

## Principper for dataejerskab og deling af DFDG data

Created by [Rune Gram-Madsen \(STAR\) \[Administrator\]](#), last modified on [Oct 26, 2016](#)

- [1 Dataejerskab og ansvar](#)
  - [1.1 Metadataejerskab og modellering](#)
  - [1.2 Indholdsmæssigt dataansvar](#)
  - [1.3 Delte masterdata](#)
  - [1.4 Dataejerskab](#)
- [2 Tilbudte snitflader mod DFDG data](#)
  - [2.1 Webservicesnitflader \(SOAP/XML format\)](#)
  - [2.2 Dataload/ETL](#)
- [3 Detaljerede principper for data i DFDG](#)
  - [3.1 Grundlæggende principper](#)
  - [3.2 Principper for sikring af konsistens i de delte masterdata](#)
  - [3.3 Principper for dataload](#)
  - [3.4 Principper for forretningsunderstøttelse](#)
  - [3.5 Standarder for dataentiteter](#)

Følgende beskriver principper for data i DFDG og deling af disse data med omkringliggende aktør-systemer.

Principperne skal opfattes som **generelle** retningslinjer til deling af data mellem DFDG og de omkringliggende applikationer. Forretningsmæssige krav og specielle behov kan undtagelsesvist fordrø andre tekniske løsninger.

### Dataejerskab og ansvar

#### Metadataejerskab og modellering

Såvel DFDG som fødesystemerne beskriver deres data via metadata, metadata er her syntaks og semantik. Den der har ansvar for at definere data, har også ejerskab til metadata. **Metadataejerskab** indebærer pligt til og ansvar for at sikre metadataes forretningsmæssige integritet, begrebsafklaring og definition samt overholdelse af stavekonventioner.

Hvor ejerskabet til data i DFDG er entydigt, så er metadataejerskabet i DFDG blandet. Metadata i DFDG stammer fra mange kilder, og det betyder, at det er andre myndigheder, der er ansvarlige for boniteten i disse metadata.

Ansvar for metadata kan principielt hidrøre fra:

- Andre statslige myndigheder, fx Digitaliseringsstyrelsen, SKAT, Kriminalforsorgen, sundhedsvæsenet, klassifikationer for uddannelser etc.
- Fælles myndighedsmetadata på beskæftigelsesområdet (som kan være aftalt mellem parterne)
- STAR-metadata (Jobnet, Jobbing, MitJobkompas, VITAS med flere)
- Kommuner
- A-kasser

### Indholdsmæssigt dataansvar

Aktørerne på beskæftigelsesområdet har forskellige ansvar for at oprette nye instanser, også kaldet forekomster, af data og holde dem opdateret eller slette dem. Så når en borger via Jobnet foretager en tilmelding, opretter han en ny tilmeldingsinstans, og udfylder den med en værdi, der fortæller, at han nu er tilmeldt som arbejdssøgende og måske også ledig. Hvis en borger er ledig og ønsker at modtage en offentlig forsørgelsesydelse, har han pligt til at tilmelde sig et jobcenter. Denne forpligtelse er udtryk for, at borgeren har et ansvar for, at data om hans tilmeldingsstatus er opdateret. Gør borgeren det ikke, overgår retten til at opdatere borgerens data til myndighederne.

**Indholdsmæssigt dataansvar** rummer således forpligtelse til og ansvar for, at data er retvisende, dvs. at de altid er opdateret inden for de frister og regler, der måtte gælde.

### Delte masterdata

Kommuner, a-kasser (og borgere) har pligt til og ansvar for at opdatere data om borgere ud fra forretningsmæssige regler, der gælder for disse data. Samtidig har både kommuner, a-kasser, stat og borgere behov for adgang til data. For at dette kan realiseres på en hensigtsmæssig måde udstilles data via delte datakataloger som **delte masterdata**, hvor DFDG er et af de centrale datakataloger på beskæftigelsesområdet.

### Dataejerskab

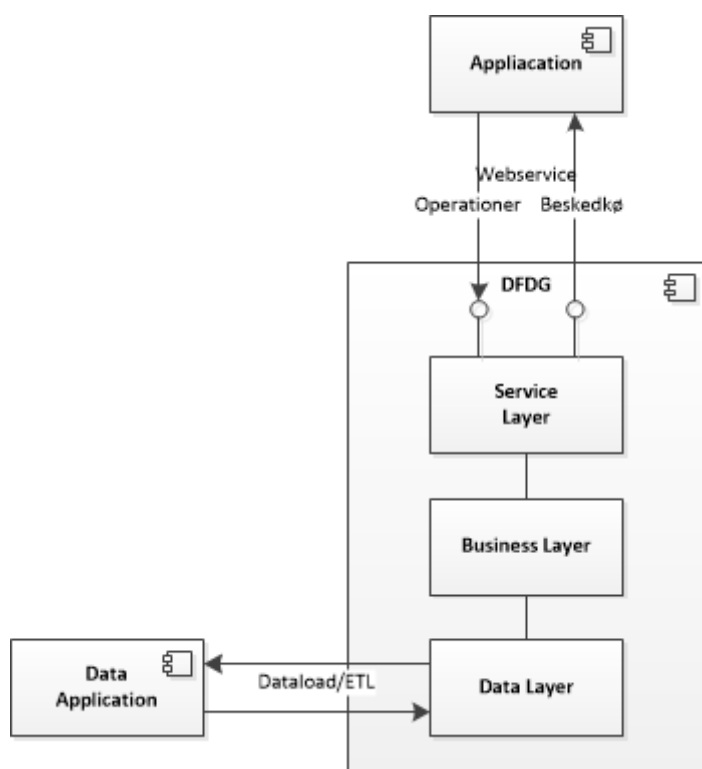
**Dataejerskab** betragtes fra en teknisk vinkel. **Teknisk dataejerskab** rummer **ansvar** for, at data valideres, lagres hensigtsmæssigt, så de er sikret mod uberettigedes adgang, mod tab eller uønsket sletning, sammenblanding med andre data mv., og så de kan tilgås efter behov af berettigede, lige som ejerskabet rummer ansvar over for persondata- og registerlovgivning.

### Tilbudte snitflader mod DFDG data

DFDG tilbyder følgende tekniske snitflader mod de omkringliggende aktør-systemer:

- Webservicesnitflader (SOAP/XML format)
  - Forretningservices
  - Dataservices (CRUD)
  - Besked-kø (WSRM lignende kø)
- Dataload/ETL
  - Dataload via staging-database

- Dataload via filer (XML/CSV)



Figur 1: To niveauer af snitflader. Webservices og dataload/ETL

### Webservicesnitflader (SOAP/XML format)

DFDG tilbyder et antal webservices, som kan anvendes af aktør-systemerne, således at sagsbehandler og aktørsystem har on-line og direkte adgang til de delte master data. Services kan inddeles i to typer:

- Forretningsservices

Forretningsservices er webservices, som udfører en eller flere forretningsoperationer, der kan resultere i opdatering af flere underliggende data-entiteter, typisk via CRUD operationer. Forretningsoperationer vil typisk implicit forudsætte et bestemt procesforløb

- Dataservices (CRUD)

Dataservices er webservices der har fokus på at opdatere/publicere [\[1\]](#) data på forskellige aggregationsniveauer

- Besked-kø, (WSRM)

Aktør-systemerne kan via en besked-kø modtage hændelser omkring opdatering af data-entiteter og udførelse af forretningsoperationer.

De tilbudte webservices arbejder typisk på individniveau, dvs. person, virksomhed osv. Der tilbydes tillige et begrænset antal tværgående services, til f.eks. søgning på tværs af dataentiteter eller til konsistenskontrol af aktørs gemte kopi af DFDG's delte master data.

## Dataload/ETL

DFDG tilbyder/understøtter et begrænset antal dataload/ETL snitflader, der typisk understøtter større ind og udlæsninger af data til/fra DFDG. Det er typisk hele populationer der udveksles. Data deles via:

1. databasetabeller via staging-database
2. filer (XML, evt. CSV) via FTP server

Dataload anvendes typisk til indlæsning af hele datasæt i DFDG eller ved udlæsning af BI data til analyse- eller ledelsesformål. Dataload/ETL anvendes som udgangspunkt ikke til overførsel af operationelle (online) data.

Dataload/ETL snitfladen er en komplementær snitflade, der anvendes når der er behov for overførsel af datatræk på hele populationer, eller hvor aftagerne ikke tilbyder en Webservicesnitflade.

## Detaljerede principper for data i DFDG

### Grundlæggende principper

1. DFDG skal indeholde alle relevante data om en borgers aktuelle og historiske beskæftigelsesindsatstilstand<sup>[2]</sup>, som skabes i beskæftigelsesindsatsen i kommuner, a-kasser, regioner, stat og af borger via selvbetjening. Med relevant menes data som mere end en af ovenstående aktører har behov for at kunne opdatere.
  1. Inden data placeres i DFDG skal det afklares hvem der har behov for tilgang til data og om data allerede findes et andet sted.
2. Derudover skal DFDG indeholde nødvendige data for at kunne styre og administrere data.
3. DFDG baseres på veldefinerede logiske datamodeller, og anvender så vidt muligt allerede definerede entiteter og attributter
4. Data skal inkludere det fulde revisionsspor med identifikation af, hvilke data der er registreret/slettet af hvem og hvornår.
5. Dataentiteter kan inkludere felter med sagsbehandlers begrundelse/noter for den enkelte registrering af hensyn til journalpligten og for at undgå dobbeltregistrering i aftagende systemer.
6. DFDG er og skal ikke være et komplet journalsystem. Generelle dokumenter og notater skal opbevares i de enkelte aftagersystemer.

### Principper for sikring af konsistens i de delte masterdata

1. Ved opdateringer skal alle transaktioner og data-elementer tidsstemples (evt. med sekvensnummer), således at aktører kan se og reagere på det korrekte tidsmæssige forløb af opdateringer.
2. Det er aktørsystemernes ansvar at reagere korrekt, hvis der ved opdatering af DFDG-data under et sagsbehandlingsforløb er sket opdatering af DFDG-data fra anden kilde ("samtidighedsfejl"), og aktørsystemet skal i givet fald gentage læsning af de opdaterede DFDG-data. Ændringer i lokale kopier af DFDG-data må (så vidt muligt) ikke gøres permanente, før ændringen er gennemført i DFDG. Det er det aftagende systems ansvar at sikre at DFDG opdateres korrekt.

3. Data skal kunne læses samtidigt af flere aktører, og ved opdateringer fra en aktør af flere samtidige dataentitetsinstanser, skal dataintegriteten sikres (f.eks. ved database-transaktionshåndtering[3]).
  1. På sigt arbejdes mod at tilbyde transaktioner på tværs af webservicekald
  2. Composite services tilbydes de steder hvor der er en forretningsmæssig samtidighed i registreringerne
4. Data- og forretningsservices, som tilbyder opdatering af data, skal i kvitteringen angive om opdateringen i DFDG lykkedes, eller om der opstod fejl (f.eks. ved kast af SOAP-Fault). Dvs. at aftager får svar med det samme.
5. Når en forretningshændelse, f.eks. som resultat af kald til en forretningsservice, udløses i DFDG, eller når en data-entitetsinstans opdateres, f.eks. som resultat af en dataservice, udsendes besked om hændelsen til alle relevante aktører via besked-kø med tilstrækkelig data til at modtager kan identificere hændelsen i DFDG og dermed den eller de berørte data-entiteter, herunder hvilken type af opdatering der er foretaget.
  1. Aftagere underrettes via WSRM beskeder
6. Aktør-systemer og DFDG, skal kunne danne og anvende unikke nøgler, GUID'er, som identificerer de enkelte instanser af data-entiteter og kan deles mellem alle aktører. De unikke nøgler bør dannes af det system, der også er kilde til data, sådan at samme nøgle følger data fra det skabes i et aftagersystem, til det gemmes i DFDG.
  1. I dag er de fleste services i DFDG konstrueret sådan at det er DFDG der danner nøglerne. I en overgang konstrueres services sådan at DFDG fortsat kan danne nøglen, hvis ikke nøglen medsendes af kildesystemet.

#### **Principper for dataload**

1. DFDG tilbyder dagligt opdaterede udtræk af data på individniveau fordelt på kommune- eller akassemedlemstilhørsforhold
2. Det forretningsmæssige mål med leverance af dataload er primært statistik og ledelsesinformation
3. Data leveres i samme format som det ligger i DFDG

#### **Principper for forretningsunderstøttelse**

1. DFDG skal i nødvendigt omfang stille services til rådighed, der understøtter forretningens behov for adgang til data.
2. DFDG kan implementere automatiske forretningsprocesser i det omfang det vurderes hensigtsmæssigt af hensyn til DFDG's rolle som masterdatabase.
3. DFDG stiller et begrænset antal grundsøgninger til rådighed, disse kan suppleres med avancerede søgninger i aftagersystemernes egne data.
4. Den af STAR skitserede sikkerheds- og autorisationsmodel skal efterleves (Dvs. sikring af transport af data, identifikation af brugere, trust-model etc.).
5. Ved batch-opdateringer af DFDG skal gælde de samme valideringer og regler som ved on-line opdateringer (således at konsistens ikke kompromitteres).

### Standarder for dataentiteter

1. Dataentiteter skal modelleres under hensyntagen til offentlige krav som Grunddataprogrammets modelregler: <http://arkitekturguiden.digitaliser.dk/grunddata-modelregler>
2. Arkitekturguidens krav skal efterleves: <http://arkitekturguiden.digitaliser.dk/standarder>

---

[1] Med 'publicere' menes læseadgang til data. I dag er Read operationen i de fleste CRUD implementeringer samlet i hhv. historik- og status-services.

[2] Som minimum 5 år tilbage i tid.