

Forsyningsdigitaliseringsprogrammet

Tværgående Arkitektur og Udvikling (TAU) Leverance - Fælles retningslinjer for forsyningssektorens metadata

Udarbejdet: TAU arbejdsplan 2. vedr. interoperabilitet

Medlemmer af arbejdsgruppe:

Niels Ejnar Helstrup Jensen, Senioringeniør, Energinet

Bruno Alexander Højrizi, Leder af GIS center, Aalborg Forsyning

Steen Grube, Tidl. Head of Department, Tidl. HOFOR

Steen Kramer, Senior specialist, Dansk Fjernvarme

Andreas Okkels Overgaard Thomsen, Business developer, Center Denmark

Maya Borge, Dataarkitekt, Digitaliseringsstyrelsen

Helle Pernille Hansen, Fagleder, Aarhus Vand

Peter West-Nielsen, Funktionsleder, Klimadatastyrelsen

Morten Lindegaard, Chefkonsulent, Klimadatastyrelsen

Lars Erik Storgaard, Chefkonsulent, Klimadatastyrelsen

FDP-sekretariatet

Nicklas Vandfeldt Hansen, Klimadatastyrelsen

Version af leverance: 1.0

Leverancebeskrivelse fremgår på **Forsyningsdigitaliseringsprogram.dk**



Forsyningsdigitaliseringsprogrammet

Indholdsfortegnelse

1. Indledning og formål	3
1.2 Baggrund.....	3
1.2 Formål.....	4
2. Governance og forankring	4
2.1 Governance af retningslinjer i regi af Forsyningsdigitaliseringsprogrammet	4
2.2 Forankring af retningslinjer i sektoren	4
3. Værdiskabelse.....	4
3.1 Discovery metadata.....	5
4. Anbefalet anvendelse af retningslinjerne	6
5. Anbefalinger til implementering	7
5.1 Anbefaling 1 – Udstillede data skal have metadata	8
5.2 Anbefaling 2 – Metadata skal følge DCAT-AP-DK.....	10
5.3 Anbefaling 3 – Udvalgte elementer fra DCAT-AP-DK implementeres.....	11
5.4 Anbefaling 4 – Et centralt datasætkatalog etableres	13
5.5 Anbefaling 5 – En governance for metadata etableres	16
5.6 Alle anbefalinger.....	19
6. Ordliste	20
Bilag 1	21
Bilag 2	22
Bilag 3	23

1. Indledning og formål

1.2 Baggrund

Retningslinjerne for metadata er udarbejdet af udvalget Tværgående Arkitektur og Udvikling i regi af Forsyningsdigitaliseringsprogrammet. Retningslinjerne tager afsæt i målsætning 2 om ensretning af data og digitalisering, som er fastsat af Forum for Forsyningsdata (FFD).

Forsyningsdigitaliseringsprogrammet (FDP) skal skabe et fælles grundlag for digitalisering i forsyningssektoren. Derfor er der igangsat et arbejdsprogram vedr. interoperabilitet i regi af Tværgående Arkitektur og Udvikling (TAU), hvor relevante repræsentanter fra forsyningssektoren og myndigheder er repræsenteret. Arbejdsprogramet arbejder blandt andet med leverance 3 om fælles retningslinjer for dataformater og metadata i FDP.

Retningslinjerne for metadata udspringer af TAU leverance 3, *Fælles retningslinjer for dataformater og metadata for forsynings- og produktionsdata* samt TAU leverance 4, *Fælles retningslinjer for datakvalitet*, i TAU's arbejdsprogram under FDP.

TAU anbefaler, at de udarbejdede retningslinjer udgør en fælles metode for Discovery metadata, som er metadata, der hjælper med at identificere og finde datasæt. Metadaten skal dokumentere forsyningssektorens data, der deles med anvendere. Retningslinjerne indeholder anbefalinger til de centrale metadataelementer, som alle er nødvendige for at realisere metadataets fulde værdipotentialer. Det indebærer:

1. Udstillede data skal som standard ledsages af metadata,
2. Metadata skal følge standarden DCAT-AP-DK
3. Udvalgte elementer fra DCAT-AP-DK anvendes som standard
4. Der bør etableres et centralt datasætkatalog til registrering, vedligeholdelse og udstilling af metadata.
5. Der bør etableres en governance-struktur med ansvar for vedligeholdelse og ensartet anvendelse.

Anbefaling 1–3 kan tages i anvendelse i forbindelse med udstilling af data. Den fulde værdiskabelse forudsætter etablering af et datasætkatalog for metadata samt en tilhørende governance-struktur (anbefaling 4 og 5). Dette er centralt for at understøtte kontrollen af metadata og påse at metadata etableres og vedligeholdes i overensstemmelse med retningslinjerne samt, at nye forretningsbehov eller standard-opdateringer implementeres ensartet. Uden dette mister metadata hurtigt den præcision, der gør datasæt søgbare og forståelige, og det bliver svært at håndhæve og udvikle retningslinjerne over tid. Dette uddybes yderligere i afsnit "5.6 Alle Anbefalinger"

Der vil blive arbejdet videre med anbefaling 4 og 5, herunder afklare de nærmere behov og hvordan de bedst understøttes. jf. *plan for implementering af fælles retningslinjer for metadata*.

Konkret anbefaler TAU de retningslinjer, der er beskrevet i afsnit 5: "Anbefaling til implementering".

Yderligere information om Forsyningsdigitaliseringsprogrammet kan findes på programmets hjemmeside: www.forsyningsdigitaliseringsprogram.dk

1.2 Formål

Formålet med retningslinjerne er at beskrive, hvordan forsyningsdata, som er tilgængelig for anvenderne, bør dokumenteres med metadata af forsyningssektorerne. Metadata beskriver data ved hjælp af en række informationer, f.eks. titel, dato, dataformat, brugsbetingelser, datakvalitet, hvordan data er indsamlet og behandlet mm. Metadata gør det lettere for anvendere at finde, forstå og vurdere data ud fra et givent behov. Retningslinjerne skal understøtte en mere ensartet dokumentation af de data, som er tilgængelig for anvendere, så disse får lettere ved at fremsøge, vurdere, få adgang til og genanvende data i forsyningssektoren og andre sektorer.

2. Governance og forankring

2.1 Governance af retningslinjer i regi af Forsyningsdigitaliseringsprogrammet

I regi af Forsyningsdigitaliseringsprogrammet fungerer TAU som rådgivende organ, og kan gennemføre review og tilpasninger af godkendte fælles retningslinjer i FDP.

Det anbefales, at de udarbejdede retningslinjer revideres årligt, således at de løbende kan blive opdateret ud fra erfaringer, ændrede behov i forsyningssektoren og/eller opdateringer til standarden for metadata.

2.2 Forankring af retningslinjer i sektoren

Brancheorganisationerne vil bidrage til at understøtte udbredelse og forankring af retningslinjerne i deres respektive sektorer gennem egne kanaler, fx branchestandarder, hjemmesider, nyhedsbreve og konferencer mv.

3. Værdiskabelse

Ved udstilling af data, er metadata helt central for at maksimere muligheden for at både finde og anvende data. Metadata er "data om data" – den ekstra kontekst, der fortæller, hvad et datasæt egentlig indeholder, og hvordan det kan anvendes. Metadata giver en struktureret beskrivelse af et datasæt, som gør det muligt for en bruger at finde data og vurdere, om de er nyttige for vedkommende. Metadata i sig selv skal udarbejdes på en måde, der maksimerer muligheden for at finde og bruge dem – her spiller brugen af standarder for metadata en væsentlig rolle.

Metadata er det første, anvendere møder, når de søger efter data. De hjælper dem med at finde data og fortæller dem om de datasæt, de har fundet: Hvad de indeholder, hvem der har produceret dem, hvem der vedligeholder dem, og hvordan de kan bruges. Dette er vigtige beskrivelser af data, som giver anvenderne mulighed for at træffe en informeret beslutning om, hvorvidt data er egnet til den påtænkte brug. Metadata gør derved ikke bare data mere synlige og nemmere at finde, men er helt central for at en udstilling af data kan omsættes til reel værdi ved at understøtte en forbedret tilgængelighed og anvendelighed af data, sikre bedre datahåndtering, højere datakvalitet og forbedrede tjenester.

Et konkret eksempel på eksisterende metadata for data i forsyningssektoren er fx Energinets data om elforbruget opgjort per netområde. Disse metadata er tilgængelige i Energinets datakatalog:

<https://www.energidataservice.dk/tso-electricity/ConsumptionPerGridarea>.

Udfordringer med manglende, uensartede og inkonsistente metadata på tværs af forsyningssektoren besværliggør fx deling, analyse, serviceudvikling og giver øgede omkostninger, da dataanvendere bruger ressourcer på at lede efter datasæt, opbygge lokale "mini-kataloger" og vedligeholde dem løbende. For at afhjælpe disse udfordringer og understøtte en mere effektiv digitalisering på tværs af sektoren, anbefaler arbejdsprojektet fælles retningslinjer for metadata, herunder behov for systemunderstøttelse i form af et datakatalog og en tilhørende governancemodel.

Metadata er værdiskabende ved:

- At gøre det lettere for brugerne at finde data.
- At fortælle brugerne, hvordan de får adgang til data.
- At fortælle brugerne om eventuelle betingelser for genbrug af data.
- At identificere datas oprindelse fra kilde til udstilling, hvilket gør dataene gennemskuelige for brugerne.
- At give brugerne en nøjagtig oversigt over data.
- At vise datakvalitet for at informere om beslutninger om datas egnethed og begrænsninger.

3.1 Discovery metadata

Retningslinjerne beskæftiger sig med Discovery metadata, som referer til metadata, der hjælper med at identificere og finde datasæt, som er relevante for et bestemt behov eller interesse. Det giver vigtig information, der gør det muligt for brugere at opdage og forstå, hvilke data der er tilgængelige og hvor de er opbevaret. Discovery metadata er en mere overordnet beskrivelse og dokumentation af data med hensyn til indhold, formål, livscyklus, adgang, vilkår og kontaktoplysninger mm. Der kan også være registreret metadata på record/række niveau i selve data – eksempelvis et registreringstidspunkt.

Metadata bør altid udformes med fokus på slutbrug og genbrug af data, herunder at data kan anvendes flere gange og til forskellige formål samt henvende sig til en bred vifte af potentielle anvendere af data. Metadata skal holdes opdaterede, hvis der sker ændringer i de data, de beskriver. Oprettelse af metadata er dermed ikke en engangsopgave, men en vedligeholdelsesopgave. Uden opdatering af metadata risikeres det, at metadata mister deres værdi, relevans og troværdighed. Et eksempel på en ændring som kræver opdatering af metadata er, at data bliver tilgængelige i et nyt filformat eller i en nyoprettet API. Et andet eksempel kan være at organisationen, som ejer data, implementerer en ny licens for brug af data. Denne nye licens skal fremgå af metadata.

I 2016 blev »FAIR Guiding Principles for scientific data management and stewardship« offentliggjort i Scientific Data. Metadata udgør F'et i FAIR¹ - data er 'Findable, Accessible, Interoperable and Reusable' Principperne har til hensigt at give retningslinjer for at forbedre datas søgbarhed (Findable), tilgængelighed (Accessible), kombinerbarhed (Interoperable) og genanvendelighed (Reusable). FAIR principper anses som god praksis indenfor arbejdet med datainfrastrukturer.

¹ <https://www.go-fair.org/>

Udviklingen af et stærkere datafundament på tværs af forsyningssektoren bør omfatte udarbejdelse af relevant metadata i tråd med EU's data space dagsorden baseret på internationale standarder. Dette skal gøre det lettere at identificere data, herunder oplysninger om blandt andet datas placering og indhold.

Værdien ved metadata er også italesat i EU's "Dataforordningen"², som har til formål at fremme et fælles europæisk data space, hvor data kan deles på tværs af sektorer og lande på en sikker og pålidelig måde. Dataforordningen definerer metadata som "en struktureret beskrivelse af indholdet eller anvendelsen af de data, hvormed søgningen i eller anvendelsen af disse data lettes".

Det skal bemærkes, at det er en anerkendt præmis, at arbejdet med metadata er under den såkaldte "høste-så problematik", som henviser til den situation, hvor de omkostninger og gevinster, der er forbundet med en indsats ikke nødvendigvis er ensartet fordelt mellem de forskellige aktører. Der vil derfor være omkostninger forbundet med at indtaste og vedligeholde metadata for én aktør (udstiller), hvorimod en anden aktør (anvenderne) høster gevinsten ved at få overblik og viden om data.

4. Anbefalet anvendelse af retningslinjerne

Retningslinjerne henvender sig primært til IT- og data-arkitekter, der designer og implementerer IT- og data-løsninger til deling af data med anvendere i forsyningssektoren, men det er tilstræbt, at retningslinjerne også kan læses af andre.

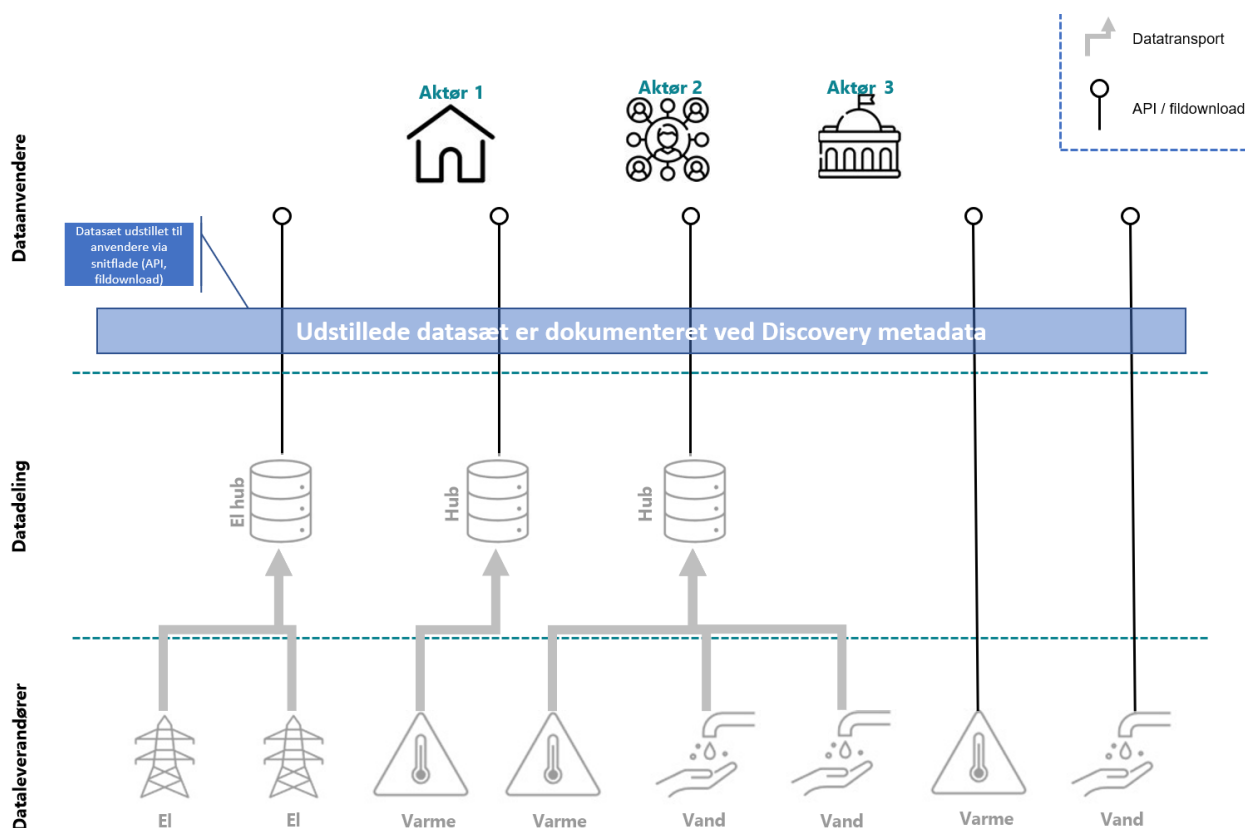
Retningslinjerne omhandler discovery metadata for udstillede datasæt. Formålet er at sikre, at anvendere både kan finde relevante data og få den nødvendige information om datasættet for at kunne anvende det korrekt. Retningslinjerne udgør ikke en datamodel for feltbeskrivelser af de data, der transporteres fra forsyningsvirksomheder til de aktører, som udstiller data for brugerne. Discovery metadata handler således om at gøre datasæt lette at finde og forstå, og således ikke om at definere selve datamodellen inde i systemerne. Dette er illustreret i figuren nedenfor.

Retningslinjerne udgør en fælles metode for Discovery metadata, som dokumenterer de data, som udstilles til anvenderne – eksempelvis via API og eller fildownloads.

Retningslinjerne udgør ikke en fælles referenceramme for metadata forstået som datamodel eller "feltbeskrivelser" af de data, som transporteres fra de enkelte forsyningsselskaber til de aktører (eksempelvis en hub), som opsamler, evt. aggregere og processere og udstiller data til anvenderne.

Discovery metadata's indplacering i en overordnet arkitektur er på det plan, hvor data er udstillet til anvenderne via hubs eller direkte fra forsyningsselskabet til anvendere som illustreret i figur 1.

² <http://data.europa.eu/eli/reg/2023/2854/oj>



Figur 1. Fælles referenceramme for metadata gælder for de data, som udstilles til anvenderne. Disse data kan være udstillet i API og eller fildownloads i datasæt fra en hub eller direkte fra en dataleverandør. Fælles er, at de udstillede data er dokumenteret med metadata i henhold til retningslinjerne. Derimod skal de data, som transporteres fra de enkelte dataleverandører (forsyningsvirksomheder) til en udstillingsplatform (en hub) ikke dokumenteres med metadata efter disse retningslinjer.

5. Anbefalinger til implementering

TAU anbefaler, at de udarbejdede retningslinjer udgør en fælles metode for Discovery metadata, som dokumenterer forsyningsbranchens data, der deles med anvendere. Såfremt retningslinjerne ikke anvendes, bør det overvejes, hvorfor der afviges fra retningslinjerne.

Retningslinjerne kan anvendes efter behov i forbindelse med udstilling af data. Den fulde værdiskabelse forudsætter imidlertid etablering af et datasætkatalog (IT-system) for metadata samt tilhørende governance der sikre den løbende vedligeholdelse og ensartethed på tværs. Dette beskrives nærmere i afsnit "5.6 Alle anbefalinger".

En aktiv udbredelse af retningslinjerne anbefales derfor først at ske, såfremt der er aftalt en finansieringsmodel for datasætkatalog og governance, og etableringen kan igangsættes.

Det overordnede rationale er, at de data, som udstilles, skal være dokumenteret og deklareret i en sådan grad, at anvenderne kan:

- Fremsøge data
- Vurdere datas egnethed ift. et behov
- Få adgang til data
- Få indsigt i vilkår og evt. begrænsninger i at anvende data mm.

Dokumentationen bør være struktureret ensartet på tværs af udstillingsplatforme (eksempelvis hubs) og på tværs af forsyningssektorerne.

Retningslinjerne indeholder derfor anbefaling til:

- At en række af relevante metadataelementer fra en international standard anvendes
- At der implementeres ét fælles datasætkatalog til adgang til metadata
- At der etableres en governance til sikring af at retningslinjerne dels vedligeholdes og dels efterleves.

Nedenfor følger de enkelte anbefalinger sammen med en begrundelse for hver enkelt.

5.1 Anbefaling 1 – Udstillede data skal have metadata

Arbejdssporet anbefaler, at dataudstillere i og omkring forsyningssektoren forpligter sig til, at en dataudstilling skal være dokumenteret med Discovery metadata efter nærværende fælles retningslinjer.

Rationale

Det overordnede rationale er, at de data, som udstilles, skal være dokumenteret og deklareret i en sådan grad, at anvenderne kan fremsøge data, vurdere datas egnethed ift. et behov, få adgang til data, få indsigt i vilkår og evt. begrænsninger i at anvende data mm.

Forsyningssektorens forretningsbehov for metadata skal opfyldes. I nedenstående tabel 1 fremgår de informationer, som er vurderet som dem, der giver [tilstrækkelig]forretningsmæssigt værdi. For ”anbefaling 3 – Udvalgte elementer fra DCAT-AP-DK implementeres” er kravniveauet for metadata for forsyningssektorens udstillede data specificeret jf. tabel 2.

Rækken af informationer kan med tiden udvides (eller indskrænkes), hvis relevante aktører har behov for at dokumentere udstillede data med flere (eller færre) informationer. Ændringen skal ske i den etablerede governance.

Informationerne ses i tabel 1 med navn, beskrivelse og hvilket behov de enkelte informationer understøtter. Eksempler på metadatainformationer fremgår også.

Metadata-information	Beskrivelse	Forretningsbehov og eksempel
Titel	Det eller de ord som navngiver datasættet.	Identificerer datasættet og giver anvenderen en umiddelbar viden om data. Eksempel: Forbrugsdata fra Æblerøds Forsyning
Beskrivelse	En tekstbaseret beskrivelse af datasættets formål og indhold.	Giver anvenderne viden om datasættet mht. formål og indhold vha. en tekstuel beskrivelse. Eksempel: Forbrugsdata beregnet ud fra måledata for vand og fjernvarme i Æblerøds Forsynings forsyningsområde. Data omfatter stamdata om installationspunkter og forbrugstal opgjort på årligt basis.
Ejer	Den organisation der har det juridiske ansvar for det samlede datasæt.	Giver anvenderne viden om hvem der er ejer data. Eksempel: Æblerød Forsyning A/S
Udgiver	Den aktør som primært er ansvarlig for at gøre datasættet tilgængelig	Giver anvenderne viden om hvem der udstiller data – udgiveren (en hub) være en anden end ejeren (forsyningsselskab). Eksempel: Data HUB
Kontakt-punkt		Giver anvenderne mulighed for at komme i kontakt med aktør for dialog om datasættet.

	Kontaktoplysninger som kan anvendes til at indsende spørgsmål eller kommentarer om datasættet.	Eksempel: kontakt@aebleroedforsyning.dk
Nøgleord	Et nøgleord fra en defineret liste.	Nøgleord giver anvenderne mulighed for at søge efter datasæt tilhørende en kategori eller et emne. Eksempel: Vand, Fjernvarme
Oprindelse	Udsagn omkring datasættets proveniens i form af udsagn om eventuelle aktørskift som er centrale for autenticiteten, integriteten og fortolkningen af ressourcen.	Giver anvenderne viden om hvordan data er indsamlet/skabt/etableret. Ydermere hvilke processer data har gennemgået. Det er også her anvenderne kan få viden om, hvilken kvalitet data har og eksempelvis hvordan eventuelle outliers og huller i data er behandlet. Eksempel: Måledata er fjernaflæst og er input til beregning af forbrugsdata. I tilfælde af manglende måledata bruges lineær interpolation til beregning hvor de – til den manglende – to nærmeste måledata bruges som input.
Udgivelsesdato	Den dato hvor datasættet først blev formelt udgivet	Anvenderne får herigennem en viden om alderen på datasættet. Eksempel: 2020-02-02
Opdateringsfrekvens	Med hvilken frekvens datasættet opdateres.	Viden om opdateringsfrekvens kan anvenderne bruge til at beslutte hvor ofte de skal hente de nye data til deres systemer. Eksempel: https://publications.europa.eu/resource/authority/frequency/ANNUALY
Dokumentation	En webside eller et dokument som dokumenterer strukturen i datasættet.	Mere detaljeret information om data – såsom en oversigt over feltbeskrivelser eller et objekttypekatalog er centralt når en anvender eksempelvis vil etablere integrationer mellem data. Eksempel: https://datahub.dk/aebleroedforsyning/data/dataset2/feltbeskrivelse
Adgangsinformation	En URL som giver adgang til en distribution af datasættet.	Anvenderne kan få adgang til data – eksempelvis til et API som kan direkte importeres i et anvendelsesystem eller til fildownload hvorfra data kan hentes. Eksempel: https://api.energidaservice.dk/dataset
Brugsbetingelser	Den brugsbetingelse distributionen er gjort tilgængelig under.	Brugsbetingelser gør det lettere for anvenderne at overskue, om man må bruge data til et givent formål. Information om betingelser i metadata er derfor med til at fremme brugen af data. Eksempel: CC BY 4.0
Personoplysningskategori	En relation til en bestemt personoplysningskategori.	Indeholder data personoplysninger kan det have betydning for, hvordan data må behandles. Oplysning om persondata er derfor relevant i metadata for at gøre anvenderne opmærksom på om anvendelse af data kræver lovhjemmel mm. Eksempel: Personoplysninger (ikke-følsomme personoplysninger)
Geografisk dækning	Et geografisk område som datasættet dækker.	Hvis en anvender er interesseret i at vide "hvad gælder her" såsom "hvilke data findes for kommune X" så er geografisk dækning den information i metadata som denne analyse kan udføres på. Eksempel: Æblerød Kommune [POLYGON((8.094 57.751,15.151 57.751,15.151 54.569,8.094 54.569,8.094 57.751))]
Format	Filformatet for distributionen.	Anvenderne får viden om det filformat som data er tilgængelig i. Det er brugbart ift. opsætning af integrationer mm. Eksempel: CSV, JSON
Yderligere information	En webside som giver brugeren adgang til datasættet, dets distributioner og/eller	Dataejeren kan have yderligere information om datasættet som anvenderne med fordel kan søge ned i.

	yderligere information såsom case-beskrivelser, vejledninger mm	Eksempel: https://www.energidataservice.dk/tso-electricity/DatahubMeasuringPointStatistics
Sprog	Det sprog der er anvendt i datasættet.	Hvis datasættet indeholder tekster er det relevant for anvenderne at få viden om hvilket sprog teksterne er skrevet i. Eksempel: http://publications.europa.eu/resource/authority/language/DAN

Tabel 1: Metadatainformationer som er relevante til at beskrive forsyningsdata.

I forhold til førnævnte behov for at metadata holdes opdateret, så er det værd at bemærke, at det ikke er alle ovenstående metadatainformationer som nødvendigvis skal ændres. Nogle informationer vil være statiske gennem datasættets levetid. Eksempelvis sprog anvendt i datasættet. Mens andre informationer kan være nødvendige at opdatere, hvis der sker ændringer i eller omkring datasættet. Eksempelvis hvis organisationen skifter navn, hvis data bliver opdateret med anden frekvens osv.

5.2 Anbefaling 2 – Metadata skal følge DCAT-AP-DK

Arbejdssporet anbefaler, at DCAT-AP-DK³ vælges som standard for metadata for forsyningssektorens udstillede data.

Rationale

Ved at dokumentationen er struktureret ensartet får anvenderne nemmere ved at sammenligne de tilgængelige datasæt. DCAT-AP er af EU Kommissionen nævnt som udgangspunkt for metadata til brug i EU Data Spaces⁴. I regi af Den Fællesoffentlige Digitale Arkitektur (FDA) har Digitaliseringsstyrelsen udarbejdet en dansk profil af DCAT-AP i tæt samarbejde med en fællesoffentlig følgegruppe, som repræsenterer både kommuner, regioner og statslige myndigheder. Profilen hedder DCAT-AP-DK og indeholder for danske forhold relevante tilpasninger. DCAT-AP-DK anvendes i det fælles offentlige datasætkatalog datavejviser.dk.

En metadatastandard, herunder et aftalt sæt regler og retningslinjer for, hvordan metadata (data om data) skal struktureres, navngives, beskrives såsom DCAT-AP-DK giver en fælles (international) ramme til at beskrive datasæt. Den kan ses som en specifikation for de elementer, som skal fremgå af metadata mht. hvad elementerne skal hedde, hvilke type information (tekst, dato mm.) de skal indeholde og om de er obligatoriske eller frivillige at udfylde. Der findes en række forskellige internationale standarder for metadata.

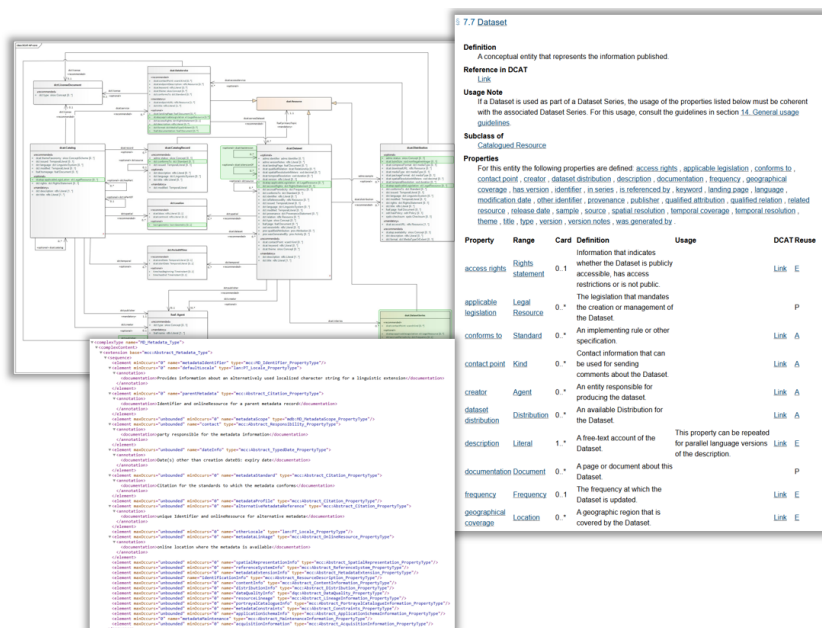
Et fællestræk ved de forskellige standarder er, at de bruger UML til at beskrive strukturen, relationerne og indholdet af metadata. Ud over UML vil de enkelte elementer typisk være beskrevet mere dybdegående mht. definition, beskrivelse, eksempel og en kilde mm.

Standarden fastlægger typisk en række minimumselementer som altid skal fremgå i metadata. Sådanne elementer omtales som obligatoriske. Derudover vil der være anbefalede og/eller frivillige elementer. Standarden er typisk tilgængelig som et pdf-dokument og/eller en online version. Endelig vil der som oftest være et eller flere tilhørende schemaer (XML) til implementeringer. Figur 2 illustrerer materiale, der typisk hører til en standard.

³ Læs om DCAT-AP-DK: <https://digst.github.io/DCAT-AP-DK/releases/v.2.0/>

⁴ 2023-10-04 Webinar on DCAT-AP for data spaces (<https://interoperable-europe.ec.europa.eu/collection/semantic-support-centre/event/webinar-dcat-ap-data-spaces>)

Standarder vedligeholdes i en fastlagt governance, som typisk indbefatter indsamling af ændringsønsker, udarbejdelse af revideret udgave og høringer af den reviderede udgave inden offentliggørelse. Det er derfor vigtigt at henvise til standardens versionsnummer, når den anvendes, eller der henvises til den.



Figur 2. Standarder for metadata omfatter typisk UML-diagrammer, objektkatalog med definitioner, krav, type, eksempler mm. og tilhørende schemaer (XML) til implementeringer.

5.3 Anbefaling 3 – Udvalgte elementer fra DCAT-AP-DK implementeres

Arbejdssporet anbefaler, at de udvalgte elementer fra DCAT-AP-DK (dem som vurderes som mest værdiskabende for forsyningssektoren) udvælges og implementeres til metadata.

Rationale

DCAT-AP-DK indeholder mange elementer til dokumentation af data. For ikke at dels overbebyrde den aktør som udstiller data og dels "drukne" anvenderne i dokumentation udvælges det sæt af minimums-elementer som ud fra erfaringer med metadata vurderes at give mest værdi. Listen med elementer kan ændres hvis behovet opstår. En ændring bør ske i den etablerede governance.

Tabel 1 (som fremgår i anbefaling 1 på side 8) indeholder de metadata-informationer, som er identificeret som de mest værdiskabende til at beskrive de udstillede data. I dette afsnit følger, hvordan informationerne mappes ind til DCAT-AP-DK.

Mapningen fremgår af tabel 2 (se nedenfor), og det bemærkes, at der er flere informationer med i tabel 2 end i tabel 1 fordi DCAT-AP-DK kræver det – eksempelvis er information om "Genstand for registrering" en obligatorisk information og skal udfyldes med "datasæt, distribution eller dataservice". Tabel 2 har kolonnerne:

- "Metadatainformation" som stammer fra tabel 1,
- "DCAT-AP-DK v.2.0.2" som viser, hvilket felt i standarden som metadata-informationen mappes til,
- "Multiplacitet" som viser om feltet skal have én værdi, flere værdier eller om det kan være tomt, og "Kravniveau", som fortæller om information skal eller bør gives.

Kravniveauet er specificeret for metadata for forsyningssektorens udstillede data. Kravniveau adskiller sig derfor for nogle elementer fra den oprindelige DCAT-AP-DK. Det er vurderet, at kravniveau er passende for forsyningssektoren, og at det repræsenterer en passende balance mellem den opgave dataejere skal udføre og den værdi det skaber for anvenderne .

Kravniveau:

- Obligatorisk (Mandatory): informationen SKAL angives
- Anbefalet (Recommended): informationen BØR angives, hvis den er tilgængelig

Multiplicitet:

1: En og kun en værdi må angives.

0..1: Ingen eller én værdi må angives.

0..*: Ingen eller flere værdier må angives.

1..*: En eller flere værdier må angives.

	Metadatainformation	DCAT-AP-DK v.2.0.2	Multiplicitet	Kravniveau
Dataset	Titel	dct:title	1	Obligatorisk
	Beskrivelse	dct:description	1	Obligatorisk
	Ejer	dct:rightsHolder	0..1	Obligatorisk
	Udgiver	dct:publisher	0..1	Anbefalet
	Kontaktpunkt	dcat:contactPoint	0..*	Anbefalet
	Nøgleord	dcat:keyword	0..*	Anbefalet
	Oprindelse	dct:provenance	1	Obligatorisk
	Udgivelsesdato	dct:issued	0..1	Anbefalet
	Opdateringsfrekvens	dct:accrualPeriodicity	0..1	Anbefalet
	Dokumentation	foaf:page	1	Anbefalet
	Adgangsinformation	dcat:distribution	1..*	Obligatorisk
	Personoplysningskategori	dcat-ap-dk:personaldatacategory	1	Obligatorisk
	Geografisk dækning	dct:spatial	0..1	Anbefalet
	Yderligere information	dcat:landingPage	0..*	Anbefalet
	Sprog	dct:language	0..*	Anbefalet
	Genstand for registrering*	rdf:type	1	Obligatorisk
	Identifikation	dct:identifier	1	Obligatorisk
distribution	Adgangsadresse	dcat:accessURL	1..*	Obligatorisk
	Format	dct:format	0..1	Obligatorisk
	Beskrivelse	dct:description	0..*	Anbefalet
	Brugsbetingelser	dct:license	1	Anbefalet
contactPoint	Navn	vcard:hasOrganizationName	1	Obligatorisk
	Email	vcard:hasEmail	1	Obligatorisk

Telefon	vcards:hasTelephone	0..*	Anbefalet
---------	---------------------	------	-----------

Tabel 2: Mapning til den anbefalede metadatastandard, DCAT-AP-DK med angivelse af multiplicitet og kravniveau.

*Metadataelementerne knytter sig til DCAT-AP-DKs Katalogpost.

I bilag 1 er en tabel med udelukkende de obligatoriske metadatainformationer.

I bilag 2 er et RDF/XML-eksempel på metadata for forsyningsdata i henhold til DCAT-AP-DK.

Læs mere om DCAT-AP-DK her: <https://digst.github.io/DCAT-AP-DK/releases/v.2.0/docs/>

5.4 Anbefaling 4 – Et centralt datasætkatalog etableres

Arbejdssporet anbefaler, at der etableres et centralt datasætkatalog, et IT-system, hvori metadata for udstillede forsyningsdata kan indtastes eller indlæses, vedligeholdes og udstilles fra.

Rationale

Ved at oprette et centralt datasætkatalog kan anvenderne tilbydes ét sted at søge informationer om forsyningsdata. Det skaber værdi for anvenderne ikke at skulle søge bredt på internettet efter relevante data. Ydermere vil det fritage de aktører som udstiller data selv at skulle implementere og vedligeholde et datasætkatalog til egne data. Der vil i så fald være tale om et "domænespecifikt katalog" for forsyningssektoren.

Et domænespecifikt katalog er et katalog, hvor organisationer, hvis datasæt er registreret i kataloget, indgår i et interessefællesskab. Dette kunne eksempelvis være et fællesskab om forsyningsdata.

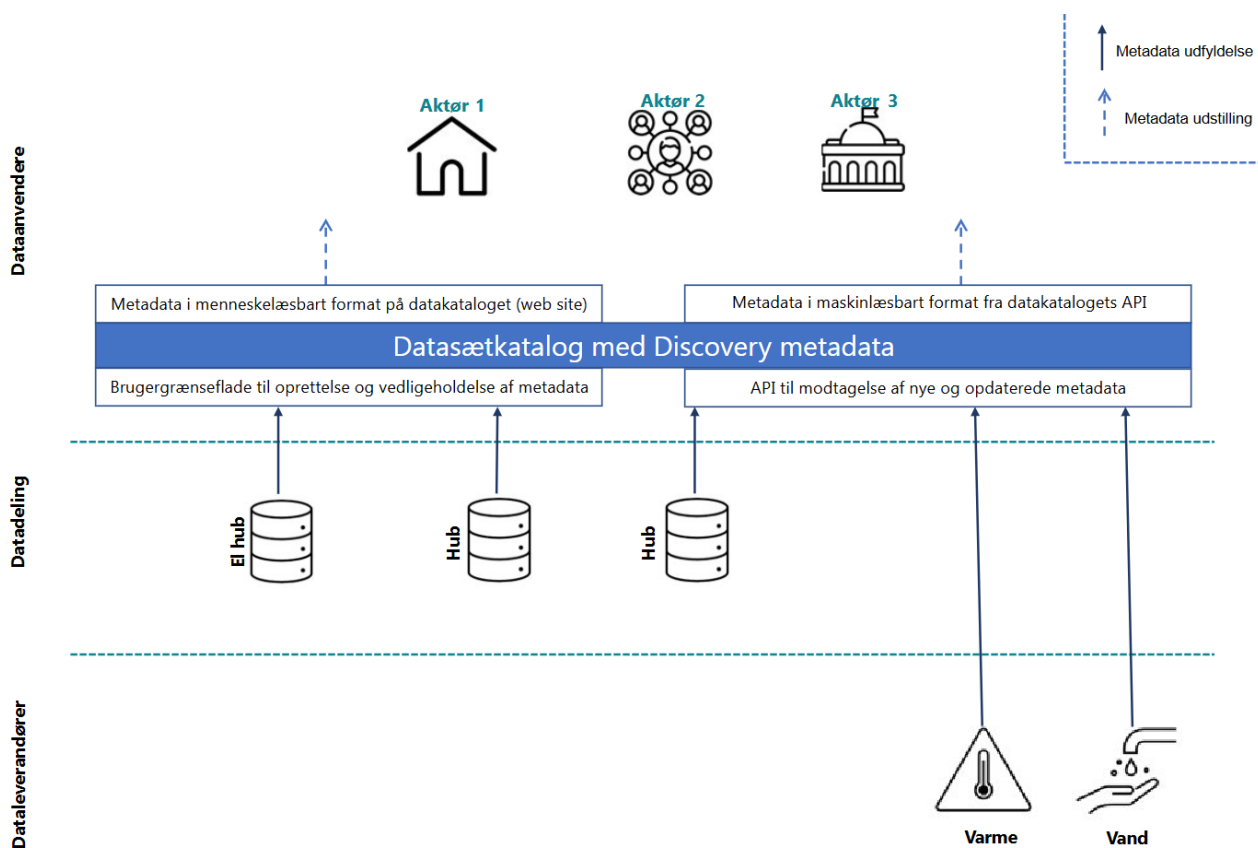
Domænespecifikke datasætkataloger, som fx EIOverblik, har ofte registreringsbehov, der er særligt tilpasset det pågældende domæne. Disse behov vil formentligt ligge uden for det, der dækkes af anvendelsesprofilen, men domænespecifikke kataloger vil stadig kunne benytte profilen dels som udgangspunkt for en udvidet profil, dels til at lade de domænespecifikke informationer indgå i et større nationalt katalog.

Et fælles datasætkatalog for forsyningsdata bør have defineret følgende overordnede forretningskrav til systemunderstøttelse:

- Understøtter DCAT-AP-DK.
- Kan opdatere metadata via GUI⁵ og via API.
- Kan udstille metadata via GUI og via API.

Figur 3 illustrerer placeringen af et datakatalog i et systemlandskab, hvor data udveksles mellem leverandører, delingsplatforme og anvendere. Der er ét fælles datasætkatalog, hvor de udstillede forsyningsdata vedligeholdes og udstilles med Discovery metadata. Metadata kan vedligeholdes og udstilles via en brugergrænseflade og en API

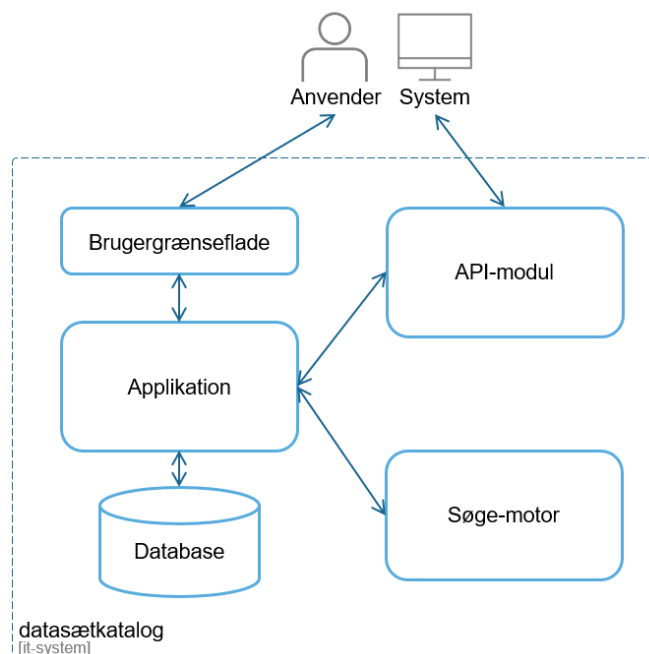
⁵ Graphical User Interface (grafisk brugergrænseflade). Det er en visuel måde for brugere at interagere med en computer eller en enhed. I stedet for at bruge tekstbaserede kommandoer, bruger en GUI ikoner, knapper, vinduer og andre grafiske elementer, som brugeren kan klikke på, trække og slippe for at udføre handlinger.



Figur 3. Indplacering af et fælles datasætkatalog med Discovery metadata for de udstillede data for forsyningsdata i samme overordnede arkitektur som vist i figur 1. Metadata kan vedligeholdes og er udstillet via brugergrænseflade og API-snitflade.

Overordnet set består et datasætkatalogsystem af moduler til 1) applikation, 2) database, 3) søgefunktionalitet, 4) API-tilgang til hhv. modtagelse og udstilling af metadata. Se figur 4 som en konceptuel arkitekturtegnning over komponenter i et datasætkatalogsystem.

Applikationen vil typisk tilbyde en brugergrænseflade hvor brugeren kan indtaste metadata. Brugergrænsefladen kan konfigureres sådan, at den er tilpasset domænets behov for metadata. Der vil være en validator tilknyttet brugergrænsefladen som vil give besked til brugeren, hvis de indtastede metadata ikke overholder den bagvedliggende standard. I databasen lagres selve metadata.



Figur 4. Konceptuel arkitektur over komponenter i et datasætkatalogsystem.

En væsentlig funktionalitet er API snitfladerne. De giver mulighed for at høste metadata fra andre datasætkataloger – eksempelvis hvis en hub vælger at implementere eget datasætkatalog. Med høstefunktionen opnås gevinsten i, at metadata kun skal indtastes og vedligeholdes ét sted.

API snitfladerne giver også muligheder for at andre datasætkataloger kan høste fra forsyningssektorens fælles datasætkatalog. Det kan eksempelvis være det fælles offentlige datasætkatalog, datavejviser.dk. Herved bliver forsyningssektorens datasæt synlige for en endnu bredere skare af anvendere, som kan resultere i en øget brug af forsyningsdata.

Et datasætkatalogsystem er et IT-system som alle andre IT-systemer kræver investering i etablering og efterfølgende drift.

Der findes en række forskellige software løsninger til metadata, både proprietære og open source. Her nævnes kort to open source produkter, CKAN og GeoNetwork som allerede i dag anvendes i fælles offentlige datakataloger - henholdsvis i datavejviser.dk og geodata-info.dk.

CKAN (Comprehensive Knowledge Archive Network) er et open-source software, der bruges til at oprette, administrere og dele datasæt, især store mængder åbne data. Det er designet til at hjælpe organisationer med at gøre deres data tilgængelige og søgbare for offentligheden, forskere og udviklere.

CKAN er især populært blandt offentlige myndigheder, forskningsinstitutioner og organisationer, der ønsker at fremme åbne data og datadeling. Det bruges blandt andet til at skabe åbne data-portaler, hvor brugere kan søge, få adgang til og downloade datasæt.

GeoNetwork er et open-source software til datasætkataloger, som organisationer anvender til at organisere, indekser, finde og dele metadata for geodata. GeoNetwork understøtter en række internationale standarder for metadata såsom ISO 19115 og DCAT-AP for metadata.

Hvorvidt der skal bygges en ny IT-løsning til datasætkataloget til dette formål eller om det kan håndteres ved at udbygge eller tilpasse eksisterende løsninger skal afdækkes yderligere som led i en specifikation om

behov og krav til systemet. Dette vil også være grundlaget for at kunne udregne, hvad omkostningerne vil være til at etablere et datasætkatalog.

5.5 Anbefaling 5 – En governance for metadata etableres

Arbejdssporet anbefaler, at der etableres en governance-struktur til vedligeholdelse af retningslinjerne og til at sikre fremdrift i udfyldelsen af metadata og ensartethed i metadata på tværs af forsyningssektoren.

Rationale

Arbejdssporet vurderer, at en stærk governance om forpligtigende samarbejder er nødvendigt for at ressourcer til udfyldelse og vedligeholdelse af metadata allokeres hos de relevante aktører. Erfaringen viser også, at en governance er nødvendigt for at sikre ensartethed i metadata sådan, at de er sammenlignelige på tværs af udstillingsplatforme og sektoren generelt. Ydermere viser erfaringen, at metadatastandarder ændrer sig over tid, hvilket kan have betydning for retningslinjerne. En governance-struktur til vedligeholdelse af retningslinjerne er derfor også nødvendig.

En stærk governance-struktur, som skal sikre at 1) metadata bliver indtastet og vedligeholdt og 2) en redaktionel fastsat linje efterleves for at metadata på tværs af dataudstillingen og sektorerne er ensartede og dermed sammenlignelige for anvenderne, og 3) at datasætkataloget driftes og holdes opdateret mht. komponenter og funktionalitet.

I figur 5 ses roller og ansvar i den anbefalede governancemodel for metadata. Overordnet set indgår disse roller:

- 1) Dataudstillere: Den aktør som udstiller data til anvenderne er ansvarlig for at vedligeholde metadata i det fælles datasætkatalog. Det kan enten ske ved at indtaste metadata direkte i datasætkatalogets brugergrænseflade eller via lokalt datasætkatalog, hvorfra det fælles katalog kan høste.
- 2) Anvender: Den aktør som søger efter forsyningsdata i det fælles datasætkatalog. Vha. metadata fremsøges og vurderes datasæt ift. anvenderens behov. Anvenderen kan søge og se metadata i det fælles datasætkatalog eller høste metadata til eget IT-system og foretage søgningen her.
- 3) Operatør: Den aktør som vedligeholder datasætkataloget og den bagvedliggende IT-infrastruktur (applikationsservere, databaseservere, søgekomponent, brugerdatabase, API-snitflader mm)
- 4) Kontrollør: Den aktør som kontrollerer metadata og påser at dataudstillere etablerer metadata. Kontrollen skal sikre, at metadata overholder retningslinjerne for metadata, og at metadata har en passende detaljeringsgrad svarende til de øvrige metadata i kataloget.

”Operatør- og kontrollør”-rollerne kan tilhøre samme organisation, men det er ikke et krav, at de skal gøre det. Det kan dog være hensigtsmæssigt, at rollerne indgår som en del af en fastlagt governance-struktur, hvor relevante faglige aktører og myndigheder er involveret.

Det vurderes hensigtsmæssigt at ”kontrollør”-rollen varetages af en neutral ikke-kommerciel aktør med ekspertise inden for metadata og metadatastandarder og i tæt samspil med de relevante aktører med faglig viden på domæneområde. Kontrolløren bør fungere som en uafhængig instans, der sikrer:

- At dataudstillere udarbejder metadata i overensstemmelse med de fastsatte retningslinjer,

- at metadata har et detaljeringsniveau, der er konsistent med øvrige poster i kataloget, samt
- at retningslinjerne løbende vedligeholdes.

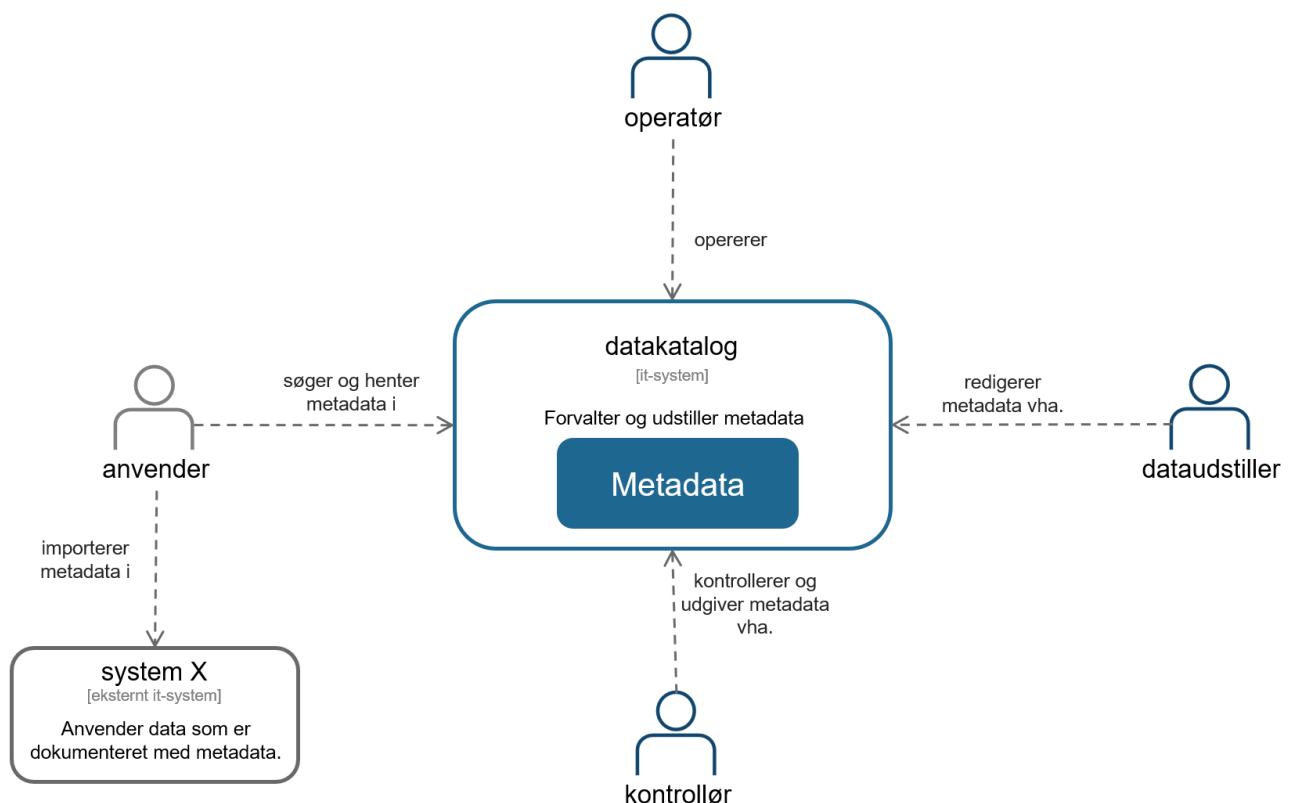
Vedligeholdelsen af retningslinjerne omfatter typisk indsamling og vurdering af ændringsønsker, udarbejdelse af reviderede udgaver samt gennemførelse af høringsprocesser inden offentliggørelse. Kontrollørdelen af governance-strukturen kan derfor fx bestå af følgende repræsentanter fra:

- Forsyningssektoren,
- anvendere af forsyningsdata,
- ansvarlig enhed med ekspertise i metadata og gældende metadatastandards.

”Operatør” - rollen skal derimod varetages af én aktør og bør ikke deles mellem flere. Rollen indebærer ansvar for drift og vedligeholdelse af datasætkataloget samt den underliggende it-infrastruktur. En relevant myndighed med erfaring i drift og vedligeholdelse af datasætkataloger som ”operatør” vil kunne sikre stabil og sikker drift af kataloget og den underliggende IT-infrastruktur (applikationsservere, databaseservere, søgekomponenter, brugerdatabase, API-snitflader m.m.).

Vedrørende omfanget af ressourcer peger erfaringer fra eksisterende governancemodeller med kontrollørrollen på et kvart til halvt årsværk. Det endelige ressourceforbrug afhænger selvsagt af omfanget af metadata og kvaliteten af disse inden de sendes til kontrol.

Omfanget af ressourcer der vil påkræves fra de forskellige roller, vil komme an på de specifikke krav og behov, der vil være til datasætkataloget, som vil blive afdækket som led i det videre arbejde jf. Plan for implementering af fælles retningslinjer for forsyningssektorens metadata.



Figur 5. Roller og ansvar i den anbefalede governancemodel for metadata. Den aktør som udstiller data til anvenderne i omverdenen er ansvarlig for at indtaste og vedligeholde metadata for de udstillede datasæt. For at sikre ensartethed ift. understøtte bedre sammenlignelighed af data, etableres en kontrolrolle som skal kontrollere og godkende metadata før det offentliggøres. En særskilt rolle er operatøren af datasætkataloget.

Hvilke konkrete aktører der bedst kan påtage sig ovenstående roller for et datasætkatalog i forsyningssektoren bør dog fastlægges som led i den egentlige specificering af løsningen for datasætkataloget. Men opmærksomheden bør rettes mod eksisterende fællesoffentlige datasætkataloger med en governance, hvor rollerne er implementeret.

Eksempelvis driver Klimadatastyrelsen geodata-info.dk, som er et datasætkatalog for offentlige geodata. Her har Klimadatastyrelsen dels rollen som operatør af IT-systemet bag datasætkataloget (baseret på GeoNetwork, som er et open source-softwareprodukt, der bruges til at opbygge og drive metadata-kataloger for geodata.) og dels rollen som kontrollør af metadata, som de enkelte dataejende organisationer vedligeholder.

Aktørerne i geodata-info.dks governancemodel:

- 1) Dataudstillere: En dataejende organisation, som enten er forpligtiget via lov eller ønsker af egen vilje at have deres geodata dokumenteret med metadata – eksempelvis Slots- og Kulturstyrelsens geodata om Beskyttede sten- og jorddiger
- 2) Anvender: En aktør som søger efter geodata til en specifik use-case – eksempelvis en byggesagsbehandler i en kommune som skal have adgang til geodata om Beskyttede sten- og jorddiger for at vurdere afstandskrav til en tilbygning
- 3) Operatør: Klimadatastyrelsens driver IT-systemet bag geodata-info med hensyn til servere til applikation og database, selve applikationen (GeoNetwork), søgemotor, API og andet støtte software
- 4) Kontrollør: Klimadatastyrelsen påser at de enkelte dataudstillere udfylder og vedligeholder metadata i et anvender-perspektiv tilstrækkelig grad

5.6 Alle anbefalinger

Erfaringen viser med al tydelighed, at det fulde værdiskabende potentiale ved metadata først nås, når alle ovennævnte anbefalinger følges og føres ud i livet. Det er ikke tilstrækkeligt, at kun én eller en delmængde af anbefalingerne følges, som fx at anvende DCAT-AP-DK, hvis metadata samtidig ikke er tilgængelige i et fælles datasætkatalog, eller hvis den fornødne governance-struktur ikke er etableret til sikring af ensartethed i metadata.

Et fælles datakatalog for metadata og en klar governance-struktur er en central del af selve fundamentet for at omsætte udstillede data til reel værdi. Et fælles datakatalog for metadata udgør den fælles indgang til og fungerer som det operationelle fundament for hele forvaltningen og udstillingen af metadata. Uden dette skal dataanvendere selv bruge ressourcer på at lede efter datasæt, opbygge lokale "mini-kataloger" og vedligeholde dem løbende. Ensartethed vil gradvist afvige, da der ikke er kontrol over om metadata udfyldes korrekt og ensartet. Både dataudstillere og- anvendere vil ende med at bruge væsentlige ressourcer på parallelle katalogløsninger, uden at nogen høster de fælles fordele.

Samtidig kan governance ikke fungere uden et centralt katalog at virke igennem. Governance-strukturen skal gøre det muligt at kontrollere metadata og påse at dataudstillere etablerer og vedligeholder metadata i overensstemmelse med retningslinjerne samt, at nye forretningsbehov eller DCAT-AP-DK-opdateringer implementeres ensartet. Det forudsætter en fælles platform, hvor kvaliteten kan kontrolleres, indholdet valideres og retningslinjerne håndhæves.

Uden et fælles datakatalog for metadata reduceres governance til et regelsæt uden mekanismer til opfølgning. Med et datakatalog for metadata kan det derimod sikres, at metadata løbende vedligeholdes og ajourføres. Det betyder, at governance ikke bare handler om at sætte rammerne, men også om at understøtte den daglige drift og holde metadata korrekte, ensartede og anvendelige over tid.

6. Ordliste

Metadata: For at kunne understøtte anvendelsen af data er det nødvendigt med oplysninger om, hvordan man har indsamlet data, hvad et datasæt i udgangspunkt kan anvendes til, hvem der har ansvaret for datasættet, og hvad kvaliteten af det er. Det hedder under ét metadata og er en slags varedeklaration.

Discovery metadata: Refererer til metadata, der hjælper med at identificere og finde datasæt, som er relevante for et bestemt behov eller interesse. Det giver vigtig information, der gør det muligt for brugere at opdage og forstå, hvilke data der er tilgængelige, hvor de er opbevaret, og hvordan de kan tilgås. Kort sagt besvarer discovery metadata spørgsmålet: "Hvilke datasæt indeholder den type data, jeg er interesseret i?"

FAIR: er et sæt retningslinjer for datastyring og -videnskab, der sigter mod at gøre data Findable, Accessible, Interoperable og Reusable . Det er ikke en standard i sig selv, men snarere et sæt principper, der skal vejlede datahåndtering for at maksimere dataværdien og -genbrugen.

UML: Står for Unified Modeling Language (forenet modelleringssprog). Det er et standardiseret, generelt modelleringssprog, der bruges til at visualisere, specificere, konstruere og dokumentere artefakter af softwareintensive systemer. Selvom det primært bruges i softwareudvikling, kan det også anvendes til at modellere forretningsprocesser og andre ikke-software systemer.

XML: Står for eXtensible Markup Language. Det er et markeringssprog designet til at transportere og lagre data. I modsætning til HTML, som definerer hvordan data skal vises, definerer XML hvad dataene er.

Schema: Samling af formelle definitioner som specificerer hvor man markerer indholdet i et XML-dokument.

DCAT-AP: Står for Data Catalog Vocabulary - Application Profile. Det er en europæisk profil af DCAT, som er en W3C-standard til metadata om datasæt.

DCAT-AP-DK: Bygger videre på den europæiske DCAT-AP standard, men tilføjer danske termer, klassifikationer og retningslinjer.

GUI: Står for Graphical User Interface (grafisk brugergrænseflade). Det er en visuel måde for brugere at interagere med en computer eller en enhed. I stedet for at bruge tekstbaserede kommandoer, bruger en GUI ikoner, knapper, vinduer og andre grafiske elementer, som brugeren kan klikke på, trække og slippe for at udføre handlinger.

API: Application Programming Interface. Et API er en snitflade til software, der kan benyttes ved udvikling af softwareapplikationer. I dag benyttes API ofte synonymt med web-API, således at det betegner en snitflade, der kan tilgås over internettet, fx med HTTP-protokollen.

CKAN: Comprehensive Knowledge Archive Network er et open-source software, der bruges til at oprette, administrere og dele datasæt, især store mængder åbne data. Det er designet til at hjælpe organisationer med at gøre deres data tilgængelige og søgbare for offentligheden, forskere og udviklere.

GeoNetwork: Er et open-source software til datasætkataloger, som organisationer anvender til at organisere, indeksere, finde og dele metadata. GeoNetwork understøtter en række internationale standarder for metadata såsom ISO 19115 og DCAT-AP for metadata.

Bilag 1

Oversigt over de obligatoriske metadatainformationer

	Metadatainformation	Beskrivelse	DCAT-AP-DK v.2.0.2	Kravniveau
Dataset	Titel	Det eller de ord som navngiver datasættet.	dct:title	Obligatorisk
	Beskrivelse	En tekstbaseret beskrivelse af datasættets formål og indhold.	dct:description	Obligatorisk
	Ejer	Den organisation der har det juridiske ansvar for det samlede datasæt.	dct:rightsHolder	Obligatorisk
	Oprindelse	Udsagn omkring datasættets proveniens i form af udsagn om eventuelle aktørskift som er centrale for autenciteten, integriteten og fortolkningen af ressourcen.	dct:provenance	Obligatorisk
	Adgangs-information	Angiver en tilgængelig distribution af datasættet	dcat:distribution	Obligatorisk
	Personoplysningskategori	En relation til en bestemt personoplysningskategori.	dcat-ap-dk:personaldatacategory	Obligatorisk
	Genstand for registrering*	Angiver om det er et datasæt, en datatjeneste eller katalog som er genstand for registreringen.	rdf:type	Obligatorisk
	Identifikation	Den primære identifikator for datasættet, for eksempel en URI eller anden identifikator i kontekst af kataloget	dct:identifier	Obligatorisk
dcat:distribution	Adgangs-adresse	Angiver en URL som giver adgang til en distribution af datasættet.	dcat:accessURL	Obligatorisk
	Format	Filformatet for distributionen.	dct:format	Obligatorisk
dcat:contactPoint	Navn	Navn på organisation	vcad:hasOrganization-Name	Obligatorisk
	Email	Email-adresse	vcad:hasEmail	Obligatorisk

Bilag 2

Eksempel på metadata i RDF/XML iht. DCAT-AP-DK for udstillet JSON data i API. Metadata-informationerne stammer fra tabel 1 i anbefaling 3. Informationerne er opdigtet til retningslinjerne her, og det er derfor ikke et eksempel på eksisterende metadata.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dct="http://purl.org/dc/terms/"
  xmlns:dcat="http://www.w3.org/ns/dcat#"
  xmlns:dcat-ap-dk="https://data.gov.dk/model/profile/dcat-ap-dk/"
  xmlns:foaf="http://xmlns.com/foaf/0.1/"
  xmlns:locn="http://www.w3.org/ns/locn#"
  xmlns:vcard="http://www.w3.org/2006/vcard/ns#"

  <foaf:Document rdf:about="https://datahub.dk/aebleroeedforsyning/data/dataset2">
    <rdf:type rdf:resource="http://www.w3.org/ns/dcat#Dataset"/>
    <dct:title xml:lang="da">Forbrugsdata fra Eblærøds Forsyning</dct:title>
    <dct:description xml:lang="da">Forbrugsdata beregnet ud fra måledata for vand og fjernvarme i Eblærøds Forsynings forsyningsområde. Data omfatter stamdata om installationspunkter og forbrugstal opgjort på årligt basis.</dct:description>
    <dct:issued rdf:datatype="http://www.w3.org/2001/XMLSchema#date">2025-04-29</dct:issued>
    <dct:identifier>https://datahub.dk/aebleroeedforsyning/data/dataset2</dct:identifier>
    <dcat:keyword xml:lang="da">Vand</dcat:keyword>
    <dcat:keyword xml:lang="da">Fjernvarme</dcat:keyword>
    <dct:language>
      <dct:LinguisticSystem rdf:about="http://publications.europa.eu/resource/authority/language/DAN"/>
    </dct:language>
    <dcat:landingPage rdf:resource="https://datahub.dk/aebleroeedforsyning/data/dataset2"/>
    <dct:spatial>
      <dct:Location>
        <locn:geometry rdf:datatype="http://replacement.io/ont/geosparql#wktLiteral">POLYGON ((8.094 57.751,15.151 57.751,15.151 54.569,8.094 54.569,8.094 57.751))</locn:geometry>
      </dct:Location>
    </dct:spatial>
    <dct:accrualPeriodicity rdf:resource="http://publications.europa.eu/resource/authority/frequency/ANNUAL"/>
    <dct:provenance>
      <dcat:ProvenanceStatement>
        <rdfs:label xml:lang="da">Måledata er fjernaflæst og er input til beregning af forbrugsdata. I tilfælde af manglende måledata bruges lineær interpolation til beregning hvor de - til den manglende - to nærmeste måledata bruges som input.</rdfs:label>
      </dcat:ProvenanceStatement>
    </dct:provenance>
    <dct:documentation>
      <foaf:page rdf:resource="https://datahub.dk/aebleroeedforsyning/data/dataset2/feltbeskrivelse"/>
    </dct:documentation>
    <dcat-ap-dk:personalDataCategory xml:lang="da">Personoplysninger (ikke-følsomme personoplysninger)</dcat-ap-dk:personalDataCategory>
    <dct:rightsHolder>
      <foaf:Organization>
        <foaf:name>Eblærød forsyning A/S</foaf:name>
      </foaf:Organization>
    </dct:rightsHolder>
    <dct:publisher>
      <foaf:Organization>
        <foaf:name>Data HUB</foaf:name>
      </foaf:Organization>
    </dct:publisher>
    <dcat:contactPoint>
      <vcard:Organization>
        <vcard:hasEmail rdf:resource="support@datahub.dk"/>
        <vcard:hasOrganizationName>Data HUB Support</vcard:hasOrganizationName>
        <vcard:hasTelephone>+4512345678</vcard:hasTelephone>
      </vcard:Organization>
    </dcat:contactPoint>
    <dcat:distribution>
      <dcat:Distribution rdf:about="https://datahub.dk/distribution/dataset2">
        <dct:description xml:lang="da">Datasæt som JSON med stamdata om installationspunkter og forbrugstal opgjort på årligt basis</dct:description>
        <dcat:accessURL>
          <rdfs:Resource rdf:about="http://api.datahub.dk/dataset2"/>
        </dcat:accessURL>
        <dct:license rdf:resource="http://creativecommons.org/publicdomain/zero/1.0/">
        <dct:format>
          <dct:MediaTypeOrExtent rdf:about="http://publications.europa.eu/resource/authority/file-type/JSON"/>
        </dct:format>
      </dcat:Distribution>
    </dcat:distribution>
  </foaf:Document>
</rdf:RDF>
```

Bilag 3

Eksempel på metadata i html-visning i datasætkatalog. Energinets data om elforbruget opgjort per netområde i datakataloget datavejviser.dk.

Hjem / Energinet / Consumption per Grid Area

Consumption per Grid Area

The residual consumption is calculated hour by hour. It consists of the total consumption of the individual grid area deducted the consumption of all remote meter reading customers.

A key between grid area number and name can be extracted from the dataset [Change of Electricity Suppliers Per Grid Area](https://www...)

Mere ▾

Destinationsside ⓘ

https://www.energidataservice.dk/tso-electricity/ConsumptionPerGridarea

Udgivelsesdato ⓘ

2020-01-01

Ændringsdato ⓘ

2025-05-01

Opdateringsfrekvens ⓘ

daglig

Dækningsperiode ⓘ

2016-01-01 -

Distribution ⓘ

API	JSON	Adgangs-URL	+
Filudtræk	CSV	Download-URL	+
Filudtræk	JSON	Download-URL	+
Websidevisning	HTML	Adgangs-URL	+
Filudtræk	XLSX	Download-URL	+

Datatjeneste ⓘ

Energi Data Service API	JSON	CSV	XLSX	+
-------------------------	------	-----	------	---

Vilkår for brug af data

+

Datakvalitet og ophav

+

Øvrige metadata

-

Identifikator ⓘ	energidataservice_19_ConsumptionPerGridarea
Emne ⓘ	EU: Energi
Nøgleord ⓘ	Energi el elektricitet elforbrug elforsyning elforsyningsstatistik energi energidata energistatistik forbrug af el forbrug af elektricitet
Geografisk område ⓘ	Danmark

Udgiver ⓘ

ENERGINET

Energinet

https://energinet.dk/

På Energi Data Service udstiller Energinet offentligt tilgængelige energidata til fri afbenyttelse. Energinet er en selvstændig offentlig virksomhed under Klima-, Energi- og Forsyningsministeriet. Energinet ejer og udvikler el- og gasnet i Danmark for at indpasse mere vedvarende energi, opretholde forsynings sikkerhed og sikre lige markeds adgang til nettene.

Datasætkontakter

Kontaktpunkt

Energinet

energidata@energinet.dk

Skaber ⓘ

Energinet

Se alle udgivers datasæt



Energidataservice

Læs mere om dette og lignende datasæt på Energi Data Service. Alle har adgang til at bruge denne service. Hvad enten du er forbruger, iværksætter, forsker eller virksomhed, så stiller Energi Data Service data til rådighed, så du får en bedre forståelse af energisystemet og kan udvikle nye tjenester til glæde for samfundet og den grønne omstilling.

Datasæt endpoints

RDF-repræsentationer af dette datasæt.

JSON-LD

RDF/XML

Turtle

Notation3

23