



Styrelsen for Dataforsyning
og Infrastruktur

Transitionsguide for fildownloads samt etablering og vedligeholdelse af kopiregister

Drift og modernisering af Datafordeleren

Maj 2024

Version 0.9 – 23-05-2024



Indholdsfortegnelse

1	Introduktion.....	3
1.1	Begreber	3
2	Fildownloads.....	4
2.1	Typer af fildownloads	5
2.2	Filformater	6
2.3	Udtræksnumre.....	6
2.3.1	Sammenhæng mellem total- og deltaudtræk	7
2.3.2	Spring i udtræksnumre	7
2.4	Navngivning af fildownloads.....	8
2.5	Skema for generering af fildownloads.....	9
2.6	Sådan hentes fildownloads.....	10
2.6.1	Metode: GetAvailableFiledownloads	10
2.6.2	Metode: GetFile	12
3	Etablering og vedligeholdelse af kopiregister ved brug af fildownloads	14
4	Transition fra filudtræk til fildownloads	16
4.1	Indhold i filer.....	16
4.1.1	Forskelle i GML og GeoPackage filudtræk og fildownloads	17
4.2	Geografisk filtrering i fildownloads.....	18
4.3	Anvendereksempel på transition fra filudtræk til fildownloads.....	19
4.3.1	Situationsbeskrivelse.....	19
4.3.2	Transition.....	19
4.4	Anvendereksempel på transition med udgangspunkt i kopiregister.....	21
4.4.1	Situationsbeskrivelse.....	21
4.4.2	Transition.....	21
5	Bilag 1: FAQ.....	22
6	Bilag 2: Overførsel af filudtræk til fildownloads	23

1 Introduktion

Dette dokument er en guide i anvendernes transition fra de abonnementsbaserede filudtræk over til entitetsbaserede fildownloads.

Dokumentet er opdelt i to dele:

- 1) Første del af dokumentet beskriver hvad fildownloads er, samt hvordan funktionaliteten anvendes.
- 2) Anden del af dokumentet forklarer forskelle, samt hvordan anvendere i forskellige situationer kan komme i gang med at anvende fildownloads.

Dokumentet er tiltænkt følgende læsere:

- 1) Eksisterende anvendere af Datafordeleren der henter filudtræk på Datafordeleren, som ønsker at komme i gang med at bruge fildownloads i stedet for.
- 2) Nye anvendere, der vil i gang med at bruge fildownloads på Datafordeleren.

1.1 Begreber

I dette dokument bruges begreberne i Tabel 1. Bemærk at begreberne også er yderligere forklaret i løbet af dokumentet med eksempler.

Begreb	Beskrivelse
Nuværende Datafordeler	
Filudtræk:	Filudtræk er en metode at få udtræk af data fra Datafordeleren. Man kan som anvender logge på selvbetjeningen og oprette et abonnement på et specifikt filudtræk – evt. med brugerdefinerede parametre, hvorefter filen bliver genereret og er tilgængelig på en FTP-server. Filudtræk er beskrevet her .
Moderniserede Datafordeler	
Fildownload:	Fildownloads beskrives i denne vejledning.
Entitet:	En entitet er en fysisk repræsentation af et registerobjekt som findes i Grunddatamodellen . De total- og deltdownloads der kan hentes fra Datafordeleren, indeholder hver især data bestående af en enkelt entitet.
Totaldownload:	Et totaldownload er et udtræk af alle data for en enkelt entitet og refererer til en enkelt fil.
Deltadownload:	Et deltadownload er et udtræk af alle data der har ændret sig, siden forrige deltadownload blev lavet, for en enkelt entitet. Et deltadownload består af en enkelt fil.
Totaludtræk:	Et totaludtræk er et udtræk af <i>alle</i> data i et register eller en delmængde af entiteter fra et register. Et totaludtræk er udgjort af en eller flere totaldownloads. Et totaludtræk kan således bestå af en eller flere filer med tilsvarende entiteter, alt afhængigt af hvor stor en delmængde af et register der er tale om. Bemærk at denne definition kun gælder for fildownloads
Deltaudtræk:	Et deltaudtræk er et udtræk af alle data der har ændret sig, siden forrige deltaudtræk blev lavet, i et register eller en delmængde af entiteter fra et register. Et deltaudtræk er udgjort af et eller flere deltdownloads. Et deltaudtræk kan således bestå af en eller flere filer med tilsvarende entiteter, alt afhængigt af hvor stor en delmængde af et register der er tale om.

Tabel 1: Begrebsliste.

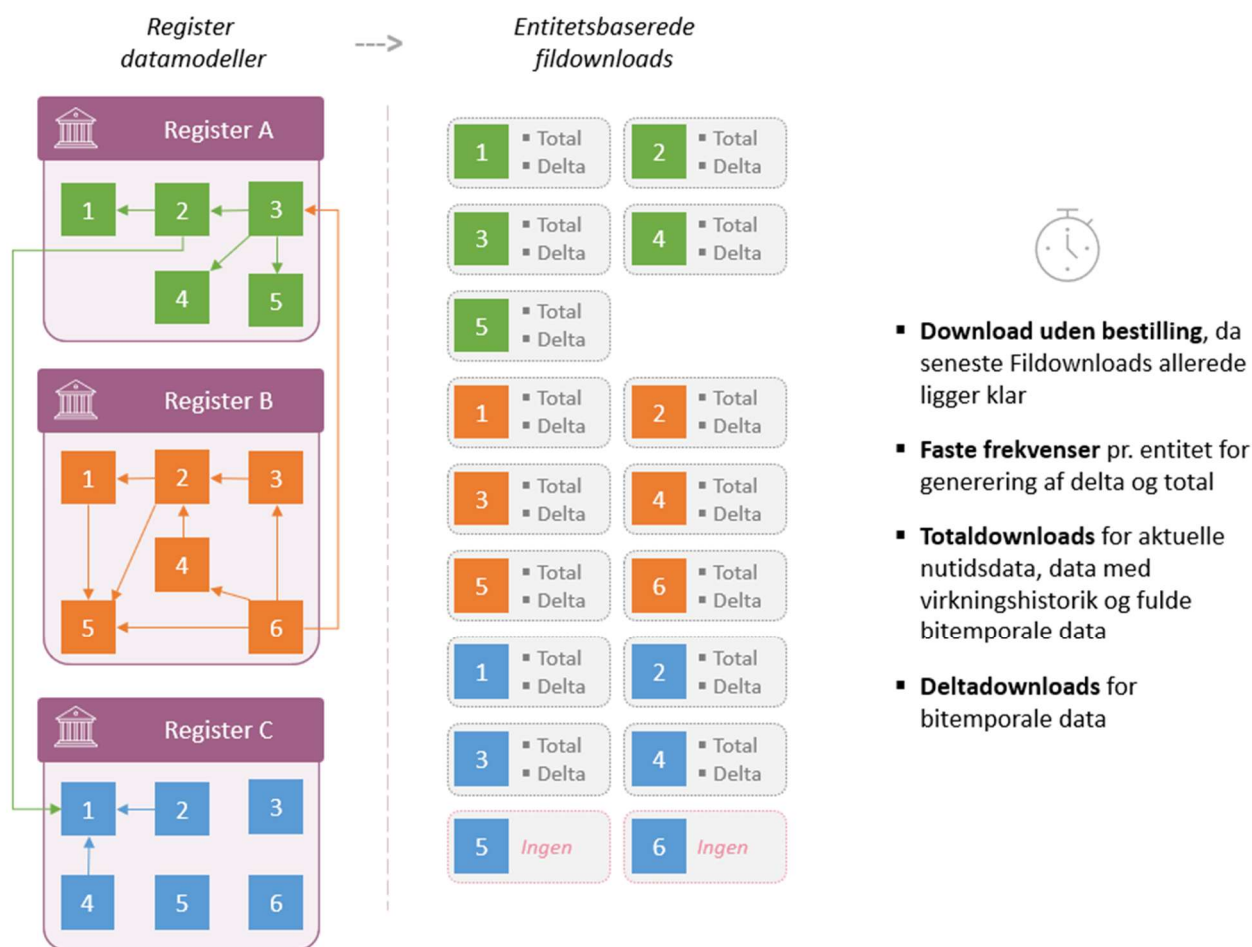
2 Fildownloads

Fildownloads er en funktionalitet på Datafordeleren til udstilling af data, men skal ikke misforstås som værende det samme som de abonnementsbaserede filudtræk, der også er på Datafordeleren. Fildownloads tilbydes som *prædefinerede entitetsbaserede fildownloads* og baserer sig på de datamodeller som registrene har på Datafordeleren.

En entitet svarer til en tabel i Grunddatamodellen som er defineret i registrenes datamodel:

<https://grunddatamodel.datafordeler.dk/>. En entitet kan eksempelvis være en "Adresse" fra DAR, en "Bygning" fra BBR eller et "Jordstykke" fra MAT. Det at fildownloads er *entitetsbaserede* betyder at et fildownload indeholder data for en enkelt type af entitet. Der leveres entitetsbaserede fildownloads på tabulære data og ikke for rasterdata (billeddata). Tabulære data forstås som strukturerede data der kan opbevares i tabeller som f.eks. adresser men ikke billedtyper som skærmkort.

Fildownloads genereres løbende i bestemte intervaller for en række registre, hvorefter filerne kan hentes gennem et REST-API, der udstiller de genererede fildownloads. Det eneste krav til at hente fildownloads er, at man har en tjenesteb bruger med brugernavn og adgangskode. Fildownloads bruger ikke abonnementer, da de allerede ligger klar til at blive hentet. REST-API'et beskrives nærmere i afsnit 2.6, mens skema for generering samt frekvenser for generering beskrives nærmere i afsnit 2.5.



Figur 1: Opsummering af fildownloads-funktionaliteten.

En opsummering hvad fildownloads består af, kan ses i Figur 1. Figuren viser 3 registre:

- **Register A**, som har 5 entiteter. Registerets entiteter refererer til hinanden indbyrdes indenfor registeret og har også en enkelt reference til en entitet i register C udenfor registeret.
- **Register B**, som har 6 entiteter. Registerets entiteter refererer til hinanden indbyrdes indenfor registeret og har også en enkelt reference til en entitet i register A udenfor registeret.
- **Register C**, som har 6 entiteter. Registerets entiteter refererer til hinanden indbyrdes indenfor registeret.

Registrenes entiteter har alle fildownloads, som både består af totaldownloads og deltdownloads. Det skal her bemærkes at Register C har fravalgt at der bliver genereret fildownloads for deres entitet nr. 5 og 6 i dette eksempel.

2.1 Typer af fildownloads

Datafordeleren udstiller flere typer af entitetsbaserede fildownloads. Der findes 2 typer af fildownloads, nemlig totaldownloads og deltdownloads (se Tabel 1). Datafordeleren opbevarer ét totaldownload pr. type pr. entitet, mens der opbevares deltdownloads 14 dage fra de blev genereret.

Hvis der genereres fildownloads for en entitet, bliver der altid genereret både totaldownloads og deltdownloads. Udover at der findes flere typer af fildownloads, findes der også flere typer af filtreringer på den data som de genererede fildownloads indeholder. Dette skyldes at delmængder af data på Datafordeleren er bitemporal og at der dermed kan være flere forekomster af det samme objekt over tid i to tidsdimensioner¹.

- **Fuldt bitemporalt fildownload:** Et fuldt bitemporalt fildownload inkluderer alt, hvad angår både registreringstid og virkningstid.
- **Temporalt fildownload:** Et temporalt fildownload filtrerer registreringsdimensionen fra, men beholder virkningsdimensionen.
- **Fildownload med aktuelle data:** Både registreringstiden og virkningstiden er sat til at være på det pågældende tidspunkt fildownload blev lavet.

Det er muligt at kombinere en type fildownload med en type af bitemporal filtrering og filformat. En opsummering af, hvilke fildownloads der eksisterer kan ses i Tabel 2.

Type af fildownload	Type af bitemporal filtrering	Typer af filformater
Totaldownload	Bitemporal	JSON, GML og GPKG
Totaldownload	Temporal	
Totaldownload	Aktuel	
Deltadownload	Bitemporal	

Tabel 2: Oversigt over understøttede typer af fildownloads på Datafordeleren.

¹ <https://confluence.sdfi.dk/display/DML/Grunddata+-+bitemporalitet>



2.2 Filformater

Tabel 3 nedenfor viser, hvilke registre der har fildownloads på deres entiteter, samt hvilke filformater der er tilgængelige for registrenes entiteter. Et kryds i en række i tabellen indikerer at filformatet generelt er understøttet for det givne registers entiteter.

Register	JSON	GPKG	GML
BBR	X		
DAR	X		
EBR	X		
MAT2	X	X	X
DAGI	X	X	X
DHM Oprindelse	X	X	X
DHM Højdekurver	X	X	X
Danske Stednavne	X	X	X
GeoDanmark Vektor	X	X	X
FIKSPUNKT	X	X	X

Tabel 3: Oversigt over understøttede filformater per register.

Det skal bemærkes det ikke er alle entiteter for et register der har fildownloads i GPKG og GML-format, selvom dette er indikeret i Tabel 3. Der skal gøre sig følgende gældende for at en entitet understøttes i GPKG eller GML-format:

- GPKG-formatet understøttes på en entitet hvis entiteten indeholder 0 eller 1 geometrier.
- GML-formatet understøttes på en entitet hvis entiteten indeholder minimum 1 geometri.

2.3 Udtræksnumre

Udtræksnumre er en fortløbende talrække der beskriver tidsmæssig sammenhæng mellem data i et register. Hver gang der genereres fildownloads for et register, påsættes alle fildownloads ved samme generering det samme udtræksnummer. Dette bevirker at fildownloads med det samme udtræksnummer indeholder konsistente data indenfor registerets egen data. Denne egenskab er særlig relevant hvis man som anvender er interesseret i at sammensætte data fra flere entiteter i et register.

Det er her vigtigt at understrege, at udtræksnumrene kun er fortløbende *per register*, og at hvert register derfor har sin egen fortløbende talrække, der er uafhængig af andre registre. Udtræksnumre kan derfor ikke garantere referentiel integritet på tværs af registrenes data, da dette afhænger af registrenes indlæsningstidspunkter, samt indlæsningsfrekvens.

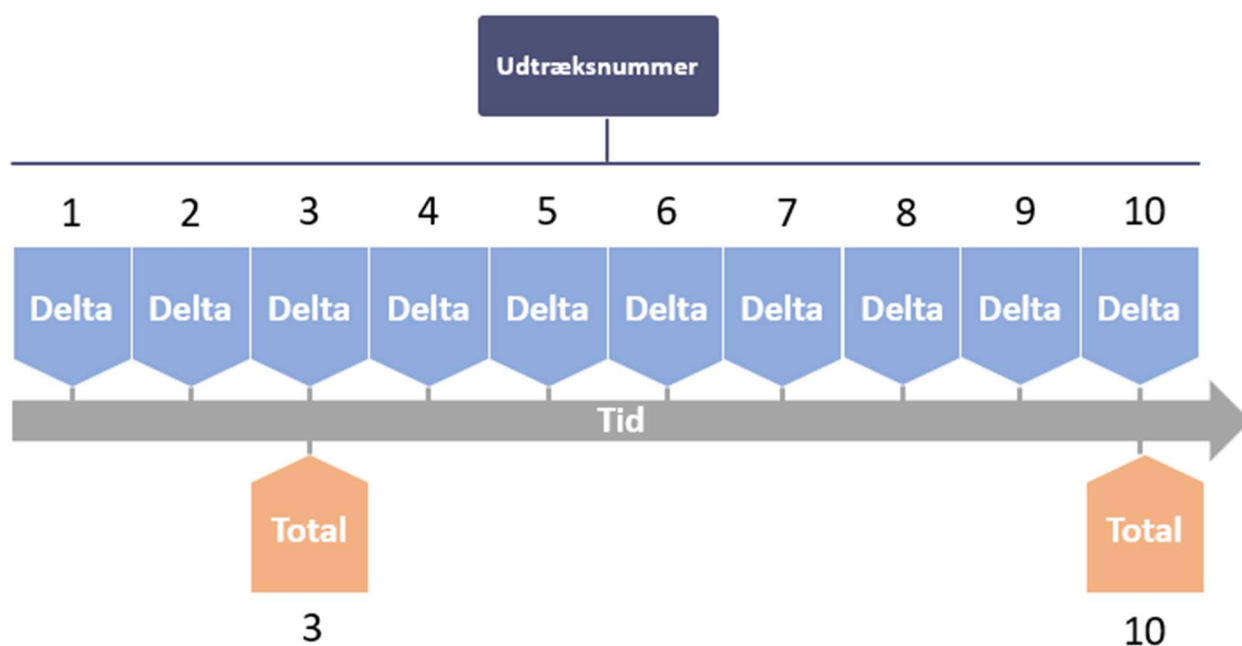


2.3.1 Sammenhæng mellem total- og deltaudtræk

Det er muligt at vedligeholde en lokal kopi af registrenes data som findes på Datafordeleren ved at hente totaludtræk – og herefter løbende hente deltaudtræk.

Kopien etableres først lokalt, ved at hente totaldownloads for den/de entiteter der er interesse for, for at skabe totaludtrækket. Idet totaludtrækket er hentet, skal det bemærkes hvilket udtræksnummer totaludtrækket har, da udtræksnummeret skal bruges til at identificere hvilke deltadownloads der er relevante at hente, for at skabe det samlede deltaudtræk der kan opdatere den initiale kopi af data. Et totaldownload og deltadownload med det samme udtræksnummer indeholder data fra samme generation. Dette betyder at deltaudtræk med samme udtræksnummer som et totaludtræk indeholder dermed ingen data som ikke også er indeholdt i totaludtrækket. Det udtræksnummer der skal bruges til at opdatere kopien af data, er dermed det næste udtræksnummer i talrækken.

Et eksempel på dette kan ses i Figur 2, som viser en række genererede deltaudtræk, samt et genereret totaludtræk. Hver generering har resulteret i, at der enten er blevet genereret et deltaudtræk og/eller et totaludtræk. Det kan ses på figuren, at der eksisterer både et totaludtræk, samt et deltaudtræk med udtræksnummeret 3 og 10. Såfremt det ønskes at etablere en kopi af data fra entiteten, kan totaludtrækket med udtræksnummeret 3 først hentes – og så herefter hente deltaudtrækkene med udtræksnummer 4, 5 og 6 for at vedligeholde datakopien. Det er således ikke nødvendigt at hente deltaudtrækket med udtræksnummer 3, da deltaudtrækket blev genereret på samme tid som totaludtrækket.



Figur 2: Overblik over hvordan udtræksnumre hænger sammen mellem total- og deltaudtræk.

2.3.2 Spring i udtræksnumre

Dette afsnit er særligt relevant for anvendere, der implementerer programmel, som anvender udtræksnumre til at indhente fildownloads automatisk. Der kan i sjældne tilfælde forekomme et spring i udtræksnumre. Dette skyldes, at der eksisterer et nyere udtræksnummer for det pågældende register, og at generering af fildownloads, derfor, vil bruge det nyere udtræksnummer for at sikre konsistente data inden for registeret.

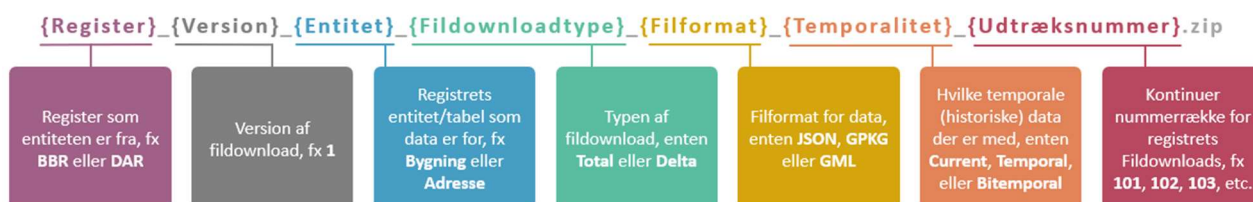


Processen for indhentning af deltaudtræk bør, på baggrund af ovenstående, følge følgende logik:

- 1) Tjek hvilket udtræksnummer der er påsat senest indhentede deltaudtræk.
- 2) Indhent deltaudtræk med udtræksnumre der er større end det sidst indhentede deltaudtræks udtræksnummer.
 - a. Dette sikrer at alle nyere deltaudtræk altid hentes uden at der springes nogle over, selvom der sker spring i udtræksnummeret på deltaudtrækkene.

2.4 Navngivning af fildownloads

De fildownloads der genereres på Datafordeleren følger alle en navnestandard, der hjælper med at beskrive hvad indholdet af filen er. Navnestandarden ser ud som set på Figur 3.



Figur 3: Overblik over navnestandard på fildownloads.

Et fildownload som følger navnestandarden kan derfor eksempelvis have navn som følgende:

1. **DAR_1_Adresse_TotalDownload_JSON_Bitemporal_10.zip**
2. **DAR_1_Adresse_DeltaDownload_JSON_Bitemporal_11.zip**
3. **BBR_1_Bygning_TotalDownload_JSON_Temporal_123.zip**
4. **MAT_1_Lodflade_TotalDownload_GPKG_Current_7.zip**
5. **DAGI_1_Kommuneinddeling_TotalDownload_GML_Current_42.zip**

Man kan ud fra navngivningen af ovenstående fildownloads udlede følgende:

1. Fildownload nr. 1 indeholder et fuldt bitemporalt totaldownload af Adresse-entiteten som ligger i DAR. Filen der downloades er en ZIP-fil, men ZIP-filen indeholder en JSON-fil. Filen har udtræksnummer 10.
2. Fildownload nr. 2 indeholder et fuldt bitemporalt deltadownload af Adresse-entiteten som ligger i DAR. Filen der downloades er en ZIP-fil, men ZIP-filen indeholder en JSON-fil. Deltafilen har udtræksnummer 11 og ville derfor indeholde alle ændringer der var sket for Adresse-entiteten siden udtræksnummer 10. Denne deltafil ville derfor kunne bruges til at opdatere en datakopi der havde indlæst fildownload nr. 1, som har udtræksnummer 10.
3. Fildownload nr. 3 indeholder et temporalt totaldownload af Bygning-entiteten som ligger i BBR. Filen der downloades er en ZIP-fil, men ZIP-filen indeholder en JSON-fil. Filen har udtræksnummer 123.
4. Fildownload nr. 4 indeholder et totaldownload med aktuelle data af Lodflade-entiteten som ligger i MAT. Filen der downloades er en ZIP-fil, men ZIP-filen indeholder en GPKG-fil. Filen har udtræksnummer 7.
5. Fildownload nr. 5 indeholder et totaldownload med aktuelle data af Kommuneinddeling-entiteten som ligger i DAGI. Filen der downloades er en ZIP-fil, men ZIP-filen indeholder en GML-fil. Filen har udtræksnummer 42.

2.5 Skema for generering af fildownloads

Skemaet i Tabel 4 viser, hvornår der genereres fildownloads for registre på Datafordeleren. Bemærk at et register kan have frabedt sig at få genereret fildownloads for bestemte entiteter i deres datamodel og vil således ikke blive genereret efter skemaet. Lige så snart et fildownload er genereret, vil det være tilgængeligt igennem REST-API'et, som udstiller fildownloads. REST-API'et beskrives nærmere i afsnit 2.6.

Register	Type af fildownload	Genereres hvornår (mellem kl. 3 og 6 om natten)
EBR	Totaldownload	Lørdag
DAR	Totaldownload	Lørdag
BBR	Totaldownload	Lørdag
MAT2	Totaldownload	Søndag
DAGI	Totaldownload	Søndag
GeoDanmarkVektor	Totaldownload	Søndag
Danske Stednavne	Totaldownload	Mandag
DHM Oprindelse	Totaldownload	Mandag
DHM Højdekurver	Totaldownload	Mandag
FIKSPUNKT	Totaldownload	Mandag
Alle registre	Deltadownload	Hver nat

Tabel 4: Oversigt over hvornår fildownloads bliver genereret.

Såfremt der genereres et deltadownload, men at der ikke er sket nogle dataændringer siden sidste gang der blev genereret et deltadownload for en pågældende entitet, vil deltadownloadet være tomt. Hvis der genereres et totaldownload uden at der er sket dataændringer siden sidste gang, det blev genereret, vil det nye totaldownload i JSON-, GML- eller GPKG-format have et ældre udtræksnummer i sit filnavn end den zip-fil den er indeholdt i.

Eksempelvis er det muligt at hente en fil med følgende navn:

- **DAR_1_Adresse_Total_JSON_Bitemporal_10.zip**

Her er udtræksnummeret på ZIP-filen 10, men ZIP-filen kan godt indeholde en JSON-fil med følgende navn:

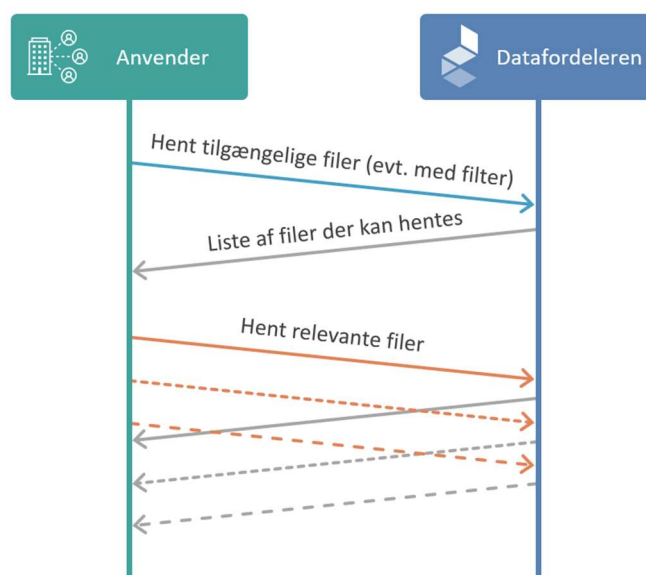
- **DAR_1_Adresse_Total_JSON_Bitemporal_9.json**

JSON-filen har et lavere udtræksnummer end ZIP-filen, da der ikke er sket nogen ændringer i data siden sidste totaludtræk blev genereret for Adresse-entiteten. Såfremt der heller ikke sker nogen ændringer i Adresse-entiteten inden næste generering, vil ZIP-filen få udtræksnummer 11, mens den indeholdte JSON-fil stadig vil have udtræksnummer 9.



2.6 Sådan hentes fildownloads

Fildownloads på Datafordeleren tilgås via et REST-API. Dette afsnit beskriver, hvordan det vil være muligt at danne et overblik over hvilke fildownloads der kan hentes, samt hvorledes relevante fildownloads herefter hentes. REST-API'et indeholder to metoder som kan kaldes. REST-API'et kræver at man først har oprettet en tjenestebruger med brugernavn og password.



Figur 4: Overblik over samlet anvendelsesmønster for fildownloads.

Figur 4 viser hvordan REST-API'et overordnet set anvendes. Her kaldes metoden "GetAvailableFileDownloads" først for at danne et overblik over mulige filer der er tilgængelige, hvorefter relevante filer hentes, ved at kalde metoden "GetFile".

2.6.1 Metode: GetAvailableFileDownloads

Metode: GetAvailableFileDownloads	
URL	https://<env>-api.datafordeler.dk/FileDownloads/GetAvailableFileDownloads?
HTTP-metode	GET
Headere i forespørgsel	Content-Type: application/json
Format	JSON
Returværdier	• Returnerer HTTP 200 – OK, ved succes. • Returnerer HTTP 400 – Bad Request, ved angivelse af forkerte parametre. • Returnerer HTTP 404 – Not Found, hvis den angivne ressource ikke kan findes. • Returnerer HTTP 500 – Internal Server Error, hvis der er sket en ukendt fejl.
Adgang	Tjenestebruger med brugernavn og password.
Dato- og tidsformat	Datoformat følger ISO 8601 Tidszonen er UTC



2.6.1.1 Parametre

Navn	Type	Beskrivelse	Valgfri?
Register	String	Begrænser responsen til kun at indeholde fildownloads fra det angivne register. Følgende registre kan angives: • DAR • DAGI • BBR • DHMOprindelse • DHMHoejdekurver • MAT (Der refereres her til Matriklen 2) • EBR • FIKSPUNKT • DS (Danske Stednavne) • GeoDKV	Ja
Version	String	Begrænser responsen til kun at indeholde fildownloads der er i den angivne version. Parameteret kan kun angives hvis "Register" parameteret er angivet.	Ja
Entity	String	Begrænser responsen til kun at indeholde fildownloads der er af samme type som den angivne entitet. Parameteret kan kun angives hvis "Register" parameteret er angivet.	Ja

2.6.1.2 Returværdier

Metoden returnerer metadata om fildownloads, der følger strukturen beskrevet i Tabel 5.

Felt	Beskrivelse
Filename	Navnet på filen som kan downloades.
Register	Angiver hvilket register data i filen tilhører.
EntityName	Angiver navnet på registerentiteten som filens data tilhører.
Frequency	Angiver hvornår/hvor ofte der bliver genereret en ny fil af samme type med nyt udtræksnummer.
TypeOfDownload	Angiver om filen er et totaldownload eller et deltadownload.
TypeOfData	Angiver hvilken bitemporal filtrering der er anvendt på data i filen. Denne kan enten være "Bitemporal", "Temporal" eller "Current".
Version	Angiver underlæggende version af filens data.
GenerationNumber	Angiver udtræksnummeret.
PointInTime	Angiver tidspunkt for registrerings- og virkningstid der blev anvendt ved generering af fildownload. Feltet er kun sat hvis der er tale om et Temporalt eller Current fildownload. Datoformat følger ISO 8601 Tidszonen er UTC



GenerationTime	Angiver tidspunktet for hvornår filen blev genereret. Datoformat følger ISO 8601 Tidszonen er UTC
ExpirationDate	Angiver hvornår filen ikke længere udstilles på Datafordeleren. Datoformat følger ISO 8601 Tidszonen er UTC
ContainedFileFormat	Angiver hvilket filformat ZIP-filen indeholder.
OutputFileFormat	Angiver filformatet på selve filen som kan downloades.

Tabel 5: Overblik over returverdier for GetAvailableFileDownloads metoden.

2.6.1.3 Eksempler på brug af metoden

Se tilgængelige fildownloads for DAGI:

<https://api.datafordeler.dk/FileDownloads/GetAvailableFileDownloads?Register=DAGI&username=xxxx&password=yyyy>

Se tilgængelige fildownloads for "Bygning"-entiteten:

<https://api.datafordeler.dk/FileDownloads/GetAvailableFileDownloads?Register=BBR&Entity=Bygning&username=xxxx&password=yyyy>

Se tilgængelige fildownloads i version 1 i EBR:

<https://api.datafordeler.dk/FileDownloads/GetAvailableFileDownloads?Register=EBR&Version=1&username=xxxx&password=yyyy>

2.6.2 Metode: GetFile

Metode: GetFile	
URL	https://<env>-api.datafordeler.dk/FileDownloads/GetFile?
http-metode	GET
Headere i forespørgsel	Content-Type: application/json
Format	JSON
Returværdier	• Returnerer http 200 – OK, ved succes. • Returnerer HTTP 400 – Bad Request, ved angivelse af forkerte parametre. • Returnerer HTTP 404 – Not Found, hvis den angivne ressource ikke kan findes eller hvis brugeren ikke har adgang til den angivne fil. • Returnerer HTTP 500 – Internal Server Error, hvis der er sket en ukendt fejl.
Adgang	Tjenestebruger med brugernavn og password.
Dato- og tidsformat	Datoformat følger ISO 8601 Tidszonen er UTC



2.6.2.1 Parametre

Navn	Type	Beskrivelse
Filename	String	Specificerer en specifik fil til download. Dette parameter kan kun angives alene.
Register	String	Specificerer, hvilket register der ønskes download fra. Hvis ikke parametrene 'LatestTotalForEntity' og 'Type' angives, bliver alle registerets totaldownloads downloadet. Følgende registre kan angives: • DAR • DAGI • BBR • DHMOprindelse • DHMHoejdekurver • MAT (Der refereres her til Matriklen 2) • EBR • FIKSPUNKT • DS (Danske Stednavne) • GeoDKV
LatestTotalForEntity	String	Henter det seneste totaldownload for den specificerede entitet. 'Type' parameteret skal angives, såfremt dette parameter er angivet.
Type	String	Angiver hvilken type fildownload der skal downloades for entiteten. 'LatestTotalForEntity' parameteret skal angives, såfremt dette parameter er angivet.
Format	String	Angiver hvilket filformat der ønskes. Hvis 'LatestTotalForEntity' parameteret er angivet, skal dette parameter også angives.
Version	String	Angiver hvilken version fildownloadet skal være i. Såfremt dette parameter ikke angives, vælges den nyeste version.

2.6.2.2 Returværdier

Denne metode returnerer altid en eller flere filer ved HTTP 200 - OK.

2.6.2.3 Eksempler på brug af metoden

Hent en specifik deltafil fra Danske Stednavne registeret:

https://api.datafordeler.dk/FileDownloads/GetFile?Filename=DS_V1_Vej_DeltaDownload_gml_Bitemporal_2335.zip&username=xxxx&password=yyyy

Kaldet returnerer en enkelt fil i en ZIP-fil.

Hent totaldownloads for alle entiteter i BBR:

<https://api.datafordeler.dk/FileDownloads/GetFile?Register=BBR&username=xxxx&password=yyyy>

Kaldet returnerer flere filer i en samlet ZIP-fil.

Hent totaldownload med aktuelle data i GML-format for en bestemt entitet i Fikspunkt-registeret:

<https://api.datafordeler.dk/FileDownloads/GetFile?Register=FIKSPUNKT&LatestTotalForEntity=HoejdefikspunktFO&Type=Current&Format=gml&username=xxxx&password=yyyy>

Kaldet returnerer en enkelt fil i en ZIP-fil.

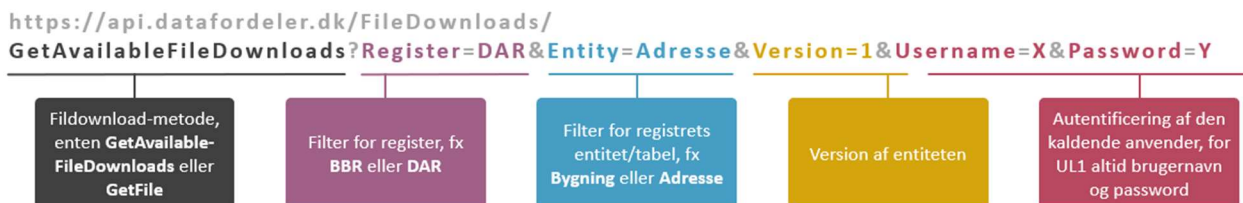


3 Etablering og vedligeholdelse af kopiregister ved brug af fildownloads

Dette afsnit beskriver hvordan en anvender kan etablere et kopiregister lokalt ved at hente totaldownloads og herefter vedligeholde det ved at hente deltdownloads. Afsnit 2.4 beskriver kort hvordan dette gøres, mens dette afsnit tager de introducerede koncepter fra indeværende vejledning og bruger dem i dette eksempel.

Et kopiregister er en hel kopi eller en delmængde af et registers data. Der kan være flere årsager til at have et kopiregister, men en af årsagerne er, at data er tilgængeligt indenfor anvenderens egen virksomhed og mindsker egen afhængighed til opslag på Datafordeleren. Det er dog vigtigt at understrege at kopiregisteret skal opdateres løbende for at beholde sin værdi, såfremt aktuelle data er nødvendige i virksomhedens brugsscenarie.

Etablering af et kopiregister initieres ved først at hente et totaldownloads af relevante entiteter fra Datafordeleren. For at finde ud af, hvilke filer der er tilgængelige, benyttes GetAvailableFileDownloads-metoden. Dette gøres ved hjælp af følgende endpoint strukturering:



Ovenstående kald vil returnere følgende (forsimplet):

```
[
  {
    "Filename": "DAR_1_Adresse_Total_JSON_Bitemporal_101.zip",
    "Register": "DAR",
    "EntityName": "Adresse",
    "Version": "1",
    "GenerationNumber": "101",
    "GeneratedTimestamp": "2024-03-06T22:10:50.123456Z",
    "ExpirationDate": "2024-03-13T22:10:50.123456Z",
    "FileDownloadType": "Total",
    "TypeOfData": "Bitemporal",
    "GenerationFrequency": "Generated Friday night"
  },
  {
    "Filename": "DAR_1_Adresse_Delta_JSON_Bitemporal_101.zip",
    "Register": "DAR",
    "EntityName": "Adresse",
    "Version": "1",
    "GenerationNumber": "101",
    "GeneratedTimestamp": "2024-03-06T22:10:50.123456Z",
    "ExpirationDate": "2024-03-20T22:10:50.123456Z",
    "FileDownloadType": "Delta",
    "TypeOfData": "Bitemporal",
    "GenerationFrequency": "Generated every night"
  }
]
```

I ovenstående ses, at der både er et total- og deltautræk tilgængeligt. Når man vil etablere sit kopiregister, er det, som tidligere beskrevet, det samlede totaludtræk man er interesseret i.



I dette scenarie repræsenterer filen "DAR_1_Adresse_Total_JSON_Bitemporal_101.zip" det samlede totaludtræk. Man vil derfor benytte sig af GetFile-metoden, som eksemplificeret nedenfor, til at hente filen:

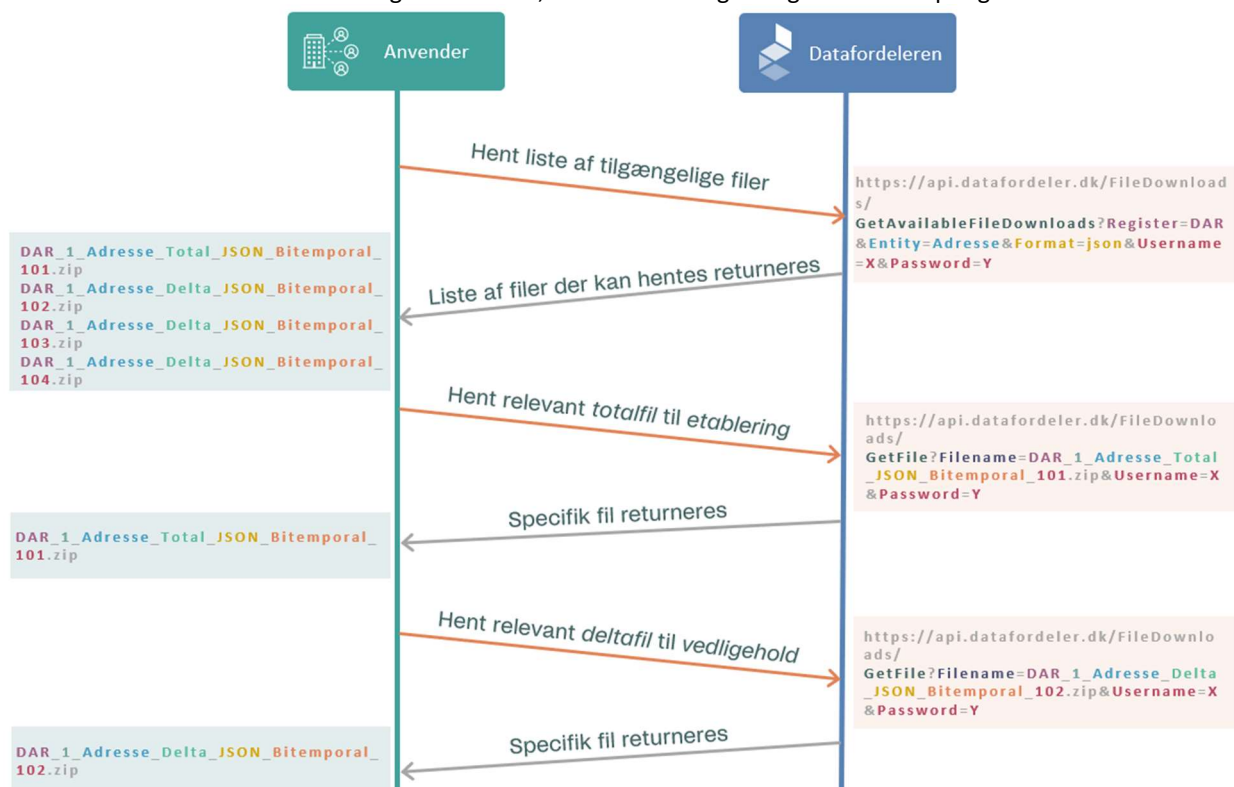


Man kan også hente det seneste totaludtræk via følgende metode:



Vedligeholdelse af kopiregisteret opnås ved brug af deltaudtræk, hvilket sikrer, at kopiregisteret kontinuerligt opdateres med de nyeste data. De tre efterfølgende filer "DAR_1_Adresse_Delta_JSON_Bitemporal_102.zip", "DAR_1_Adresse_Delta_JSON_Bitemporal_103.zip" og "DAR_1_Adresse_Delta_JSON_Bitemporal_104.zip" repræsenterer i dette eksempel hver især deres eget deltaudtræk, bestående af entiteterne tilsvarende det totaludtræk der tidligere var hentet. Hver fil indeholder ændringer siden det forrige deltaudtræk og har et unikt udtræksnummer (henholdsvis 102, 103 og 104, som vist i Figur 5), der fungerer til at identificere hver iteration og sikrer korrekt anvendelse af data.

Som beskrevet ovenfor, viser følgende illustration i Figur 5 hvorledes total-og deltaudtræk benyttes mellem Datafordeleren og anvenderen, til at etablere og vedligeholde et kopiregister.



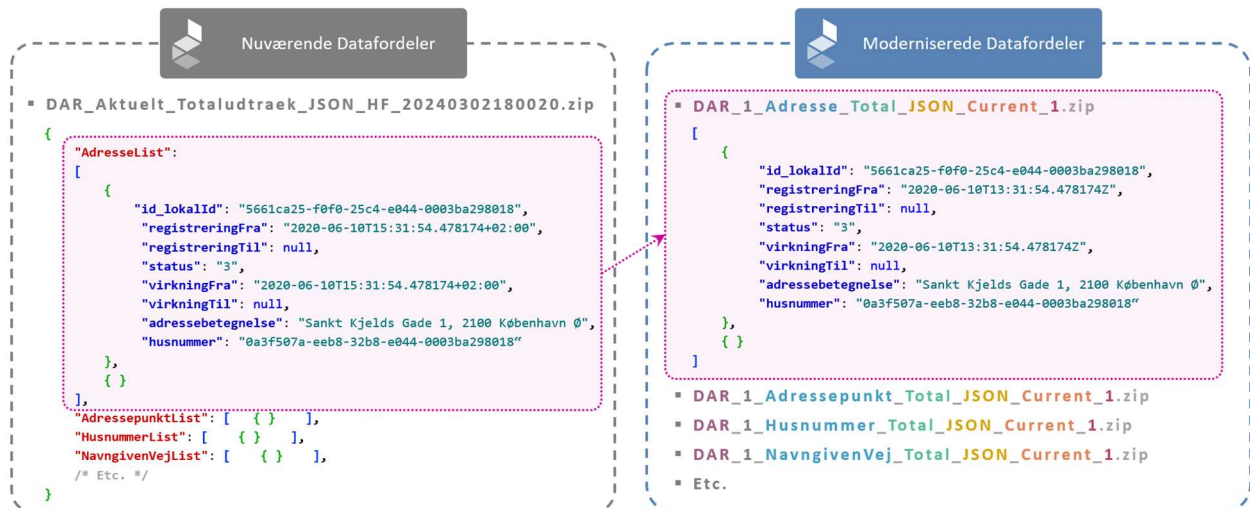
Figur 5: Illustration af etablering og vedligeholdelse af kopiregister ved brug af total- og deltaudtræk.

4 Transition fra filudtræk til fildownloads

Dette afsnit beskriver forskelle mellem abonnementsbaserede filudtræk og prædefinerede entitetsbaserede fildownloads, samt hvordan man som anvender af filudtræk, kan komme i gang med at bruge entitetsbaserede fildownloads i stedet for. Det skal bemærkes at filudtræk og fildownloads driftes parallelt i en periode, så anvendere af Datafordeleren har mulighed for at skifte til fildownloads.

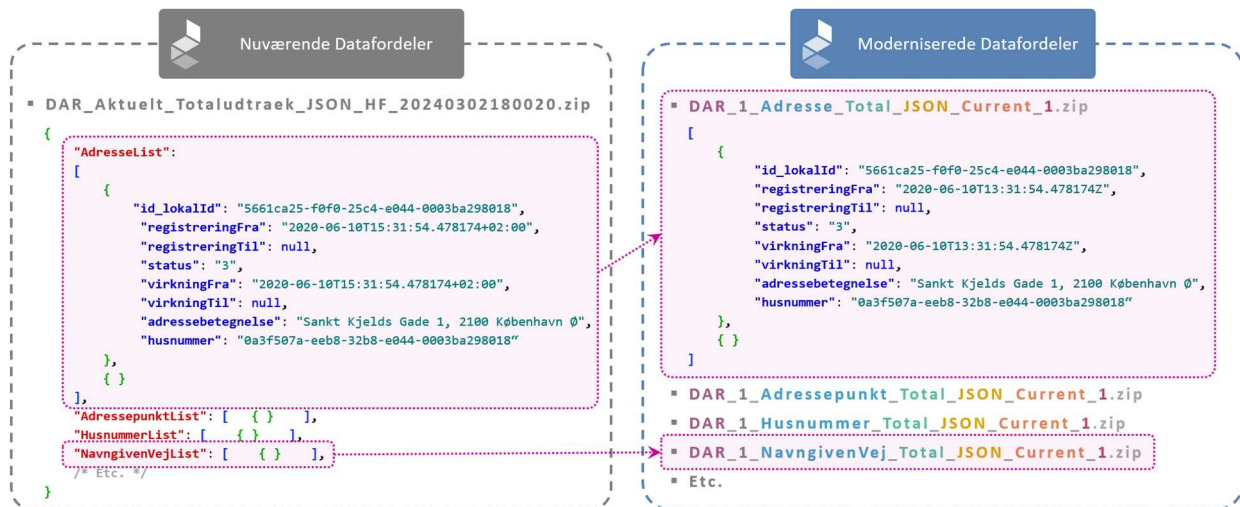
4.1 Indhold i filer

Totaludtræk på den Nuværende Datafordeler indeholder normalt hele registerets data, hvor alle registerets entiteter blev fordelt i lister og derefter samlet i én samlet fil. Dette kan ses i venstre side i Figur 6, som indeholder 4 lister med entiteter i. Den Moderniserede Datafordeler tager disse lister der hver især indeholder en entitet og opdeler så disse i hver deres filer i stedet for. Dette kan ses i højre side i Figur 6. Det er altså det samme data der udstilles i filerne sammenlagt, men hvor forskellen er at fildownloads hver især indeholder en delmængde af det samlede registers datasæt.



Figur 6: Sammenligning af data i filudtræk og fildownloads.

Den Moderniserede Datafordeler udstiller på denne måde den samme data i fildownloads, som bliver udstillet i filudtræk på den Nuværende Datafordeler, dog med den ændring, at filerne er opdelt per entitet.



Figur 7: Sammenligning af data i filudtræk og fildownloads.

Da fildownloads bliver genereret på entitetsniveau, betyder det også at hele entiteten udstilles, og at der ikke filtreres felter fra i det udstillede data, da dette er det samme som det der indlæses på Datafordeleren. Hvis man som anvender gerne vil vide præcis hvilke felter der eksisterer i fildownloads for en given entitet, skal man konsultere registrenes Dataleverancespecifikation (DLS), der i dens Bilag 3 (Opdatering) beskriver hvilke felter hvert registers entiteter består af. Registrenes DLS'er kan hentes [her](#).

4.1.1 Forskelle i GML og GeoPackage filudtræk og fildownloads

I forbindelse med implementering af eksport af GML og GeoPackage fildownloads, er der taget en række design-beslutninger som er forskellige fra de valg som blev truffet da eksport af GML og GeoPackage filudtræk blev designet. Dette resulterer i følgende nævneværdige forskelle i de genererede filer, hvilket er beskrevet i dette afsnit.

Generelle forskelle der gælder for begge filtyper:

- I filudtræk eksporteres dato-felter i dansk tid. I fildownloads eksporteres dato-felter i UTC.
- I filudtræk kan der eksistere et felt som indeholder Datafordelerens opdateringstidspunkt for den pågældende database-række. Dette er kun inkluderet i et filudtræk, hvis registeret eksplicit har bestilt dette i deres DLS. Dette felt er altid inkluderet i fildownloads.

Fildownloads udstiller data per-entitet, med samme datamodel som registrene sender til Datafordeleren. Hvis et register i deres DLS for filudtræk har valgt at omdøbe kolonner, lave sammenstillinger mellem tabeller, har beregnede felter, eller lign., vil disse ikke være at finde i fildownloads. Der kan dermed forekomme flere/færre felter i fildownloads, og nogen felter kan have andre navne.

GeoPackage-specifikke forskelle:

- I filudtræk eksporteres dato-felter til en kolonne defineret med typen DateTime, og er begrænset til millisekund præcision. I fildownloads eksporteres dato-felter til en kolonne defineret med typen TEXT, og er begrænset til mikrosekund præcision, hvor datoen er skrevet i ISO8601 format.

I tilfælde af at registre har brugt karakterer i deres felt- eller entitet-navne som ikke er tilladt af GeoPackage standarden, renses navnene i fildownloads på følgende måde: æ/ø/å udskiftes til hhv. ae/oe/aa, og alle andre ugyldige karakterer udskiftes med _ (en understreg). Dette kan forårsage forskelle, hvis den manuelt definerede rensning som er brugt i filudtræk, ikke er udført på samme måde.



GML-specifikke forskelle:

- Der er en række mindre forskelle i XSD-skemaerne imellem filudtræk og fildownloads. F.eks. udstilles GML fildownloads med det pågældende registers *namespace*, hvorimod filudtræk bruger Datafordelerens *namespace*. Herudover indeholder fildownload XSD-skemaer alle registerets annotationer og type-restriktioner, hvilket ikke er inkluderet i filudtræk XSD-skemaerne.
- I filudtræk eksporteres dato-felter med op til millisekund præcision. I fildownloads eksporteres dato-felter med op til mikrosekund præcision.
- I tilfælde af at registre har brugt karakterer i deres felt- eller entitet-navne som ikke er tilladt i XML-standardens, renses navnene i fildownloads ved at skrive disse med XML LocalName encoding reglerne. Dette kan forårsage forskelle, hvis den manuelt definerede rensning som er brugt i filudtræk, ikke er udført på samme måde.

GML "srsName" attributten skrives i filudtræk med den korte form "EPSG:25832" eller "EPSG:4326". I fildownloads bruges URN versionen af EPSG-definitionen, dvs. "urn:ogc:def:crs:EPSG::25832" og "urn:ogc:def:crs:EPSG::4326".

4.2 Geografisk filtrering i fildownloads

Følgende afsnit beskriver, hvordan man, som anvender, kan postprocessere fildownloads på forskellige geografiske egenskaber. Man kan lave geografisk filtrering på registerdata ved at filtrere på de egenskaber der ses i Tabel 6. For data, som kan filtreres geografisk, gælder det, at et datapunkt er associeret med:

- En afledt geografisk identifikator, eksempelvis kommunekode. Et datapunkt kan filtreres på baggrund af denne kode.
- En geometri, eksempel en polygon med EPSG referencekode. Polygonen kan sammenlignes med en referencegeometri, såsom en kommune, og på baggrund af det geometriske overlap kan datapunktet filtreres.

Såfremt et datapunkt ikke har en reference til nogen af ovennævnte, kan det være tilfældet at datapunktet har en reference til en entitet der er associeret med enten en geometrisk identifikator eller geometri. Det er derfor ikke altid muligt direkte at foretage geografisk filtrering på en entitet uden først at behæfte entiteten med en geografisk identifikator først. Bemærk også at tabellen kun dækker de registre der har fildownloads. Der er i moderniseringens Udviklingsleverance 1 ikke implementeret generisk geografisk filtrering på fildownloads, idet det ville kræve, at Datafordeleren gør sig adskillige antagelser om, hvordan data anvendes i den anden ende. Såfremt disse antagelser ikke stemmer overens med den faktuelle brug, kan Datafordeleren fejlagtigt lave geografiske filtreringer, der er anderledes end anvenderens faktiske behov.

Register	Felt
BBR	Kommunekode
EBR	Kommunekode
MAT	Kommunekode
GeoDanmark Vektor	GeoDanmark kommunepolygon, kommunekode
DAGI	DAGI kommunepolygon
DHM Højdekurver	DAGI kommunepolygon

Tabel 6: Overblik over hvilke geografiske egenskaber der skal filtreres på for forskellige registre.



4.3 Anvendereksempel på transition fra filudtræk til fildownloads

Dette eksempel tager udgangspunkt i en anvender der allerede bruger Datafordeleren og henter data fra DAR igennem filudtræk.

Afsnittet giver et eksempel på, hvordan en eksisterende anvender på Datafordeleren kan overgå fra at bruge filudtræk i sit system til at bruge fildownloads. Afsnittet påbegyndes med en situationsbeskrivelse efterfulgt af konkrete trin der beskriver transitionen.

4.3.1 Situationsbeskrivelse

Dette eksempel antager at en anvender allerede har etableret en kopi af DAR og skal i gang med at bruge fildownloads.

Anvenderen har først etableret sin datakopi af DAR ved at opsætte et abonnement på et brugerdefineret totaludtræk fra DAR med følgende parametre:

- SincePrevious = True
- RegistreringFra = 2021-03-01
- VirkningFra = 2021-03-01
- Format = JSON
- Dagligt udtræk

Anvenderen henter filudtræk fra Datafordeleren via sin SSH2-tjenestebruger dagligt og indlæser deltafilen i sin datakopi.

4.3.2 Transition

For at anvenderen kan komme i gang med at bruge fildownloads, skal anvenderen vælge mellem 2 muligheder:

Mulighed A) Anvenderen ønsker at etablere en ny datakopi med fildownloads.

Mulighed B) Anvenderen ønsker at indlæse data oveni en eksisterende datakopi med fildownloads.

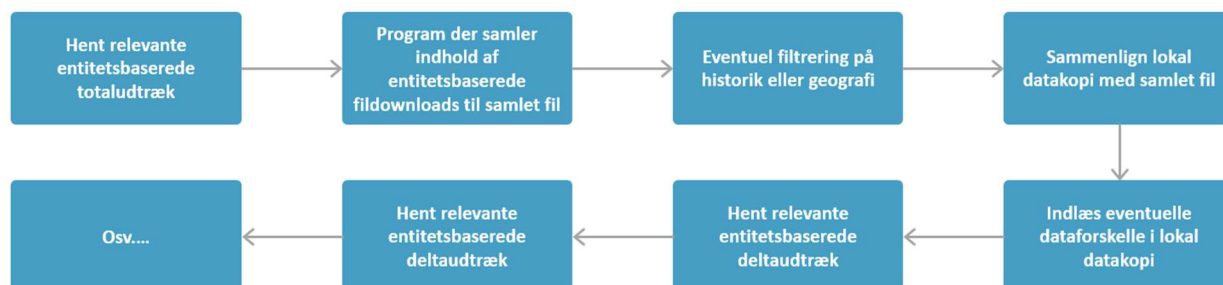
Anvenderen gør herefter følgende

- 1) Anvenderen opretter en tjenestebruger med brugernavn og password på selvbetjeningen.
- 2) Anvenderen kalder metoden GetAvailableFileDownloads med sin nye tjenestebruger som følger:
 - a. <https://api.datafordeler.dk/FileDownloads/GetAvailableFileDownloads?Register=DAR&username=xx&password=yyyy>
- 3) Da anvender nu ved, hvilke fildownloads der eksisterer, kalder anvender metoden GetFile, for at hente totaldownloads for alle entiteter i DAR:
 - a. <https://api.datafordeler.dk/FileDownloads/GetFile?Register=DAR&username=xxxx&password=yyyy>
- 4) Anvender er nu i besiddelse af et totaludtræk i DAR. Anvenderen kan nu vælge at sammensætte alle filerne til én stor fil og postprocessere denne, eller postprocessere filerne enkeltvis. Postprocesseringen filtrerer data fra, der har RegistreringFra og VirkningFra ældre end 2021-03-01.
- 5) Dette skridt varierer, alt efter om anvenderen valgte mulighed A eller B, som er beskrevet ovenfor:

- a. Hvis anvenderen valgte mulighed A), skal anvenderen etablere sin lokale datakopi ved at indlæse data fra filerne. I dette eksempel har totaludtrækket fra DAR udtræksnummer 14.
- b. Hvis anvenderen valgte mulighed B) i stedet, skal anvender finde forskellen mellem sin lokale datakopi og den data der findes i totaludtrækket fra DAR som anvenderen netop har hentet.

Anvender sætter herefter en vedligeholdelse op af sin datakopi. Vedligeholdelsen består i at hente deltadownloads ind fra DAR, som genereres dagligt. Anvenderen bemærker, at totaludtrækket har udtræksnummer 14, hvilket betyder at anvenderen skal hente deltadownloads fra DAR med udtræksnummer 15 og højere for at sørge for at datakopien holdes opdateret. Anvenderen kalder igen GetAvailableFileDownloads for at se, hvilke deltafiler der findes for DAR.

- c. <https://api.datafordeler.dk/FileDownloads/GetAvailableFileDownloads?Register=DAR&username=xx&password=yyyy>
- 6) Anvenderen ser at der findes deltadownloads for samtlige entiteter i DAR med udtræksnumrene 15 og 16. Anvenderen ved derfor, at det vil være nødvendigt at hente alle deltadownloads for alle entiteter i DAR med de ovenstående udtræksnumre. Dette gør anvenderen ved at kalde GetFile for alle entiteterne i DAR med både udtræksnumre 15 og 16:
 - a. <https://api.datafordeler.dk/FileDownloads/GetFile?Filename=filename&username=xxxx&password=yyyy>
- 7) Anvenderen holder løbende øje med genereringen af DAR deltadownloads der genereres hver dag og sørger for at hente disse og indlæse dem i deres lokale datakopi.
- 8) Anvenderen har nu sat sit system op til at bruge fildownloads.



Figur 8: Opsummering af transitionsskridt for en anvender.

Transitionen i ovenstående skridt er opsummeret i Figur 8.



4.4 Anvendereksempel på transition med udgangspunkt i kopiregister

Dette eksempel tager udgangspunkt i en anvender som allerede har etableret en kopi af BBR, som skal i gang med at bruge fildownloads.

4.4.1 Situationsbeskrivelse

I dette eksempel henter anvenderen daglige totaludtræk hvor registreringshistorikken ikke er nødvendig, og derfor filtreres fra. Anvenderen har sat følgende abonnement op:

- SincePrevious = False
- RegistreringFra = NOW()
- Format = JSON
- Dagligt udtræk

Anvenderen bruger en tjenestebruger med brugernavn og password til at tilgå filudtræk.

4.4.2 Transition

- 1) Anvenderen kalder metoden GetAvailableFileDownloads med sin eksisterende tjenestebruger som følger:
 - a. <https://api.datafordeler.dk/FileDownloads/GetAvailableFileDownloads?Register=BBR&username=xxxx&password=yyyy>
- 2) Da anvender nu ved hvilke fildownloads der eksisterer, kalder anvender nu metoden GetFile for at hente totaldownloads for alle entiteter i BBR:
 - a. <https://api.datafordeler.dk/FileDownloads/GetFile?Register=BBR&username=xxxx&password=yyyy>
- 3) Anvender er nu i besiddelse af et totaludtræk af BBR, heriblandt et totaludtræk med temporale data, dvs. registreringshistorikken er filtreret fra. Anvenderen kan nu vælge at sammensætte alle filerne til én stor fil og postprocessere denne, eller postprocessere filerne enkeltvis.
- 4) Anvenderen etablerer herefter den samme vedligeholdelsesproces som blev beskrevet i afsnit 4.3.2. Da anvenderen her ikke er interesseret i registreringshistorikken, skal anvender dog huske at postprocessere deltdownloads og filtrere denne historik fra, inden den data indlæses oveni den lokale datakopi.



5 Bilag 1: FAQ

- 1) Er min IP-adresse stadig godkendt?
 - a. Ja det er den.
- 2) Hvilken zone udstilles fildownloads i?
 - a. Fildownloads udstilles i zone 0.
- 3) Hvilken type bruger skal jeg bruge for at hente fildownloads?
 - a. En tjenestebruger med brugernavn og password er påkrævet for at tilgå fildownloads.
- 4) Hvordan hentes fildownloads?
 - a. Dette sker ved at kalde REST-API'et der udstiller fildownloads. Se afsnit 2.6 for en uddybende beskrivelse.
- 5) Hvilke tidspunkter bliver fildownloads genereret?
 - a. Dette sker som udgangspunkt mellem kl. 3 og 6 om natten. Se skemaet i afsnit 2.5 der giver et overblik over hvornår der genereres fildownloads for hvert register.
- 6) Er der fildownloads for alle registre?
 - a. Nej, det er der ikke. Det er kun de registre som står anført i Tabel 3 som får fildownloads i de filformater der står angivet i tabellen.



6 Bilag 2: Overførsel af filudtræk til fildownloads

Filudtræk	Tilsvarende fildownloads
BBR-Aktuelt-Totaludtræk_HF	BBR_TotalDownload_Current for alle entiteter i BBR
BBR-Totaludtræk_HF	BBR_TotalDownload_Bitemporal for alle entiteter i BBR
DAGI10MULTIGEOM	DAGI_TotalDownload_Current for alle entiteter hvor skala er 1:10.000
DAGI10MULTIGEOMHIST	DAGI_TotalDownload_Bitemporal for alle entiteter hvor skala er 1:10.000
DAGI10MULTIGEOMHIST_HF	DAGI_TotalDownload_Bitemporal for alle entiteter hvor skala er 1:10.000
DAGI10MULTIGEOM_HF	DAGI_TotalDownload_Current for alle entiteter hvor skala er 1:10.000
DAGI2000MULTIGEOM	DAGI_TotalDownload_Current for alle entiteter hvor skala er 1:2.000.000
DAGI2000MULTIGEOM_HF	DAGI_TotalDownload_Current for alle entiteter hvor skala er 1:2.000.000
DAGI250MULTIGEOM	DAGI_TotalDownload_Current for alle entiteter hvor skala er 1:250.000
DAGI250MULTIGEOM_HF	DAGI_TotalDownload_Current for alle entiteter hvor skala er 1:250.000
DAGI500MULTIGEOM	DAGI_TotalDownload_Current for alle entiteter hvor skala er 1:500.000
DAGI500MULTIGEOM_HF	DAGI_TotalDownload_Current for alle entiteter hvor skala er 1:500.000
DAR-Aktuelt-Totaludtræk_HF	DAR_TotalDownload_Current for alle entiteter i DAR
DAR-Totaludtræk_HF	DAR_TotalDownload_Bitemporal for alle entiteter i DAR
DHMhoejdekurver025	DHM_referencekurve025_TotalDownload_Temporal
DHMhoejdekurver025_HF	DHM_referencekurve025_TotalDownload_Temporal
DHMhoejdekurver050	DHM_formkurve05_TotalDownload_Temporal
DHMhoejdekurver050_HF	DHM_formkurve05_TotalDownload_Temporal
DHMhoejdekurver250	DHM_formkurve25_TotalDownload_Temporal
DHMhoejdekurver250_HF	DHM_formkurve25_TotalDownload_Temporal
DHMhoejdekurverekstrepunkter	DHM_kote05_TotalDownload_Current DHM_kote25_TotalDownload_Current
DHMhoejdekurverekstrepunkter_HF	DHM_kote05_TotalDownload_Current DHM_kote25_TotalDownload_Current
DHMoprindelseKorrektion	DHM_korrektionsegne_Bitemporal med GeoDK Soe-entiteten
DHMoprindelseKorrektion_HF	DHM_korrektionsegne_Bitemporal med GeoDK Soe-entiteten



DHMoprindelsePunktoprindelse	DHM_punktoprindelse_Bitemporal
DHMoprindelsePunktoprindelse_HF	DHM_punktoprindelse_Bitemporal
DKstednavneBearbejdedeNohist	DS_TotalDownload_Current for alle entiteter
DKstednavneBearbejdedeNohist_HF	DS_TotalDownload_Current for alle entiteter
EBREjendomsbeliggenhedSimpel_HF	EBR_TotalDownload_Current for alle entiteter
EBREjendomsbeliggenhed_HF	EBR_TotalDownload_Bitemporal for alle entiteter
GeoDanmark60	GeoDKV_TotalDownload_Temporal for alle entiteter
GeoDanmark60HIST	GeoDKV_TotalDownload_Bitemporal for alle entiteter
GeoDanmark60HISTKommune_HF	GeoDKV_TotalDownload_Temporal for alle entiteter med geografisk filtrering på relevant kommune
GeoDanmark60HISTRegion_HF	GeoDKV_TotalDownload_Temporal for alle entiteter med geografisk filtrering på relevant region
GeoDanmark60HIST_HF	GeoDKV_TotalDownload_Bitemporal for alle entiteter
GeoDanmark60_HF	GeoDKV_TotalDownload_Temporal for alle entiteter
GeoDanmark60bruger	GeoDKV_TotalDownload_Bitemporal for alle entiteter
GeoDanmark60bruger_HF	GeoDKV_TotalDownload_Bitemporal for alle entiteter
MATkortdataGaeldendeDKComplete_GML	MAT_Totaldownload_Current for alle entiteter med tilknyttet jordstykke pr. entitet
MATkortdataGaeldendeDKComplete_GPKG	MAT_Totaldownload_Current for alle entiteter med tilknyttet jordstykke pr. entitet
MATkortdataGaeldendeKommuneComplete_GML	MAT_Totaldownload_Current for alle entiteter med tilknyttet jordstykke pr. entitet og med geografisk filtrering
MATkortdataGaeldendeKommuneComplete_GPKG	MAT_Totaldownload_Current for alle entiteter med tilknyttet jordstykke pr. entitet og med geografisk filtrering
MATregisterdataTotalUdtraek	MAT_TotalDownload_Bitemporal for alle entiteter