

## *Styregruppen for data og arkitektur*

# Review-rapport for: Overblik over offentlige data - datasætkatalog

### Indhold

Arkitekturreview af overblik over offentlige data - datasætkatalog	2
Reviewgrundlag	2
Projektresume	2
Indstilling	3
Anbefalinger	3
Anbefalinger til det nuværende projekt	3
Tværgående anbefalinger	5

## Arkitekturreview af overblik over offentlige data - datasætkatalog

Arkitekturreviewet af datasætkataloget er udført på baggrund af projektets fremsendte materialer: Kravspecifikation til udbudsmateriale, sag vedr. opstart af projekt til styregruppen for data og arkitektur, samt screenshots af nuværende pilotløsning. Reviewet er udført i overensstemmelse med model for reviews, godkendt af styregruppen for data og arkitektur maj 2017 og er at betragte som et designreview. Reviewboardet og deltagere er listet i nedenstående tabel:

<b>Reviewboard:</b>	Gerner Zinglersen, KL
	Rune Reimann Petersen, KL
	Heidi Vanparys, SDFE
	Allan Kristensen, Miljøstyrelsen
<b>Sekretariat for 8.1:</b>	Sarah Kirkeby Danneskiold-Samsøe, sek.
	Michael Philip Poulsen, sek.
	Mads Hjort, arkitekt
	Lars Thomsen, arkitekt
<b>Projektdeltagere:</b>	Zenia Nørregaard, projektleder
	Thoke Graae Magnussen, arkitekt

## Reviewgrundlag

Udgangspunktet for reviewet af Datasætkataloget udgøres af hvidbog om arkitektur for digitalisering, herunder i særdeleshed hvidbogens principper og regler for arkitektur godkendt i styregruppen for data og arkitektur på dennes møde 24. maj 2017. Principperne er gengivet nedenfor.

1. *Arkitektur styres på rette niveau efter fælles rammer (styring)*
2. *Arkitektur fremmer sammenhæng, innovation og effektivitet (strategi)*
3. *Arkitektur og regulering understøtter hinanden (jura)*
4. *Sikkerhed, privatliv og tillid sikres (sikkerhed)*
5. *Processer optimeres på tværs (opgaver)*
6. *Gode data deles og genbruges (information)*
7. *It-løsninger samarbejder effektivt (applikation)*
8. *Data og services leveres driftssikkert (infrastruktur)*

## Projektresumé

Projektet skal etablere en løsning, så det er nemt for initiativerne i digitaliseringsstrategien at beskrive deres datasæt på en ensartet måde, og som automatisk samler disse datasæt-beskrivelser i et fællesoffentligt datasætkatalog. Løsningen skal anvendes af digitaliseringsstrategiens initiativer til dokumentation og registrering af datasæt, som de anvender, ønsker at anvende eller har dataejerskab over. På sigt ses løsningen anvendt bredt til dokumentation af datasæt i den offentlige sektor. Løsningen skal alene give et overblik over hvilke datasæt, der findes – ikke distribuere selve datasættene. Løsningen forventes baseret på internationale standarder og løsninger herfor (DCAT-

AP standarden fra EU og CKAN-løsningen) samt særligt erfaringerne fra OpenData DK samarbejdet, VirkData og geodata-info.dk.

## Indstilling

*Det er reviewboardets vurdering, at arkitekturen for datasætkataloget er i overensstemmelse med principper og regler for den fællesoffentlige arkitektur, som fremstillet i hvidbog om fællesoffentlig digital arkitektur, givet at reviewboardets anbefalinger følges.*

5 ud af 8 af hvidbogens principper vurderes i gult. Dette skal overordnet ses som et udtryk for, at projektet er på vej i den rigtige retning, men at projektet bør følge boardets anbefalinger ift. det endelige design af løsningen i relation til den fortsatte overensstemmelse med hvidbogens principper og regler.

Niveau	Vurdering
Styring	Fuldt opfyldt
Strategi	Fuldt opfyldt
Jura	Fuldt opfyldt
Sikkerhed	Delvist opfyldt
Opgaver	Delvist opfyldt
Information	Delvist opfyldt
Applikation	Delvist opfyldt
Infrastruktur	Delvist opfyldt

Reviewbordet har udarbejdet **13** anbefalinger, givet i denne review-rapport.

10 anbefalinger er til det nuværende projekt. Projektet anmodes om at tage stilling til disse anbefalinger i en handlingsplan.

Hertil er udarbejdet 3 tværgående anbefalinger, hvortil sekretariatet laver beslutningsoplæg til styregruppen for data og arkitektur.

## Anbefalinger

Reviewet af datasætkataloget har identificeret en række anbefalinger, der fremstår i to kategorier:

1. Anbefalinger til det nuværende projekt: Herunder fremstår anbefalinger til projektet i dets nuværende og kommende faser, som det er præsenteret for reviewboardet.
2. Tværgående anbefalinger: Disse anbefalinger identificeres af reviewboardet som centrale og relevante for projektets fremtidige succes, men samtidig af en sådan karakter, at disse udfordringer ikke kan løses af projektet isoleret set.

For så vidt angår anbefalinger i kategori 1 anmodes projektet om at imødegå disse ud fra følg-eller-forklar princippet i deres bemærkninger til review-rapporten samlet i en handlingsplan. For anbefalinger i kategori 2, udarbejder sekretariatet for initiativ 8.1 beslutningsoplæg. Disse beslutningsoplæg indgår ligeledes i styregruppen for data og arkitekturs behandling af reviewet.

## Anbefalinger til det nuværende projekt

### 1. Det anbefales, at projektet øger samarbejdet med ejere af andre (datasæt-)kataloger.

En række myndigheder har på nuværende tidspunkt allerede etableret datasætkataloger, fx som følge af INSPIRE-krav på SDFE's område. Datasætkataloget skal eksistere i samspil med disse kataloger, og reviewboardet anbefaler derfor et øget samarbejde med eksisterende katalogejere, herunder med fokus på at undgå, at en registrant skal indtaste metadata i mere end ét katalog, og med fokus på validering mellem kataloger.

### 2. Det anbefales, at projektet udarbejder et roadmap for kommende faser.

Udarbejdelse af et fællesoffentligt datasætkatalog er en kompleks opgave, både ift. styring og teknik. Reviewboardet anbefaler, at projektet fokuserer på den nuværende opgave med etablering og styring af et simpelt katalog. Hertil bør udarbejdes et roadmap for kommende udviklinger af kataloget, eksempelvis med inddragelse af anvendelsesmuligheder ift. borger.dk og linked data.

### 3. Det anbefales, at sammenhængen til referencearkitektur for brugerstyring, herunder indvirkning på udbudsstillelse af komponenter, afklares.

Referencearkitektur for brugerstyring er normsættende for fællesoffentlige løsninger. Reviewboardet finder derfor, at datasætkataloget bør overveje, om der skal være en selvstændig brugerstyringskomponent af registranter i løsningen, eller om NemLogin kan anvendes.

### 4. Det anbefales, at projektet afklarer, hvorledes løsningen forholder sig til relevant lovgivning.

Projektet nævner en række lovmæssige relationer i sin kravspecifikation, men forholder sig ikke konkret til, hvordan datasætkataloget evt. afløfter lovmæssige krav eller understøtter digitaliseringsklar lovgivning. Reviewboardet anbefaler derfor, at den lovmæssige tilknytning afklares.

### 5. Det anbefales, at projektet analyserer de tværgående brugerrejser, som løsningen vil indgå i samt følgende krav til udarbejdelse af dokumentation og vejledninger til løsningen.

Datasætkataloget ønsker at kunne udgøre en platform, hvor alle offentlige datasæt kan registreres, hvilket giver en potentiel meget stor brugergruppe. Af dette følger, at kravene til forståelsen af brugsscenarier (brugerrejser) øges samt udarbejdelse af dokumentation og understøttende vejledninger til brugerne. Reviewboardet anbefaler derfor, at tværgående brugerrejser og følgende krav til dokumentation og vejledninger inddrages som et prioriteret element i projektet.

### 6. Det anbefales, at dataflows kortlægges.

Reviewboardet finder, at projektet bør kortlægge de dataflows, hvori datasætkataloget indgår. Herved opnås bedre forståelse for nødvendige mappinger og standardanvendelser i forbindelse med udveksling mellem og høst af andre kataloger samt forståelse for de barrierer, der kan være ved indtastning i datasætkataloget. En central komponent i denne kortlægning er valideringsmuligheder

allerede ved indtastningstidspunktet for slutbrugeren uagtet at denne kan indtaste i en anden løsning, som datasætkataloget blot høster fra.

#### **7. Det anbefales, at der gennemføres en modelreview af datasætkataloget.**

Der er i forbindelse med datasætkataloget udarbejdet en første kernemodel, der vurderes meget detaljeret og antages at befinde sig på niveau 1 ift. de fællesoffentlige begrebs- og datamodelregler. Det er reviewboardets anbefaling, at denne model tilrettes til en egentlig kernemodel med et forretningsmæssigt fokus, og at der i tillæg hertil udarbejdes en begrebsmodel. Reviewboardet anbefaler, at begge disse modeller gennemgår et modelreview.

#### **8. Det anbefales, at løsningens mapninger tydeliggøres.**

For at kunne høste og evt. indgå i et økosystem med øvrige kataloger er mapninger mellem metadatastandarder et centralt element i datasætkatalogløsningen. Det er reviewboardets erfaring, at arbejdet med mapninger er komplekst, og at der er principielle beslutninger om måden at mappe på, herunder om der mappes til en fælles standard fra alle løsninger, eller om der laves bilaterale mapninger mellem løsninger.

Reviewboardet anbefaler, at projektet behandler stillingtagen til måden at mappe på, samt tydeliggør de mapninger, der forventes at indgå i løsningen.

#### **9. Det anbefales, at brugssituationerne hvor retningslinjer for webservices anvendes præciseres.**

Projektet henviser i kravspecifikationen til retningslinjer for webservices på en række krav. Reviewboardet finder, at denne henvisning er på et meget generelt niveau, og at der kan være en række brugssituationer, hvor retningslinjer for webservices ikke er egnede.

Det anbefales derfor, at det præciseres i hvilke brugssituationer, relateret til krav, at retningslinjer for webservices skal bringes i anvendelse.

#### **10. Det anbefales, at rationaler for ønskede formater, herunder RDF, ekspliciteres.**

Projektet stiller i sin kravspecifikation en række krav til formater. Det er reviewboardets vurdering, at rationalerne for valget af netop disse formater ikke fremgår klart af materialet. For at sikre gode designmæssige valg anbefaler boardet derfor, at rationaler for krav kortlægges og ekspliciteres.

### **Tværgående anbefalinger**

#### **11. Det anbefales, at et udvidet mandat for anvendelse af et fremtidigt datasætkatalog afklares snarest muligt.**

Datasætkataloget og dets anvendelse rummer mange muligheder, der kun forøges med udbredelsen af anvendelsen. For nuværende er mandatet begrænset til digitaliseringsstrategiens initiativer, men det er reviewboardets opfattelse, at datasætkataloget rummer langt større potentiale. Skal

mandatet for datasætkataloget udvides kræves det imidlertid også, at governance, vedligehold og ejerskab tænkes ind fra start.

Reviewboardet anbefaler derfor, at mulighederne for et udvidet mandat for anvendelsen af datasætkataloget snarest kortlægges, herunder også følgekrav til governance, vedligehold og ejerskab.

**12. Det anbefales, at identifikation af datasæt i projekter indgår fast som en del af et arkitekturreview.**

Datasætkatalogets nytte er stærkt sammenhængende med antallet af datasæt, der er indeholdt i kataloget. For at sikre populationen af datasætkataloget anbefaler reviewboardet derfor, at identifikation af datasæt indgår som fast element i alle arkitekturreviews.

**13. Det anbefales, at næste revision af hvidbog om fællesoffentlig digital arkitektur overvejer, hvordan forretningsbegrebet kan bringes i anvendelse.**

Reviewboardet finder, at der i forhold til en vurdering af et projekt som datasætkataloget på baggrund af hvidbogens principper og regler, mangler en måde at adressere forretningen på. Boardet anbefaler derfor, at dette behov afdækkes og evt. inkluderes i næste revision af hvidbogen.