



Referansearkitektur for KUDAF

04.10.2022

Innhold

1	Sammendrag	5
2	Introduksjon	6
2.1	Bakgrunn	6
2.1.1	Føringer, pågående og videre arbeid	6
2.1.2	Forankring	6
2.1.3	Målgruppe	6
2.1.4	Hva er en referansearkitektur	6
2.1.5	Emneknagger	7
2.2	Hva er KUDAF?	8
3	Visjon for datafellesskap i kunnskapssektoren	9
4	Konsepter i datafellesskapet	18
4.1	Hva er data?	18
4.1.1	Produksjonsdata	18
4.1.2	Statistikdata	18
4.1.3	Analyserbare data	18
4.1.4	Rapporter	19
4.1.5	FAIR	20
4.2	Hva er et domene?	20
4.3	Hva er Begrep?	20
4.4	Hva er Metadata?	21
4.5	Hva er Datakvalitet?	22
4.6	Hva er et Dataprodukt?	24
4.7	Hva er Kunnskapssektorens datadelingsplattform?	27
4.8	Evner for datadelingsplattformen	27
5	Virkemidler i datafellesskapet	28
5.1	Orden i eget hus	28
5.2	Verdikjeder	30
5.4	Informasjonsforvaltning	31
5.5	Kompetansebygging	34
5.6	Veiledninger	35
5.7	Prosesser	36
5.8	Fasilitere samarbeid	39
5.9	Koordinering og samordning gjennom felles forum	40
5.9.1	Rådgivende juridisk forum	40



5.9.2	Felles teknisk fagforum.....	40
5.9.3	Fagnettverk for «Orden i eget hus».....	41
5.10	Utvikle evner hos aktørene	42
5.10.1	Evner hos koordinerende aktør	42
5.10.2	Felles krav som stilles til aktørene – datatilbyder og datakonsument	46
5.10.3	Evner hos datatilbyder	48
5.10.4	Evner hos datakonsument.....	51
5.11	Roller	54
5.12	Tjenester og Verktøy	57
5.13	Komponentene i datadelingsplattformen	58

ENDRINGSLOGG

Versjon	Dato	Endring	Produsent	Godkjent
0.1		Internt arbeidsdokument	Heidi Bergh-Hoff, Håkon Sønderland m. fl.	
0.2	22.09.21	Første, foreløpige versjon som legges frem for prosjektstyret.	Heidi Bergh-Hoff, Håkon Sønderland m. fl.	
0.3		Innarbeidet kommentarer fra prosjektstyret og andre deler av KUDAF forprosjektet	Heidi Bergh-Hoff, Håkon Sønderland	
0.4	14.10.21	Akseptert endringsloggen og avstemt med styringsdokumentet.	Heidi Bergh-Hoff, Håkon Sønderland	
0.5	14.10.21	Innarbeidet kommentarer fra prosjektstyret og interne behov hovedsakelig for avstemming mellom leveranser fra prosjektet	Heidi Bergh-Hoff, Håkon Sønderland	
1.0	29.10.21	Endret navn på KUDAF-plattformen til Kunnskapssektorens datafellesskap	Per Hovde	
1.1	05.09.22	Flyttet innhold kapittelet «Visjon for datafellesskap i kunnskapssektoren» tilbake til	Heidi Bergh-Hoff	



Versjon	Dato	Endring	Produsent	Godkjent
		referansearkitekturen etter at innholdet ble fjernet fra styringsdokumentet.		



DISTRIBUSJONSLOGG

Versjon distribuert	Dato	Navn
0.2	22.09.21	Prosjektstyret
0.4		
1.0	29.10.21	Kunnskapsdepartementet



1 Sammen drag

Referansearkitekturen kan betraktes som en veiledning til helhetlig datadeling i kunnskapssektoren og skal bidra til å øke samhandling og støtte bruksområdene for data. Den gir retningslinjer for hvordan den tekniske infrastrukturen skal bygges (løsningsarkitekturen) og hvordan informasjonsforvaltning skal drives gjennom tydelige roller og ansvar samt veiledning for å utvikle og utnytte nødvendige evner.

Utviklingen av evner organiseres i en reise fra en applikasjonssentrisk arbeidsmåte og over til en data og kunnskapsdrevet tilnærming. Reisen starter med å forstå og utnytte egne data for egne mål, og utvikler seg gjennom fasene til å inkludere forbedret samhandling med andre for å skape innsikt og gevinst i kunnskapssektoren og i samfunnet generelt.

Referansearkitekturen veileder aktørene i hvilke tiltak de skal foreta seg, og i hvilken rekkefølge, for å oppnå en effektiv «reiserute». Gevinstene fra kunnskapssektorens datafellesskap (KUDAF) oppnås underveis på denne reisen. Referansearkitekturen tar utgangspunkt i en visjon for datafellesskapet beskrevet i kapittel 3 og inneholder:

- En «reisebeskrivelse» til målbildet i form av en virksomhetsstige for aktørene i Kunnskapssektoren
- Definisjoner på sentrale konsepter som inngår i målbildet
- En beskrivelse av virkemidlene brukt for å oppnå målbildet



2 Introduksjon

2.1 Bakgrunn

Målet med denne referansearkitekturen er å være et grunnlag for å etablere deling av data for videre- og sekundærbruk i kunnskapssektoren. Orden i eget hus er en forutsetning for å lykkes med å dele data og skape en datadrevet kunnskapssektor. Derfor er det et viktig delmål i dette arbeidet å tilrettelegge for orden i eget hus for samtlige aktører i kunnskapssektoren.

Økt deling av data er en forutsetning for å oppnå Kunnskapsdepartementet sin ambisjon om at arbeidet i departementet, underliggende direktorater og virksomheter innenfor departementets sektoransvar – utdanning, forskning og integrering – skal være datadrevet og kunnskapsbasert.

Referansearkitekturen skal tilrettelegge for å etablere et økosystem, som definert i dette forprosjektet, i kunnskapssektoren for deling av data til videre- og sekundærbruk, i og utenfor sektoren. Det er lagt vekt på at de data som skal deles er for bruk til analyse, politikkutvikling, forskning, statistikk og rapportering.

Det vises ellers til styringsdokumentet for bakgrunn og mål for KUDAF.

2.1.1 Føringer, pågående og videre arbeid

Referansearkitekturen som er nedfelt i dette dokumentet tar utgangspunkt i pågående arbeid og førende strategier og prinsipper nasjonalt, innen sektoren og internasjonalt.

2.1.2 Forankring

Referansearkitekturen er utarbeidet i et samarbeidsprosjekt med prosjektdeltagelse fra Direktoratet for høyere utdanning og kompetanse (HK-dir), Lånekassen, Digitaliseringsdirektoratet (Digdir), Utdanningsdirektoratet (Udir) og Sikt. Styringsgruppen består av medlemmer fra HK-dir, Sikt, Udir, Forskningsrådet, Lånekassen, Digitaliseringsdirektoratet, NAV, Statistisk Sentralbyrå og Kunnskapsdepartementet.

2.1.3 Målgruppe

Målgruppen for referansearkitekturen er primært IT-beslutningstagere, IT-arkitekter og tekniske prosjektledere. Deler av dokumentet er også relevant for informasjonsforvaltere, dataprodukteiere, domeneansvarlige, tjenesteeiere, tjenesteansvarlige, utviklere og datakonsumenter.

2.1.4 Hva er en referansearkitektur

Digitaliseringsdirektoratet sin definisjon på referansearkitektur¹ er lagt til grunn for beskrivelsen av referansearkitektur: «Referansearkitekturer gir veiledning til utforming av arkitekturer og løsninger, ut fra erfaring eller som målbilder». Referansearkitekturer er brukt for å løse avgrensede, men gjentakende problemstillinger. Et eksempel fra den analoge virkeligheten er at de fleste dører er rektangelformede. Avvik fra denne normen er kostnadsdrivende fordi en må spesialbestille, og har en tendens til å skape problemer for både de som skal bygge huset og de som skal bruke døren etterpå. Bruk av referansearkitekturer bidrar til et felles språk, konsistent implementering av tekniske løsninger og etterlevelse av felles standarder. Referansearkitekturer kan ha ulike detaljeringsnivå, fra overordnet til svært spesifikke. I denne konteksten beskriver en referansearkitektur logiske strukturer og

¹ <https://www.digdir.no/samhandling/referansearkitekturer/2131>



begrepsapparatet som gjelder innenfor ett spesifikt område på et overordnet nivå. Referansearkitekturen kan også gi eksempler på logiske tjenester, komponenter og hvordan interaksjon skal foregå mellom disse².

Generelt kan en referansearkitektur beskrives på mange ulike abstraksjonsnivåer, fra spesifikk til mer generelle. Basert på de funn som er avdekket gjennom forprosjektet, blant annet modenhetsanalysen for orden i eget hus, inneholder denne referansearkitekturen beskrivelser av overordnede virkemidler og tiltak som skal bidra til en datadrevet sektor og samfunn.

På øverste nivå beskriver man de forretningsmessige mål for området, endringsreisen for å oppnå målbildet, sentrale konsepter i tilnærmingen og virkemidlene brukt for å oppnå målbildet. Virkemidlene inkluderer tekniske komponenter i en datadelingsplattform for KUDAF. Plattformens komponenter og ønskede egenskaper er identifisert, men ytterligere kravspesifisering av plattformen er utsatt til programgjennomføringen. En referansearkitektur for datadeling i kunnskapssektoren gir både virksomheter og leverandører informasjon om felles virkemidler og rammeverk som gjelder for utviklingen av en datadrevet sektor.

2.1.5 Emneknagger

I dette dokumentet har vi brukt emneknagger for evner som er aktuelle for aktørene som er omtalt her å inneha. Disse er basert på veikartet som er referert i kapittel 5 Virkemidler i datafellesskapet, og ser slik ut (eksempel):

**Orden i eget
hus**

² https://www.ehelse.no/standardisering/standarder/referansearkitektur-for-datadeling/_/attachment/inline/33e8a3a6-2061-47db-bfd2-08b0e09bb108:d812f427533dfb987463cc459b35b981f5096505/Referansearkitektur%20for%20datadeling.pdf

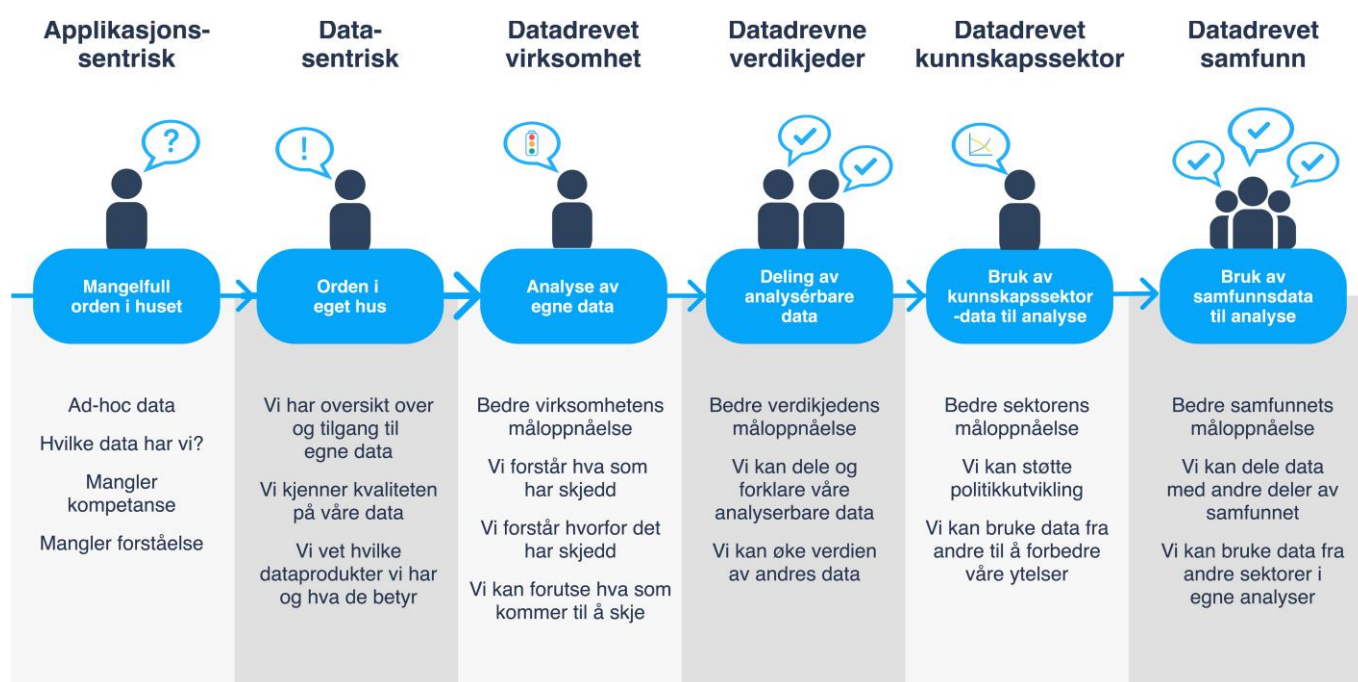


2.2 Hva er KUDAF?

KUDAF er et program som skal utvikle et sett av virkemidler for å realisere et økosystem for deling av data i kunnskapssektoren for videre- og sekundærbruk

Dette dokumentet beskriver de viktigste virkemidlene som KUDAF vil bruke.

Koordinerende aktør har ansvaret for å utvikle aktørene i sektoren gjennom disse virkemidlene. Av virkemidlene er kompetansebygging, veiledning, fasilitering og en datadelingsplattform blant de viktigste.



Figur 1 Kunnskapssektorens vei til et datadrevet samfunn



3 Visjon for datafellesskap i kunnskapssektoren

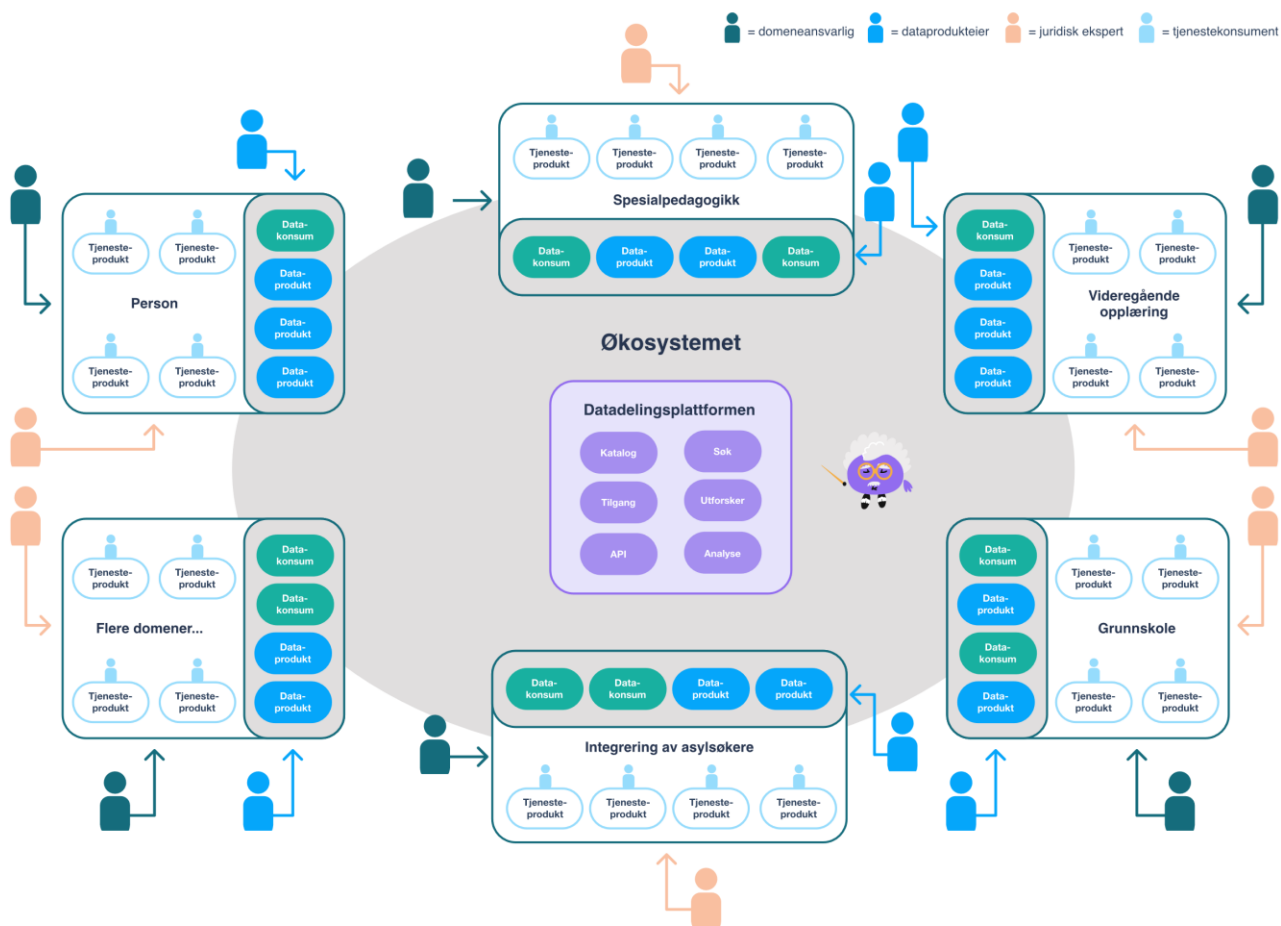
Referansearkitekturen baserer seg på en visjon for datafellesskap i kunnskapssektoren der aktører samhandler i et økosystem for deling av data. Sektorens sentrale prosesser gir føring for standardisering og informasjonsforvaltning i domener. Verdiskaping skjer i og mellom domenene i en samhandlingsmodell.

3.1 Økosystem

I arbeidet med referansearkitekturen for KUDAF har vi tatt utgangspunkt i en idé om samhandling mellom ulike parter i en større helhet. En sentral visjon fra konseptutredningen er et «økosystem for deling av data» som benytter datadelingsplattformen som omtalt tidligere. Deltagende aktører utfyller hverandre i økosystemet og samhandlingen skaper større verdi enn summen av det de enkelte aktørene kan oppnå hver for seg. Offentlige aktører i kunnskapssektoren, studenter, forskere, tjenestetilbydere, med flere, skal både skape, tilby, bearbeide og konsumere data på nye måter som gir alle insentiv og gevinst. Effektmålene for KUDAF gjør det mulig for virksomheter som deltar i økosystemet å forbedre innsikt fra egne og andres data. Denne innsikten bidrar til forbedret virksomhetsstyring og måloppnåelse som fremmer samfunns målet.

I økosystemet ser vi for oss de forskjellige aktørene i ulike konstellasjoner kalt domener (se seksjon 3.3). Økosystemet skal skapes sammen med en koordinerende aktør med overordnet ansvar for å legge til rette for samhandling om deling og bruk av data i kunnskapssektoren som stimulerer arbeidet med verdikjeder av data, og en leverandør av den underliggende infrastrukturen som inngår i datadelingsplattformen.





Figur 2: Et økosystem for datadeling i og mellom domener.

Funksjoner som er domenespesifikke, kalles her tjenesteprodukter. På samme måte som tjenesteproduktene, det vil si det de domenespesifikke systemene leverer internt i domenet, blir dataprodukter levert inn i økosystemet en viktig leveranse fra et domene. Dataprodukter tilgjengeliggjøres og deles mellom domenenene i økosystemet. Dataprodukteiere har ansvaret for at dataprodukter økosystemet etterspør blir tilbudt og at de har den kvalitet som forventes.

Domenene kan opptre både som datatilbydere (tilbyr dataprodukt) og datakonsumenter (bruker dataprodukter), som oftest i kombinasjon.

Datadelingsplattformen består av felles teknisk infrastruktur, prosesser og tjenester for digital samhandling som utnytter infrastrukturen, og et sett med felles spilleregler og standardiserte grensesnitt mellom aktørene som inngår i økosystemet, nedfelt i referansearkitekturen. Nasjonale felleskomponenter er benyttet i størst mulig grad og suppleres etter behov av tilleggsfunksjonalitet i kunnskapssektoren.

Den koordinerende aktøren har en nøkkelrolle i økosystemet som motivator, veileder og fasilitator for informasjonsforvaltning og deling av data.

Kunnskapssektoren består av aktører som alle vil bli påvirket av en transformasjon fra applikasjonssentrisk til datadrevet for sin virksomhet. De fleste av disse aktørene vil være datatilbydere, datakonsumenter, eller begge deler i et økosystem for kunnskapssektoren. For disse aktørene vil informasjonsforvaltning være en sentral bidragsyter for å oppnå Orden i eget hus og en tilnærming til en datadrevet virksomhet.

De viktigste aktørene i et økosystem for datadeling i kunnskapssektoren er:

- **Datatilbyder**
En datatilbyder er en aktør som tilgjengeliggjør dataprodukter til konsumenter som tilfredsstiller vilkårene for å kunne få utlevert etterspurte datasett. Alle virksomheter som har data som er gjort tilgjengelig for deling med andre aktører er datatilbydere.
- **Datakonsument**
En datakonsument er en aktør som kan benytte data fra en datatilbyder, enten internt hos aktøren eller eksternt hos en annen aktør. En datakonsument har forståelse for hvilke data det er behov for og kan begrunne bruken av data juridisk og faglig.
- **Koordinerende aktør**
Koordinerende aktør har en viktig rolle for å tilrettelegge for datadeling i sektoren, som å definere spilleregler og standardiserte grensesnitt mellom aktørene som inngår i økosystemet, og veilede aktørene i arbeidet med å bidra inn i det felles økosystemet. Koordinerende aktørs mandat og rolle er beskrevet i dokumentet «Mandat for koordinerende aktør», og blir derfor ikke detaljert nærmere her.

3.2 Kunnskapssektorens sentrale prosesser

Kunnskapssektoren jobber i prosesser for å utføre bruksnære funksjoner innen utdanning, forskning og integrering, samt administrasjon av disse funksjonene.

Administrasjons- og støtteprosesser skal i størst mulig grad være standardiserte for å jobbe effektivt, og for å skape felles forståelse innenfor disse domenene. For å oppnå dette skal begreper benyttet i prosessene og attributter knyttet til disse begrepene også standardiseres. Her vil sentrale begreper brukt i kunnskapssektoren være definert, som for eksempel «student» og «ansatt». Felles forståelse av disse begrepene skaper også grunnlag for samhandling i bruksnære prosesser, for eksempel når studenter og ansatte fra ulike institusjoner deltar i studieemner eller forskningsprosjekter. Denne tilnærmingen er i tråd med eksisterende harmoniseringsprosesser i sektoren.

Bruksnære lærings-, forsknings- og integreringsprosesser skal fremme innovasjon gjennom dynamiske og lokale arbeidsprosesser. Disse prosessene vil være knyttet til en brukskontekst, og spesifikke begreper for denne brukskonteksten vil defineres.

Data som benyttes i standardiserte, administrative arbeidsprosesser er knyttet til der hvor prosessen kjøres. I dag skjer dette ofte lokalt hos en institusjon. For eksempel ønsker vi at en ansettelsesprosess bruker en standardisert definisjon for «ansatt» som alle institusjoner forstår likt, mens persondata knyttet til ansettelse av en bestemt person hos en institusjon vil forvaltes hos institusjonen som ansetter personen. Noen prosesser er felles for flere institusjoner. For eksempel i prosessen for samordnet opptak til høyere utdanning bør persondata om søkere forvaltes av fellestjenesten SO. Da kan deltagende institusjoner få tilgang til opplysninger om søkere som er relevante for dem i prosessen.

Behovene for data i bruksnære funksjoner er annerledes. Lærings- og forskningsprosesser vil typisk ønske en bred tilgang til lærings- og forskningsressurser produsert i andre deler av egen institusjon eller i andre institusjoner nasjonalt og internasjonalt. Det kan også være ønskelig å publisere resultater fra disse prosessene bredt. Prosesser innen læring og forskning vil derfor ønske god støtte til å dele ressurser på tvers av institusjonene. Aktører som jobber med integrering, kan også ønske å dele beste praksis bredt. Eksempler på tjenester som muliggjør slik datadeling er Nasjonalt Vitenarkiv og eventuelle digitale læringsressurser.



Forståelse av behovene til kunnskapssektorens sentrale prosesser er brukt for å utforme forvaltning av informasjon i domener som beskrevet i neste seksjon ^{3 4}.

3.3 Informasjonsforvaltning i domener

Kunnskapssektorens arbeidsprosesser støttes av informasjon. Referansearkitekturen foreslår at ansvar for forvaltning av denne informasjonen skal plasseres i en faglig kontekst som vi kaller et «domene». Innen et domene definerer et fagområde sine begreper i det vokabularet som benyttes der. For eksempel vil et begrep som «Søker» ha en beskrivelse, forståelse og sine egne data i et studieadministrativt domene. Domener representerer dermed en avgrensning av interessentene i en kontekst som må bli enige om definisjon og bruk av begreper.

Referansearkitekturen foreslår at aktører innen et fagområde som skal benytte standardiserte, administrative prosesser deler et informasjonsdomene med felles ansvar for begrepene som støtter de standardiserte prosessene. Referansearkitekturen foreslår også at aktører som deltar i bruksnære prosesser deler et informasjonsdomene med begrepene som støtter bruksområdet. Informasjonsdomener vil typisk omfatte flere aktører, og aktørene vil typisk delta i flere informasjonsdomener.

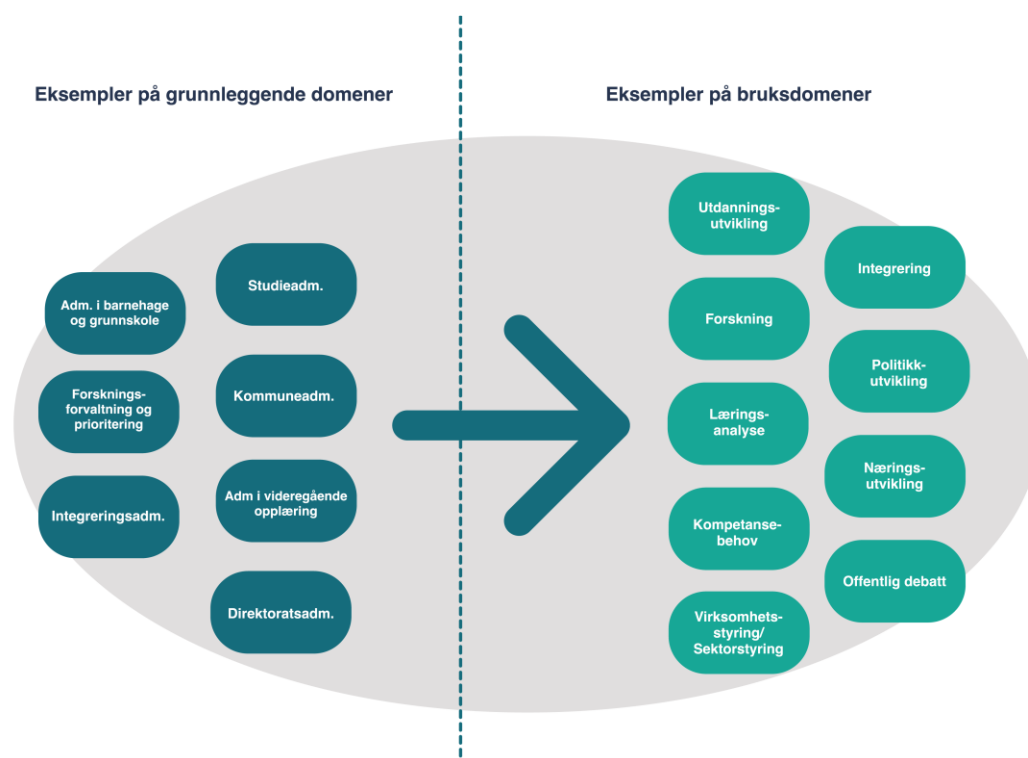
Arbeidet med domener vil kreve samarbeid og koordinering mellom aktørene. Her vil koordinerende aktør måtte spille an aktiv rolle i fasilitering og veiledning. For eksempel å kunne identifisere hvilke domener en aktør opererer i og hva som hører inn under hvert enkelt domene. Det er naturlig å tenke seg at for eksempel et økonomidomene vil kunne se annerledes ut og omfatte forskjellige deler av virksomheten hos aktører som opererer i ulike deler av sektoren. Her vil felles fora og koordinerende aktør måtte bidra til å standardisere begreper, definisjoner og formater for dataprodukter i domene.

Referansearkitekturen beskriver konsepter og virkemidler for å oppnå datadeling mellom domener. I tillegg til organisasjonsutvikling og informasjonsforvaltning, defineres datadelingsmønstre gjennom oppslag, publisering av hendelser og flere som foreløpig ikke er spesifisert. Vi grupperer domene som enten grunnleggende domener, typisk med ansvar for administrative prosesser, eller domener utviklet for spesielle brukstilfeller som avbildet under.

³ Modellen er basert på arbeid med valg av operasjonelle modeller beskrevet i «Enterprise Architecture as Strategy» av Ross, Weill og Robertson, 2006

⁴ Modellen beskrevet i seksjonen her er utarbeidet for og i samarbeid med aktører i høyere utdanning og forskningssektorene. Vi tror modellen er også relevant for grunnutdanning og integrering, men vi utsetter til programgjennomføringen å diskutere modellen med aktører i disse sektorene.





Figur 3: Domener i kunnskapssektoren

De *grunnleggende domenerne* inneholder data som støtter de sentrale prosessene for administrasjon og styring av læring, forskning og integrering hos aktørene i kunnskapssektoren.

Disse domenerne inneholder blant annet *masterdatakilder* med informasjon om, for eksempel, studenter, elever, ansatte, innvandrere, barn, forskningsprosjekt, læringsobjekter, med flere. Disse kildene er definert ut fra en dyp forståelse av sektorenes funksjon og vil typisk være forholdsvis stabile over tid. Dataprodukter definert i grunnleggende domener blir gjerne gjenbrukt av bruksnære domener.

De *bruksnære domenerne* er definert til å støtte en spesifikk brukernær kontekst. Domener for brukskontekst inkluderer for eksempel situasjoner der politisk administrasjon vurderer effekten av tiltak, studenter evaluerer emner, der studenter får automatisert tilbakemeldinger for å forbedre læringsprosessen (læringsanalyse) og der deltagere i et forskningsprosjekt samler data til analyse. De bruksnære domene kan befinne seg både innenfor og utenfor sektoren.

Dataprodukter definert her skal være spesifikke og skreddersydd til brukskonteksten.

Tilnærmingen med bruk av domene er basert på faglitteratur om Data Mesh⁵.

3.4 Samhandling

Når datatilbydere og datakonsumenter deler data, skjer det som en del av en digital samhandling. For at samhandlingen skal være vellykket, må tilbydere og konsumenter sikre at de handler i henhold til loven (juridisk samhandlingsevne), at aktørene har avklart forventninger til hverandre (organisatorisk samhandlingsevne), at datatilbyder og konsumenter har samme forståelse av dataenes betydning

⁵ Fra «How to Move Beyond a Monolithic Data Lake to a Distributed Data Mesh» av Zhamak Dehghani



(semantisk samhandlingsevne) og at de tekniske løsningene som utfører datadeling fungerer sammen slik de skal (teknisk samhandlingsevne). Rammeverk for digital samhandling vist under utdyper disse samhandlingsevner ⁶.



Figur 4: Rammeverk for digital samhandling

Modellene som er utgangspunkt for referansearkitekturen er laget for å ivareta juridisk-, organisatorisk-, semantisk- og teknisk samhandlingsevne og hjelper dermed de som benytter modellene til å lykkes med digital samhandling.

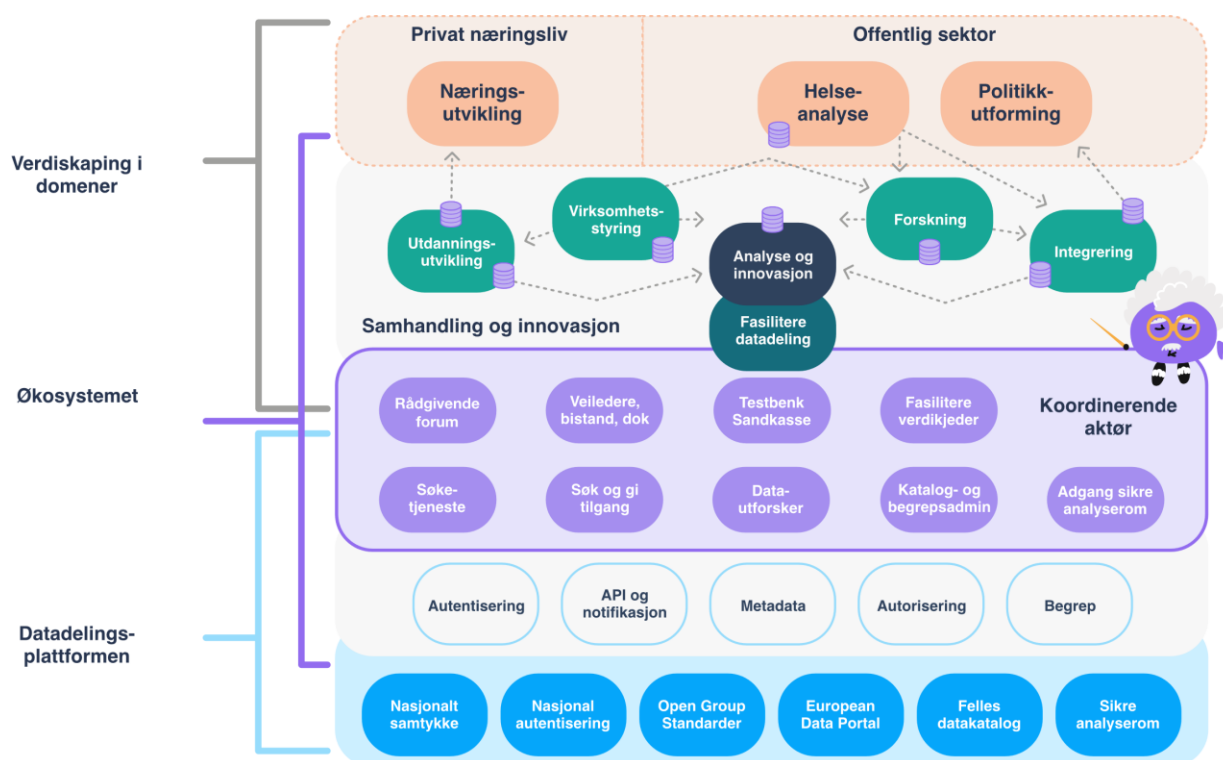
3.4.1 Samhandlingsmodell

Samhandlingsmodellen vi vil benytte i KUDAF er illustrert under. Den er basert på Digitaliseringsdirektoratets samhandlingsmodell for datadeling⁷. Digitaliseringsdirektoratets modell omhandler i stor grad deling av data for tjenesteproduksjon. Samhandlingsmodellen i KUDAF er basert på ønsket om å kunne utnytte data til analyse og innovasjon. Modellen baserer seg derfor på deling av dataprodukter mellom virksomheter innenfor et økosystem i kunnskapssektoren. For å kunne dele slike dataprodukter har vi i dette dokumentet skissert en rekke virkemidler som inngår i modellen. Beskrivelsene av dataprodukt og virkemidlene i datafellesskapet er beskrevet i henholdsvis kapittel 4 og kapittel 5. Modellen har tre nivåer, Verdiskaping, Økosystem og Plattform. Nivåene er beskrevet under modellen.

⁶ <https://www.digdir.no/nasjonal-arkitektur/rammeverk-digital-samhandling/2148>

⁷ <https://www.digdir.no/datadeling/datadeling-krever-samhandling/2379>





Figur 5: KUDAFs samhandlingsmodell for deling av data for analyse og innovasjon

3.4.1.1 Verdiskapingsnivået

Det finnes mange ulike og overlappende samhandlingsklynger med nærliggende egenverdi for deltakerne:

- Økosystemer i samfunnet, som i sektoren eller innen fagområder
- Verdikjeder
- Gjenbruk av datasett uten noe form for integrasjon, som kan være datasett fra kunnskapssektorens økosystem og andre steder.

3.4.1.2 Kunnskapssektorens økosystem

Virksomhetene bidrar med analysedataprodukter til gjenbruk og viderebruk. Gjennom virkemidlene i modellen kan datakonsumenter ta i bruk dataproduktene og skape verdi. Kunnskapssektorens økosystem er en del av et større og helhetlig økosystem gjennom bruk av nasjonale og internasjonale felleskomponenter.

3.4.1.3 Datadelingsplattformen

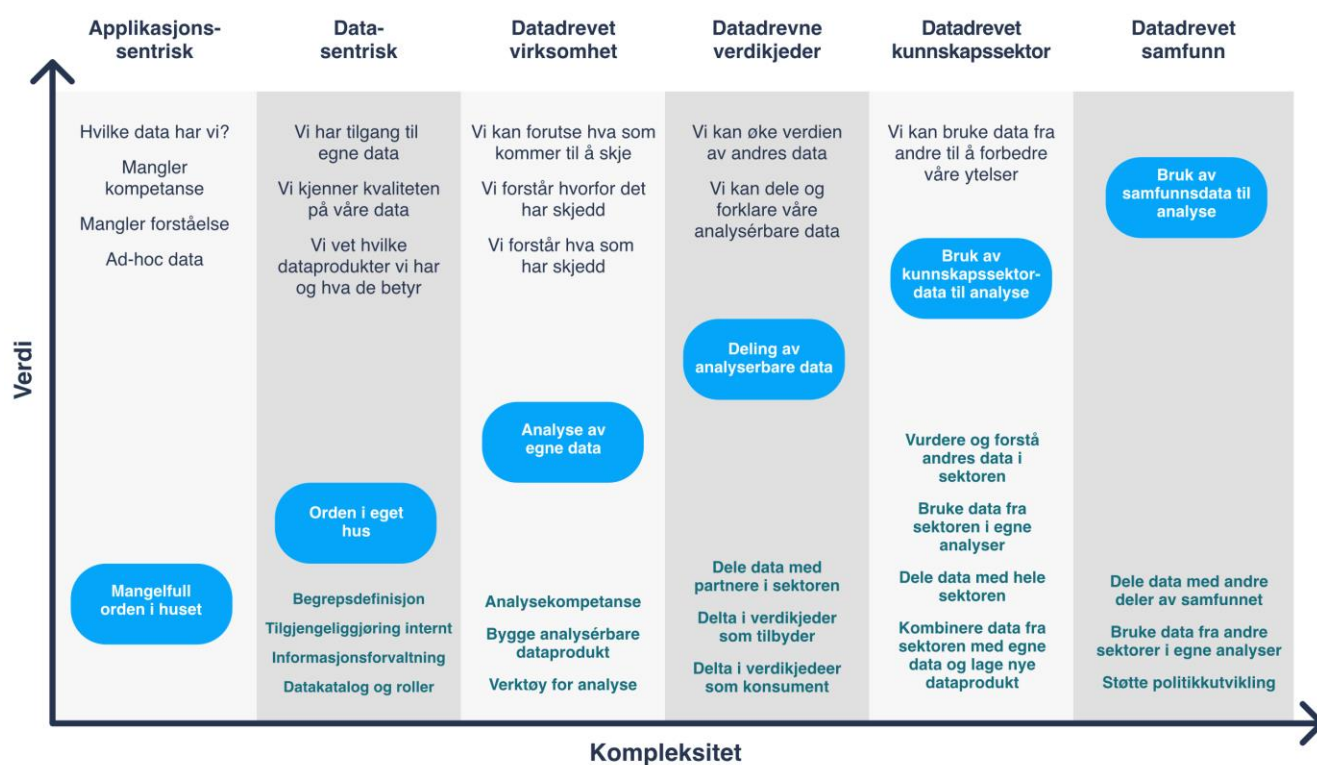
Mange aktører bidrar inn mot plattformen som gir evne til å kunne utveksle dataprodukter. Dette kan inkludere aktører utenfor sektoren, som Digdir, Altinn, Open Group og andre. I tillegg til felles teknisk infrastruktur inkluderer plattformen prosesser og tjenester som gjør aktørene i stand til å samarbeide om deling av data.

3.5 Virksomhetsstigen

En virksomhetsstige beskriver reisen for å løfte aktørene i kunnskapssektoren og dermed hele sektoren mot et datadrevet samfunn. Figur 6 viser utviklingen fra innadrettet virksomhetsfokus til deltakelse i et



datadrevet samfunn for aktører i kunnskapssektoren. Plataer er brukt for å illustrere graden av verdi for aktørene og sektoren. Aktørene har forskjellig utgangspunkt og vil befinne seg på ulike nivå, kanskje også internt i virksomheten. Modenhet kan variere mellom ulike datasett benyttet i en virksomhet på tilsvarende måte som at noen skuffer i et hus kan være rotete mens andre er velorganisert der hvor vi skal skape Orden i eget hus. Der hvor datasett er forstått, dokumentert, godt forvaltet og bearbeidet til analyserbare dataprodukter er modningsnivået høyere enn der hvor datasett er uten disse karakteristikka. Med en smidig tilnærming kan deler av virksomheten, for eksempel innen et domene, ha kommet lenger enn andre, i deling eller analyse av egne data. Hensikten med stigen er å illustrere evner virksomheten kan ha, og hva det kan innebære av gevinster for virksomheten og sektoren av å kunne utnytte tilgjengelig data.



Figur 6: Virksomhetsstigen for en datadrevet kunnskapssektor



4 Konsepter i datafelleskapet

For å kunne forstå sentrale konsepter i KUDAF, er det nødvendig å definere hva vi mener når vi snakker om de. Prosjektet har kommet frem til at de konseptene som er beskrevet under er de viktigste å forstå. Uten en felles forståelse for disse konseptene er det vanskelig å kunne arbeide videre med datadeling i sektoren.

4.1 Hva er data?

Målsetningen for arbeidet med KUDAF er en datadrevet, kunnskapsbasert kunnskapssektor. Det sentrale begrepet for å utvikle dette er «data». Data som begrep er brukt i mange sammenhenger, ofte uten en klar definisjon. For å forstå måten vi bruker begrepene data, datasett og dataprodukt, vil vi her først prøve å definere forskjellige typer data, og til slutt i seksjon 4.6 definere «dataprodukt», som er det sentrale identifikatoren vi bruker om data som skal deles i kunnskapssektoren.

4.1.1 Produksjonsdata

Produksjonsdata er data som er nødvendig for tjenesteproduksjon i en virksomhet. Det inkluderer oppgaver og prosesser knyttet til brukertjenester som virksomheten har ansvaret for. Disse data må være tilgjengelige for hyppig og effektiv tilgang og være lagret på en pålitelig og sikker måte. Slike data skal være korrekte, pålitelige, godt dokumenterte og administrert.

Slike data er ofte i strukturert form og lagret i relasjonsdatabaser eller lignende, men omfatter også ustrukturerte data slik som dokumenter. I noen tilfeller vil det i sett av slike data finnes nøkkelinformasjon, slik som for eksempel personnummer, som gjør at data kan kobles til andre sett av data. I mange tilfeller vil ikke nøkkelinformasjonen som er knyttet til disse datasettene gi noen mening utenfor systemet, domenet eller virksomheten der disse produksjonsdataene finnes.

I kunnskapssektoren er produksjonsdata knyttet til den operasjonelle tjenesteproduksjonen for virksomheter i sektoren. Det kan være saksbehandlingsdata, persondata, økonomidata og så videre.

4.1.2 Statistikkdata

Statistikkdata er data som er kommet frem som et resultat av en prosess som kan innebære sammenstilling, kalkulering eller analyse av andre data, ofte produksjonsdata. Statistikkdata kan benyttes internt i virksomheten som utarbeider den og av eksterne, til styring, offentlig innsyn, politikkutforming, forskning og publisering. En viktig grunn for å aggregere data er å unngå å offentliggjøre sensitive data som kan knyttes til enkeltpersoner eller grupper av personer.

I de fleste tilfeller inneholder slike statistikkdata ingen meningsfull nøkkelinformasjon. Det gjør de ute av stand til å bli koblet til andre datasett eller dataprodukter.

I kunnskapssektoren blir statistikkdata produsert internt i virksomhetene og i enkelte virksomheter, slik som direktorater, som har et spesielt ansvar for dette. Eksempler på dette kan være data som HKDIR behandler i DBH eller data som sendes til SSB.

4.1.3 Analyserbare data

Analyserbare data er data på et format som gjør de spesielt egnet til bruk i forskning, analyser og sammenstillinger. Slike data er ofte forhåndsbehandlet slik at de er på gitte, standardiserte format, egnet for behandling i relevant analyseprogramvare. Analyserbare data akkompagneres av lett tilgjengelige metadata som skal sikre at data (tilblivelse, kodeverk struktur, kvalitet og forbehold) kan forstås, og



inneholder dessuten nøkkelinformasjon (for eksempel pseudonymisert person-id) slik at de kan kobles med andre dataprodukter.

Slike data inneholder i sin enkleste form, i tillegg til nøkkelinformasjonen, kun én faktaopplysning, ofte kalt variabel, pr analyseenhet. For eksempel lønn for hver person eller antall årsverk for hvert institutt.

Der det er hensiktsmessig, og hvor flere faktaopplysninger om et sett enheter er samlet inn samtidig (for eksempel i en survey eller en konkret innrapportering), organiseres i datasett med flere faktaopplysninger/variabler knyttet til hver enhet.

En slik organisering av datasett med mange variabler kan gi økt brukervennlighet i analysesammenheng, men samtidig redusere fleksibilitet med tanke på kobling av data fra ulike datasett.

Andre dimensjoner som kan finnes i eller i nær tilknytning til data er metadata som handler om når, hvor, hvorfor, hvem, hvordan, hva for dataproduktet. Dette kalles kontekstinformasjon i 6 dimensjoner: W5H

Who - Hvem er dataene om?

What - Hva måler/beskriver dataene?

When - Hvilket tidspunkt/hvilken tidsperiode gjelder dataene?

Where - For hvilket område (region, virksomhet, etc.) gjelder dataene?

Why - Hvorfor er dataene samlet inn? Formål

How - Hvordan ble dataene samlet inn? (måling, registrering, spørreskjema, måleinstrument, etc.)

Kontekstinformasjon er sentralt for å forstå, forvalte, foredle og formidle data. Data som mangler informasjon om kontekst, er i realiteten verdiløse.

Analyserbare data inneholder nøkkelinformasjon som inneholder nøkler som identifiserer data, for eksempel ned på personnivå. Data kan være pseudonymisert, eller syntetiske, med syntetiske nøkler slik at de ikke like enkelt skal kunne kobles til enkeltpersoner. For å kunne koble slike data sammen må de gjennom en prosess for pseudonymisering der alle data som inngår gis koblebare, syntetiske nøkler. Slike data må fremdeles beskyttes mot misbruk da det eksisterer risiko for at tilleggsinformasjon vil gjøre det mulig å koble disse data til personer.

En kraftigere virkemiddel for å hindre kobling av data til enkeltpersoner er anonymisering som tradisjonelt har handlet om å fjerne informasjon fra data, gjerne ved hjelp av aggregeringer (bruk fylke eller landsdel istedenfor kommune, aldersgrupper istedenfor alder, inntektsgrupper istedenfor inntekt). Anonymisering er også ofte en manuell prosess, og hvilken informasjon man kan fjerne, avhenger av brukstilfellet.

I kunnskapssektoren blir analyserbare data produsert og konsumert av flere aktører, for eksempel NSD og SSB. Eksempler på bruk er NSD sin interne bruk og microdata.no. KUDAF vil tilby analyserbare data og disse vil kunne brukes av alle.

4.1.4 Rapporter

Data i form av rapporter består av statistikkdata, ofte i kondensert form, og tekstlige fremstillinger av resultater, tiltak og prosesser. Typisk er rapporter utarbeidet som et resultat av utredninger eller periodisk rapportering, som for eksempel årsrapporter.

Data i rapporter er som regel lite tilgjengelige for annen bruk enn det de var tiltenkt og kan ikke enkelt gjenbrukes.



I kunnskapssektoren blir det produsert rapporter for mange forskjellige formål, som for eksempel årsrapporter. KUDAF vil kunne levere data som grunnlag for slike rapporter.

4.1.5 FAIR

Et sett av prinsipper for god tilrettelegging for viderebruk av data, og spesielt forskningsdata, er etablert i de internasjonale FAIR-prinsippene, et akronym basert på ordene findable, accessible, interoperable, og reusable^{8 9}. FAIR-prinsippene er rettet både mot å redusere tekniske og administrative hindre for tilgang og gjenbruk, og mot at data skal tilfredsstillе ulike kvalitetskrav. De legger også stor vekt på at data befinner seg i et digitalt økosystem, og at de må kunne håndteres maskinelt.

I KUDAF-sammenheng blir det sentralt å følge FAIR-prinsippene. Siden det er lett å glemme, er det spesielt viktig å huske «reusable», siden deling av data er et hovedmål for tiltaket.

4.2 Hva er et domene?

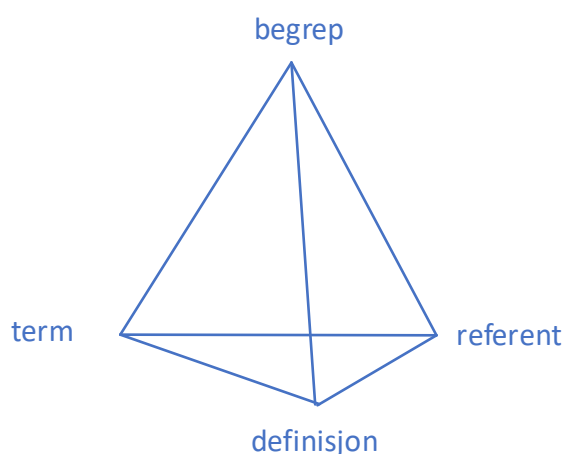
Et identifiserbart fagområde eller en avgrensbar faglig kontekst kaller vi et «domene».

Innen et domene vil et fagområde kunne definere sin informasjon i en informasjonsmodell og sine begreper i det vokabularet som benyttes der. For eksempel vil et begrep som «Søker» ha en beskrivelse, forståelse og sine egne data i et studieadministrativt domene.

En aktør i kunnskapssektoren vil ha virksomhet innenfor flere domener, for eksempel personaladministrasjon og læringsadministrasjon.

På samme måte kan det være domener som omfatter flere institusjoner gjennom standardiserte prosesser og begreper samt fellesløsninger, som felles økonomisystem eller studentadministrasjon.

4.3 Hva er Begrep?



Figur 7 Dimensjonene i begrepsarbeidet

⁸ <https://www.force11.org/group/fairgroup/fairprinciples>

⁹ <https://www.uio.no/for-ansatte/arbeidsstotte/forskningsstotte/forskningsdata/fair-data/>



«Begrep er en abstraksjon som inneholder kjennetegn som beskriver et fenomen»
«Begrep er mental forestilling om et konkret eller abstrakt fenomen i den virkelige verden»

For å kunne omtale et begrep trenger man termer, som er det som setter navn på et begrep, og definisjoner, som er beskrivelser som avgrenser begrepet. Referent henviser til en bestemt forekomst av et begrep. For mer om hva et begrep er henvises til [Termlosen](https://data.norge.no/specification/termlosen/)¹⁰ fra Språkrådet og Digitaliseringsdirektoratet.

I forprosjektet til KUDAF er det igangsatt et eget arbeid med Begrepskoordinering, og for arbeidet med definisjon av begreper og termer i kunnskapssektoren henvises til dette arbeidet.

Det er imidlertid verd å merke seg at begrepsarbeidet har en del kontaktpunkter mot referansearkitekturen.

Sentrale tema som må drøftes er blant annet:

- Omforent bruk av begreper er en forutsetning for at vi klarer å forstå data likt (semantisk samhandling). Begrepsarbeidet kan ikke være løsrevet fra løsninger og arkitektur.
- Å etablere begreper er en av de prosessene som må være en del av fellesforum for aktørene.
- Begreper er sentralt for å gi mening til dataprodukt.
- Må kunne styre begrepsbruk i metadata som beskriver datasett tilgjengeliggjort maskinelt
- Når begreper skal iverksettes, må de være nedfelt i felles infrastruktur (metadatakatalog, API-management etc)
- Hvordan skal vi ta i bruk begreper for maskinell bruk, bruk i søk og for «aktive metadata»?

4.4 Hva er Metadata?

Metadata betyr «data om data». Metadata defineres som data som inneholder kontekstinformasjon som beskriver data. Metadata oppsummerer basisinformasjon om data slik at det skal bli lettere å spore og jobbe med disse dataene. Eksempler på metadata kan være:

Hvordan data oppsto
Hensikten med data
Tidspunkt når data oppsto
Hvem som laget data
Hvor data oppsto
Hvilke standarder som er brukt i forbindelse med definisjon av data
Størrelser for data
Kvalitetsmål for data
Semantisk beskrivelse, knytning til begrep og term
Syntaktisk beskrivelse, knytning til kode og regler

Et dagligdags eksempel på metadata kan være når du får informasjon om et musikkstykke som spilles, inneholdende artistnavn, navn på låt, utgivelsesår, hvilket selskap den kom ut på og så videre.

Metadata kan inneholde **W5H** knyttet til forskningsdata og dessuten understøtte **FAIR**, som beskrevet tidligere.

¹⁰ <https://data.norge.no/specification/termlosen/>



Metadata blir ofte holdt i en datakatalog, der beskrivelser av datasett kan finnes. Et eksempel på dette er den nasjonale «Felles datakatalog» på data.norge.no¹¹.

Metadata kan benytte standarder for beskrivelse av datasett og datakataloger. Et eksempel på det er «Standard for beskrivelse av datasett, datatjenester og datakataloger», DCAT-AP-NO, som er den standarden som Felles datakatalog benytter.

Metadatabeskrivelser på datasett- og katalognivå er nødvendig, men normalt ikke tilstrekkelig for å forstå og tolke data fullt ut. Enkelte metadatastandards som DDI¹² har støtte for detaljerte beskrivelser av variabler/data-elementer på svært presist nivå (koding/representasjon, begreper, referanser til klassifikasjonsverk, datakvalitet, dataformat (tall, tekst, datoformat)), etc.

4.5 Hva er Datakvalitet?

Datakvalitet

En viktig målsetning for arbeidet med informasjonsforvaltning hos aktørene er å forbedre kvaliteten på data. Et av de viktigste virkemidlene for å få til bedre datakvalitet er å øke forståelsen av hva datakvalitet er. Datakvalitet defineres slik av denne referansearkitekturen¹³:

Dataene har evnen til å støtte de informasjonsformål de er tiltenkt, og at de oppfyller formålets krav til:

- Korrekthet/Nøyaktighet (accuracy)
- Kompletthet/Fullstendighet (completeness)
- Oppdaterte/Aktualitet (timeliness)
- Konsistens (consistency)

Det finnes i tillegg en rekke dimensjoner man kan legge til for å sikre kvalitet i data, men vi har her valgt å holde oss til de vi mener er viktigst.

Vi deler kvalitetsbehovet i to hoveddimensjoner:

- Hos aktøren: Forbedre forståelse og arbeidet med kvalitet i data som aktøren har ansvar for
- Delbar dataprodukt: Forstå og kunne beskrive kvalitet for de delbare dataproduktene

Hos aktøren vil arbeidet med informasjonsforvaltning og «orden i eget hus» innebære arbeid med å forstå, forbedre og dokumentere datakvalitet i de data aktøren allerede har ansvar for. Her vil denne referansearkitekturen og veiledning fra koordinerende aktør være støtte til dette arbeidet.

Delte dataprodukt vil ha varierende krav basert på tiltenkt bruk av dataproduktet. Dersom dataproduktet skal deles for et bestemt formål, for eksempel i en verdikjede, kan kvalitetsmål avtales mellom tilbyder og konsument. Det viktigste vil være å dokumentere formål og kvalitet på slik måte at det er mulig å

¹¹ <https://data.norge.no/>

¹² Data Documentation Initiative - <https://ddialliance.org/>

¹³ Definisjonen vår er konsistent med Skatteetaten sin definisjon i felles datakatalog "innebærer at data skal være korrekte, komplette, oppdaterte og konsistente og har evnen til å støtte de informasjonsformål de brukes til"



vurdere i tilfelle det er ønske om annen bruk av dataproduktet. Eksempel på dette kan være dataprodukt levert til et forskningsprosjekt, der det er viktigere for forskere å få et komplett datasett enn at kvaliteten er sikret 100%. I et slikt tilfelle kan datakonsumenten, i dette tilfellet forskere, være i stand til på egenhånd å vurdere og sikre kvaliteten i bruk. I tilfelle gjenbruk av et slikt dataprodukt må det komme frem i dokumentasjonen av dataproduktet at dette er forutsetninger og formål det er tenkt brukt til.

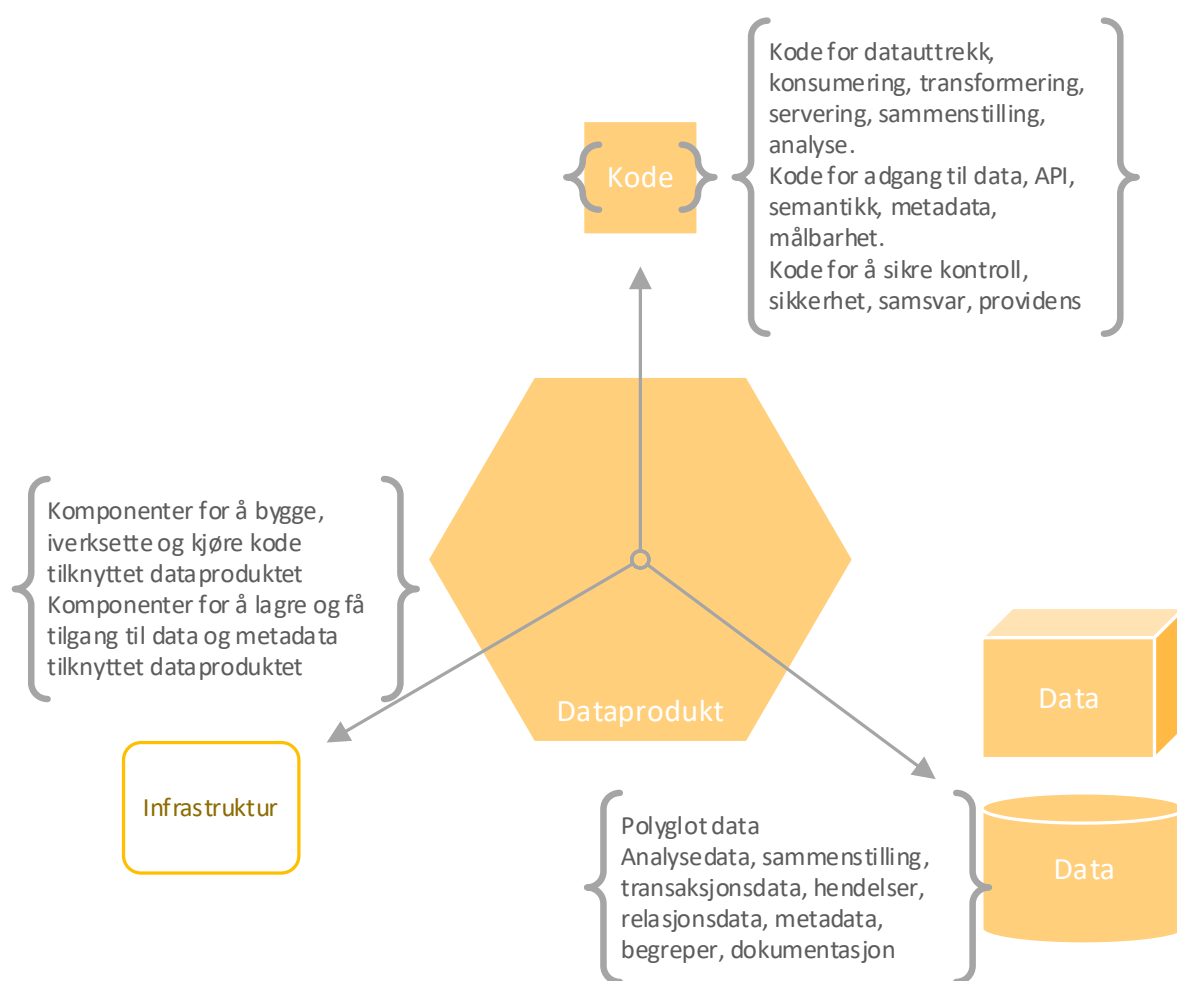
I andre tilfeller er der et dataprodukt skal gjøres generelt tilgjengelig, og tilbyder ikke nødvendigvis har kontroll på hvem som bruker dataproduktet eller til hva. Da blir det viktig for konsument å forstå hvilken kvalitet dataproduktet har, og hvilke valg er tatt underveis i produksjon av datasettet, for eksempel kriterier for at data blir inkludert. Det finnes norske og internasjonale standarder for hvordan kvalitet skal beskrives. Det foreslås her at kvalitetsbeskrivelser i metadatakatalogen baserer seg på den norske [DQV-AP-NO](https://data.norge.no/specification/dqv-ap-no/)¹⁴ (Norsk applikasjonsprofil av DQV) utarbeidet av Digitaliseringsdirektoratet.

¹⁴ <https://data.norge.no/specification/dqv-ap-no/>



4.6 Hva er et Dataprodukt?

Et dataprodukt som definert i denne referansearkitekturen er basert på idéer fra [Data mesh](#)¹⁵. Dataprodukter produseres i kontekst av et domene. Dataprodukter er de delbare ressursene som bor i domenet og håndteres av de som har ansvaret for data i det gitte domenet¹⁶. Dataproduktet må inneholde mer enn bare datasettet. Dataprodukter må inneholde metadata for å beskrive kontekst, infrastruktur for å tilby domeneuavhengig teknologi for å skape, bearbeide, lagre og tilgjengeliggjør dataprodukter og kode for å uttrykke hvordan data i dens kontekst skal brukes i domenene. Dataprodukter er hovedbyggeløss («architectural quantum»¹⁷ som brukt i data mesh¹⁶) som brukes for å handtere analysedata. For å benytte slike dataprodukter må aktørene som deltar i utvekslingen av data samarbeide og bli enige om spilleregler og standardisering. Data skal være desentralisert og domene-sentrisk, der interoperabilitet baserer seg på felles standarder og semantisk forståelse. Dette skal muliggjøre bruk av data for analyse og innsikt, statistikk, beslutningstaking og politikkutvikling. Figuren under er en illustrasjon av hva et dataprodukt kan inneholde.



Figur 8 Dimensjonene i et Dataprodukt

¹⁵ <https://martinfowler.com/articles/data-monolith-to-mesh.html>

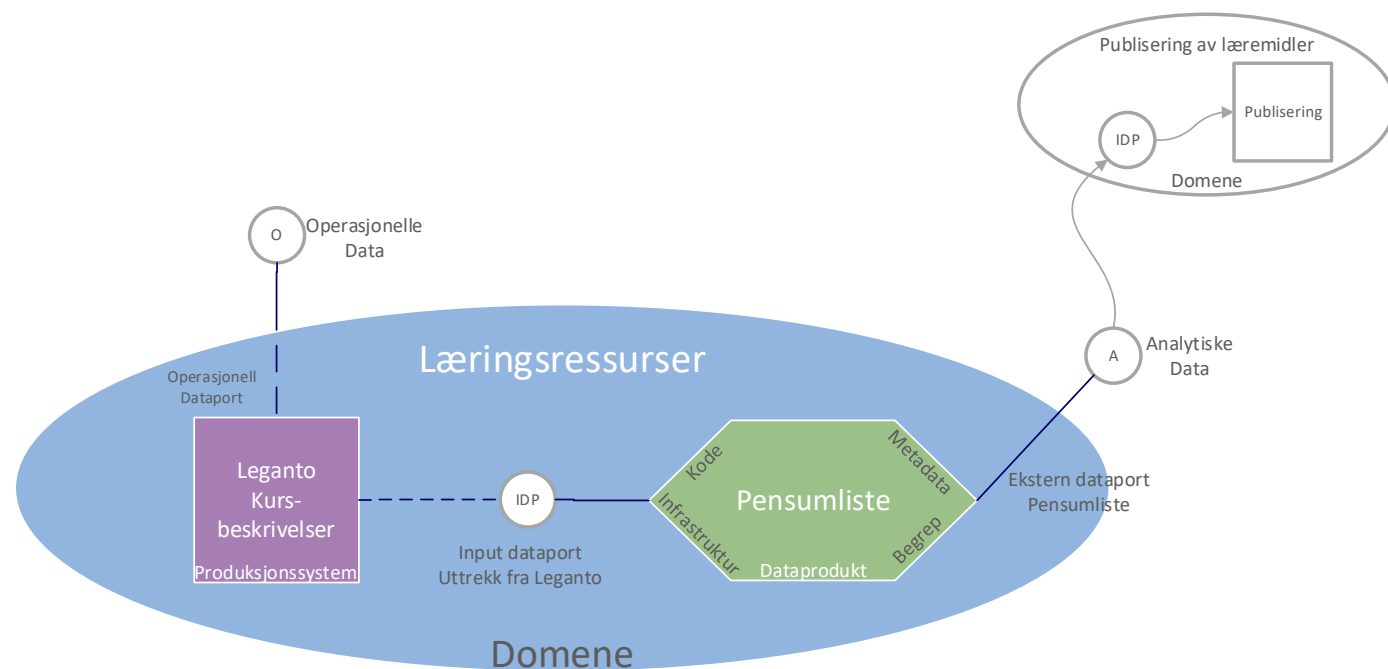
¹⁶ <https://martinfowler.com/articles/data-mesh-principles.html#TheGreatDivideOfData>

¹⁷ <https://www.oreilly.com/library/view/building-evolutionary-architectures/9781491986356/ch04.html>

Datasettet i dataproduktet skal kunne være såkalt polyglot. Det betyr at det kan være i mange forskjellige formater, alt etter behov og hensikt. For analysedata som skal brukes i forskning, som for eksempel data til microdata.no¹⁸, er det mer hensiktsmessig å følge de prinsippene som er lagt for disse dataene og standardisere format for data og metadata. Vi innser at det er et stykke vei å gå for at alle data for analyse skal være på et slikt veldefinert format, så det er viktig at definisjoner av metadata og dataprodukter kan gjøres enklere og likevel brukbare for verdikjeder som favner flere aktører.

Eksempel på dataprodukt:

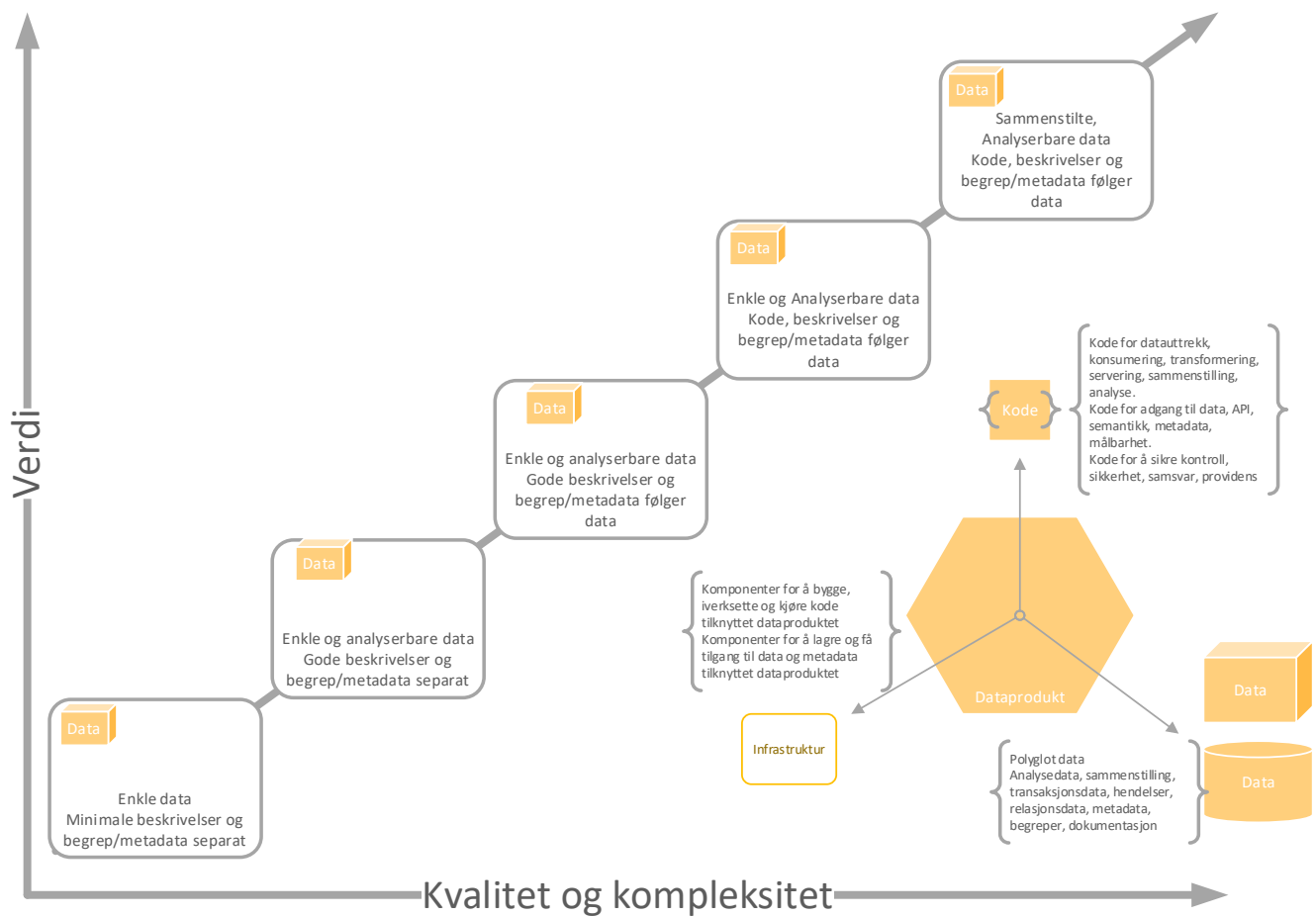
Forlagsbransjen har bruk for å vite hvilke bøker som skal brukes i pensum på universiteter og høyskoler. Denne informasjonen er tilgjengelig i systemet Legato i form av pensumlister. Et dataprodukt som skal gi den nødvendige informasjonen sammenstilt til forlagene kan utvikles basert på data som finnes i Legato. I tillegg til selve datasettet inneholder dataproduktet metadata som beskriver datasettet, dets opprinnelse, kvalitet, begreper og mer. Kode i dataproduktet sørger for uthenting fra Legato, sammenstilling av data i et analyserbart format og publisering av data til en ekstern dataport. Dataproduktet bruker en infrastruktur som er tilrettelagt for å utvikle og idriftsette. Eksempelet er basert på en eksisterende løsning, men videreutviklet som et tenkt sammenstilt analytisk dataprodukt.



Figur 9 Pensumliste som eksempel på dataprodukt

Figuren under viser en mulig utviklingsstige mot gode, fullt utviklede dataprodukter. Dette er ikke ment som en lineær vei som alle må følge, men mer en indikasjon på verdi som kan oppnås når et dataprodukt utvikles underveis.

18 <https://microdata.no/>



Figur 10 Verdi og kvalitet for dataprodukter

4.7 Hva er Kunnskapssektorens datadelingsplattform?

I KUDAF utgjør plattformen som vi har valgt å kalle Kunnskapssektorens datadelingsplattform et sett med prosesser, tjenester og teknologiske komponenter som skal sette aktørene i kunnskapssektoren i stand til å finne, utforske og dele data seg imellom.

Plattformen skal benytte nasjonale standarder, prosesser og nasjonale felleskomponenter der det er mulig. Dersom det er mangler i forhold til behov innen kunnskapssektoren, vil KUDAF bistå i utviklingen av de nasjonale løsningene der det er mulig. Dersom KUDAF bruker egne løsninger vil disse i alle tilfeller utvikles slik at de kan samhandle og utveksle data med, de nasjonale løsningene.

Prosessene, tjenestene og de teknologiske komponentene i plattformen kan iverksettes sentralt, men også som prosesser, tjenester eller komponenter som hver enkelt aktør kan benytte.

Plattformen skal gi aktørene redskap for å kunne utvikle nødvendige evner og dele data.

Koordinerende aktør blir i stor grad ansvarlig for å ta frem prosesser, tjenester og teknologiske komponenter i plattformen og tilby de til aktørene, både datatilbydere og datakonsumenter.

4.8 Evner for datadelingsplattformen

Det er en rekke evner som datadelingsplattformen må tilby. Her er en liste med noen av disse:

- Kunne utnytte nasjonale prosesser, standarder og teknologiske komponenter

- Kunne tilby veiledning og fasilitering

- Kunne fasilitere verdikjeder

- Kunne tilby administrasjon av begreper og kataloger

- Kunne tilby søknad om tilgang og kunne gi tilgang til dataprodukter

- Kunne søke etter dataprodukter

- Kunne tilby sikker identitets- og tilgangsstyring

- Kunne flytte data inn og ut av sikre analyserom uten å ha tilgang til de selv

- Kunne tilby verktøy som datatilbydere kan bruke for å skape metadata og datasett

- Kunne støtte tilbyder med å publisere dataprodukt

- Kunne tilgjengeliggjøre dataprodukt etter forespørsel

- Kunne støtte konsument ved å tilby utforskning og testing av datasett

- Kunne få tilgang til dataprodukt

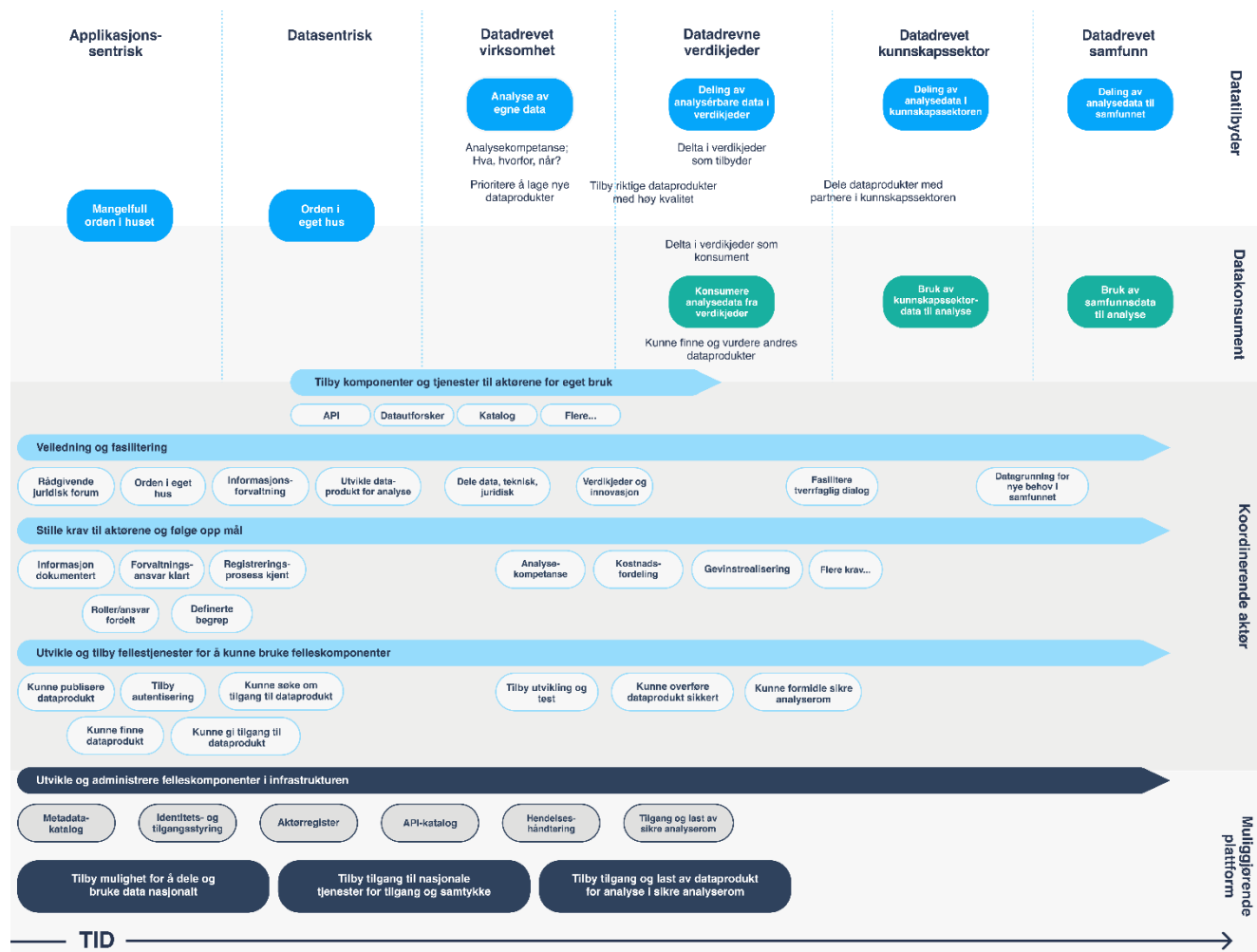
- Kunne innhente dataprodukt



5 Virkemidler i datafellesskapet

Styringsdokumentet beskriver en visjon (kapittel 3) og en strategi for gjennomføring (kapittel 8). Visjonen inneholder en virksomhetsstige som består av plataer i en reise mot et datadrevet samfunn. Figuren under viser et veikart for KUDAF der evner og virkemidler er som er nødvendig for å nå plataene i virksomhetsstigen er skissert.

I dette kapittelet i referansearkitekturen beskriver vi de viktigste virkemidlene for å kunne gjennomføre strategien og oppnå visjonen for KUDAF.



Figur 11 Veikart for KUDAF

5.1 Orden i eget hus

Orden i eget hus

Orden i eget hus er pålagt statlige virksomheter gjennom Digitaliseringsrundskrivet og dekker generelle behov også utenfor KUDAF. Arbeidet med Orden i eget hus i sektoren er både en forutsetning og en kritisk suksessfaktor for KUDAF.

Orden i eget hus skal forenkle etterlevelse av lovverk i offentlige virksomheter (GDPR/Personvernlovgivningen, Forvaltningsloven, Arkivloven, Offentlighetslova m.fl.). Alle disse forutsetter at man har kontroll på data og dokumenter virksomheten forvalter. Referansearkitekturen legger til grunn at Orden i eget hus skal bli organisert som et kontinuerlig arbeid og en måte å drive informasjonsforvaltning¹⁹ på.



Figur 12 Nødvendige steg for å skape Orden i eget hus (fra Digitaliseringsdirektoratet)

Forprosjektet, i samråd med prosjektstyret, besluttet å gjennomføre en modenhetsanalyse for å få oversikt over nåsituasjonen for Orden i eget hus i sektoren.

I denne modenhetsanalysen ble det gjort en vurdering av virksomhetenes modenhet i forhold til gjennomføring av Orden i eget hus. Denne vurderingen viser at modenheten for Orden i eget hus er gjennomgående lav i sektoren, og at det er nødvendig med et generelt løft.

Modenhetsanalysen gir pekere for hvordan veien videre med orden i eget hus kan planlegges og gjennomføres. I KUDAF organiseres i hovedsak dette arbeidet gjennom arbeidsgruppen for Informasjonsforvaltning. Det er likevel slik at det er viktig å påpeke at det er stor avhengighet mellom tiltak for informasjonsforvaltning og tiltak som anbefales i denne referansearkitekturen.

I modenhetsanalysen påpekes spesielt tre områder der det er behov for å utvikle virksomhetene i kunnskapssektoren for å understøtte arbeidet med orden i eget hus, og dermed legge grunnlaget for deltagelse i datadeling gjennom KUDAF. Utarbeidelse og anvendelse av en referansearkitektur skal, sammen med koordinerende aktør, bidra til å understøtte disse områdene.

De tre områdene er (jfr. eget dokument *Mandat for koordinerende aktør*):

¹⁹ <https://www.digdir.no/informasjonsforvaltning/kva-er-informasjonsforvaltning/2116>

1. Kompetanse
Referansearkitekturen skal beskrive mekanismer for å oppnå omforent forståelse, beste praksis og felles løsninger i sektoren. Videreutvikling og bruk av disse mekanismene bygger kompetanse hos aktørene.
2. Forståelse
Referansearkitekturen skal bidra i felleskap med Informasjonsforvaltning til å øke forståelsen for orden i eget hus
3. Ressurser
Arbeid med referansearkitekturen skal, ved koordinerende aktør og i samarbeid med aktørene;
 - bidra med menneskelige ressurser og utvikling av ressurser hos aktørene
 - beskrive og utvikle verktøy, tjenester og komponenter som aktørene kan ta i bruk
 - beskrive og utvikle en felles datadelingsplattform

5.2 Verdikjeder

Verdikjeder er tiltenkt en viktig rolle som metode for å analysere, gjennom en verdikjedeanalyse, tilrettelegge og gjennomføre datadeling i kunnskapssektoren. Gjennom å identifisere og arbeide med verdikjeder kan aktørene i sektoren raskere og smidigere få gevinster av datadeling. Arbeidet med verdikjeder vil gi innsikt i eksisterende data, bruk av disse data, utfordringer og gevinster som kan oppnås. En verdikjede vil i sin natur være begrenset til et mindre antall aktører og omhandle data innenfor et mindre omfattende område. Det vil allikevel kunne skape verdier langt utenom det konkrete området som omfattes av verdikjeden.

Verdier som skapes ved å gjennomføre verdikjedarbeid:

- Konkret innsikt og oversikt over det problemområdet som verdikjeden omfatter
- Finne utfordringer og løsninger i verdikjeden
- Skape konkret løsning for datadeling for verdikjeden
- Utvide innsikt hos aktørene for data som omfattes av verdikjeden
- Være drivende for arbeidet med begreper, metadata og dataprodukter hos aktørene
- Bedre kompetanse hos aktørene om arbeidet med begreper og metadata
- Øke kompetanse hos aktørene når det gjelder hjemler og juridiske problemstillinger
- Skape teknisk innsikt og kompetanse for datadeling for analyse hos aktørene

Eksempler på verdikjeder finnes i dokumentet *Oppsummering av gjennomførte verdikjeder i KUDAF forprosjekt*.

Hvordan kunnskapssektoren kan jobbe med verdikjeder er beskrevet nærmere i dokumentet *Metodebeskrivelser for verdikjedarbeid*.



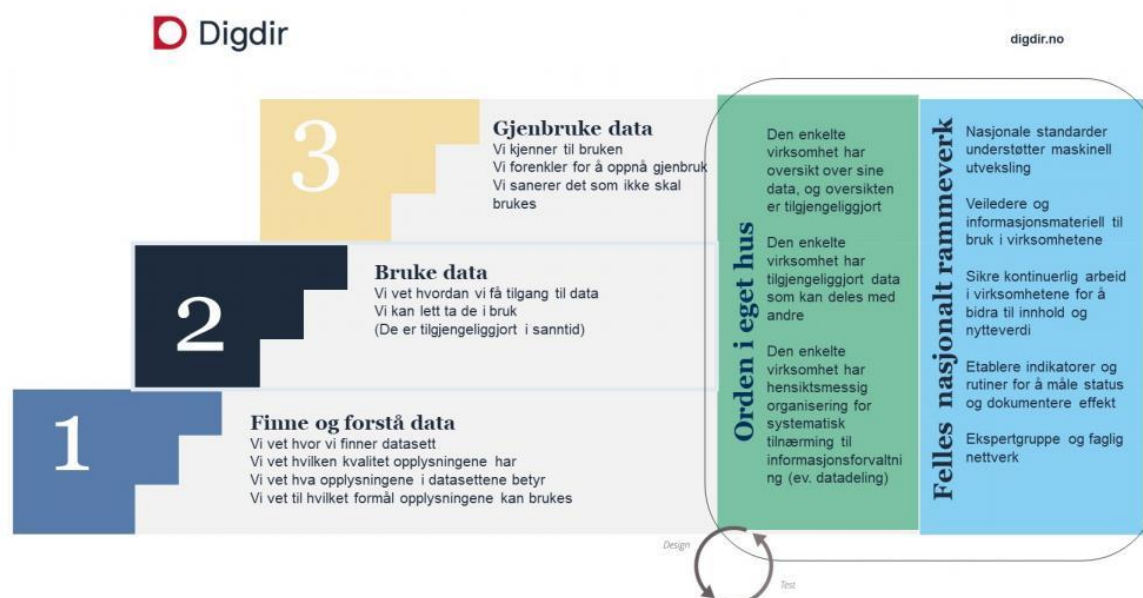
5.4 Informasjonsforvaltning

Informasjonsforvaltning

Informasjonsforvaltning er en sentral del av Orden i eget hus, og er essensielt for at virksomhetene i sektoren skal bli datadrevne.

Informasjonsforvaltning i kunnskapssektoren skal bygge på Digitaliseringsdirektoratets definisjoner og prinsipper:

«Informasjonsforvaltning betyr et helhetlig syn på aktiviteter, verktøy og andre tiltak for å sikre best mulig kvalitet, utnytting og sikring av informasjon i en virksomhet. Organiseringa av informasjonen skal være systematisk og henge sammen med virksomheten sine arbeidsprosesser.»



Figur 13 Felles mål bilde for informasjonsforvaltning (fra Digitaliseringsdirektoratet)

Prinsipper og rammeverk for informasjonsforvaltning som er beskrevet av Digitaliseringsdirektoratet gjentas ikke her^{20 21}.

I arbeidet med referansearkitekturen for KUDAF har vi funnet det hensiktsmessig å bygge på føringene fra digitaliseringsdirektoratet med føringar for harmoniserte forvaltningsprosesser, roller og ansvar samt kvalitets og sikkerhets mekanismer tilpasset deling av dataprodukter for analyseformål. Figuren under

²⁰ <https://www.digdir.no/informasjonsforvaltning/rammeverk-informasjonsforvaltning/2118>

²¹ <https://www.digdir.no/informasjonsforvaltning/kva-er-informasjonsforvaltning/2116>

beskriver disse føringer som overordnet evner. Definisjoner for disse evnene er beskrevet som muliggjørende evner i Referansearkitekturen for deling av data i høyere utdanning og forskning.²²



Figur 14 Sentrale evner innen informasjonsforvaltning

Føringene i vår kontekst for dataprodukter som skal deles i kunnskapssektoren er som følger:

Klarhet i hvem som forvalter hvilke data

Dataprodukter forvaltes der dataproduktet oppstår. Det gjelder også om dataproduktet er en sammenstilling av andre dataprodukter.

Aktørene i kunnskapssektoren skal kunne følge data tilbake til kilde og forvalter av dataprodukter.

Harmoniserte forvaltningsprosesser

Gjenbruk og standardisering av prosesser for informasjonsforvaltning vil være et viktig virkemiddel.

Sikkerhet

Identitets- og tilgangsstyring

Kunne overføre dataprodukt sikkert

Kunne formidle sikre analyserom

Vesentlige aspekter av sikkerhet, slik som identitets- og tilgangsstyring er beskrevet i [«Referansearkitektur for deling av data i høyere utdanning og forskning»](#)²³. I tillegg må KUDAF utrede hvordan sikker overføring av dataprodukter til sikre analyserom kan gjennomføres.

Datakvalitet

²² [https://unit-norge.github.io/unit-ra/main/unit-ra-datadeling-](https://unit-norge.github.io/unit-ra/main/unit-ra-datadeling-tiln%C3%A6rming/Tiln%C3%A6rming%20til%20utvikling%20av%20referansearkitekturen.html#_muliggj%C3%B8rende_kapabiliteter)

[tiln%C3%A6rming/Tiln%C3%A6rming%20til%20utvikling%20av%20referansearkitekturen.html#_muliggj%C3%B8rende_kapabiliteter](https://unit-norge.github.io/unit-ra/main/unit-ra-datadeling-tiln%C3%A6rming/Tiln%C3%A6rming%20til%20utvikling%20av%20referansearkitekturen.html#_muliggj%C3%B8rende_kapabiliteter)

²³ [https://unit-norge.github.io/unit-](https://unit-norge.github.io/unit-ra/main/B%C3%B8ker/Referansearkitektur%20for%20deling%20av%20data%20i%20h%C3%B8yere%20utdanning%20og%20forskning.html)

[ra/main/B%C3%B8ker/Referansearkitektur%20for%20deling%20av%20data%20i%20h%C3%B8yere%20utdanning%20og%20forskning.html](https://unit-norge.github.io/unit-ra/main/B%C3%B8ker/Referansearkitektur%20for%20deling%20av%20data%20i%20h%C3%B8yere%20utdanning%20og%20forskning.html)



Datakvalitet

Hva er Datakvalitet?, er beskrevet i dette dokumentet. KUDAF skal hjelpe aktørene i Kunnskapssektoren med å oppnå bedre kvalitet i sine data gjennom veiledning, et kompetanseløft og å tilby tjenester og komponenter.



5.5 Kompetansebygging

Manglende kompetanse, som årsak til lite utvikling av orden i eget hus og datadeling i kunnskapssektoren, var et av de viktigste funnene i modenhetsundersøkelsen. Kompetansebygging hos alle aktører i kunnskapssektoren er derfor en av de viktigste tiltakene som KUDAF ønsker å bidra til å gjennomføre. Kompetansen må utvikles hos koordinerende aktør, datatilbydere og datakonsumenter, og gjøres tilgjengelig for aktører utenfor sektoren og allmenheten.

Virksomhetsstigen (se kapittel 2.2 Hva er KUDAF?) og veikartet (se kapittel 5 Virkemidler i datafellesskapet) viser oversikter over hvilke evner det må utvikles kompetanse innen hos aktørene. Kapittel 5.10 Utvikle evner hos aktørene, går nærmere inn på disse evnene.

Det må bygges fagnettverk for å fange opp behov for kompetanseheving og å spre kompetanse som tilegnes (se kapittelet om felles forum).

Kompetanse er en ressurs som krever kontinuerlig forbedring samtidig som behovene endres kontinuerlig. Aktørene i sektoren må forvalte sine ressurser på en slik måte at de har oversikt over hvilken kompetanse de innehar og hvilken de trenger å utvikle. Koordinerende aktør skal være en ressurs for å veilede i kompetansebygging.



5.6 Veiledninger

Veiledninger blir en viktig del av leveransene fra KUDAF. I tillegg til videreutvikling av denne referansearkitekturen er det behov for å utvikle en rekke veiledere for aktørene og andre interesserte parter som oppdateres fortløpende med beste praksis opparbeidet i økosystemet. Utvikling av veiledere og kompetanse henger sammen og støtter hverandre.

Prosjektet baserer seg på mest mulig gjenbruk, for eksempel av veilederen for Orden i eget hus fra Digitaliseringsdirektoratet. Erfaringer og beste praksis fra verdikjedeanalyser og andre kilder kan deles gjennom veiledninger.

Også for utvikling av veiledere vil et fagnettverk være sentralt for å avdekke behov og dele kunnskap.

Eksempel på veiledere som kan utvikles:

- Hvordan identifisere verdi i eksisterende data
- Hvordan drive datautforskning i prosjekter
- Hvordan planlegge prosjekter for å jobbe datadrevet
- Tips til prioritering av dataprodukter ut fra gevinstpotensiale
 - «Top down» fra strategier og mål
 - «Bottom up» fra brukerhistorier
- Hva kan vi kaste?
- Dataprodukter: Hvordan gå fra Ad-hoc -> repeterbare -> automatiserte
- Begrepskoordinering
- Produksjon av dataprodukt
- Veikart for modningsprosessen
- Vær varsom plakat for konsumenter som sammenstiller data inn i nye dataprodukter



5.7 Prosesser

For aktørene i kunnskapssektoren er det en rekke oppgaver som må løses for å gå fra en applikasjonssentrisk tankegang til å understøtte en datadrevet sektor. For å hjelpe til med denne overgangen vil KUDAF definere prosesser som kan gjennomføres hos aktørene og i sektoren. Koordinerende aktør er ansvarlig for å utvikle disse prosessene og stimulere til at aktørene tar de i bruk. Et fagnettverk og felles forum vil være et viktig verktøy for å identifisere, utvikle og prøve ut disse og ytterligere prosesser i sektoren. I referansearkitekturen vil vi nevne noen av dem.

Forslag til prosesser som skal etableres:

Prosess	Hva inngår	Hvem berører det	Roller
Identifisere og forstå data i kontekst Informasjon dokumentert	Utarbeide praksis for å identifisere hvilke data virksomheten behandler, og hva de betyr. Forstå dette i kontekst av: domene som data inngår i, formål for å innhente, bruke og dele data, knyttet til tiltak som data brukes for	Datatilbyder	Dataprodukteier, domeneansvarlig
Å gå fra data til delbare dataprodukt Kunne publisere dataprodukt Kunne finne dataprodukt	Basert på identifiserte og forståtte data kunne utarbeide dataprodukt i form av data og metadata, samt kode for uthenting fra eksisterende systemer og leveranse til deling i for eksempel API. Sørge for dataminimering der det er nødvendig, kunne stole på felles komponenter, som sikre analyserom, der det ikke er mulig	Datatilbyder	Dataprodukteier, domeneansvarlig
Å dele data på en smidig måte gjennom automatisering og standardisering	Basert på standarder og ferdige komponenter kunne smidig levere dataprodukt raskt og med tilstrekkelig kvalitet	Datatilbyder	Dataprodukteier, utviklere, leverandører
Begrepsarbeidet Definerte begrep	Identifisere og definere begreper, termer og definisjoner i egen virksomhet. Kunne finne og samordne begreper på tvers av sektoren.	Datakonsument, Datatilbyder	Dataprodukteier



Å sikre kvalitet Datakvalitet	Kunne sikre, måle, overvåke og automatisere kvalitet i dataproduktene. Sørge for kvalitet i innhenting, lagring og uthenting av data. Sikre at kvaliteten på data er målbar ved å sørge for å gjennomføre målinger. Tilrettelegge for overvåking og automatisere denne, slik at avvik varsles tidligst mulig	Datatilbyder	Dataprodukteier
Identifisere, forstå og kunne analysere/aggregere egne data og analysevariabler Analysekompetanse	Bygge videre på identifiserte og forståtte data slik at virksomheten kan gjøre egne analyser, forstå hvilke data som kan inngå i delbare dataprodukter som er egnet for analyse og aggregering	Datatilbyder, Datakonsument	Dataprodukteier
Kunne knytte metadata, begrep og analysevariabler til dataprodukter	Bygge videre på begrepsarbeidet og kunne knytte begreper til metadata og dataprodukter. Identifisere analysevariabler i dataprodukter og beskrive disse	Datatilbyder, Datakonsument	Dataprodukteier
Kunne dokumentere analyserbare dataprodukter Analysekompetanse	Lage standardisert dokumentasjon som sikrer kvalitet og muliggjør gjenfinning av dataprodukter for analyse	Datatilbyder	Dataprodukteier
Tilbakemeldinger på datakvalitet Datakvalitet	Standard prosess for tilbakemeldinger fra datakonsumenter på kvaliteten av dataprodukter	Datatilbyder, Datakonsument	Dataprodukteier
Livssyklus for dataprodukt	Definere hvordan dataprodukter oppstår, utvikles, brukes og avvikles i et livsløp	Datatilbyder	Dataprodukteier



Livssyklus for API	Definere hvordan API oppstår, utvikles, brukes og avvikles i et livsløp	Datatilbyder	Dataprodukteier
Livssyklus for datamodeller	Definere hvordan datamodeller oppstår, utvikles, brukes og avvikles i et livsløp	Datatilbyder	Dataprodukteier
Livssyklus for begreper	Definere hvordan begreper oppstår, utvikles, brukes og avvikles i et livsløp	Datatilbyder, Datakonsument	Dataprodukteier
Å avklare kostnadsfordeling	Prosess for å avklare kostnadsfordeling mellom tilbyder, konsument og koordinerende aktør og eventuelt andre, som for eksempel Kunnskapsdepartementet	Datatilbyder, Datakonsument, Koordinerende aktør, andre	Dataprodukteier, domeneansvarlig, ledelse, eier

Tabell 1 Forslag til prosesser som skal etableres



5.8 Fasilitere samarbeid

Fasilitere tverrfaglig dialog

Det er viktig å erkjenne at det i arbeidet med å øke datadeling i kunnskapssektoren vil være en skjevfordeling mellom innsats og utbytte for datatilbydere og datakonsumenter. Det vil i mange tilfeller ikke være slik at en datatilbyder på kort sikt har utbytte av å legge til rette for bruk av sine data. Erfaringer i arbeidet med verdikjedeanalyser i forprosjekter bekrefter dette, sammen med det faktum at aktørene ikke vet/forstår at de trenger hverandres data. Underveis så vi at produsent og konsument trenger en tredjepart som kobler og får dem til å snakke sammen om de riktige tingene.

Det er derfor nødvendig at det jobbes med motivasjon og fasilitering fra eier og fra koordinerende aktør slik at forståelsen for og gjennomføringen av slik datadeling blir mulig. Det er naturlig å tenke seg en forholdsvis lang periode der det må benyttes intensiver, krav, motivasjon og andre former for fasilitering for å opparbeide en datadrevet sektor som fungerer.

Det vil være behov for at koordinerende aktør eller felles forum i sektoren hjelper til med å harmonisere behovsforståelse mellom datatilbydere og datakonsumenter.

Koordinerende aktør vil være sentral for å identifisere, analysere og utvikle verdikjeder i samarbeid med aktørene. Disse verdikjedene vil være viktige for å vise samfunnsverdi og å utvikle prosesser, tjenester og komponenter som kan gjenbrukes i sektoren.



5.9 Koordinering og samordning gjennom felles forum

Å skape en datadrevet kunnskapssektor, der deling av data er en naturlig del av kulturen, er ikke mulig uten samarbeid mellom aktørene. De beste løsningene blir skapt av de som er berørt. Mange fagfelt blir berørt av en slik endring av praksis. Kunnskapssektorens egne fagfelt selvsagt, men også juridiske og IT-faglige endringer skal gjennomføres. Det er derfor ønskelig å organisere fora der aktørene kan bidra med, og få, faglig støtte for sine virksomheter.

5.9.1 Rådgivende juridisk forum

Rådgivende juridisk forum

Erfaringer fra offentlig sektor der man ønsker å dele data mellom virksomheter viser at det ofte er ulike juridiske hindre eller ulike tolkninger av regelverket som forsinke eller stopper slike prosesser. Hjemler for innhenting og oppbevaring av data kan gi rom for tolkning eller direkte hindre annen bruk av data. Virkelige eller uklare krav fra personvernlovgivningen kan skape ytterligere usikkerhet. Det kan være utfordrende for virksomhetene å få avklart og fjernet slike hindre.

I datadeling er det alltid flere enn en part. Det betyr at partene kan ha ulike oppfatninger eller tolkninger av hvordan juridiske konflikter kan løses. Slike problemer kan forsinke eller stoppe datadeling, selv om slik deling kan gi en stor samfunnsmessig gevinst.

Det er derfor besluttet å opprette et rådgivende juridisk forum i forbindelse med gjennomføringen av KUDAF. Forumet skal kunne gi juridiske råd til aktørene i kunnskapssektoren og, dersom partene ønsker det, kunne finne løsninger der konflikter oppstår.

Det rådgivende juridiske forumet skal bestå av jurister fra aktører i ulike deler av kunnskapssektoren. Se eget dokument *Mandat for Rådgivende juridisk forum* for nærmere beskrivelse.

5.9.2 Felles teknisk fagforum

Dele data teknisk

Samarbeid på tvers av virksomheter kan organiseres i tilknytning til fagfelt, eller om du vil domener, som deles av flere.

Fagfora kan utvikles for å dele erfaringer, identifisere verdikjeder og finne løsninger på felles problemer.

Disse fora kan omfatte domener på tvers eller generelle problemstillinger, slik som de IT-faglige. Koordinerende aktør får en viktig rolle i å identifisere og samordne disse faglige felles fora.

IT-faglige fora vil, i tillegg til fagaktørene, inkludere de aktørene som utvikler og forvalter fellesløsninger for kunnskapssektoren. Dette kan også omfatte kommersielle leverandører til sektoren. En av oppgavene til slike fellesfora er å identifisere og prioritere utviklingen av tjenester og verktøy for å hjelpe aktørene (se under).



5.9.3 Fagnettverk for «Orden i eget hus»

Orden i eget hus

En viktig forutsetning for en datadrevet kunnskapssektor er arbeidet med Orden i eget hus. Som modenhetsanalysen har vist er det et stykke vei å gå for aktørene. Kunnskapsdepartementet har i sine tildelingsbrev lagt vekt på at virksomhetene er forpliktet til å jobbe med Orden i eget hus. Arbeidet med Orden i eget hus vil derfor ha et spesielt fokus fra KUDAF og koordinerende aktør.

KUDAF vil, som en del av sin leveranse, sammen med koordinerende aktør, opprette et nettverk av fagressurser fra virksomhetene som skal jobbe med Orden i eget hus.



5.10 Utvikle evner hos aktørene

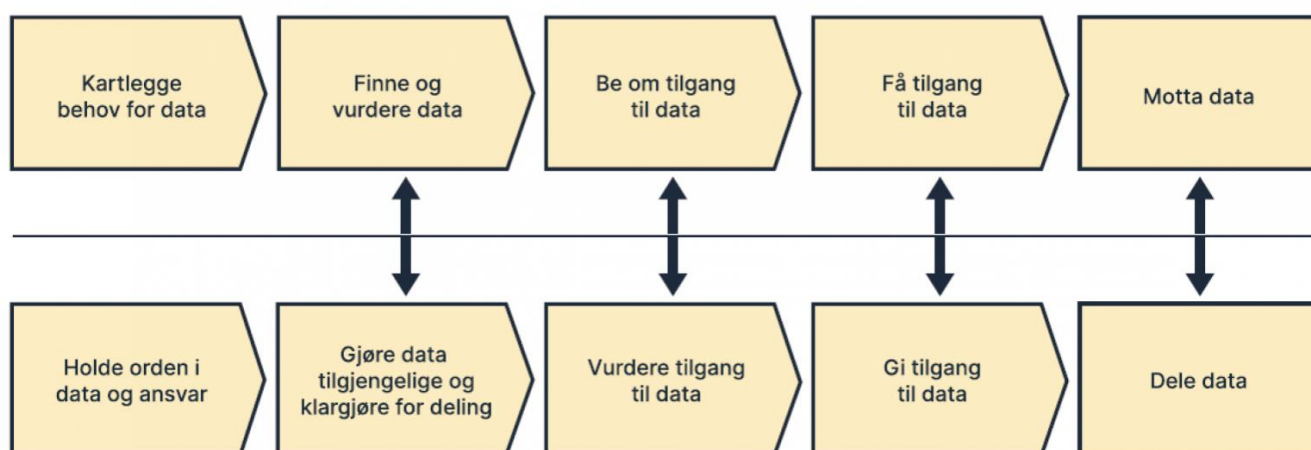
For at kunnskapssektorens aktører skal kunne nå målet om en datadrevet sektor, må det utvikles en rekke evner, også kalt kapabiliteter. Evner i denne sammenhengen handler om mer enn bare kompetanse, det må også være mål, ledelse og organisasjon på plass, slik at oppgaver kan løses i praksis. Inkludert i dette vil være samordnende aktiviteter for leverandørstyring.

I denne sammenhengen definerer vi tre hovedtyper aktører: Datatilbyder, datakonsument og koordinerende aktør. I tillegg til disse har vi andre fasiliterende aktører, som for eksempel leverandører av IT-løsninger til sektoren, eller tjenesteleverandører, som sky- og driftstjenester.

I det følgende vil vi skissere evner som vi har identifisert som viktige for aktørene å inneha. Dette er ikke ment som endelig liste, og vi antar at behovet for nye evner vil materialisere seg etter hvert som sektoren får erfaringer å bygge på.

Vi tar utgangspunkt i verktøykassen for deling av data fra Digitaliseringsdirektoratet avbildet i Figur 15 og beskrevet i referansen.²⁴ Verktøykassen beskriver stegene i datadelingsprosessen for datatilbydere og datakonsumenter, samt samhandlingen mellom disse.

Prosess for konsumenten, som skal få data fra andre



Prosess for tilbyderen, som skal gi data til andre

Figur 15 Verktøykassen for deling av data fra Digitaliseringsdirektoratet

Vi supplerer verktøykassen med evner som skal utvikles gjennom KUDAF for å støtte gjenbruk og viderebruk av data til analyseformål i et fellesskap. Noen evner er spesifikke for aktøren (koordinerende aktør, tilbyder og konsument) og noen evner er felles for alle.

5.10.1 Evner hos koordinerende aktør

Koordinerende aktør innehar en unik rolle i utviklingen av et økosystem for data i kunnskapssektoren. Aktøren skal ikke bare ta frem, utvikle og igangsette felleskomponentene i datadelingsplattformen, men

²⁴ <https://www.digdir.no/datadeling/nasjonal-verktoykasse-deling-av-data/2243>



skal også sette de andre aktørene i stand til å delta i samarbeidet. For dette kreves det at aktøren innehar evner innen mange fagfelt utover det IT-tekniske. Aktøren må bidra med undervisning, fasilitering, veiledning, ledelse, samordning av prosesser, motivasjon, rådgivning og juridisk kompetanse, for å nevne noe.

Evner som skal gi verdi til aktørene

For at kunnskapssektorens aktører skal kunne nå målet om en datadrevet sektor, må det utvikles en rekke evner (gjerne omtalt som kapabiliteter) hos koordinerende aktør. Evnene skal sette koordinerende aktør i stand til å hjelpe aktørene i sektoren med å løse oppgaver for å skape verdi av data. Evner i denne sammenhengen handler om mer enn bare kompetanse, det må også være mål, ledelse og organisasjon på plass, slik at oppgaver kan løses i praksis. I denne sammenhengen definerer vi tre hovedtyper aktører; datatilbyder²⁵, datakonsument²⁶ og koordinerende aktør. I tillegg til disse tre aktørene har vi også tjenesteleverandører, som leverer underliggende tjenester i den tekniske infrastrukturen, i tillegg til leverandører av skytjenester og driftstjenester.

I det følgende vil vi skissere evner som vi har identifisert som viktige for koordinerende aktør å inneha for å kunne ivareta behovene hos de andre aktørene. Dette er ikke ment som endelig liste, og vi antar at behovet for nye evner vil materialisere seg etter hvert som sektoren får flere erfaringer å bygge på.

Koordinerende aktør innehar en unik rolle i utviklingen av et økosystem for data i kunnskapssektoren. Aktøren er ansvarlig for å etablere og forvalte felleskomponentene i datadelingsplattformen, og skal bidra til å sette de andre aktørene i stand til å bidra i samarbeidet. For dette kreves det at aktøren innehar evner innen mange fagfelt, også utover det IT-tekniske. Aktøren må bidra med opplæring, fasilitering, veiledning, ledelse, samordning av prosesser, harmonisering, motivasjon, rådgivning og juridiske avklaringer.

Det følgende er identifiserte evner som koordinerende aktør må inneha eller kunne opparbeide.

Inneha nødvendig kompetanse for å oppnå KUDAFs målsetninger, og kunne dele denne med aktørene

Koordinerende aktør må opparbeide kompetanse innen en rekke felt, både IT-tekniske og andre, slik at denne kompetansen kan deles med de andre aktørene. Aktøren må inneha kompetanse, og evnen til å formidle den, innen datadeling (IT-teknisk), begrepskoordinering, digital transformasjon, informasjonssikkerhet og personvern, informasjonsforvaltning, orden i eget hus, juridisk samordning, verdikjeder og mange andre områder.

Det må forventes at koordinerende aktør må kunne tilgjengeliggjøre personer eller team med relevant kompetanse som kan stilles til rådighet for aktørene i sektoren i kortere eller lengre perioder. Dette vil

²⁵ <https://www.digdir.no/datadeling/dele-data-med-andre/2252>

²⁶ <https://www.digdir.no/datadeling/bruke-data-fra-andre/2251>



kunne være representanter fra ulike deler av sektoren som engasjeres for en tidsavgrenset periode for å løse konkrete utfordringer. Koordinerende aktør må kunne utarbeide undervisningsopplegg, inkludert materiale og veiledere som gjøres tilgjengelig for aktørene i sektoren.

5.10.1.1 Å kunne veilede aktørene

Dele data, teknisk,
juridisk

Koordinerende aktør må ha tilgang til ressurser og en organisasjon som kan veilede aktørene i sektoren. Koordinerende aktør må kunne utnytte ressurspersoner fra ulike deler av sektoren med spesialkompetanse innenfor sine områder. Veiledning må være organisert slik at det er mulig for aktørene å få rask bistand til å løse sine utfordringer.

5.10.1.2 Å fasilitere verdikjeder og fagområdesamarbeid

Verdikjeder og
innovasjon

Fasilitere
tverrfaglig dialog

En viktig oppgave for koordinerende aktør, spesielt de første årene av arbeidet med KUDAF, blir å hjelpe aktørene i sektoren i å identifisere, analysere og utvikle verdikjeder for data. Koordinerende aktør må inneha spesialkompetanse på verdikjedeutvikling og en organisert tilnærming til å utvikle seg selv og aktørene slik at de er i stand til drive slik utvikling. Etter hvert som verdikjeder blir identifisert og videreutviklet, vil det kunne dukke opp behov for tettere samarbeid mellom aktører innenfor avgrensede områder og formål med bruken av data. Dette kan betraktes som et «lokalt økosystem» innenfor et delområde (f.eks. integrering) for å utforske hvordan nye målsetninger kan oppnås. Her kan koordinerende aktør ha en rolle som initiativtaker og fasilitator for dette samarbeidet. Koordinerende aktør må derfor ha evne til å kunne gjennomføre slike tiltak.

5.10.1.3 Å kunne fasilitere rådgivende juridisk forum

Rådgivende
juridisk forum

Koordinerende aktør må kunne organisere og fasilitere det juridiske forumet, der aktørene i sektoren kan få hjelp til å avklare juridiske utfordringer som fortolkning av regelverk og sondering av mulighetsrommet for endringer i regelverk når dette står i veien for potensielt store samfunnsøkonomiske gevinster. Koordinerende aktør må derfor ha tilgang til ressurser med høy kompetanse på gevinst- og samfunnsøkonomisk analyse ved deling og bruk av data på tvers av virksomheter i sektoren og utenfor sektoren, for å forberede saksunderlag for det juridiske forumet.

5.10.1.4 Å kunne fasilitere og veilede begrepsarbeid og koordinering

Koordinerende aktør må ha kompetanse på og kunne veilede aktørene om arbeid med begreper og termer. Aktørene i sektoren skal kunne få hjelp til å utvikle sine egne begreper og termer i de domene de



representerer. Koordinerende aktør må kunne fasilitere koordinering og om mulig harmonisering av begreper og termer mellom domene og mellom aktørene. Koordinerende aktør får et spesielt ansvar for å ta frem teknologiske løsninger for å ivareta deling og samhandling for begreper og termer i sektoren, og å sikre at begreper og termer samordnes med bruk av metadata i slike løsninger.

**Fasilitere
tverrfaglig dialog**

5.10.1.5 Utvikle og administrere felleskomponenter i datadelingsplattformen

Metadata-
katalog

Identitets- og
tilgangsstyring

Aktørregister

API-katalog

Hendelses-
håndtering

Tilgang og last
av sikre
analyserom

Koordinerende aktør må inneha evne til å organisere og gjennomføre utvikling og administrasjon av felleskomponenter i datadelingsplattformen (se referansearkitektur kapittel «Komponentene i datadelingsplattformen»). Dette inkluderer å inneha den nødvendige kompetansen på disse komponentene og nasjonale felleskomponenter.

5.10.1.6 Utvikle og tilby fellestjenester for å kunne bruke felleskomponenter

API

Datautforsker

Katalog

Flere...

Tilby utvikling
og test

Kunne overføre
dataprodukt sikkert

Kunne formidle sikre
analyserom

Mens felleskomponentene i den tekniske infrastrukturen utgjør de sentrale delene av datadelingsplattformen, er det viktig å bygge fellestjenester på toppen av disse som gjør disse evnene lettere tilgjengelige for aktørene i sektoren og andre som ønsker tilgang til dataprodukter fra kunnskapssektoren. Koordinerende aktør må ha evne til å utvikle og administrere disse tjenestene slik at de kan tilbys til aktørene. Dette kan være tjenester som gir mulighet for å søke om tilgang og gi slik tilgang, datautforsker, søketjeneste, utvikling, testmiljø og sandkasser, begrepskatalog og metadatakatalog for å nevne noen. Disse tjenestene vil bygge på felleskomponenter i datadelingsplattformen og også gi adgang til å bruke nasjonale felleskomponenter.

5.10.1.7 Utvikle og tilby komponenter og tjenester som aktørene kan bruke i sine egne løsninger

API

Datautforsker

Katalog

Flere...

Kunne publisere
dataprodukt

Tilby
autentisering

Kunne finne
dataprodukt

Kunne søke om
tilgang til
dataprodukt

Kunne gi tilgang til
dataprodukt

Tilby utvikling
og test

Koordinerende aktør må ha evne til å utvikle, administrere og tilby komponenter, verktøy og tjenester som beskrevet i referansearkitektur kapittelet «Virkemidler i datafelleskapet / Tjenester og Verktøy».



5.10.1.8 Tilby mulighet for å dele og bruke data nasjonalt

Tilby mulighet for å dele og bruke data nasjonalt

Tjenester og løsninger som tilbys i datadelingsplattformen og til aktørene skal i den grad det er mulig bruke nasjonale felleskomponenter og gjøres tilgjengelig gjennom disse der det er mulig. På samme måte skal data som er gjort tilgjengelig gjennom nasjonale løsninger eller i andre deler av offentlig sektor kunne gjøres tilgjengelig for aktørene i kunnskapssektoren. Koordinerende aktør skal ha evnen til å sikre at data kan deles og brukes nasjonalt.

5.10.1.9 Tilby tilgang til nasjonale tjenester for tilgang og samtykke

Tilby tilgang til nasjonale tjenester for tilgang og samtykke

Koordinerende aktør skal tilby bruk av nasjonale fellestjenester for tilgang og samtykke på samme måte som for andre nasjonale tjenester. Aktøren må ha kompetanse og evne til å utvikle tjenester som kan utnytte disse nasjonale komponentene. Eksempel på slik løsninger er Maskinporten, ID-porten og Altinn Samtykke.

5.10.1.10 Tilby tilgang til bruk av analyser i, og sikker overføring av dataprodukter til sikre analyserom

Kunne overføre dataprodukt sikkert

Kunne formidle sikre analyserom

Tilgang og last av sikre analyserom

Tilby tilgang og last av dataprodukt for analyse i sikre analyserom

Sikre analyserom er en tjeneste som tilbys av noen aktører i og utenfor sektoren. Koordinerende aktør skal kunne tilby tjenester slik at aktørene i sektoren kan få tilgang til slike sikre analyserom og mekanismer for å kunne sikkert overføre data fra datatilbydere inn i det sikre analyserommet. Slike tjenester må kunne tilby å sette opp avtaler mellom aktør som skal bruke det sikre analyserommet og leverandøren. I tillegg til å sikre dataoverføring, bør koordinerende aktør tilby mulighet for pseudonymisering av sensitive data på en måte som sikrer at kobling av data fremdeles er mulig.

5.10.2 Felles krav som stilles til aktørene – datatilbyder og datakonsument

KUDAF-programmet og koordinerende aktør vil iverksette mange tiltak som skal hjelpe aktørene med å forbedre sine virksomheter. Mange av disse er beskrevet i denne referansearkitekturen og i andre dokumenter fra prosjektet. Målet er at det skal være mulig for aktørene å gjennomføre gradvise og smidige forbedringer mot målet med orden i eget hus og en datadrevet kunnskapssektor.

Det er likevel slik at det vil stilles krav til aktørene i denne prosessen. KUDAF som prosjekt eller koordinerende aktør kan ikke gjennomføre de endringene som er nødvendig hos hver enkelt aktør. Det er det bare aktørene selv som kan gjøre. Det vil være mål og krav om måloppnåelse i forbedringsprosessene. Endelig formulering av disse kravene vil kunne være en oppgave for noen av de felles fora som skal etableres (se over).



Vi har likevel valgt å ta med noen sentrale eksempler på slike krav her:

Kontroll på egne begreper

Definerte begrep

Forstå hvilke sentrale begreper som brukes i domenene i virksomheten og hvilke forhold de har til hverandre. Uten en slik forståelse er det umulig å relatere egne data til andres og kunne forklare hva egne data betyr.

Dokumentert og tilgjengeliggjort oversikt over egne data

Informasjon dokumentert

Når man har kontroll over begreper og forstår hva det betyr, må virksomheten bruke dette til å dokumentere de data som finnes i domenene og gjøre dette tilgjengelig. Tydelig og god dokumentasjon lett tilgjengelig vil gjøre det mulig å bruke data som finnes i organisasjonen på nye og bedre måter, både internt i virksomheten og for andre i sektoren.

Velorganisert informasjonsforvaltning

Forvaltnings- ansvar avklart

En forutsetning for å kunne oppnå målene i Orden i eget hus er en god og velfungerende informasjonsforvaltning. Virksomhetene må organisere sin informasjonsforvaltning slik at de nødvendige tiltak kan gjennomføres. Det betyr at informasjonsforvaltningsfunksjonen må ha den nødvendige gjennomslagskraft for å utvikle virksomheten for å bli datadrevet.

Utviklet teknisk kompetanse innen informasjonsforvaltning

Informasjons- forvaltning

I tillegg til å ha organisatoriske gjennomslag for informasjonsforvaltning må også virksomheten kunne gjennomføre de tekniske tiltak som er nødvendige. Det betyr å inneha teknisk kompetanse om informasjonsforvaltning, for eksempel kan de være kompetanse tilknyttet standarder, metadatahåndtering, verktøyintegrasjon, teknisk prosessstøtte, data pipelines og data engineering.

Systematisk tilnærming til dataforvaltning



Virksomheten har kompetanse og avsatte ressurser for operativ forvaltning av data gjennom dens livssyklus. Relevante aktiviteter vil være tilknyttet innsamling, validering, lagring, bearbeiding, beskyttelse, tilgjengeliggjøring, bevaring, verifisering, kvalitetsmåling, minimering, og destruering.

Utviklet kompetanse om dataprodukter og analysedata/analyserbare data

Analyse- kompetanse

I en datadrevet virksomhet vil det være viktig å ha kontroll over egne dataprodukter og analysedata. Virksomhetene må derfor sikre kompetanse for å kunne forstå dataprodukter og kunne bygge og forvalte skalerbar infrastruktur og datastrømmer for hensiktsmessig tilgang til data for ulike bruk.

Utviklingskompetanse

Tilby utvikling og test

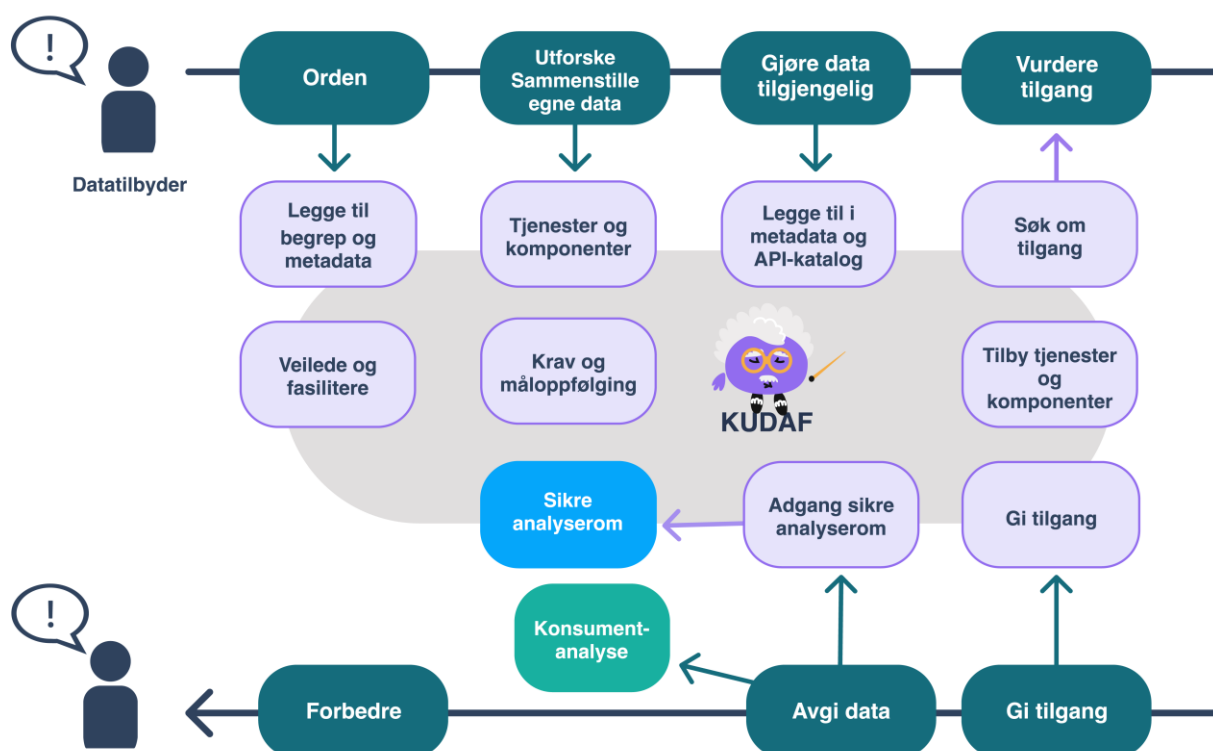
I en datadrevet virksomhet vil det å kunne utvikle egne dataprodukter bli en betydelig suksessfaktor. Det blir derfor viktig å kunne levere slike med kvalitet og innen rimelig tid. For at virksomhetene skal kunne levere dataprodukter internt og eksternt vil det være behov for å utvikle ny kompetanse for eksisterende IT-utviklere og i noen tilfeller tilføre ny IT-utviklingskapasitet. For de virksomhetene eller i de domene der man benytter systemer utviklet eksternt, enten det er «standardsystemer» eller tilpassede, så må det vurderes om leverandøren av disse systemene kan levere de nødvendige dataproduktene, eller om det er mer hensiktsmessig å utvikle slike selv.

5.10.3 Evner hos datatilbyder

Datatilbyderne er kanskje de som står foran de største utfordringene når det gjelder deltagelse i en datadrevet kunnskapssektor. Likevel er det gevinster som kan oppnås for aktørene som skal tilby sine data. Gjennom å kjenne sine egne data og finne nye måter å utnytte de på kan aktørene finne bedre måter å utnytte ressurser, få ny innsikt og forbedret styring. I samarbeid med andre aktører kan virksomhetene oppnå å løse sine samfunnsoppdrag på en bedre måte.

Figuren under illustrer noen av de evnene som datatilbyderen må inneha i en prosessflyt for deling av data. En del av disse evnene er beskrevet i seksjonen Evner for datadelingsplattformen.





Figur 16 Evner for deling av data hos en datatilbyder

Det følgende er identifiserte evner som datatilbyder må inneha eller kunne opparbeide.

5.10.3.1 Vite hvilke data virksomheten har (orden i eget hus)

Arbeidet med tilrettelegging for «orden i eget hus» og informasjonsforvaltning foregår parallelt med arbeidet med denne referansearkitekturen. I kapittelet «Virkemidler i fellesskapet – Orden i eget hus» beskrives dette nærmere. Vi vil derfor bare sette lys på noen av de evnene som datatilbyder skal utvikle i dette arbeidet her.

Orden i eget hus

Vite hvilke dataprodukter domenet har; For de domene som virksomheten omfatter må det identifiseres og forstås hva som er de viktige dataproduktene som kan leveres.

Informasjon dokumentert

Vite hvordan dataprodukter henger sammen (gjerne knyttet til variabler); Data i domener og på tvers av domener er gjerne knyttet sammen med variabler. Virksomheten må ha kunnskap om hvordan de forskjellige dataproduktene henger sammen, og hvordan sammensatte dataprodukter kan produseres

Kunne publisere dataprodukt

Vite hvem som har ansvar for dataene i dataprodukter (inkludert domener og roller); Virksomhetene må identifisere hvem som har ansvaret for data i domener og ha personer i roller i domeneene som har ansvar for disse. I dette dokumentet omtaler vi disse som dataprodukteier og domeneansvarlig. I tillegg bør virksomheten ha juridisk kompetanse som kan vurdere slike ansvarsforhold.

Roller/ansvar fordelt

Vite hva data brukes til (formål); Virksomheten må ha kunnskap om hvilke formål som data er samlet inn for å fylle og hvilke sekundærformål som data brukes til.

Orden i eget hus

Ha kvalitetsmål knyttet til formål; Forskjellige formål for bruk av data vil kreve ulike grad av kvalitet på disse data. For hvert formål som data skal brukes til må virksomheten kunne sette og følge opp mål for kvaliteten på disse data.

Datakvalitet

Kjenne kvaliteten til data; Virksomheten må kunne måle og kunne vurdere kvaliteten i data som virksomheten behandler.

Datakvalitet

Kunne forvalte dataprodukter med riktig kvalitet; Virksomheten må være i stand til å måle, følge opp og sikre at dataprodukt som leveres har riktig kvalitet over tid

Forvaltningsansvar avklart

Ha versjonskontroll på API, datasett, begreper og datamodeller;

Flere krav...



5.10.3.2 Være i stand til å tilby data til andre

Virksomheten må ha evnen til å kunne produsere dataprodukter. Dataprodukter er sammensatt av data, kode for ekstraksjon og sammenstilling og kode for sikker levering. Virksomheten må kunne

vurdere hvem som skal få adgang til dataproduktene. De må kunne vurdere ansvar for viderebruk av dataproduktene.

**Kunne gi tilgang til
dataprodukt**

5.10.3.3 Kunne prioritere. lage nye og endre eksisterende dataprodukter

Mange endringer vil kunne skje i forhold til de data en virksomhet bruker. Det vil være nødvendig for virksomheten å kunne lage nye dataprodukter basert på ny og oppdatert informasjon, nye behov og nye formål for data. Virksomheten må ha evnen til å kunne analysere definisjon og bruk av dataprodukter: Hva skjedde? Hvorfor skjedde det? Hva kommer til å skje? Hva burde vi gjøre?

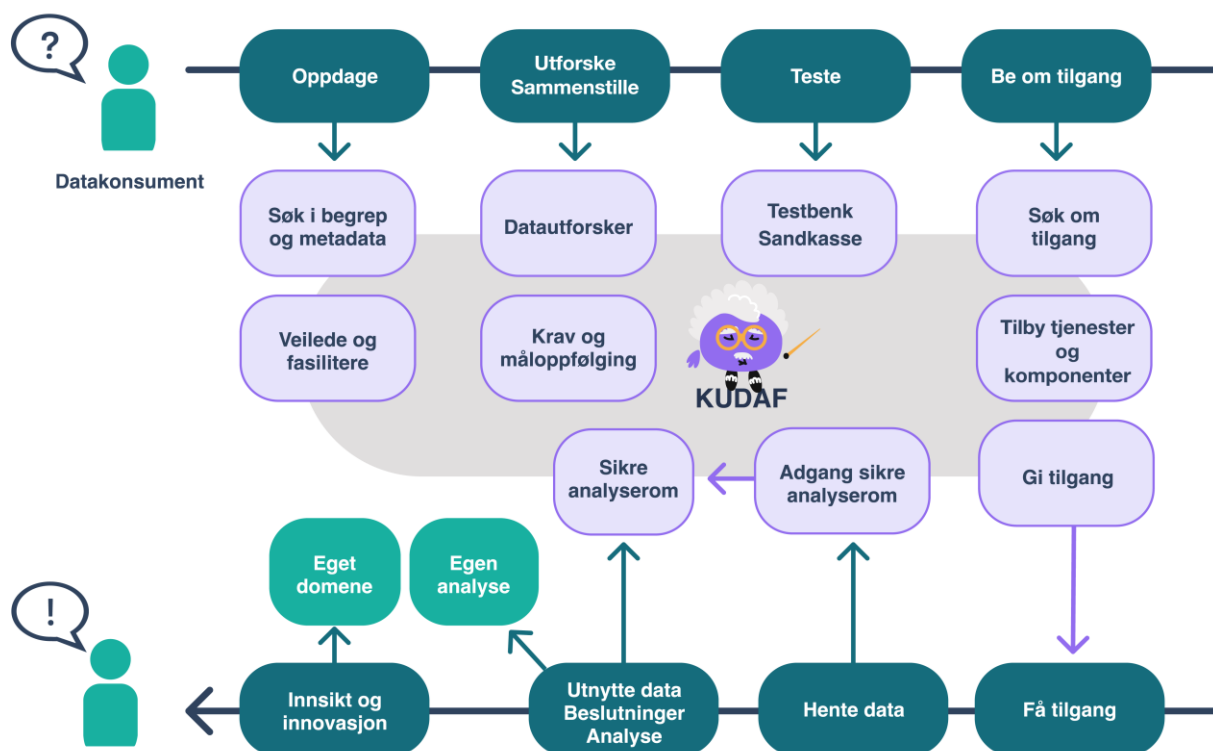
**Kunne publisere
dataprodukt**

5.10.4 Evner hos datakonsument

Datakonsumenter som skal bruke dataprodukter fra andre aktører i eller utenfor sektoren må inneha eller tilegne seg en rekke evner (kapabiliteter). I tillegg til de IT-faglige er det flere andre evner som må være til stede.

Figuren under illustrer noen av de evnene som datakonsumenten må inneha i en prosessflyt for deling av data. En del av disse er detaljert i seksjonen Evner for datadelingsplattformen.





Figur 17 Evner for deling av data hos en datakonsument

5.10.4.1 Kunne finne datasett man ønsker seg

Datakonsumenten må være i stand til å kunne søke i innhold i dataprodukt til å avgjøre om dette er de data som aktøren ønsker.

Datakonsumenten må kunne finne riktige dataprodukt ut ifra ønsket tiltak som datakonsumenten ønsker å gjennomføre.

**Kunne finne
dataprodukt**

5.10.4.2 Forstå dataproduktet, se relevans for eget bruk

Datakonsumenten må kunne forstå og tolke metadata, begrep og dokumentasjon av dataprodukt for å kunne avgjøre relevans for eget bruk.

Flere krav...

5.10.4.3 Kunne gjøre juridisk vurdering av bruk av dataprodukter

Datakonsumenten må kunne vurdere om dataprodukter er av en slik art at virksomheten har lovlig rett til tilgang til og bruksrett for dem.

Flere krav...

5.10.4.4 Kunne føre dialog med datatilbydere om nye dataprodukt som det er behov for

Datakonsumenter må kunne identifisere og beskrive dataprodukter som de har behov for. I dialog med datatilbyder må de kunne utarbeide spesifikasjon for, utvikle og teste nye dataprodukt.

Flere krav...

5.10.4.5 Kunne bruke datasett

Analyse- kompetanse

27

Vi vet hva som skjedde, hva bør vi gjøre med det?

Vi vet hvorfor det skjedde, hva bør vi gjøre med det?

Vi tror vi vet hva som kommer til å skje, hvordan kan vi best håndtere det?

Dataene anbefaler en handling, er vi enige?

Ta ansvar for etisk bruk av dataprodukt

5.10.4.6 Ansvarliggjøre konsument for bruk av data

KUDAF-programmet, i samarbeid med aktørene, vil utvikle en «Vær varsom» plakat for bruk av data.

Aktører som konsumerer data, må være kjent med innholdet på en slik plakat og forstå konsekvensene av eget bruk av andres data. Datakonsumenten må inneha evne til egne juridiske vurderinger i forhold til innhenting, bruk og publisering av data.

5.10.4.7 Kunne sammenstille data fra forskjellige kilder i nye datasett

Analyse- kompetanse

For å kunne utnytte data fra forskjellige kilder må datakonsumenter inneha kompetanse for å forstå slike sammenhenger og kunne sikre kvalitet i nye dataprodukter fra slike sammenstillinger.

27 STAGES OF DATA ANALYTICS MATURITY (ADAPTED FROM DAVENPORT & HARRIS 2007 / GARTNER 2012:

[https://www.researchgate.net/figure/STAGES-OF-DATA-ANALYTICS-MATURITY-ADAPTED-FROM-DAVENPORT-HARRIS-2007-GARTNER-](https://www.researchgate.net/figure/STAGES-OF-DATA-ANALYTICS-MATURITY-ADAPTED-FROM-DAVENPORT-HARRIS-2007-GARTNER-2012_fig1_325285755)

2012_fig1_325285755



5.11 Roller

Vi har identifisert de ønskede roller listet her. Videre detaljering av rollene vil skje i hovedprosjektet. Rollene for dialogpartnere innen tjenesteproduksjon er dokumentert i [«Referansearkitektur for deling av data i høyere utdanning og forskning»²⁸](#).

Roller/ansvar fordelt

Dataprodukteier (både IT og forretning må være dekket her)

Dekker domeneforståelse av innhold, formål, analysekompetanse, utførende juridisk.

Rolle	Dataprodukteier
Beskrivelse av rollen	Ansvarlig for en eller flere dataprodukter som er delbare ressurser i et domene. Dataproduktet inneholder datasettet, metadata, infrastruktur og kode som skal følge dataproduktet.
Ansvars-områder og oppgaver	Formål og veikart for dataproduktet. Dialog med dataprodukt-konsumentene og balansering av de ulike behov som datakonsumentene har. Livssyklus til datasettet levert, inklusivt når endringer innføres og skjemaer og data avvikles. Begrepene som er definert i datasettet og for riktig gjenbruk av begreper definert i andre dataprodukter. Metadata som beskriver datasettet, inklusivt formål for datasettet Ansvarlig for at metodikken og standardene for å beskrive dataproduktet følges Definisjon av «key performance indicators» (KPIs) for dataproduktet, f.eks. tid for datakonsumentene å oppdage og benytte dataproduktet. Utvikle og lede et team som skal ha utviklings og datakompetanse for å utvikle og levere dataprodukter fra domenet. Teamet må ha kompetanse på å levere infrastruktur, kode og datasett for dataprodukter som skal brukes i analyse.

²⁸ <https://unit-norge.github.io/unit->

[ra/main/B%C3%B8ker/Referansearkitektur%20for%20deling%20av%20data%20i%20h%C3%B8yere%20utdanning%20og%20forskning.html](https://unit-norge.github.io/unit-ra/main/B%C3%B8ker/Referansearkitektur%20for%20deling%20av%20data%20i%20h%C3%B8yere%20utdanning%20og%20forskning.html)



	Sørge for at deling av opplysninger er i samsvar med loven, med særlig vekt på taushetsplikt og personvern.
Typiske roller/stillinger som rollen skal samarbeide med	Domeneansvarlig, datakonsument, tjenesteansvarlig, dataforvalter, Juridisk forum i KUDAF, Nasjonal ressurscenter for deling av data

Domeneansvarlig

Rolle	Domeneansvarlig
Beskrivelse av rollen	Har ansvar for aktiviteter og tiltak innen domenet for å sikre riktig kvalitet, utnytting og sikring av informasjon i domenet.
Ansvars-områder og oppgaver	Være prosessdriver for informasjonsforvaltning Følge med at begrepene blir utarbeidet etter retningslinjer i domenet. Passe på at forvaltningsprosessen blir fulgt og at begrepene har riktig status i forhold til begrepsforvaltningsprosessen. Ha oversikt over helheten i domenet og bidra til koordinering, harmonisering og godkjenning av innhold, inklusiv samordning av konsumenter med sammenfallende behov og eksisterende begreper i begrepskatalogen. Publisering av begreper i begrepskatalogen Drive opplæring knyttet til forvaltning av informasjon i domenet.
Typiske roller/stillinger som rollen skal samarbeide med	Dataprodukteier, Datakonsument, Tjenesteansvarlig, Dataforvalter

Juridisk ekspert med ansvar for regelverksutvikling

Rolle	Juridisk ekspert med ansvar for regelverksutvikling
Beskrivelse av rollen	Oppfølging og videreutvikling av regelverk knyttet til deling av data slik at lovverket muliggjør kunnskapsbaserte beslutninger og effektiv offentlig forvaltning på tvers av offentlige organisasjoner og sektorer. Sentrale lover som forvaltningsloven, personopplysningsloven og særlovgivning på det enkelte området skal videreutvikles etter behov.
Ansvars-områder og oppgaver	???



Typiske roller/stillinger som rollen skal samarbeide med	Domeneansvarlig, dataprodukteier, datakonsument, tjenesteansvarlig, Juridisk forum i KUDAF, Nasjonalt ressurscenter for deling av data
--	--

Dialogpartnere innen tjenesteproduksjon (utgangspunkt for diskusjonene våre):

Tjenesteeier
Tjenesteansvarlig
Dataforvalter



5.12 Tjenester og Verktøy

For at aktørene i kunnskapssektoren skal settes i stand til å dele sine data er avgjørende at de har de riktige verktøyene tilgjengelig. I tillegg til ledelse, organisatoriske forbedringer og kompetanse (se over), må det også være tilgjengelig tjenester og verktøy som hjelper aktørene i gang. Et av de viktigste virkemidlene som KUDAF og koordinerende aktør skal bidra med er utvikling og tilgjengeliggjøring av tjenester og verktøy som aktørene kan ta i bruk.

Forprosjektet har avdekket behov for noen tjenester og verktøy. Aktørene selv, gjennom felles fora, og i samarbeid med koordinerende aktør, må identifisere nye behov og prioritere hvilke som bør utvikles først.

Tjenester kan realiseres som tjenester hos en tjenesteleverandør, for eksempel en sky-leverandør, eller hos koordinerende aktør.

Verktøy kan realiseres i form av utviklingsmiljø, komponenter eller programvare som aktørene kan ta i bruk i sitt eget miljø, enten det er lokalt hos aktøren eller hos en sky-leverandør.

Forslag til tjenester og verktøy:

Hva	Tjeneste	Verktøy/Komponent	Eksempel/funksjonalitet
Dataprodukt for analyse fra data for tjenesteproduksjon Kunne publisere dataprodukt	bygge dataprodukter, inneholdende API, data, metadata og kode	komponenter til å bygge dataprodukter, inneholdende API, data, metadata og kode	Microtjenester API gateway
Skape forslag til aktive metadata Informasjonsforvaltning	skaper forslag til metadata som kan bearbeides og justeres videre	komponenter som skaper forslag til metadata som kan bearbeides og justeres videre	
Lokal metadatakatalog Metadata-katalog	Metadatakatalog som tjeneste	Programvare eller komponenter installert lokalt for metadatakatalog	begrepsdefinisjoner i domene, koblet mot sentral metadatakatalog
Enkel analyse av data Analyse-kompetanse	tjenester som kan selvbetjenes for å gjøre analyser av egne eller andres data	Programvare eller komponenter installert lokalt for å lage enkle analyser av egne eller andres data	NSD dataplattform SSB DAPLA-skydataplattform
Sikkerhet Identitets- og tilgangsstyring	Tjenester for identitets og tilgangsstyring	Komponenter for identitets og tilgangsstyring	

Tabell 2 Tjenester og verktøy



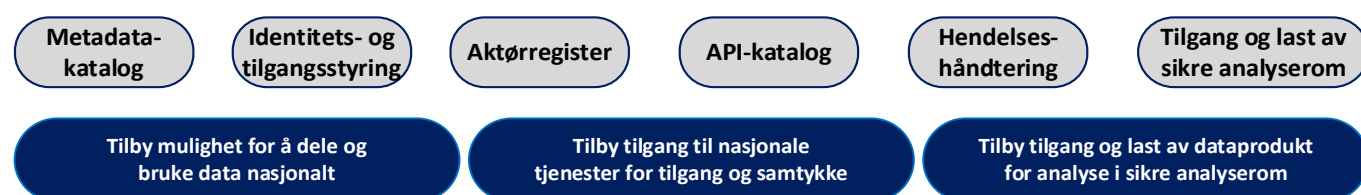
5.13 Komponentene i datadelingsplattformen

Datadelingsplattformen skal i første versjon være basert på datadelingsplattformen som spesifisert i [«Referansearkitektur for deling av data i høyere utdanning og forskning»](https://unit-norge.github.io/unit-ra/main/B%C3%B8ker/Referansearkitektur%20for%20deling%20av%20data%20i%20h%C3%B8yere%20utdanning%20og%20forskning.html)²⁹.

Det vil være nødvendig å videreutvikle og utvide denne for de behov som KUDAF har i et hovedprosjekt/program. Den refererte referansearkitekturen er rettet inn mot behov som er identifisert for Universitet, Høgskoler og Forskning (UHF). Den skal i tillegg dekke utveksling av data for andre, mer operasjonelle, behov enn KUDAF sitt område.

Et spesifikt område som ikke er dekket, er tilgang og lasting av data til Sikre analyserom.

For å oppnå målene for KUDAF vil det være behov for å utvikle komponenter som går ut over det som er spesifisert i dokumentet det er linket til her. En del av disse er nevnt tidligere i dette dokumentet. Det kan forventes at behov for ytterligere komponenter blir avdekket i prosjektet/programmet sitt gang.



²⁹ [https://unit-norge.github.io/unit-](https://unit-norge.github.io/unit-ra/main/B%C3%B8ker/Referansearkitektur%20for%20deling%20av%20data%20i%20h%C3%B8yere%20utdanning%20og%20forskning.html)

[ra/main/B%C3%B8ker/Referansearkitektur%20for%20deling%20av%20data%20i%20h%C3%B8yere%20utdanning%20og%20forskning.html](https://unit-norge.github.io/unit-ra/main/B%C3%B8ker/Referansearkitektur%20for%20deling%20av%20data%20i%20h%C3%B8yere%20utdanning%20og%20forskning.html)

