eller konfiguration, der skal rettes efterfølgende.

Dette betyder også, at der i dag er et begrænset antal testmiljøer, og at disse er forholdsvis statiske i konfigurationen. Den enkelte registermyndighed har udfordringer med at ændre og konfigurere testmiljøerne til at understøtte specifikke testbehov, da de enkelte Miljøer benyttes af flere registermyndigheder samtidigt.

Registermyndighederne stiller hver især testdata til rådighed, men processen for levering og indlæsning af testdata i testmiljøerne er manuel og tidskrævende. Dette medfører, at testdata ofte er mangelfulde, samt at kun nogle få registermyndigheder leverer testdata.

Den moderniserede Datafordeler

Den moderniserede Datafordeler skal kunne stille et vilkårligt antal testmiljøer til rådighed, og øget automatisering skal gøre det muligt hurtigt at få yderligere testmiljøer samt at nedlægge disse igen, når de ikke skal bruges længere.

Den moderniserede Datafordeler skal anvende tre typer af testdata:

- **Produktionslignende:** Udsnit af data fra produktionsmiljø, kan indeholde persondata
- **Registerspecifikke:** Testdata designet til at dække alle situationer, der kunne opstå i et registers data. Denne type testdata skal gøre det nemt for registermyndighederne at teste op imod andre registre.
- **Syntetiske:** Testdata specialdesignet til afprøvning af den moderniserede Datafordeleren, for eksempel i form af performance eller svartider. Indeholder aldrig persondata.

Ved oprettelse af testmiljøer skal SDFI frit kunne vælge, hvilken type testdata der skal stilles til rådighed i testmiljøet. Disse testdata skal til enhver tid nemt og uden Netcompanys deltagelse kunne opdateres med en ny version. Registermyndighederne skal nemt kunne levere testdatasæt.

3.5.2 Backlog

ID	Beskrivelse
05.01	Replikering og oprettelse af testmiljøer skal foregå automatisk.
05.02	Applikationen i den moderniserede Datafordeler skal simpelt kunne kopieres og flyttes som helhed eller opdelt til forskellige testmiljøer.
05.03	Det skal være muligt at kopiere data fra et testmiljø til et andet testmiljø.
05.04	Som datafordelermyndighed eller registermyndighed skal man i testmiljøer selv kunne oprette, indlæse samt slette testdata (produktionslignende, registerspecifikke samt syntetiske), således at man under udvikling kan teste den moderniserede Datafordeler på forskellige testdata.
05.05	Som datafordelermyndighed eller registermyndighed skal man i testmiljøer kunne teste tredjepartsudviklet kode, således at man under udvikling kan teste kvaliteten af koden.
05.06	Som bruger skal man mod testmiljøet kunne teste på kommende releases af tjenester, således at man kan teste, at egne løsninger virker, inden de pågældende releases lægges i produktionsmiljøet.

[Tilbage til oversigt](#)

3.6 MB6: Øget brug af automatiseret test

3.6.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Den afgivende leverandør har opbygget en samling af automatiske test, der anvendes til regressionstest og til daglig test i den afgivende leverandørs udviklingsmiljø i forbindelse med udviklingsopgaver, så det opda- ges, hvis nyudvikling, ændringer eller fejlrettelser har introduceret Fejl eller afledte Fejl i andre komponenter.

SDFI har ikke adgang til den afgivende leverandørs værktøjer til gennemførelse af automatiserede test. Det er således alene den afgivende leverandør, der kan gennemføre automatiserede test og har adgang til test- resultater og testrapporter.

De automatiserede test kan også kun afvikles i den afgivende leverandørs udviklingsmiljø og kan derfor ikke anvendes i de øvrige testmiljøer til at sikre, at opsætning og konfiguration af den nuværende Datafordeler med tilhørende tjenester er korrekt.

Den moderniserede Datafordeler

Der ønskes i den moderniserede Datafordeler et automatiseret test setup, der understøtter det agile samar- bejde og forenkler gennemførelsen af test ved at automatisere regresionstest.

3.6.2 Backlog

ID	Beskrivelse
06.01	Den moderniserede Datafordeler skal afvikle regresionstest som automatiseret test.
06.02	Som datafordelermyndighed har man behov for hyppige regresionstest af den moderniserede Da- tafordeler, således at datafordelermyndigheden opnår den bedst mulige kvalitetssikring.
06.03	Som datafordelermyndighed og registermyndighed har man behov for regresionstest, således at datafordelermyndighed og registermyndighed opnår et mindre træk på forretningsressourcer.
06.04	Datafordelermyndigheden har behov for, at det automatiserede test-setup genererer fyldestgø- rende testrapportering, således at datafordelermyndigheden kan træffe beslutning om idriftsæt- telse af releases.
06.05	Som datafordelermyndighed skal man kunne afvikle automatiseret test, således at datafordeler- myndigheden kan understøtte hyppige releases.
06.06	Som datafordelermyndighed skal man have indsigt i konfigurationen samt indholdet af de automa- tiserede tests, således at der opnås gennemsigtighed i, hvordan der testes.
06.07	Som datafordelermyndighed har man brug for, at det automatiserede test-setup integrerer mod datafordelermyndighedens testværktøjer, således at testresultater automatisk kan tilgås af data- fordelermyndigheden.

[Tilbage til oversigt](#)

3.7 MB7: Optimering af hændelser samt hændelsesgenerering

3.7.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Den nuværende Datafordeler er konfigureret til at danne hændelsesbeskeder i det fællesoffentlige beskedformat for hændelsesbeskeder om grunddata ud fra oplysninger i dataopdateringer. Det blev på et tidligt tidspunkt besluttet i grunddataprogrammet, at der skulle anvendes det fællesoffentlige beskedformat, og at dette var det eneste format, hændelsesbeskeder skulle være i. Ændringer til hændelsesbeskedformatet vil skulle ske i den etablerede governancestruktur for grunddata.

Den nuværende Datafordeler understøtter to typer hændelser:

- **Datanære hændelser**, der genereres på baggrund af en ændring i data. Hændelsesbeskeder genereres af den nuværende Datafordeler ved, at der på dataopdateringer er påhæftet en markering for hændelsesbeskeder. Denne markering påhæftes af registermyndigheden. Markering angiver, hvilken type datanær hændelsesbesked der skal genereres: Insert, Update eller Delete. Internt i den nuværende Datafordeler forstås og håndteres disse som tre uafhængige hændelser, der kan abonneres individuelt på. Anvenderorganisationer vil derfor se en dataopdatering som en af de tre forskellige hændelser (Insert, Update, Delete).
- **Forretningsnære hændelser**, der dannes alene på baggrund af en forretningsmæssig aktivitet, har som udgangspunkt ikke nogen kobling til dataopdateringer. Forretningsnære hændelsesbeskeder genereres af registermyndigheden, og den nuværende Datafordeler benyttes alene som distributionsmekanisme til brugerne. Forretningsnære hændelser anvendes pt. ikke.

brugere abonnerer på hændelser for at kunne modtage eller hente hændelsesbeskeder. Brugere skal i dag oprette et abonnement på hændelser for at kunne hente/modtage hændelsesbeskeder. Der gives kun adgang til hændelsesbeskeder, der er genereret efter oprettelsestidspunkt for abonnementet. Brugere kan således ikke tilgå hændelser fra før oprettelsen af abonnementet. Datanære hændelsesbeskeder kan eksempelvis bruges som notifikationer på opdateringer af data i deres kopiregistre.

brugere kan tilgå hændelsesbeskeder i grunddataprogrammets beskedformat og kan således flytte disse hændelsesbeskeder ind i egne systemer baseret på dette format.

Den nuværende Datafordeler indeholder mulighed for simpel filtrering af hændelsesbeskeder, men denne filtrering dækker kun en begrænset del af hændelsesbeskedens attributter.

Det opleves, at hændelsesbeskeder ved den nuværende Datafordeler kan være længe undervejs og ikke understøtter nær-realtidsbehovet. Det har derfor ikke været muligt for anvenderorganisationerne at etablere nær-realtidsopdaterede kopiregistre, hvor opdatering af data trigges af en datanær hændelsesbesked.

Den moderniserede Datafordeler

Det ønskes at forenkle og forbedre funktionalitet i relation til generering og udstilling af hændelser.

3.7.2 Backlog

ID	Beskrivelse
07.01	Som bruger ønsker man værktøjsunderstøttelse til validering af og fejlsøgning på, at hændelsesabonnement er opsat korrekt i forhold til de dataopdateringer, der ønskes hændelsesbeskeder for, således at brugere hurtigt kan identificere forkert opsætning af abonnement.
07.02	Som bruger skal man inden for rammerne af den moderniserede Datafordeler kunne specificere indhold i hændelsesbeskeder, således at brugere kan få hændelsesbeskeder, der passer ind i egne systemer.
07.03	Som bruger skal man kunne vælge format for hændelsesbeskeder, således at brugere kan få hændelsesbeskeder, der passer ind i egne systemer.
07.04	Som bruger skal man kunne tilgå hændelsesbeskeder uden først at have oprettet et abonnement, således at brugere kan benytte hændelsesbeskederne i egne systemer uden af være begrænset af tidspunktet for oprettelsen af et abonnement.
07.05	Som bruger skal man kunne filtrere hændelsesbeskeder, således at brugere kun modtager de hændelsesbeskeder, der er relevante for dem.
07.06	Som bruger skal man kunne filtrere hændelsesbeskeder på registreringstid og virkningstid, således at brugere kun modtager de hændelsesbeskeder, der er tidsmæssigt relevante for dem.
07.07	Som bruger skal man kunne få hændelsesbeskeder på grundlag af, at virkningstiden indtræffer, således at brugere kan igangsætte forretningsprocesser til rette tid.
07.08	Som bruger skal man kunne få hændelsesbeskeder på alle objekttyper, således at brugere kan igangsætte forretningsprocesser relateret til hver objekttype.

[Tilbage til oversigt](#)

3.8 MB8: Optimering af logning og overvågning via logning

3.8.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Den nuværende Datafordeler har etableret logning. Den centrale log på tværs af alle løsningskomponenter er baseret på standardprogrammet LogPoint. Der måles eksternt via en egenudviklet svartidsmåleklient, der måler og logger uafhængig af LogPoint.

I relation til logning i den nuværende Datafordeler opleves det, at:

- Det er svært at få overblik over opdelingen i komponenter og flow af data gennem den nuværende Datafordeler.
- Det er svært at anvende logning til overvågning og tjek af driftsstatus for den nuværende Datafordeler, herunder til fejlfinding.
- Udtræk af rå logs ikke er mulig.
- Det er svært at anvende LogPoint til analyser og statistikker af logdata.
- LogPoints kapacitet modsvarer ikke brugen af den nuværende Datafordeler.
- Tilgængeligheden af LogPoint er svingende.

Den moderniserede Datafordeler

SDFI ønsker adgang til detaljeret logning af den moderniserede Datafordeler for bl.a. at understøtte følgende forretningsmæssige behov:

- Driftsovervågning
- Fejlfinding
- Forbrugstal, herunder fakturering
- Forretningsudvikling

Dashboards, der benyttes til visning af driftsstatus, skal bl.a. automatisk modtage og vise status på den moderniserede Datafordelers vigtigste moduler, herunder forbrugstal mv. Dashboards er vigtige i forbindelse med driftsovervågning af infrastrukturen og applikationer. Den moderniserede Datafordeler skal automatisk kunne levere detaljerede rapporter.

3.8.2 Backlog

ID	Beskrivelse
08.01	Som registermyndighed skal man kunne tilgå logning vedrørende egne data, således at man har det fornødne grundlag for sporing, statistik samt brugstal, fakturering af forbrug og auditering.
08.02	Som datafordelermyndighed skal man kunne tilgå logning, således at man har det fornødne grundlag for sporing, statistik samt brugstal og auditering.
08.03	Som datafordelermyndighed skal man kunne tilgå detaljerede logs, således at datafordelermyndigheden kan deltage i fejlsøgningen.
08.04	Som datafordelermyndighed skal man kunne visualisere udvalgte driftslogs grafisk, således at visualiseringen gøres tilgængelig for alle brugere.
08.05	Som datafordelermyndighed skal man kunne konfigurere visualisering af driftslogs, således at det kan målrettes brugerne.
08.06	Den moderniserede Datafordeler skal stille logdata til rådighed og udstille disse overfor SDFI og den relevante registermyndighed.
08.07	Den moderniserede Datafordeler skal sikre sporbarhed, således at sikkerhedsbrud kan dokumenteres.
08.08	Den moderniserede Datafordeler skal logge rolletildeling og brugerstyring.
08.09	Den moderniserede Datafordeler skal kunne anvende logdata i et overvågningssystem, hvor der kan opstilles regler for f.eks. alarmer.
08.10	Den moderniserede Datafordeler skal på baggrund af logdata automatisk kunne oprette Incidents, når visse tærskelværdier overskrides. Tærskelværdier aftales nærmere mellem Kunde og Leverandør.
08.11	Som bruger skal man kunne abonnere på beskeder om driftsstatus, således at man som bruger er informeret om den moderniserede Datafordelers driftstilstand i realtid.

[Tilbage til oversigt](#)

3.9 MB9: Optimering af distribution af geodata

3.9.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Den nuværende Datafordeler udstiller geografiske data i alle typer af tjenester. Dette er:

- Geotjenester
- REST-tjenester
- Hændelsesbeskeder
- Filudtræk, som kan distribueres i 12 forskellige formater

Geodata i den nuværende Datafordeler bliver håndteret og udstillet med forskelle i indlæsning, datamodel, fremsøgningslogik, tjenestedesign, navnekonvention mm. afhængig af registermyndighedens specifikation.

Den nuværende Datafordeler har performanceproblemer ved udstilling af geodata, dvs. svartider er længere end sammenlignelige udstillinger af geodata og derfor for lange i forhold til behov hos anvenderorganisationer, f.eks. når de anvender OGC- eller REST-baserede tjenester som datakilde i GIS-klienter.

Performanceproblemer gør sig især gældende for WMTS-tjenester, hvor der i den nuværende Datafordeler ikke er nogen prægenerering uden egentlige brugerforespørgsler. Dette giver svartidsproblemer i GIS-klienter, hvis brugeren zoomer på eller panorerer rundt på et kort.

Det opleves også, at den nuværende Datafordeler ikke kan møde efterspørgsel efter nye metoder i distribution af geodata. Dette kan f.eks. være nye tiltag på OGC-området, som det i det nuværende setup tager meget lang kalendertid at få implementeret.

Som et specifikt eksempel på denne problematik kan nævnes Styled Layer Descriptor (SLD - <https://www.ogc.org/standards/sld>), som ikke er understøttet grundet det tekniske setup på den nuværende Datafordeler. SLD efterspørges af brugere af geodata, men behovet kan ikke opfyldes pt.

Der opleves usmidige metodikker til skalering og implementering af varianter af eksisterende geografiske tjenester for hurtigt at kunne tilpasse til ændrede behov og udvikling i markedet for anvendelse af geodata.

Den moderniserede Datafordeler

Den moderniserede Datafordeler skal indeholde en løsning til distribution af geodata, der begrænser kompleksiteten i Datafordeleren samt sikrer ensartethed i design, implementering og drift af geodata.

3.9.2 Backlog

ID	Beskrivelse
09.01	Som registermyndighed skal man kunne specificere, hvilke versioner af OGC-standarder for en geografisk tjeneste, der skal anvendes, således at flest mulige anvenderorganisationer kan anvende tjenesten.
09.02	Den moderniserede Datafordeler skal i det omfang, det understøtter SDFIs Forretningsmæssige mål og behov samt giver den mest optimale økonomiske og teknologiske løsning, anvende standardprogrammel, herunder open source, i udstilling og distribution af geografiske tjenester.
09.03	Datafordelermyndigheden ønsker at der anvendes internationale standarder. Geografiske tjenester skal overholde standarder fra Open Geospatial Consortium (OGC).

[Tilbage til oversigt](#)

3.10 MB10: Optimering af indlæsning af data

3.10.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Datafordelerens indlæsningsfunktionalitet er et af de mest kritiske elementer i Datafordelerens forretningsarkitektur: Data skal kunne indlæses hurtigt og sikkert, og indlæsningen skal fungere stabilt, således at der f.eks. kan sagsbehandles på baggrund af data umiddelbart efter, at disse er indberettet i de bagvedliggende forvaltningssystemer.

På baggrund af indlæsningens kritikalitet for forretningen, vægtes indlæsningen højt ved beregningen af Servicemålet for den nuværende Datafordelers driftseffektivitet.

Performance

Den nuværende Datafordeler modtager data fra registre på såkaldte replikeringskanaler, som både kan bestå af SOAP-tjenester (også kalde SOA-porte) og FTP-overførsler (også kaldet batch porte). Fælles for begge teknologier er formatet for udveksling af data mellem register og den nuværende Datafordeler. Protokollen samt de specifikationer, der er påkrævet for at data kan udveksles på en replikeringskanal, er specificeret i en Data Leverance Specifikation (DLS). En DLS indeholder også information om andre emner, f.eks. tjenester, men i relation til indlæsninger af et register med tabulære data er følgende informationer væsentlige:

- Datamodel for registret. Både indlæsningsmodel og udstillingsmodel.
 - Indlæsningsmodel (Master load). Leveres som XMI-fil.
 - Udstillingsmodel (Publish). Leveres som XMI-fil.
 - Specialiseret udstillingsmodel (specialised publish). Aftales pr. gang.
 - For de fleste registre anvendes samme model for både overførsel, indlæsning og udstilling.
- Skema-fil for udveksling af data. Leveres som en XML-skemafil og anvendes for både SOAP-baserede og FTP-baserede overførsler på replikeringskanalerne.
- Angivelse af sikkerhed for data (zone 0 eller zone 5).
- Specifikation af Hændelsesbeskeder, da disse genereres som en del af indlæsning. Se moderniseringsbehov "MB7: Optimering af hændelser samt hændelsesgenerering" for detaljer omkring hændelsesgenerering.

For registre med tabulære data består en opdateringspakke af en XML-fil, der beskriver operationer på databaserækker (oprettelse, opdatering eller sletning) samt markering af eventuel generering af tilhørende hændelsesbeskeder.

Der benyttes et sekvensnummer til at styre opdateringsrækkefølgen. Registermyndigheden påfører en opdateringspakke det næste nummer i sekvensen. Sekvensnummeret styres af den nuværende Datafordeler, hvor det næste ledige sekvensnummer kan hentes på en tjeneste. Når den nuværende Datafordeler er klar til at indlæse næste pakke på en replikeringskanal, bliver det pågældende sekvensnummer gyldigt, og pakken med dette sekvensnummer vil blive behandlet.

SOA-opdateringer har i dag et Servicemål på otte sekunder. Svartiden beregnes, fra opdateringen modtages af SOAP-tjenesten, og opdateringens sekvensnummer er gyldigt, til opdateringen er gennemført i den nuværende Datafordeler. Denne tidsgrænse er for den nuværende Datafordeler fremkommet ved, at der i forbindelse med udarbejdelse af kontraktmaterialet i 2013 blev afdækket volumen for forventet antal SOA-opdateringer pr. register. Den maksimale volumen for SOA-opdateringer var her 10.000 pr. døgn, hvilket kunne håndteres med et Servicemål på otte sekunder. Her var der i øvrigt ikke taget hensyn til kødannelse for opdateringer i løbet af døgnet, men regnet med en jævn fordeling. Udviklingen har dog vist, at det nuværende behov er langt større end 10.000 SOA-opdateringer pr. døgn, endda op til en faktor 10 i forhold hertil, hvilket ikke kan understøttes af den nuværende Datafordeler. Her skal det yderligere bemærkes, at for nogle registre er opdateringsmønsteret, at en SOA-opdatering indeholder opdateringer til flere dataobjekter og instanser af disse (f.eks. 100 databaserækker). Dette er dels grundet anvendelse af bitemporalitet, dels grundet registrets forretningsmæssige opdateringsmønster af data på den nuværende Datafordeler.

Batch-opdateringer har i dag et Servicemål på 300 sekunder og angiver, hvor hurtigt den nuværende Datafordeler skal starte indlæsning. Svartiden beregnes, fra opdateringen er færdigoverført til den nuværende Datafordeler, og sekvensnummeret for opdateringen er gyldigt, til opdateringen er påbegyndt i systemet. Da en batch-opdatering kan indeholde mange opdateringer, er der ikke Servicemål for, hvornår indlæsning skal være gennemført.

Indlæsning på den nuværende Datafordeler understøtter både push- og pull-mønstre, hvor push understøttes gennem FTP-overførsler af opdateringer fra registre eller ved kald til SOAP-tjenesten for benyttelse af SOA-opdateringer. Pull-mønsteret, hvor den nuværende Datafordeler henter opdateringer fra registermyndighedens system, anvendes ikke af nogen registre i dag. Der er ingen myndigheder, der udstiller data eller forventer at udstille data på den nuværende Datafordeler, der har tilkendegivet, at de vil anvende pull-mønsteret.

Totalindlæsninger var oprindeligt tiltænkt kun at skulle gennemføres for et register i forbindelse med etablering af dette på den nuværende Datafordeler, dvs. initial-indlæsningen. Efterfølgende har der vist sig behov for at gennemføre totalindlæsninger for registre, der er etableret på den nuværende Datafordeler, f.eks. for at rette mange datafejl eller for at fysisk fjerne data, der ikke længere skal distribueres.

Totalindlæsninger af et eksisterende register sker ved, at alle eksisterende rækker først markeres som slettede (soft delete), dvs. at der sættes et sletteflag på hver række i databasen. Der er således tale om en logisk og ikke fysisk sletning. Herefter indlæses alle data fra registret. En totalindlæsning tager i dag op til flere dage at færdiggøre. Tjenester har ingen data at udstille under totalindlæsning, hvorfor disse reelt ikke er tilgængelige, så længe totalindlæsningen pågår. Da grunddataregistrene anvender bitemporalitet, vokser totalindlæsninger i volumen over tid, og det tager længere og længere tid at indlæse et fuldt register.

Den oplevede nedetid for tjenester vil derved også blive længere. Der arbejdes i øjeblikket på to tiltag til at forbedre denne situation:

- Et tiltag, der fjerner brugen af soft delete, hvilket vil nedsætte tidsforbrug på totalindlæsning.
- Et tiltag, der indlæser totalindlæsninger i et sæt skyggetabeller, og ved færdig indlæsning skifter produktionstabeller ud med skyggetabeller, hvorved den oplevede nedetid på tjenester kan reduceres til sekunder frem for dage/timer.

Disse to tiltag forventes at være fuldt implementeret i den nuværende Datafordeler i den version, som Netcompany har overtaget fra den afgivende leverandør under denne Kontrakt.

Der har været situationer, hvor en registermyndighed har vurderet, at det, der skal rettes via en totalindlæsning, har en så vigtig forretningsmæssig betydning, at registermyndigheden valgte at lukke alle tjenester, indtil totalindlæsningen var gennemført (f.eks. ved sikkerheds-incidents). Lignende situationer forventes at opstå fremadrettet.

Det skal her bemærkes, at der ikke genereres hændelsesbeskeder ved totalindlæsninger, så eventuelle opdateringer, der er indeholdt i en totalindlæsning, vil ikke blive udsendt som hændelsesbeskeder. Dette betyder, at anvenderorganisationer, der benytter hændelsesbeskeder til at igangsætte forretningsprocesser eller vedligeholde en kopi af data, mangler information om opdateringer ved totalindlæsninger.

Indlæsning af geodata (rasterfiler) via ovenstående replikeringskanaler er en udfordring i forhold til performance grundet de store datamængder, som raster-geodata omfatter. Dette har været løst ved at anvende eksterne harddiske som udvekslingsmedie mellem registret og den nuværende Datafordeler. Der er en igangværende opgave, der har til formål at kikke på at anvende online fil-tjenester som udvekslingsmekanisme i stedet for de fysiske harddiske.

Kvalitet og stabilitet

Valideringer, der foretages på dataopdateringer i forbindelse med indlæsning i den nuværende Datafordeler, er begrænset til tekniske valideringer, dvs.:

- Skemavalidering, hvor det valideres, at en opdateringspakke overholder det skema, der er aftalt anvendt for dataopdateringer fra det pågældende register.
- Validering af rækkefølge, hvor der valideres på sekvensnummer for opdateringspakke, der modtages via FTP eller som SOA-opdatering, for at sikre, at opdateringspakker indlæses i den korrekte rækkefølge, og at der ikke er huller i sekvensen af opdateringer.

Indlæsning i den nuværende Datafordeler er et langt og komplekst forløb, hvor alle aktiviteter skal gennemføres, for at en dataopdatering modtaget fra et register kan markeres som gennemført. Først herefter kan indlæsning af næste dataopdatering fra registret påbegyndes. Samtidig er indlæsning designet og implementeret som "one size fits all", dvs. både SOA-opdateringer og FTP-opdateringer håndteres på samme måde, når man ser bort fra selve den tekniske modtagelse.

De aktiviteter, der gennemføres i forbindelse med indlæsning af dataopdateringer, er:

- Skemavalidering
- Validering af sekvensnummer
- Indlæsning af dataopdatering til Master-Load-database placeret i sikkerhedszone 5
- Generering af hændelsesbeskeder på baggrund af hændelsesbesked-markering i opdateringspakken Generering pr. abonnement
- Kopiering fra Master-Load-database til Master-Publish-database i både sikkerhedszone 5 og sikkerhedszone 0. Kopieringen sker parallelt.
- Replikering af data fra Master-Publish- til Publish-Slave-databaserne (sker parallelt)
- Commit af dataopdatering
- Frigivelse af hændelsesbeskeder
- Markering af dataopdatering som gennemført, hvorefter næste dataopdateringspakke kan blive indlæst

For SOA-opdateringer gennemføres hele dataopdateringen i et commit-scope. For batch-opdateringer (FTP) anvendes et større commit-scope, men ved mange opdateringer i samme fil bliver disse opdelt i forskellige commit-scopes. Replikering til publish-databaser sker på tabelniveau frem for på forretningsobjektniveau, hvorved forretningsobjekter kan blive inkonsistente, hvis en del af opdateringen ligger i et commit-scope og den anden del af opdateringen ligger i et andet commit-scope. Afledt af denne problemstilling lukkes CPR-tjenester i den periode, hvor der indlæses CPR-opdateringer, for at sikre, at det er valide og konsistente data, der udstilles på CPR-tjenesterne.

I ovenstående flow er der som nævnt kun indbygget parallelitet for kopiering/replikering til publish-databaser, hvilket sætter en række begrænsninger på indlæsningskapaciteten.

Den nuværende Datafordeler har funktionalitet til at håndtere situationer for data, hvor der ikke er konsistens mellem den opdateringshandling (databaseoperation), der er angivet af registret, og status for databaserækken i den nuværende Datafordeler. Denne funktionalitet er betegnet som upsert og er en funktionalitet, der konfigureres aktiv/inaktiv for det enkelte register. Funktionaliteten er implementeret for at håndtere situationer, hvor indlæsning går i stå grundet forkerte opdateringer fra registre. Da databaserækker kan identificeres med nøglefelter, er funktionaliteten i korte træk:

- Hvis registret har angivet, at indhold i opdateringspakken skal gennemføres som en update af en eksisterende databaserække på den nuværende Datafordeler, men databaserækken ikke eksisterer, bliver opdatering konverteret til en insert.
- Hvis registret har angivet, at indhold i opdateringspakke skal gennemføres som en insert af en databaserække, men databaserækken allerede findes på den nuværende Datafordeler, bliver opdatering konverteret til en update.

Upsert anvendes af ganske få registre pt. Det bliver logget, når upsert er anvendt for en dataopdatering, men der er ikke implementeret funktionalitet til at give registrene besked om, hvilke specifikke opdateringer der har været anvendt upsert på. Dvs., at registrene ikke kan foretage fejlsøgning eller korrigerende handlinger i egne systemer og eventuelt fremsende nye korrekte opdateringer til den nuværende Datafordeler.

Upsert funktionalitet har introduceret en opblødning på det ellers meget skarpe og stringente dataansvar, der er mellem registermyndigheder og datafordelermyndigheden for at afhjælpe datafejl.

Erfaring fra rettelse af logiske fejl i indlæsningskomponenten viser, at grundet det komplekse indlæsningsflow er fejlsøgning vanskelig at gennemføre og meget tidskrævende. Der er samtidig ikke funktionalitet eller processer til at rette op på de datafejl, den nuværende Datafordeler eventuelt introducerer, udover at gennemføre en totalindlæsning af registrets data igen.

Den nuværende Datafordeler udstiller en række tjenester til registermyndigheder til brug for status- og fejlhåndtering for opdateringspakker. Det er her muligt maskinelt via tjeneste-kald at få aktuel status på en eller flere opdateringspakker. Hvis indlæsning af en opdateringspakke f.eks. er fejlet, vil den have status: "4 – leverance indeholdt fejl og skal genfremsendes". Den efterfølgende fejlhåndtering gennemføres primært manuelt, dvs. identifikation af den/de specifikke fejl, der gør, at indlæsning af dataopdateringspakken fejler, samt de efterfølgende fejlretningshandling. Dette omfatter, at registret skal korrigere/rette fejl og gensende en dataopdateringspakke med det oprindelige sekvensnummer og med de rettede data, hvorefter indlæsning genoptages. En sådan fejlhåndteringssituation har den afledte konsekvens, at al dataopdatering fra det pågældende register sættes på pause, indtil fejlen er rettet.

Den nuværende Datafordeler har implementeret en synkroniseringsmekanisme (proaktiv synkronisering), men denne er fravalgt af registrene, da den kræver en række snitflader hos registermyndighedernes systemer for at den nuværende Datafordeler kan validere, om datakopien i den nuværende Datafordeler er identisk med registrets data. Synkroniseringsmekanismen er beskrevet nærmere i systemdokumentationen.

Funktionaliteten til ekstraordinær synkronisering anvendes, men er en tung og ressourcekrævende proces, hvorfor den kun anvendes efter nøje overvejelser. Den ekstraordinære synkronisering består af et dataudtræk fra registret, der er markeret som værende et totalt synkroniseringsudtræk. Selve synkroniseringen i den nuværende Datafordeler foregår ved, at den nuværende Datafordeler foretager en sammenligning af data række for række og foretager kun ændringer i de eksisterende data, hvor det er nødvendigt for at sikre synkroniteten. Der genereres ikke hændelsesbeskeder på ændringer, der gennemføres.

Den ekstraordinære synkronisering kan også gennemføres som en total opdatering. Dvs., at indholdet af de eksisterende data slettes, og det fremsendte indlæses på tomme tabeller. Dette svarer til en totalindlæsning.

Overvågning af indlæsning er primært manuel og har i en række tilfælde ikke været tilstrækkelig finkornet, hvilket har betydet, at indlæsning af nogle opdateringer fra et register har været unødigt længe undervejs.

Når dataopdateringer fra et register ikke indlæses rettidigt og uden fejl, opleves dette af datafordelermyndigheden, registermyndigheden og anvenderorganisationer som om, at indlæsning på den nuværende Datafordeler er nede.

Udvalgte databaser er i dag installeret på fysiske databaser af hensyn til stabilitet samt performance. Øvrige databaser kører på virtuelle servere og skalerer horisontalt i den implementerede infrastruktur.

Den moderniserede Datafordeler

Indlæsning er en meget vigtig og central funktionalitet i Datafordeleren, hvorfor der er en række emner som ønskes dækket af dette moderniseringsbehov, herunder ønskes der bl.a. at:

- Øge kapaciteten på indlæsning, så dataopdateringer indlæses i en tilfredsstillende hastighed.
- Designet af den moderniserede Datafordeler understøtter, at indlæsning sker med en sådan hastighed, at selv store opdateringer ikke påvirker udstillingen af data.
- Reducere afhængighed mellem indlæsning af data og anden funktionalitet i den moderniserede Datafordeler, f.eks. generering af hændelsesbeskeder.
- Anvende nutidig og fremtidssikret teknologi til at modtage og indlæse dataopdateringer.
- Understøtte behov for opdatering af data i nær-realtid og realtid.
- Sikre fokus på push-model for opdatering af data.
- Totalindlæsning af et register kan gennemføres, uden at det påvirker udstilling af data, dvs., at data er tilgængelig i tjenester under indlæsning ved en total genindlæsning af data for et register.
- Totalindlæsninger kan gennemføres på timer frem for dage.
- Dimensionere løsningen fra starten af, så der på mellemlangt og langt sigt sikres tilstrækkelig indlæsningskapacitet til at håndtere flere parallelle og serielle opdateringer i nærrealtid fra registre.
- Håndtere periodevise peaks af dataopdateringer fra et eller flere registre.
- Fleksible muligheder for op- og nedskalering af indlæsningskapaciteten, uden at gå på kompromis med robusthed i indlæsningen.
- Data valideres ved indlæsning, så eventuel inkonsistens i data kan identificeres og håndteres.

3.10.2 Backlog

ID	Beskrivelse
10.01	Indlæsningen i den moderniserede Datafordeler skal være robust, således at midlertidige afbrydelser ikke fører til fejlbehæftede data og langvarige fejltilstande.
10.02	Indlæsningen i den moderniserede Datafordeler skal kunne skaleres til smidigt at håndtere et stigende antal registre, datamængder og dataopdateringer i den efterfølgende driftssituation.
10.03	Indlæsning i den moderniserede Datafordeler skal kunne skaleres, så periodevise peaks af dataopdateringer fra registre kan gennemføres på kort tid.
10.04	Indlæsning i den moderniserede Datafordeler skal kunne skaleres, så totalindlæsninger kan gennemføres på kort tid.
10.05	Som registermyndighed skal jeg kunne benytte forskellige opdateringsmønstre og teknologier på den moderniserede Datafordeler, således at jeg kan vælge det, der passer bedst i forhold til mit register i situationen.
10.06	Den moderniserede Datafordeler skal validere data ved indlæsning ud fra foruddefinerede regler om format og mulige dataværdier i en opdateringspakke, så fejlbehæftede data hurtigt findes og ikke når ud til brugerne.
10.07	Som registermyndighed ønsker jeg, at fejlbeskeder ved indlæsningsfejl er fyldestgørende og informative, således at jeg har de bedste muligheder for hurtig fejlretning.
10.08	Som registermyndighed ønsker jeg, at data kan indlæses i nær realtid, således at anvenderorganisationerne kan få den opdaterede data hurtigst muligt.
10.09	Som registermyndighed ønsker jeg, at relaterede data skal kunne indlæses samlet, således at inkonsistente data ikke udstilles.
10.10	Som anvenderorganisation ønsker jeg, at tjenester kan tilgås samtidig med, at alle former for indlæsninger gennemføres, således at jeg altid har adgang til data.
10.11	Som anvenderorganisation ønsker jeg, at udstillede data er opdaterede og konsistente, således at jeg ikke træffer beslutninger på baggrund af forældede eller inkonsistente data.
10.12	Som anvenderorganisation ønsker jeg, at de opdaterede data er udstillet i tjenesterne, når der notificeres om det i hændelsesbeskeder, således at jeg kan hente opdaterede data, når jeg modtager besked om opdateringen.
10.13	Som anvenderorganisation ønsker jeg at modtage information om forbedring af datafejl ved totalload og ændringer i data, som ikke udføres gennem opdateringspakker, således at jeg kan gennemføre nødvendige korrigerende og fejlrettende handlinger.

[Tilbage til oversigt](#)

3.11 MB11: Optimering af forespørgsels- og udstillingsværktøjer

3.11.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Stort set alle tjenester, der er udstillet på den nuværende Datafordeler, er enten specielt udviklet eller konfigureret. Konfiguration af en tjeneste sker også som konfiguration af det specialudviklede kundespecifikke programmel i den nuværende Datafordeler. Konfiguration af nye eller ændring til eksisterende tjenester kræver indgående viden og indsigt i den nuværende Datafordeler, hvorfor det løses af den afgivende leverandør.

Der er udviklet og leveret et "Udviklerrammeværk", der gør det muligt for tredjepartsleverandører at konfigurere tjenester på den nuværende Datafordeler. Afprøvning af dette rammeværk har dog vist, at brugen af dette kræver samme indgående viden og indsigt som ved den "normale" konfiguration. Samtidig er det stadig specialudviklet programmel, der udstiller tjenester. Det specialudviklede programmel er designet til, at tjenester specificeres af registermyndigheden og konfigureres af den afgivende leverandør, og understøtter kun tjenester af typen:

- REST
- Hændelser
- Filudtræk
- Geografiske tjenester

Der er ingen mulighed for, at brugere selv kan udvælge eller specificere, hvilke data de ønsker, ud over hvad der er muligt at specificere i input-parametre til tjenesterne. Bemærk, at input-parametre er fastlagt af registermyndigheden i DLS'en.

Den moderniserede Datafordeler

Den moderniserede Datafordeler skal, i det omfang det understøtter SDFIs forretningsmæssige behov samt giver bedste økonomisk og teknologiske løsning, baseres på standardprogrammel til forespørgsler og udstilling af data, således at udviklingsomkostninger minimeres. Desuden skal den moderniserede Datafordeler stille fleksible søgemuligheder til rådighed for brugerne, således at der ikke skal konfigureres nye tjenester og søgeparametre, efterhånden som nye brugerdrevne behov opstår.

3.11.2 Backlog

ID	Beskrivelse
11.01	Som registermyndighed skal man fleksibelt kunne få udstillet data med mindst mulig konfiguration af nye tjenester, således at kompleksiteten ved opsætningen af nye tjenester minimeres.
11.02	Som anvenderorganisation skal man have frihed til at skræddersy søgninger efter egne behov, således at man kan definere egne brugerdefinerede udtræk, der passer ind i anvenderorganisationens systemer.
11.03	Datafordelermyndigheden skal kunne udstille REST-tjenester, der understøtter streaming, således at behovet for filudtræk samt REST-paginerings minimeres.
11.04	Anvenderorganisationerne ønsker at modtage data fra REST-tjenester via streaming, således at data er tilgængelig fra anmodningstidspunktet.
11.05	Der skal i relevant omfang benyttes standardprogrammel, herunder open source, ved forespørgsler og udstilling af data, jf. MB3.

[Tilbage til oversigt](#)

3.12 MB12: Optimering af understøttelse af udstilling af data i nær realtid

3.12.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Forskellige registermyndigheder har forskellige behov mht. udstilling af data i nær realtid, dvs. hvor hurtigt data og opdateringer kan blive tilgængelige for brugerne. For Ortofoto-registret er det f.eks. vigtigt, at data bliver udstillet, når det er færdigt for et givent år, og herefter vil der ikke være egentlige opdateringer til det. For CVR er det f.eks. vigtigt, at opdateringer af data hurtigst muligt udstilles for brugerne i takt med, at data løbende opdateres i registret. Den nuværende Datafordeler understøtter ikke sidstnævnte behov for udstilling af opdaterede data i nær realtid tilstrækkeligt.

Den nuværende Datafordeler er konfigureret til at danne hændelsesbeskeder ud fra oplysninger i dataopdateringer, jf. situationsbeskrivelse i afsnit 3.7. Hændelsesbeskeder hænger således tæt sammen med dataopdateringer. Når et register har mange dataopdateringer, vil registermyndigheden ofte samle dem i batch-opdateringer for at undgå den kødannelse på indlæsning i den nuværende Datafordeler, der opstår, hvis de mange dataopdateringer bliver indsendt enkeltvist. Volumenmæssigt har den nuværende Datafordeler større kapacitet på batch-opdateringer end på enkeltopdateringer (SOA-opdateringer). Selvom samlingen i batch-opdateringer gør selve indlæsningen mere effektiv, er den også med til at give en forsinkelse, fra data er opdateret i registret, til dataopdateringerne er på den nuværende Datafordeler, og hændelsesbeskeder er dannet. Dette gør, at hændelser fra flere registre i dag kommer med store forsinkelser set i forhold til opdateringer af data i registrene. Hvis en registermyndighed sender opdateringer via batch hver halve time, vil hændelsesbeskeder for de ældste opdateringer være op til 30 minutter forsinket i forhold til opdateringen i selve registret.

I den nuværende Datafordeler er SOA-opdateringer karakteriseret ved, at en opdatering indeholder en enkelt eller få dataopdateringer, dvs. opdatering af et begrænset antal databaserækker for tabulære data. SOA-opdateringer er teknisk implementeret som web-services (SOAP). Batch-opdateringer er karakteriseret ved, at registret samler mange opdateringer og afleverer disse med jævne mellemrum, der kan variere fra få minutter til et døgn. Batch-opdateringer er teknisk implementeret som SFTP-baserede dataoverførsler fra registret til den nuværende Datafordeler. Se Bilag 3 (SDFIs It-miljø og situationsbeskrivelse) for nærmere detaljer.

Nær-realtidsopdateringer håndteres ved SOA-opdateringer i den nuværende Datafordeler, dvs., at indlæsning gennemføres på mellem to og otte sekunder ved kald til web-servicen med et fastsat Servicemål på maks. otte sekunder.

Generering af hændelsesbeskeder er i den nuværende Datafordeler en integreret del af indlæsning af data, jf. situationsbeskrivelse i afsnit 3.7. Det betyder, at generering og udstilling af hændelsesbeskeder gennemføres som en del af den transaktion, der indlæser data, så det sikres, at hændelsesbeskeder ikke er tilgængelige for brugerne, før de opdaterede data er tilgængelige på tjenesterne.

I dag er det registrets ansvar at sikre, at en afsendt opdatering indeholder de dataopdateringer, der er nødvendige for at opdatere et dataobjekt som helhed, og derved sikre, at der ikke opstår inkonsistens i den nuværende Datafordelers kopi af data ved opdateringen. For tabulære data betyder det, at opdateringen skal

indeholde alle de databaserækker, der er nødvendige for at opdatere det berørte dataobjekt. Det kan f.eks. inkludere databaserækker for underobjekter, hvis der er tale om et komplekst dataobjekt.

Den moderniserede Datafordeler

Generelt er der behov for, Den moderniserede Datafordeler minimerer risikoen for flaskehalse fra indlæsning af dataopdateringer til udstilling af disse, således at udstilling kan ske i nær-realtid.

3.12.2 Backlog

ID	Beskrivelse
12.01	Som bruger har man behov for, at data udstilles i nær-realtid, således at man kan dække Forretningsmæssige mål og behov.
12.02	Den moderniserede Datafordeler skal ifm. dataindlæsning sikre konsistens for de enkelte dataobjekter, der modtages opdateringer på fra registret, og hurtigt kunne udstille data til brugere.
12.03	Den moderniserede Datafordeler skal have tilstrækkelig indlæsningskapacitet til at kunne formidle data i nær-realtid.
12.04	I Den moderniserede Datafordeler skal hændelsesbeskeder udstilles/være tilgængelige, så snart dataobjektet er opdateret og tilgængeligt på tjenester.
12.05	Som anvenderorganisation er der behov for, at Den moderniserede Datafordeler skal håndtere hændelsesgenerering ved opdateringer i nær-realtid, således at hændelsesbeskeder er tilgængelige for anvenderorganisationen umiddelbart efter, dataopdateringen er indlæst i den moderniserede Datafordeler.

[Tilbage til oversigt](#)

3.13 MB13: Etablering af automatisering af konfigurerings af tjenester

3.13.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Enhver konfiguration af den nuværende Datafordeler forudsætter i dag tunge administrative og tekniske processer. Registermyndigheder og datafordelermyndighed bestiller ændringer ved at sende formelle ændringsanmodninger sammen med en udfyldt dataleverancespecifikation (herefter benævnt "DLS").

DLS er en manuelt udfyldt Excel-formularer, der indeholder beskrivelse af den ønskede konfiguration. DLS anvendes som specifikation for at konfigurere tjenesten på den nuværende Datafordeler. Denne konfiguration sker manuelt, da der ikke er værktøjer til automatisk at generere konfigurationen fra DLS'en. Findes der fejl eller uhensigtsmæssigheder under konfigurationen i DLS'en, sendes denne frem og tilbage, indtil fejl er udbedret, hvilket gør udarbejdelsen og implementeringen af de enkelte DLS'er til en tidskrævende proces.

Det er muligt at anvende udviklerrammeverket til at konfigurere ændringer til eller nye tjenester. Men når den aftestede konfiguration etableret med udviklerrammeværket skal overføres til produktionsmiljøet, skal den ovennævnte proces med DLS gennemføres på tilsvarende vis. DLS skal udfyldes manuelt, da udviklerrammeværket ikke indeholder funktionalitet til at generere denne.

Den manuelle og tunge proces kompliceres yderligere i de tilfælde, hvor en tjeneste omfatter data fra et andet registre. Her er der behov for at gennemføre en parthøring mellem registermyndighederne for at sikre, at data anvendes korrekt, f.eks. ved anvendelse af CPR-data. Det opleves, at "time-to-market" ved konfigurerings af tjenester forsinkes af de manuelle processer.

Den moderniserede Datafordeler

Der er behov for, at den moderniserede Datafordeler tilbyder, at konfigurerings samt sletning af tjenester foregår uafhængig af Netcompany gennem konfiguration og ikke egentlig udvikling. Denne konfigurerings skal kunne ske automatiseret ud fra templates eller lignende. Derved gøres procesunderstøttelse af arbejdsgange smidig og sikrer kortest muligt "time-to-market".

3.13.2 Backlog

ID	Beskrivelse
13.01	Den moderniserede Datafordeler skal ved en standardiseret proces kunne oprette, klonе, ændre og slette en tjeneste i alle miljøer (f.eks. webservice, filudtrækstjeneste, hændelsesbeskedtjeneste og attributter tilknyttet disse), således at datafordelermyndigheden opnår en økonomisk og tidsmæssig besparelse ved konfiguration af tjenester.
13.02	Datafordelermyndigheden og registermyndigheden skal ved en standardiseret proces kunne oprette, klonе, ændre og slette en tjeneste i et testmiljø, således at de kan arbejde med konfiguration af tjenester uafhængig af Netcompany.
13.03	Datafordelermyndigheden og registermyndigheden skal kunne gemme en tjenestes konfiguration, således at konfigurationer kan genbruges ved fremtidigt design af tjenester.
13.04	Datafordelermyndigheden og registermyndigheden skal ud fra gemte konfigurationer kunne oprette nye tjenester i testmiljøet, således at design og implementeringsprocessen forenkles ved udvikling af nye tjenester.
13.05	brugere skal kunne tilgå dokumentation på nye samt ændrede tjenester, inden de sættes i produktion, således at brugere tidligt kan forberede sig på nye eller ændrede tjenester.
13.06	Datafordelermyndigheden skal kunne tildele autorisation til konfiguration af tjenester i testmiljøet.
13.07	Konfiguration af nye eller ændrede tjenester skal indgå i release-processen på en automatiseret måde.
13.08	Som datafordelermyndighed og registermyndighed ønsker man en hurtig og smidig proces omkring konfiguration af tjenester, således at indsatsen primært ligger hos Netcompany, og det samlede ressourceforbrug er minimalt.

[Tilbage til oversigt](#)

3.14 MB14: Optimering af filudtræk

3.14.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Den nuværende Datafordeler tilbyder brugere tre typer filudtræk, der er defineret som:

”Prædefinerede”

Et sæt af inputparametre for et datasæt ligger fast. Disse parametre anvendes til at fremsøge i de foranderlige data for at fremstille en leverance, som kan sendes til alle abonnenter på dette datasæt. Der kan i et abonnement specificeres en daglig, ugentlig eller månedlig leverance. For udvalgte prædefinerede filudtræk leveres disse på aftalte tidspunkter.

”Prægenererede”

Der er tale om statiske datasæt (”filer”), som forbliver urørt af Netcompany i både modtagelse og hele datasættets levetid. Filer distribueres til bruger ”as-is” ved, at datasættet bliver sendt fra registrene, kanaliseret (urørt) igennem den nuværende Datafordeler og lagt ud til download. Den nuværende Datafordeler fungerer således som et repository/opbevaringssted, hvor Anvenderne henter filudtræk af datasæt, som registrene laver. Denne form for filudtræk er relevant for historiske data og ikke besøgt hyppigt.

”Brugerdefineret”

Der er tale om enten et abonnement eller en ”ad-hoc”-bestilling, der specificeres af brugeren, og hvor det genererede udtræk sendes til den specifikke bruger. Specifikationen går ud på at angive værdier for mulige input-parametre for filudtrækket. Ved abonnement produceres filudtrækket dagligt, ugentligt eller månedligt, mens det sker straks ved ”ad-hoc”-bestilling.

De prædefinerede filudtræk anvendes ofte til at opbygge kopiregistre, hvor brugerne har hurtig og stabil adgang til data fra den nuværende Datafordeler (typisk de tabulære registre), f.eks. et centralt internt system hos en myndighed, som sørger for ensartet adgang til den nuværende Datafordelers data. Det er en slags ”intern” datafordeler hos myndigheden, hvor det er vigtigt, at data er så aktuelle som muligt.

I den nuværende Datafordeler betragtes alle brugerdefinerede filudtræk som unikke. Hvis tre brugere specificerer det samme brugerdefinerede filudtræk, vil der blive genereret tre identiske udtræk. Hvis der samtidig er opsat dagligt abonnement, vil der hver dag blive genereret tre identiske udtræk.

Denne håndtering af brugerdefinerede filudtræk på den nuværende Datafordeler er både ressourcekrævende og skaber flaskehalse, når de brugerdefinerede filudtræk eksekveres, da de ofte bliver udtrukket på det samme tidspunkt. Tidspunkt for generering angives af brugeren, hvilket typisk bliver bestillingstidspunktet. Mange brugerdefinerede udtræk vil aflede det store ressourcetræk.

For de prædefinerede filudtræk er der med den afgivende leverandør aftalt et natligt udtræksvindue, hvor disse bliver genereret. Aftalen skal sikre, at de omfattede filudtræk er færdige på bestemte tidspunkter. Disse tidspunkter er afstemt med brugerne.

Den moderniserede Datafordeler

Datafordelermyndigheden har brug for, at filudtræk optimeres (prædefinerede, prægenererede og/eller brugerdefinerede), samt at registermyndigheder kan distribuere store statiske datasæt med et minimalt forbrug af ressourcer med henblik på at sikre sikker og stabil drift.

3.14.2 Backlog

ID	Beskrivelse
14.01	Som datafordelermyndighed eller registermyndighed skal man selv kunne konfigurere prædefinerede filudtræk, således at det kan gøres uafhængig af Netcompany.
14.02	Som datafordelermyndighed skal man på en nem og enkel måde kunne identificere ens eller tæt på ens brugerdefinerede filudtræk, således at der kan oprettes et tilsvarende prædefineret filudtræk, der kan anvendes i stedet af brugerne.
14.03	Den moderniserede Datafordeler skal kunne identificere aktive filudtræk, der ikke anvendes, og gøre dem inaktive.
14.04	Som bruger skal man kunne deaktivere et filudtræk, således at den moderniserede Datafordeler ikke belastes unødigt.
14.05	Den moderniserede Datafordeler skal kunne stille brugerdefineret filudtræk til rådighed senest på det af brugerne angivne leverancetidspunkt.
14.06	Som anvenderorganisation skal man gennem konfiguration kunne definere egne filudtræk, der giver fuld mulighed for at specificere, præcist hvilke data der udtrækkes, således at bitemporalitet kan understøttes.

[Tilbage til oversigt](#)

3.15 MB15: Optimering af funktionalitet til selvbetjening

3.15.1 Beskrivelse af hovedområde

Ved selvbetjening i dette punkt forstås de grænseflader i Datafordeleren, hvor brugerne kan betjene sig selv.

Situationsbeskrivelse

Den nuværende Datafordelers selvbetjening indbefatter portaler til dokumentation, selvbetjenings- samt administrationsportal.

Selvbetjenings- samt administrationsportal

Selvbetjenings- samt administrationsportal er anvenderorganisationernes og registermyndighedernes indgang til den nuværende Datafordeler og dækker blandt andet brugen af produktkatalog, bestilling af leverancer, vejledninger, FAQ, brugeradministration samt konfigurerings. Det er en webapplikation, udviklet som kundespecifikt programmel, som primært er kodet i AngularJS. Den egenudviklede kode bevirker, at applikationen i relation til den nuværende Datafordeler er krævende at vedligeholde, bl.a. navngivninger, metadata, attributter mm. Simple ændringer, som at ændre navnet på en tjeneste eller et register, skal bestilles som en udviklingsopgave hos den afgivende leverandør.

Den nuværende Datafordelers selvbetjeningsportal tilbyder i dag brugerne følgende funktionalitet:

- Begrænset listning af tilgængelige tjenester
- Ansøgningsvejledninger på datafordeler.dk (For ansøgere, der ønsker adgang til registre med fortrolige og personfølsomme data. Disse vejledninger henviser til manuelle processer hos registermyndighederne.)
- Vise metadata for data og tjenester
- Oprettelse af brugere
- Abonnement og enkelt udtræk af filudtræk
- Abonnement på hændelser

Administrationsportalen tilbyder i dag følgende funktionalitet:

- Tildeling af brugeradgang til den pågældende registermyndigheds data
- Se registermyndighedernes egne datamodeller, replikeringskanaler og tjenester
- Konfiguration af registermyndighedernes tjenester

Dokumentation

Til dokumentation anvendes standardprogrammelt Confluence samt Datafordeler.dk, der er en egenudviklet webapplikation baseret på CMS systemet Umbraco (skrevet i C#).

Confluence tilbyder i dag brugerne følgende funktionalitet:

- Dokumentation
- Brugervejledning
- User stories
- Servicekatalog
- Tværgående søgefunktion

Datafordeler.dk tilbyder i dag brugerne følgende funktionalitet:

- Produktkatalog
- Driftsinformation og support

Brugere kan ikke få hjælp og vejledning direkte i selvbetjenings- samt administrationsportalen. Hjælp og vejledning findes i en tredjepartsdokumentationsløsning, som er baseret på "Confluence". Brugere skal skifte mellem den nuværende Datafordelers administrations-portal, dokumentationsløsningen og datafordeler.dk for at se produktkataloget.

Den moderniserede Datafordeler

I den moderniserede Datafordeler ønskes det, at anvendelsen af selvbetjening opleves som et samlet hele. Derudover skal selvbetjening optimeres til understøttelse af den brugervendte funktionalitet samt minimere Datafordelerens kompleksitet over for brugeren.

3.15.2 Backlog

ID	Beskrivelse
15.01	Anvenderorganisationerne har behov for via en brugergrænseflade at følge en styret proces med hjælpefunktioner til sammensætning af URL'er til tjenester, således at det bliver nemt for anvenderorganisationerne at tilgå data fra den moderniserede Datafordeler.
15.02	Anvenderorganisationerne har behov for, at inputfelter i formularer i selvbetjenings- og administrationsportalen automatisk udfyldes med kontekstrelevant data (f.eks. tjenestebrugerens navn) således, at det letter konfigurationsarbejdet.
15.03	Anvenderorganisationerne har behov for, at inputfelter i formularer, der vises i selvbetjenings- og administrationsportalen, giver mulighed for at vælge parametre fra en liste med relevant indhold, således at det letter konfigurationsarbejdet.
15.04	Som brugere skal man kunne anvende såvel en brugergrænseflade som en tjeneste til at tilgå selvbetjeningsfunktioner på den moderniserede Datafordeleren, således at brugere kan tilgå selvbetjeningsfunktioner på den måde, der passer deres behov bedst.
15.05	Som bruger har man behov for at møde en brugervenlig og navigerbar selvforklarende grænseflade, således at man kan tilgå dokumentation i den aktuelle selvbetjeningskontekst.
15.06	Som datafordelermyndighed har man behov for at kunne administrere registermyndighedernes og anvenderorganisationernes adgang og rettigheder, således at datafordelermyndigheden kan styre adgang til data uden involvering af Netcompany.
15.07	Som registermyndighed er der behov for at kunne oprette og administrere anvenderorganisationernes adgang og rettigheder til at hente registerets data, således at registermyndighederne kan styre adgang til data uden involvering af Netcompany.
15.08	Som registermyndighed har man behov for at se en oversigt over brugere og brugeregenskaber, således at man letter administrationen ved tildeling af brugeradgange.
15.09	Som anvenderorganisationen har man behov for fra én og samme søgefunktion at kunne fremsøge den nødvendige information om registre, indhold i tjenesterne, attributterne mm., uagtet hvorvidt informationen er spredt over eventuelle underliggende selvbetjenings- og -dokumentationsplatforme, således at anvenderorganisationen kan fremsøge al information fra et og samme sted.
15.10	Som bruger har man behov for at kunne administrere brugere på samme tid, således at man nemt og hurtigt kan tildele rettigheder til mange brugere.
15.11	Som bruger skal man kunne abonnere på automatisk besked om status og fremtidige ændringer af tjenester, således at man er opdateret om mulige driftsforstyrrelser.
15.12	Som bruger skal man fra den moderniserede Datafordeler kunne ansøge om adgang til et register med fortrolige eller personfølsomme data, således at ansøgningen automatisk oprettes i og behandles fra den moderniserede Datafordelers administrationsportal.

[Tilbage til oversigt](#)

3.16 MB16: Optimering af understøttelse af kopiregistre

3.16.1 Beskrivelse af hovedområde

Ved kopiregistre forstås en lokal oprettet kopi af alle grunddata eller dele af dem, hvor man som anvenderorganisation vedligeholder kopi af registerdata med opdateringer fra Datafordeleren. Det kan være alle data fra et register, en delmængde af data fra et register eller kombinationer af data fra forskellige registre.

Situationsbeskrivelse

Den nuværende Datafordeler er oprindeligt designet ud fra princippet om, at kopiregistre ikke skal understøttes, da alle anvenderorganisationer vil hente data online via de udstillede tjenester.

Det har efterfølgende vist sig, at der er forretningsbehov og forretningsmodeller hos anvenderorganisationer, der tilsiger, at der er behov for at understøtte oprettelse og vedligeholdelse af kopiregistre.

Oprettelse af kopiregister

Oprettelse af et kopiregister er understøttet på den nuværende Datafordeler ved, at man som anvenderorganisation kan hente et totaludtræk af et eller flere registre. Dette sker ved at hente et filudtræk, der indeholder alle data for et register.

Ønsker man som anvenderorganisation flere registre, skal der hentes et filudtræk pr. register.

Vedligeholdelse af kopiregister

Vedligeholdelse af data i et kopiregister kan i princippet implementeres med to forskellige teknologier i forhold til de muligheder som den nuværende Datafordeler tilbyder:

- Modtage information om opdateringer via hændelsesbeskeder
- Modtage deltaudtræk via filudtræk, hvor deltaudtræk indeholder ændringer siden sidste deltaudtræk

Den tekniske mulighed for anvenderorganisationer for at modtage information om ændringer via hændelsesbeskeder er at opsætte abonnement på alle hændelsesbeskeder for et register. Når man derefter modtager en hændelsesbesked om ændringer til data, skal man via REST-tjenesterne hente de opdaterede data.

Denne opdateringsmodel kræver flere ting:

- Der skal opsættes abonnement på alle hændelsesbeskeder
- Anvenderorganisationen skal implementere logik til at behandle hændelsesbeskeder og hente de opdaterede data på de relevante tjenester, når der modtages en hændelsesbesked. Anvenderorganisationen skal selv styre dette.
- Registret skal have defineret hændelsesbeskeder for alle data i registret. Bemærk, at for hvert dataobjekt er der tre forskellige hændelsesbeskeder (Create, Update, Delete).
- Registret skal opmærke alle opdateringer med den rigtige type hændelsesbesked.
- Anvenderorganisationer skal have kendskab til den enkelte registermyndigheds forretningsdomæne.

Hvis ikke alle ovenstående betingelser er til stede, vil det ikke være muligt at vedligeholde et konsistent kopiregister. Det har vist sig, at der ikke er nogen registre, hvor definition af hændelsesbeskeder er komplet for alle data i registret.

Samtidig viser erfaringen, at der går hændelsesbeskeder tabt på den nuværende Datafordeler. Det er derfor ikke muligt med den nuværende Datafordeler at holde et kopiregister opdateret og konsistent ved hjælp af hændelsesbeskeder, uden at man som anvenderorganisation skal genopfriske data i kopiregistret gennem et totaludtræk.

Den nuværende Datafordeler tilbyder også vedligeholdelse af et kopiregister gennem delta-filudtræk ved at anvende "since previous". Dette gør, at man udtrækker data, der er opdateret siden sidste gang, udtrækket kørte. Afhængig af øvrige udtræksparametre, der anvendes for delta-udtrækket, kan der opstå situationer, hvor dataopdateringer ikke medtages i delta-udtrækket. Som resultat heraf vil kopi-registret ikke længere være konsistent med data på den Nuværende datafordeler.

Specielt anvendelse af parametre, der arbejder på de bitemporale dimensioner, giver udfordringer med delta-udtræk. Også her er totaludtræk midlet til at udbedre dataproblemerne. En enkelt anvenderorganisation har implementeret et setup, hvor der hentes daglige totaludtræk, og hvor man selv finder de data, der er ændret siden sidste totaludtræk. Det anses dog ikke for en holdbar løsning, at alle, der ønsker at vedligeholde et kopiregister, skal implementere dette setup.

Større anvenderorganisationer har grundet ovenstående i dag svært ved at opretholde et fuldstændigt kopiregister, og data er derfor ofte ude af sync i deres lokale kopiregister.

Den moderniserede Datafordeler

I den moderniserede Datafordeler ønskes der funktionalitet, som gør det muligt for anvenderorganisationer at etablere en nær-realtids opdateret kopi af udvalgte grunddata. Det kan være alle grunddata eller udvalgte dele. De anvendende applikationer kan være fra anvenderorganisationer eller tjenester i den moderniserede Datafordeler, der udstiller deres specifikke grunddata over for deres anvendere.

Behovet er en funktionalitet, som dels kan anvendes eksternt af anvenderorganisationer som andre tjenester i den moderniserede Datafordeler, dels kan anvendes internt i den moderniserede Datafordeler til at understøtte den ønskede modulære arkitektur. Et eksempel kunne være en WMS tjeneste, der udstiller husnumres beliggenhed. Denne tjeneste har (ud over geodata, der anvendes som baggrundskort) kun behov for en nær-realtids opdateret kopi af følgende data fra Danmarks Adresseregister (DAR): Husnummeret samt husnummerets koordinater.

Funktionalitet til oprettelse og vedligeholdelse af kopiregistre skal være en del af den moderniserede Datafordelerens modulære arkitektur og kunne anvendes internt i Datafordeleren.

Det skal være muligt at udstille både aktuelle og historiske data, så anvenderorganisationer har mulighed for at etablere en kopi af aktuelle data eller en kopi af data med historik. Funktionaliteten skal endvidere kunne anvendes til at downloade en enkelt tabel fra et grunddataregister.

3.16.2 Backlog

ID	Beskrivelse
16.01	Som anvenderorganisation ønsker man at kunne oprette og vedligeholde en lokal, eksakt kopi af data fra den moderniserede Datafordeler, således at anvenderorganisationen kan anvende data lokalt.
16.02	Som anvenderorganisation ønsker man, at lokale kopier af data løbende bliver opdateret, således at data i den lokale kopi er identiske med data i den moderniserede Datafordeler.
16.03	Som anvenderorganisation ønsker man at kunne vælge opdateringsfrekvensen mellem den moderniserede Datafordeler og den lokale kopi, således at synkronisering sker efter anvenderorganisationens behov.
16.04	Som anvenderorganisation skal man kunne udvælge delmængder af data fra den moderniserede Datafordeleren, som er relevant for den pågældende anvenderorganisation, således at anvenderorganisationen kun har de data, der er relevante for anvenderorganisationens forretning.
16.05	Som anvenderorganisation skal man kunne oprette samt vedligeholde lokale kopier af data uden domæneviden om det pågældende register, således at anvenderorganisationen kan vedligeholde kopiregistre uden domæneeksportise.
16.06	Som anvenderorganisation skal man kunne oprette samt vedligeholde lokale kopier af data med historik som i registeret eller af data valgt ud fra temporale eller bitemporale egenskaber i data, således at anvenderorganisationen kun får den grad af historik, der er relevant for forretningen.

[Tilbage til oversigt](#)

3.17 MB17: Optimering af håndtering af metadata

3.17.1 Beskrivelse af hovedområde

Situationsbeskrivelse

Den nuværende Datafordeler indeholder en metadatadatabase, og der er visning af metadata fra metadatabasen på den nuværende Datafordelers administrations-portal, jf. situationsbeskrivelse i punkt 3.15. I dag angives metadata i dataleverancespecifikationerne (DLS), dvs., at oprettelse og opdatering af metadata sker ved udfyldelse af regneark i en dataleverancespecifikation (DLS) og i datamodellerne (UML-modeller i XMI-filer). Der er kun metadata for registre, filudtræk og tjenester på den nuværende Datafordeler i et begrænset omfang.

Metadata for registre, filudtræk og tjenester bliver vedligeholdt uden for den nuværende Datafordelers metadatadatabase og udstillet på Confluence og datafordeler.dk, jf. situationsbeskrivelse i punkt 3.15. Desuden er der metadata for registre, som er udstillet på Geodata-info.dk og i Det fællesoffentlige datasætkatalog (findes på <https://datasets.catalogue.data.gov.dk/>). Begge udstiller metadata for en lang række andre data end data på den nuværende Datafordeler. Geodata-info er den nationale danske geoportal, der udstiller metadata for geodatasæt og tjenester, for både data omfattet af EU's INSPIRE-direktiv og andre data. Det fællesoffentlige datasætkatalog indeholder metadata for offentlige datasæt fra forskellige organisationer.

En ekstra detalje specifikt for filudtræk på FTP/SFTP-serveren er, at de altid er ledsaget af en JSON/XML-fil med genereret metadata om det konkrete filudtræk.

Den moderniserede Datafordeler

Metadata kan være på tre forskellige niveauer, som kan specificeres ved brug af gældende internationale standarder:

- **Discovery:** metadata til fremsøgning af tilgængelige data
- **Evaluation:** metadata til evaluering af brugbarhed, f.eks. ud fra datakvalitet
- **Use:** metadata til at tilgå data

Der er behov for, at metadata kan tilgås via tjenester, således at metadata kan hentes af eksterne metadataportaler, og at der er mulighed for at kombinere metadata fra den moderniserede Datafordeler med metadata fra andre platforme.

Den moderniserede Datafordeler skal kunne håndtere metadata for både geografiske og ikke-geografiske data. Dvs., at den moderniserede Datafordeler skal kunne håndtere registermyndigheders forskelligartede behov for at udstille metadata for brugerne på alle tre niveauer uanset typen af data. Dette således, at registermyndigheder kan få udstillet metadata for brugere uanset typen af data.

Der er behov for:

- Oprettelse og vedligeholdelse af metadata er enkelt for registermyndighederne og ikke kræver gentagne indtastninger af de samme metadata.
- At der er overensstemmelse mellem tjenestemetadata² og andre metadata.
- At metadata udstilles gennem standardtjenester i den moderniserede Datafordeler.

3.17.2 Backlog

ID	Beskrivelse
17.01	Den moderniserede Datafordeler skal kunne integreres med registermyndigheders systemer til administration af discovery metadata.
17.02	Den moderniserede Datafordeler skal give brugerne metadata på tre forskellige niveauer, som alle kan specificeres ved brug af gældende internationale standarder: <ul style="list-style-type: none"> • Discovery: til fremsøgning af tilgængelige data • Evaluation: til evaluering af brugbarhed, f.eks. ud fra datakvalitet • Use: til at tilgå data
17.03	Den moderniserede Datafordeler skal kunne håndtere metadata for både geografiske og ikke-geografiske data.
17.04	Som registermyndighed skal man kunne administrere metadata på den moderniserede Datafordeler, således at registermyndigheden ikke behøver et system til administration af metadata for data distribueret på den moderniserede Datafordeler.
17.05	Den moderniserede Datafordeler skal sikre overensstemmelse mellem tjenestemetadata og andre metadata.
17.06	Som bruger skal man let kunne finde opdaterede metadata for datasæt og tjenester på den moderniserede Datafordeler, således at brugeren hurtigt får retvisende metadata.
17.07	Den moderniserede Datafordeler skal ledsage filudtræk med oplysninger omkring de konkrete filer.
17.08	Den moderniserede Datafordeler skal udstille metadata om tjenester på en struktureret måde, som gør det muligt at implementere alternative katalogfunktioner.

[Tilbage til oversigt](#)

² tjenestemetadata findes f.eks. i GetCapabilities-svar fra tjenester med geodata, der følger nuværende OGC-standarder.