

Prosjektskisse

Felles økosystem for deling av miljø- og havdata

Utarbeidet av: Havforskningsinstituttet (HI), Stiftelsen NILU (NILU), NORCE og Meteorologisk institutt (MET)

ENDRINGSLOGG

Versjon	Dato	Endring	Godkjent
0.3	20.08.2026	Første versjon	
0.5	02.09.2026	Inkludert innspill fra referansegruppen og etablert kapittel 4.2 og kapittel 6	
0.7	16.09.2026	Inkludert innspill fra styringsgruppen og etablert kapittel 7	
0.8	25.09.2026	Inkludert innspill fra referansegruppen og etablert første fullstendig prosjektskisse	
1.0	30.09.2026	Innarbeidet kommentarer fra styringsgruppen og ferdigstilt prosjektskissen	Styringsgruppen 1.10.2026

DISTRIBUSJONSLOGG

Versjon distribuert	Dato	Navn
0.3	20.08.2026	Referansegruppen
0.7	16.09.2026	Referansegruppen
1.0	30.09.2026	Styringsgruppen

Oppsummering: Felles økosystem for deling av miljø- og havdata

Bakgrunn og formål

Prosjektskissen, utarbeidet av Havforskningsinstituttet (HI), Meteorologisk institutt (MET), Stiftelsen NILU (NILU) og NORCE etter bestilling fra Klima- og miljødepartementet (KLD) og Nærings- og fiskeridepartementet (NFD), konkretiserer etableringen av et nasjonalt, desentralisert dataøkosystem for miljø- og havdata. Formålet er å sikre at data forblir hos autoritative kilder, men blir finnbare, sammenstillbare og gjenbrukbare på tvers av sektorer og forvaltningsnivåer i tråd med EUs Data Space-prinsipper.

Problembeskrivelse

Dagens datalandskap er fragmentert og preget av institusjonelle siloer. Selv om mange data oppfyller FAIR-prinsippene, er de sjelden tilrettelagt for maskinell sammenstilling og tverrfaglig kobling (f.eks. hav–land). Dette medfører betydelig tidsbruk til manuell datahenting og uavklarheter rundt dataeierskap og lisenser og et mer krevende beslutningsgrunnlag for forvaltning, beredskap og næringsliv.

Anbefalt løsning

Det anbefales å etablere et “Dataområde for miljø- og havdata” (Data Space). Dette er vurdert til å ha større nytteverdi framfor nullalternativet og konsolidering og domeneportal-alternativet, og mer realiserbart og med vesentlig lavere risiko og kostnad enn en nasjonal løsning - samling av miljø- og havdata på ett sted. Anbefalt alternativ Dataområde for miljø- og havdata vil benytte felleskomponenter og åpne standarder. Dette betyr standardisert tilkobling av datatilbydere, felles metadatakatalog og maskinlesbare lisenser, standardiserte api'er for søk på og uthenting av data og metadata og data forblir hos datatilbyderne. Dette gir lavest gjennomføringsrisiko, begrenser kostnadene, bevarer datasuvereniteten og entydig maskinell tilgang til data. Alternativet sikrer også kompatibilitet med dataområder tatt fram gjennom blant annet EUs Green Deal Data Space satsing.

Samfunnsmessig nytteverdi

- *Styrket beslutningsgrunnlag*: Gir stat, fylker og kommuner rask, automatisert tilgang til data med kjent kvalitet for arealplanlegging, klimaregnskap, miljøovervåking og akuttberedskap.
- *Bedre utnyttelse av ressurser*: Reduserer tidsbruk og transaksjonskostnader for næringsliv, konsulenter og forvaltningsorganer ved å fjerne manuelle, ressurskrevende sammenstillinger av miljø og havdata.
- *Økt verdiskaping og gjenbruk*: Bedre utnyttelse av offentlig finansierte data samt tilrettelegging for kommersiell tjenesteutvikling og forskning.

Styring og finansiering

Prosjektet foreslås ledet av initiativtakerne med departementene som oppdragsgivere og premissgivere. Realiseringen forutsetter en forutsigbar nasjonal grunnfinansiering av drift og videreutvikling av felleskomponenter samt brukerstøtte på datadeling som metadata og FAIR-prinsipper veiledning. Det anbefales å søke medfinansiering av etableringsprosjektet fra EU-programmer som Digital Europe. Dataområdet vil suksessivt implementere brukerhistorier og det foreslås at det også finnes muligheter for å søke tilskudd til oppkobling av datatilbydere til dataområdet for spesielt samfunnsmessige lønnsomme brukerhistorier.

Prosjektets kostnad og varighet

Etableringsprosjektet vil ha 4 års varighet. Etter 1 år vil det foreligge en *proof of concept* (POC) inkludert en enkel pilot. På det tidspunktet beslutter man om man ønsker å gå videre. Etter ytterligere 2 år vil dataområdet være fullt ut operativt, men med noe redusert kapasitet. Etter 4 år vil dataområdet gå over i en driftsfase, men innrulling av nye datatilbydere og implementering av nye brukerhistorier vil fortsette.

Kostnadene for POC-fasen er anslått til 18 500 000 NOK i 2026 kroner. Dette inkluderer både personal-, reise- og infrastrukturkostnader. Etter POC-fasen oppdateres estimatene for fase 1 og 2 basert på erfaringene og videre drift.

Innhold

Oppsummering: Felles økosystem for deling av miljø- og havdata.....	2
Bakgrunn og formål.....	2
Problembeskrivelse.....	2
Anbefalt løsning.....	2
Samfunnsmessig nytteverdi.....	2
Styring og finansiering.....	3
Prosjektets kostnad og varighet.....	3
Innhold.....	4
Bakgrunn.....	8
Oppdragsbeskrivelse.....	8
Datadeling i Norge.....	8
Sentrale begrepsavklaringer.....	9
Deling av miljø- og havdata.....	11
Prosess.....	12
1 Status.....	12
1.1 Hva gjør private aktører.....	13
1.2 EU og nye datadelingsintitiativer.....	13
1.3 Digitale tvillinger.....	15
2 Problembeskrivelse.....	16
2.1 Hva er utfordringen.....	16
2.1.1 Data kan være FAIR, men likevel ikke KI-klare.....	16
2.1.2 Dataene kan være finnbare (Findable), men ikke alltid gjenbrukbare.....	17
2.1.3 Pålitelighet til datakildene.....	17
2.1.4 Ulike (meta)datastandarder og beskrivelser.....	18
2.1.5 Data suverenitet og lisensiering.....	18
2.1.6 Datakvaliteten er inkonsistent og vanskelig å håndtere.....	18
2.1.7 Finansiering og ressurser.....	19
2.2 Oppsummering.....	19
3 Effekter - gevinster og ulemper.....	21
3.1 Hvordan løser et felles dataøkosystem problemet?.....	21
3.2 Fordeler og ulemper ved et felles økosystem.....	22
3.2.1 Offentlig forvaltning.....	22
3.2.2 Forskning.....	22
3.2.3 Privat næringsliv og tjenesteutvikling.....	23
3.2.4 Andre.....	23
3.3 Oppsummering av fordeler og ulemper for ulike aktører.....	24
3.3.1 Fordeler for brukere.....	24
3.3.2 Ulemper for brukere.....	25
3.4 Brukermatrise og eksempler.....	26
4 Konsepter.....	27

4.1. Nullalternativet.....	29
4.1.1 Dataplan.....	29
4.1.2 Kontrollplan.....	29
4.1.3 Teknologiplan.....	29
4.1.4 Kompetanseoverføring og formidling.....	29
4.1.5 Konsekvenser og måloppnåelse.....	29
4.2 Konsolidering + Miljødataportal.....	31
4.2.1 Dataplan.....	32
4.2.2 Kontrollplan.....	32
4.2.3 Teknologiplan.....	33
4.2.4 Kompetanseoverføring og formidling.....	33
4.2.5 Konsekvenser.....	33
4.3 Nasjonal løsning - samling av miljø- og havdata på ett sted.....	35
4.3.1 Dataplan.....	35
4.3.2 Kontrollplan.....	36
4.3.3 Teknologiplan.....	36
4.3.4 Kompetanseoverføring og formidling.....	36
4.3.5 Konsekvenser.....	36
4.4 Dataområde (data space) for miljø- og havdata.....	37
4.4.1 Dataplan.....	39
4.4.1.1 Sikre grunnleggende interoperabilitet mellom datasett.....	39
4.4.1.2 Metadata-innhøsting - felles metadatakatalog.....	40
4.4.1.3 Kvalitetsbeskrivelse og proveniens.....	40
4.4.2 Kontrollplan.....	41
4.4.2.1 Identitet og tilgangsstyring.....	41
4.4.2.2 Kontrakt/lisenstjeneste.....	42
4.4.2.3 Innrulling og fratredelse av datatilbydere.....	42
4.4.2.4 Fellesregler for deltagelse.....	42
4.4.3 Teknologiplan.....	43
4.4.3.1 KI-vennlighet og maskinlesbarhet.....	43
4.4.3.2 Innsikt og administrasjonsgrensesnitt.....	44
4.4.3.3 Arkitekturprinsipper ut fra et IT-teknisk perspektiv.....	45
4.4.3.4 Markeds plass og verdiskaping.....	46
4.4.4 Kompetanseoverføring og formidling.....	46
4.4.5 Konsekvenser.....	47
4.5 Sammenligning av de vurderte konseptene.....	48
5 Styringsmodell for ferdig løsning.....	51
5.1 Skisse til organisering og styring.....	51
5.2 Anbefalt finansieringsmodell.....	53
6 Forutsetninger.....	54
6.1 Langsiktig og forutsigbar finansiering.....	54
6.2 Åpne data og brukerregisterering.....	54
6.3 Fortsette arbeidet med datakuratering og implementering av FAIR kompatibilitet.....	55
6.4 Åpen kildekode og gjenbrukbare komponenter.....	55

6.5 Eksisterende systemer for tilgang til data.....	55
6.6 Driftsleverandør og suverenitet.....	56
6.7 Maskinlesbarhet og KI-vennlighet.....	56
6.8 Tjenestedesign, brukerorientering og samarbeid.....	56
6.9 Kravstille bruk av dataområdet.....	57
6.10 Kompetanse.....	57
6.11 Eierskap / annet.....	57
6.12 Omfang av dataområdet.....	58
6.13 Sammenstilling av forutsetninger for Dataområdet for miljø- og havdata.....	58
7 Prosjektorganisering, fremdriftsplan og kostnadsoverslag.....	60
7.1 Prosjektorganisering.....	60
7.2 Fremdriftsplan.....	60
7.2.1 Faseoversikt.....	63
7.3 Estimer.....	66
7.4 Mulige finansieringskilder.....	66
8. Synergier med eksisterende dataløsninger.....	69
8.1 Integrering av eksisterende datadelingstjenester.....	69
8.2 Verdiøkning av eksisterende løsninger.....	69
8.3 Økt av synlighet av norske data gjennom samarbeid.....	69
8.4 Gjenbruk.....	70
9. Instituttens anbefaling.....	71
Appendix A. FAIR prinsipper.....	73
Appendix B. Brukerhistorier.....	75
B.1 Climate-Driven Air Quality & Extreme Heat Risk Assessment (CLIM-AIR).....	75
B.1.1. User Story and Core Objective.....	75
B.1.2. Required Data Sources and Domains.....	75
B.1.3. Data Space Interoperability and Technical Workflow.....	76
B.1.4. Expected Output and Societal Value.....	76
B.2 Arctic Sea Ice Dynamics & Safe Marine Operations (ARCTIC-NAV).....	77
B.2.1 User Story and Core Objective.....	77
B.2.2 Required Data Sources and Domains.....	77
B.2.3. Expected Output and Societal / Economic Value.....	78
B.3 High-Latitude Sea Ice Dynamics & Marine Ecosystem Health (OSI-MARINE).....	78
B.3.1. User Story and Core Objective.....	78
B.3.2. Required Data Sources and Domains.....	78
B.3.3. Expected Output and Societal / Economic Value.....	79
B.4 Oslofjorden — næringsavrenning fra jordbruk.....	79
B.4.1 Brukerhistorie og mål.....	80
B.4.2 Nødvendige datakilder og domener.....	80
B.4.3 Hva dataområdet løser.....	81
B.4.4 Tre konsekvenser for konseptet.....	81
B.5 Norges klimagassbalanse – fra målte konsentrasjoner til samfunnsøkonomiske konsekvenser (KLIMA-BALANSE).....	81
B.5.1 Brukerhistorie og mål.....	81

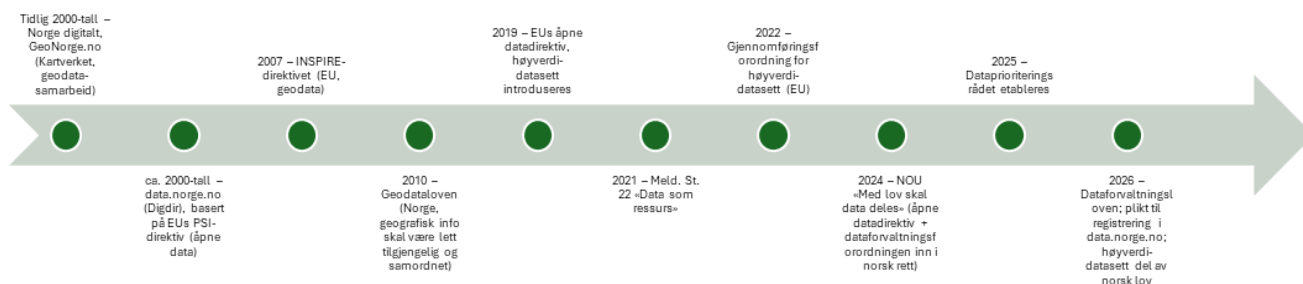
B.5.2. Nødvendige datakilder og domener.....	81
B.5.3 Hva dataområdet løser.....	82
B.5.4. Hva dataområdet ikke løser alene.....	82
B.5.5. Forventet resultat og samfunnsverdi.....	83
B.6 Oslofjorden – fra kilde på land til samlet effekt i fjorden (OSLOFJORD).....	83
B.6.1. Brukerhistorie og mål.....	83
B.6.2. Nødvendige datakilder og domener.....	84
B.6.3. Hva dataområdet løser.....	84
B.6.4. Forventet resultat og samfunnsverdi.....	85
B.7 Arendal Kommune — Forbedring av beslutningsgrunnlaget for kommunal arealforvaltning i marine områder.....	85
B.7.1. Brukerhistorie og mål.....	86
B.7.2. Nødvendige datakilder og domener.....	86
B.7.3 Hva dataområdet kan bidra til å løse.....	87
B.8 Rømt laks.....	87
B.8.1 Brukerhistorie og mål.....	88
B.8.2. Nødvendige datakilder og domener.....	88
B.8.3 Hva dataområdet kan bidra til å løse.....	88
B.9 Fra spredte naturdata til et dokumentert kunnskapsgrunnlag for arealbeslutninger..	89
B.9.1 Brukercasen i kortform.....	89
B.9.1.1. Behovet er dokumentert - og aktuelt.....	89
B.9.1.2. Fra data til beslutning - en konkret brukerreise.....	90
B.9.2 Hvilke data må kobles - og hva tilfører de?.....	90
B.9.3 Land, ferskvann, kyst og hav må kunne ses i sammenheng.....	91
B.9.4 Hva et dataområde må tilføre.....	92
B.9.5. Samme naturgrunnlag - flere anvendelser.....	92
B.9.6 Forventet merverdi.....	93
B.10 Kunnskapsgrunnlag for næringsutvikling i kystsonen – kommunal saksbehandling på tvers av land, vassdrag og fjord.....	93
B.10.1 Brukerhistorie og hovedformål.....	94
B.10.2 Eksempel på datakilder og domener.....	94
B.10.3 Forventet resultat og samfunnsnytte.....	96

Bakgrunn

Oppdragsbeskrivelse

Denne prosjektskissen konkretiserer konseptet for et felles dataøkosystem for miljø- og havdata som NILU, NORCE og MET presenterte for Forskningsrådets porteføljestyre for klima og miljø, i møte med KLD, NFD, KD, HI og Miljødirektoratet 21. januar 2026 og HI sin søknad til Forskningsrådet om finansiering av Norwegian Marine Data Space. Utgangspunktet er et føderert og desentralisert økosystem der data som hovedregel forblir hos den autoritative datatilbyderen, mens felles standarder, katalogfunksjon og regler for identitet og tilgang gjør dataene tilgjengelige på tvers av virksomhetene. Løsningen skal utvikles i tråd med europeiske data space-prinsipper og legge til rette for samhandling med Green Deal Data Space og andre relevante europeiske initiativer.

Datadeling i Norge



Det har i lang tid vært jobbet med å tilgjengeliggjøre data i Norge. Allerede tidlig på 2000-tallet startet samarbeidet Norge Digitalt, et forpliktende samarbeid knyttet til offentliggjøring av kartdata. Med Kartverket i førersetet ble kartrelaterte data samlet og tilgjengeliggjort via [GeoNorge.no](https://geonorge.no). I 2007 ble [Inspire-direktivet](#)¹ vedtatt i EU, og Norge fikk sin egen [Geodatalov](#)² i 2010. Geodataloven innebar en lovbestemmelse om at geografisk informasjon knyttet til Norge skulle være lett tilgjengelig og samordnet. Digdir sitt nettsted data.norge.no startet opp omtrent samtidig. Dette initiativet ble etablert på basis av EUs PSI direktiv om viderebruk av offentlig informasjon. Nettstedet inneholder mange typer data, og ikke kun georefererte data.

Både data.norge.no og geonorge.no er portalløsninger som har bedret tilgjengeligheten til offentlige data. Hele det offentlige Norge skal publisere metadata om deres åpne data i disse tjenestene.

I 2019 kom EU's [åpne datadirektiv](#)³. Her ble konseptet høyverdi-datasett introdusert, og det ble etablert en gjennomføringsforordning i 2022. Høyverdi-datasett er data med spesielt stor

¹ Celexnr.: 32007L0002, INSPIRE-direktivet, https://knowledge-base.inspire.ec.europa.eu/legislation/inspire-directive_en

² LOV-2010-09-03-56, Lov om infrastruktur for geografisk informasjon (geodataloven), <https://lovdata.no/dokument/NL/lov/2010-09-03-56>

³ Celexnr.: 32019L1024, <https://lovdata.no/static/NLX3/32019L1024.pdf>

verdi for samfunnet, miljøet og økonomien. Dette er blant annet geodata, jordobservasjoner og miljø, meteorologi, statistikk i form av demografiske og sentrale økonomiske indikatorer.

I 2021 kom stortingsmeldingen [Data som ressurs](#)⁴, som konkluderer med at det offentlige Norge har store mengder data, og at det her ligger et ubenyttet potensial for både næringsliv, forskning og samfunnet. Brukerne opplevde at det var vanskelig å få tak i data, og dermed også vanskelig å kapitalisere på disse.

I 2024 kom NOU'en [Med lov skal data deles](#)⁵, som utreder mulighetene for hvordan de to EU-lovene "åpne datadirektiv" og "[dataforvaltningsforordningen](#)"⁶ skal inkluderes i norsk lov. Dataprioriteringsrådet ble etablert i 2025 basert på anbefalinger i denne NOU-en. Rådet skal bidra til å få fortgang i tilrettelegging og deling av data med størst samfunnsverdi og potensiale for verdiskapning i Norge, både befolkningen og næringslivet.

Våren 2026 ble [lov om datadeling og dataforvaltning \(dataforvaltningsloven\)](#)⁷ vedtatt, og prinsippet om at offentlige data som hovedregel skal være søkbare, gratis og enkel å finne. Dermed ble det pliktig å registrere alle offentlige data i den nasjonale datakatalogen data.norge.no. Konseptet høyverdi-datasett er nå også formelt en del av norsk lov.

Det pågår et arbeid med å revidere INSPIRE-direktivet. Revideringen skal sikre at direktivet fortsatt er aktuelt, og at det tilpasses EUs åpne datadirektiv, definisjonen av høyverdidatasett og den nye norske dataforvaltningsloven. Det er forventet at EU ferdigstiller revideringen sent i 2026 eller tidlig i 2027, og at Norge deretter tilpasser geodataloven.

Sentrale begrepsavklaringer

Dataforvaltningslovens definisjon av data:

Data: enhver framstilling av handlinger, fakta eller informasjon og enhver sammenstilling av slike handlinger, fakta eller informasjon, også i form av lyd- eller bildeopptak eller audiovisuelle opptak

Videre opererer vi med følgende begreper

Miljø- og havdata: Alle type fagdata, inkludert rådata og produkter, innenfor hav- og klima- og miljøsektoren som miljø-, hav-, klima-, atmosfære- og naturdata, og koblingen mellom disse. Dataene kan ha sin opprinnelse i forvaltning, privat sektor, kommuner eller forskning.

⁴ Meld. St. 22 (2020–2021), Data som ressurs, <https://www.regjeringen.no/no/dokumenter/meld.-st.-22-20202021/id2841118/>

⁵ NOU 2024: 14, Med lov skal data deles, <https://www.regjeringen.no/no/dokumenter/nou-2024-14/id3046520/>

⁶ Cellexnr.: 32022R0868, Dataforvaltningsforordningen (DGA), <https://www.regjeringen.no/no/sub/eos-notatbasen/notatene/2021/juli/datastyringsforordningen-dga-/id2889711/>

⁷ LOV-2026-06-19-36, Lov om datadeling og dataforvaltning (dataforvaltningsloven), <https://lovdata.no/dokument/NL/lov/2026-06-19-36>

Åpne data: Data som alle har tilgang til, og som fritt kan gjenbrukes og deles videre.

Beskyttede data: Data som offentlig sektor eier, men som inneholder informasjon som ikke kan deles fritt uten videre.

Metadata: En beskrivelse av dataene som forklarer hva dataene betyr, hvordan de er samlet inn, hvilke formater og lisenser som gjelder og hvor ofte de oppdateres.

FAIR⁸: Data med veldefinerte metadata som er finnbare (Findable), tilgjengelige (Accesible), samhandlende (Interoperatible) og gjenbrukbare (Reusable)⁹, se Appendix A for detaljert informasjon.

TRUST prinsipper¹⁰: Rammeverk etablert i 2020 for å sikre digitale dataarkiver. Transparens i styring, dataregler, brukerlisenser og betingelser. Ansvarlighet (Responsibility) i form av autentisitet, integritet og etisk forvaltning av data. Prioritering av behov og krav fra brukerne (Users), og gjøre data og tjenester aksesserbare og brukervennlige. Bærekraftig (Sustainable) ved å ha en strategi for langsiktig drift og finansiering av datatjenester og datakuratering. Benytte passende og avansert teknologi for å sørge for sikker og effektiv datadeling og bruk.

Økosystem for deling av data: En samlebetegnelse på et nettverk av aktører og systemer som utveksler data seg imellom basert på felles regler. Begrepet tar ikke hensyn til hvilke regler, standarder eller teknologi som benyttes. I denne prosjektskissen benytter vi samlebetegnelsen økosystem for deling av data konsekvent før vi i kapittel 4 ser på 4 ulike former for økosystemer. Fra og med kapittel 4 benytter vi Dataområde for deling av miljø- og havdata når vi snakker om foreslått konsept. Det blir også referert til samlebetegnelsen økosystem når det er hensiktsmessig.

Dataområde (data space): En type økosystem for deling av data som består av et føderert rammeverk og digital infrastruktur for sikker og kontrollert datadeling på tvers av organisasjoner. Data forblir som hovedregel hos den autoritative datatilbyderen. Felles regler, standarder og tjenester lar brukere og systemer finne og kombinere data på tvers, uten at det etableres en ny sentral database. En dataportal eller en datakatalog gjør data synlige. Et dataområde gjør dem

⁸ Wilkinson, M., Dumontier, M., Aalbersberg, I. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* **3**, 160018 (2016). <https://doi.org/10.1038/sdata.2016.18> og <https://www.go-fair.org/fair-principles/>

⁹ Dataprioriteringsrådet har laget et forslag til norsk oversettelse av FAIR-prinsippene og ambisjonsnivåer for etterlevelse. [Forslaget](#), Veileder ved tilgjengeliggjøring av data fra offentlig sektor, er nå på høring. <https://www.digdir.no/datadeling/horing-veileder-ved-tilgjengeliggjoring-av-data-fra-offentlig-sektor/8166>

¹⁰ Lin, D., Crabtree, J., Dillo, I. *et al.* The TRUST Principles for digital repositories. *Sci Data* **7**, 144 (2020). <https://doi.org/10.1038/s41597-020-0486-7>

anvendbare hos andre, uten at eieren mister kontrollen over deling- og bruksvilkårene.

Datatilbyder: Aktør som tilgjengeliggjør data, enten på egne eller andres vegne.

Datakonsument: Bruker av data.

Dataeier: Aktør som eier dataene og bestemmer lisensiering til data.

Deling av miljø- og havdata

KLD og NFD har ansvaret for store mengder miljø- og havdata, mye av dette har høy samfunnsnytte og er å regne som høyverdi-datasett. To av departementenes underliggende enheter, HI og MET, er både viktige kilder til miljø- og havdata, og har etablert god kompetanse og infrastruktur for datadeling innenfor sine områder.

I 2024 kom [Næringsplan for norske havområder](#)¹¹. Planen fremhever viktigheten av innsamling og deling av relevante data for havet for å skape bærekraftig innovasjon. I likhet med planen peker den europeiske strategien Green Deal også på datadeling som essensielt for å sikre innovasjon og omstilling. I 2025 ble rapporten [Deling av havdata](#)¹² publisert, med fokus på deling mellom private næringsaktører og det offentlige.

Følgende strategier, konsepter, tidligere forskningssøknader og utredninger er også benyttet som bakgrunnsmateriale:

- HI sin søknad til Forskningsrådet om finansiering av Norwegian Marine Data Space
- FAIR-prinsippene
- Rammeverk og komponenter utviklet for Data Spaces i Europa, spesielt [Green Deal Data Space](#)¹³
- [KLDs kunnskapsstrategi 2025-2030](#)¹⁴
- [Fremtidens digitale Norge](#)¹⁵
- [Digitaliseringsstrategi for klima- og miljøsektoren 2025–2030](#)¹⁶
- [Fremtidens miljødata \(Menon\)](#)¹⁷

¹¹ Næringsplan for norske havområder,

<https://www.regjeringen.no/no/dokumenter/naringsplan-for-norske-havomrader/id3046882/>

¹² Deling av havdata. Innhenting og deling av havdata fra næringsaktører.

<https://www.regjeringen.no/no/dokumenter/innhenting-og-deling-av-havdata-fra-naringsaktorer/id3143237/>

¹³ <https://green-deal-dataspace.eu>

¹⁴ Klima- og miljødepartementets kunnskapsstrategi 2025–2030,

<https://www.regjeringen.no/no/dokumenter/klima-og-miljodepartementets-kunnskapsstrategi-20252030/id3097450/>

¹⁵ Fremtidens digitale Norge, Nasjonal digitaliseringsstrategi 2024–2030

<https://www.regjeringen.no/no/dokumenter/fremtidens-digitale-norge/id3054645>

¹⁶ Digitaliseringsstrategi for klima- og miljøsektoren 2025 - 2030

<https://www.regjeringen.no/no/dokumenter/digitaliseringsstrategi-for-klima-og-miljosektoren-2025-2030/id3089710/>

¹⁷ Fremtidens miljødata, Menon Economics og A-2. Menon-publikasjon nr. 153/2021,

<https://menon.no/prosjekter/fremtidens-miljodata>

- Tiltaket om å videreutvikle et effektivt og tilgjengelig miljødatasystem for forskning og forvaltning i Regjeringens langtidsplan for forskning og høyere utdanning 2023–2032

Prosess

HI og MET har fått i oppdrag å ta fram et prosjektforslag knyttet til forbedring av det samlede datadelingslandskapet innenfor miljø- og havdata. Sammen med seg har de to aktører som også har lang erfaring og store mengder miljø- og havdata, NILU og NORCE.

Utkastet til prosjektforslaget skal sendes til høring hos en rekke viktige interessenter innenfor miljø- og havområdet:

- NIVA
- NINA
- Norsk Polarinstitutt
- SINTEF Ocean
- SINTEF Digital
- IFE Halden
- Kartverket
- Miljødirektoratet
- Digitaliseringsdirektoratet
- Norges Forskningsråd

1 Status

Store mengder miljø- og havdata, både historiske data og kontinuerlige målinger, er åpent tilgjengelig og deles i dag via ulike nasjonale og internasjonale plattformer. Selv om forventningen til deling av data har eksistert like lenge for både forskning og forvaltningsdata, har forskningsdata de senere årene i praksis hatt et større krav til å deles og til å tilfredsstille FAIR prinsippene. Ved tildeling av forskningsmidler stilles krav om dataforvaltningsplaner, datadeling og FAIR-ness, blant annet for å kunne reproducere resultatene, og på den måten ettergå forskningen og bygge videre på eksisterende resultater.

Dagens landskap av datadelingsinfrastrukturer for miljø- og havdata er lite standardisert, både nasjonalt og i Europa. Noen organisasjoner har et høyt nivå av modenhet innenfor datadeling. De har gjerne gode, offentlig tilgjengelige grensesnitt eller plattformer for deling av sine data og tjenester, både nasjonalt og internasjonalt. Det finnes mange datakataloger i Norge som tjener ulike formål, ofte levert av datasentre eller datalagre (generalistiske eller domenespesifikke). Disse er ofte koblet til aggregatorer, der datasentre kobler sine ressurser på en maskinlesbar måte til større nasjonale initiativer (f.eks. data.norge.no, Geonorge.no). Autoritative datakataloger og nasjonale aggregatorer kan også understøtte større europeiske eller internasjonale tiltak, som The European Open Science Cloud (EOSC), The Ocean Data and Information System (ODIS), Sustaining Arctic Observing Networks (SAON), International Council for the Exploration of the Sea (ICES), World Meteorological Organization Information System (WIS), European Marine Observation and Data Network

(EMODnet) og mange andre. Større datadelingsgrensesnitt eller plattformer er synlige for flere mennesker, og dermed lettere å oppdage for potensielle brukere.

Noen datasett deles ikke eksternt, og gjenbruk av disse skjer mer tilfeldig, f.eks internt mellom forskere. Dette kan skyldes manglende ressurser til å dele, manglende tillit til hvordan dataene blir brukt, eller generelt manglende motivasjon for å dele.

1.1 Hva gjør private aktører

Alle de store Big Tech firmaene tilbyr gratis lagring av andres data, og tilbyr tilrettelagte plattformer for å analysere disse dataene. Hensikten er ofte å binde brukeren til deres beregnings- og utforskningstjenester. Et eksempel er Google Earth Engine. Her tilbys en stor katalog av satellittbilder og andre geodata, som du enkelt kan integrere i Googles betalbare beregningstjenester. Skal du bruke datasettene utenfor Googles plattform koster det penger, eller det går veldig tregt å laste ned dataene. Norges satellittbilder over regnskogene tilgjengeliggjøres blant annet på denne måten.

Andre private aktører har en annen tilnærming, de ønsker å tjene penger på de dataene de selv samler inn fra sine produkter. Produktene kan være beregnet på konsumermarkedet, men informasjon fra produktet samles inn og aggregeres i større plattformer og tilgang selges. Eksempler for firmaer som selger aggregerte data er leverandører av skipsinformasjonssystemer, bilprodusenter, leverandører av værinformasjon og produsenter av mobiltelefoner. Mange private aktører deler sin innsamlede data med forskningsgrupper og offentlig sektor uten betaling. Noen gir dataene fra seg uten betingelser, men det er ofte et krav om at data kun skal benyttes internt og ikke videredeles. Disse dataene finnes sjeldent i åpne datakataloger.

En tredje forretningsmodell er dataplattformer der kommersielle aktører betaler for bruk, mens forskning og utdanning slipper å betale. I tillegg kan dataprodukter og -tjenester selges som abonnementsløsninger. Et norsk eksempel på dette er HUB Ocean's [Ocean Data Platform](https://app.huboccean.earth)¹⁸, lansert i 2024. HUB Ocean er en stiftelse opprettet i 2019 av Aker Gruppen og World Economic Forum, og samarbeidet er finansiert av industrielle partnerskap, offentlige tilskudd og internasjonalt samarbeid.

1.2 EU og nye datadelingsinitiativer

Innenfor EU er det mange initiativer som bygger desentraliserte, sikre og interoperable digitale økosystemer som legger til rette for en tillitsbasert datadeling på tvers av EU.

[MeteoGate](https://meteogate.eu/)¹⁹ er en løsning utviklet av [EUMETNET](https://www.eumetnet.eu/)²⁰ som retter seg mot meteorologiske data. Her er det etablert en "one-stop-shop" for meteorologiske observasjoner, farevarsler, radardata og klimaprofiler fra de europeiske meteorologiske instituttene.

¹⁸ Ocean Data Platform, <https://app.huboccean.earth>

¹⁹ MeteoGate, <https://meteogate.eu/>

²⁰ EUMETNET, <https://www.eumetnet.eu/>

Det er også en rekke operative datadelingsøkosystemer på plass:

Copernicus Data Space Ecosystem²¹ er et åpent økosystem som gir gratis og umiddelbar tilgang til et bredt spekter av blant annet satellittdata og tjenester fra Copernicus Sentinel-opdragene (Europeiske jordobservasjonsprogrammet) knyttet til land, hav og atmosfære.

Blue-Cloud 2026²² er et europeisk Horizon-initiativ som bygger et samlet europeisk økosystem for hav, sjøer, kyst- og innlandsfarvann. Det utvider sitt første pilotprosjekt og integreres med European Open Science Cloud (EOSC) som en tematisk node, og har som mål å levere kjernedata og -tjenester til den digitale tvillingen av havet (EDITO, se nedenfor).

European Green Deal Data Space (GDDS)²³ er et initiativ fra Europakommisjonen. Det tilbyr et digitalt miljø som er utformet for å legge til rette for sikker og pålitelig deling av miljødata på tvers av EU, med mål om miljø-, biologisk mangfold-, forurensnings- og klimadata som kan støtte målene i EUs grønne giv, en helhetlig strategi for å gjøre EU klimanøytralt innen 2050. Kommisjonen driver frem etableringen av et fullt operativt GDDS gjennom **Digital Europe Programme**²⁴, med mål om å forbedre tilgangen til miljødata i hele EU.

Et av de aller nyeste økosystemene er **SAGE**²⁵ prosjektet, som vil etablere et operativt Green Deal Data Space (GDDS). Prosjektet har lagt til grunn syv brukerhistorier for datadeling og interoperabilitetsbehov. Ut fra disse behovene arbeides det med å identifisere en portefølje av komponenter som til sammen skal sikre denne typen datadeling. SAGE henter bl.a informasjon om datasett fra EUs dataportal²⁶. Usecasen EIRENE ser på luftkvalitet og helse, og henter blant annet data fra ACTRIS.

SAGE er et kortvarig prosjekt. Det er finansiert av EU og planlegges ferdigstilt i 2027/28.

EOSC - European Open Science Cloud²⁷ er et felles europeisk initiativ som gir forskere tilgang til data og tjenester på tvers av landegrenser og fagfelt. Plattformen fungerer som en felles markeds plass med datatjenester fra mange ulike aktører og gir brukerne adgang til å søke i en metadatakatalog over forskningsdata. Katalogen er i praksis en inngangsport til søk på tvers av hundrevis av europeiske metadatakataloger innenfor ulike forskningsfelt.

DSSC - Data Spaces Support Centre²⁸ er et EU-initiativ som utvikler felles retningslinjer, standarder og verktøy man kan ta i bruk for å etablere data spaces. Formålet er å gjøre det enklere for virksomheter og offentlige aktører å dele og gjenbruke data på tvers av sektorer på en sikker, interoperabel og tillitsbasert måte.

²¹ Copernicus Data Space Ecosystem, <https://dataspace.copernicus.eu/>

²² Blue-Cloud 2026, <https://blue-cloud.org/>

²³ European Green Deal Data Space (GDDS), <https://green-deal-dataspace.eu>

²⁴ Digital Europe Programme, <https://commission.europa.eu/funding-and-tenders/find-funding/eu-funding-programmes/digital-europe-programme>

²⁵ SAGE (the Data Space for a Sustainable Green Europe), <https://www.greendealdata.eu>

²⁶ Biodiversity Information System for Europe (BISE), <https://biodiversity.europa.eu>

²⁷ EOSC - European Open Science Cloud, <https://open-science-cloud.ec.europa.eu>

²⁸ DSSC - Data Spaces Support Centre, <https://dssc.eu>

1.3 Digitale tvillinger

Digitale tvillinger er interaktive, virtuelle representasjoner av en fysisk virkelighet. Tvillingene bruker store mengder data innenfor sitt fagområde for å kunne gjenspeile sine “fysiske tvillinger”. Ved hjelp av KI og avanserte modeller kan brukeren (f.eks forskere og beslutningstagere) simulere aktuelle eller potensielle fremtidsscenarier innenfor sitt fagområde. Tvillingene produserer store mengder data i form av scenarier og prognoser.

Eksempler på digitale tvillinger innenfor fagområdet miljø og hav:

Digital tvilling av havet (EDITO)²⁹ er en interaktiv plattform som muliggjør simuleringer av klimaendringer i havet og forurensing. Hensikten er å forstå, beskytte og forvalte havområdene våre på en bærekraftig måte. EDITO er infrastrukturen som fasiliteterer den digitale tvillingen av havet. Den bruker store europeiske dataaggregatorer som EMODnet, Copernicus, og andre pan-europeiske infrastrukturer som hovedkilde av data.

Den digitale modellen av jorden (Destination Earth - DestinE)³⁰: Interaktiv plattform finansiert av EU-kommisjonen og utviklet av ESA, EUMETSAT og ECMWF. Plattformen muliggjør simulering av atmosfæren, interaksjon med andre deler av jordsystemet, effekter på andre sektorer (som jordbruk, helse, transport osv) og interaksjon med menneskelig aktivitet. Hensikten er å overvåke og forutse klimaendringer, miljøendringer og ekstremvær og gi beslutningstagere og andre støtte til å håndtere hendelser og langsiktige endringer.

Det finnes også andre digitale tvillinger innenfor andre fagområder som benytter store mengder miljø- og havdata. Et eksempel er:

Digitale tvillinger av lokale urbane samfunn (EU Local digital twins LDT4SSC)³¹. Lokale digitale tvillinger (LDT) har som mål å tilby en digital kopi av europeiske byer for å optimalisere planleggingen av bymiljøer. Slike digitale tvillinger bruker data hentet fra dataøkosystemer for å konstruere en virtuell kopi av et fysisk samfunn. De lokale tvillingene benytter **European Data Space for Smart Communities** (DS4SSCCO)³² som datakilde. Dette gjør det mulig for byer å koble digitale tvillinger for urban mobilitet med sanntids, tverrdomenbaserte samfunnsdata (f.eks. energi, miljø) for prediktiv scenariomodellering. DS4SSCC tilbyr sikker og standardisert datadeling, felles standarder, datamodeller og juridiske kjøreregel for personvern. Begge er finansiert gjennom EU-programmet Digital Europe.

²⁹ EDITO - European Digital Twin Ocean, <https://www.edito.eu>

³⁰ Destination Earth - DestinE, <https://destination-earth.eu>

³¹ LDT4SSC, <https://ldt4ssc.eu>

³² DS4SSCC, <https://www.ds4sscc.eu>

2 Problembeskrivelse

Som vist i forrige kapittel er dagens datadelingslandskap svært divergent, med mange ulike løsninger og aktører. Resultatet er at det brukes mye tid på sammenstilling av data i både privat næringsliv og i forvaltningen, og at arbeidet med å etablere beslutningsunderlag for forvaltningen og for håndtering av kriser og beredskap er vanskelig. Forvaltningen og næringslivets rapporteringskrav er ofte manuelle oppgaver og grunnlagsdata brukes sannsynligvis forskjellig fra aktør til aktør. Det er også et potensial for å gå glipp av kunnskap da vi ikke klarer å effektivt sammenstille data på tvers av fagdomener.

2.1 Hva er utfordringen

Landskapet av datadelingsinitiativer som har som mål å harmonisere tilgang og bruk av miljø- og havdata, er som beskrevet i kapittel 1 omfattende. Innsatsen er ofte begrenset til enkelte datatilbydere, hvor dataene primært tjener deres oppgaver og interesseområder. Dette kan være et fagområde, eller enda mer begrenset, et konsortium eller enkelt forskningsprosjekt. Det er også blitt etablert distribuerte datastyringsrammeverk der deltakerne blir enige om visse regler for interoperabilitet for å dele og tilgjengeliggjøre data til flere fagområder og brukergrupper. Disse initiativene bygger ofte på prinsipper om åpenhet og beste praksis, hvor hovedforpliktelsen er begrenset til å tilby noenlunde interoperable metadata. Utfordringene relatert til dataenes reelle interoperabilitet blir sjelden tatt i betraktning.

2.1.1 Data kan være FAIR, men likevel ikke KI-klare

Mye av dagens data er ikke klare for KI-bruk. Konseptet FAIR-data ble formelt introdusert i 2016 (se Appendix A). Prinsippet om FAIR-kompatibilitet er et viktig fundament for all datadrevet utvikling og mange av de eksisterende dataforvaltnings- og publiseringsrammeverkene som finnes i dag er basert på FAIR-prinsippene, spesielt innen forskning. Selv om FAIR-prinsippene fortsatt er et gyldig og uunnværlig grunnlag for datadeling, er det tydelig behov for ytterligere normalisert tilnærming til både deling og konsumering av data, for at flere aktører skal kunne dele mer data.

I takt med økende bruk av KI, må det også etableres mer formelle krav for å gjøre data KI-klare. Det innebærer at informasjonen som KI-modeller får tilgang til og tolker, må være tydelig, kontekstualisert, godt strukturert og godt forvaltet. Kravene for å tilrettelegge bruk av KI kan variere, men kjernekravene inkluderer:

- Maskinlesbarhet og brukervennlighet
- Konsistent bruk av semantikk
- Bruk av veldefinerte standarder for datarepresentasjon
- Lenke til beskrivelser for å knytte aktuelle data til standard definisjonene
- Mulighet for å identifisere mangler i metadata og håndtere manglende verdier

Bruks- og tilgangsprinsipper må være på plass for å sikre at dataene ikke blir misbrukt. Sporbarhet, opphavsinformasjon og endringshistorikk er viktig for å sikre at dataene ikke er kompromittert. Disse kravene er viktige bidrag til etisk bruk av KI.

2.1.2 Dataene kan være finnbare (Findable), men ikke alltid gjenbrukbare

Selv om mye har bedret seg innenfor datadelingområdet, er det fortsatt mange uløste utfordringer med både sentraliserte og distribuerte data rammeverk. Det å finne dataene er sannsynligvis den mest og best håndterte utfordringen blant datasentere og infrastrukturer som tilbyr publisering og datadeling. Mange typer data er søkbare innenfor sine spesifikke datakataloger, gjerne gjennom metadataaggregatorer og også noen ganger gjennom generiske web-baserte søk. Generiske web-baserte søk eksponerer aktuelle data direkte i søkemotorer. Dette er kun mulig ved riktig koding av metadata.

Å finne riktig data kan være et vanskelig og tidkrevende arbeid, og selv om data blir funnet er det ofte ikke klart hvordan og om dataene kan bli brukt. Dette kan skyldes ustrukturert informasjon om tilgangsrettigheter, manglende dokumentasjon av kvalitet, usikkerhet, skala, oppdateringsfrekvens, lisens, metode og provenans, og andre begrensninger. Denne typen informasjon er svært sjeldent maskinlesbart. Mangler i metadata er heller ikke uvanlig, og uten fyldig metadata og klar semantikk er det ikke mulig å prosessere dataene uten menneskelig innblanding. I slike tilfeller kan ikke dataene integreres automatisert i beslutningssystemer, KI-baserte prosesser eller i menneskelige vurderinger i forvaltningsbeslutninger.

For å kunne gå fra rene data til verdidøkende tjenester må dataene tilfredsstille visse krav både med hensyn til protokoller for aksessering, og kvaliteten på innholdet i dataene. Slike verdidøkende tjenester kan være alt fra enkle visualiseringsverktøy til mer avanserte applikasjoner og beregningsmodeller. Noen ganger tilbyr ikke disse tilgang til metadatainformasjonen i det hele tatt. Likeledes hender det at aksesseringspunktet ikke opererer i henhold til veldefinerte aksessprotokoller. Det er dermed ikke mulig å forstå om aksesseringspunktet tilbyr nedlastning, om det er en html-side eller om dataene strømmes. Det oppstår da situasjoner at selv om dataene er lesbare, vil de ikke gi noen mening for konsumenten. Selv i de tilfellene hvor standardiserte mekanismer for å aksessere data finnes, er det ikke uvanlig å komme over ikke-standardiserte API-er eller usikre protokoller som overfører data uten engang å tilby kryptering. Mangler i metadata, inkludert endringshistorikk, provenans og sitering i tillegg til manglende eller dårlig standardisering, mangel på kvalitetsinformasjon og konsistente beskrivelser kan medføre store problemer når data fra flere forskjellige kilder trenger å bli integrert i større verdikjeder og prosesser.

2.1.3 Pålitelighet til datakildene

Troverdigheten til datakildene er et viktig aspekt som er i ferd med å bli mer og mer aktuelt . Antall tilbydere av data som er tilgjengelig via internett er overveldende, og det er blitt essensielt å kunne gjenkjenne de originale datakildene, og til å kunne evaluere troverdigheten til dataene som blir tilbudt. Om dataene flyttes rundt på tvers av plattformer og ingen formelt garanterer for integriteten, kan dataene bli endret, miste sin sporbarhet og dermed også sin troverdighet. Hovedårsaken til dette er en manglende klar definisjon og håndtering av datasuverenitet og reguleringer, det vil si manglende evne til å kontrollere, administrere og bestemme over egne data.

TRUST-prinsippene definerer viktige grunnkomponenter som dataarkiver bør ivareta og som sikrer god datadeling og dataforvaltning over tid. På samme måte som det å opprettholde FAIR data over tid, trenger vi trygge digitale dataarkiver med klar styring, stabil og pålitelig infrastruktur, og i tillegg et sett av regler for å supportere datatilbydere og datakonsumentene. Selv om mange av dataarkivene har offisielle sertifikater som CoreTrustSeal, nestorSeal, International Oceanographic Data and Information Exchange Quality Management Framework (IODE-QMF) eller ISO 16363, er fortsatt mange dataarkiver ikke sertifisert, og mangler også dokumentasjon for hvordan data blir forvaltet over tid, noe som påvirker påliteligheten.

2.1.4 Ulike (meta)datastandarder og beskrivelser

Det finnes mange forskjellige metadatastandarder. De er ofte skreddersydd for sitt domenes spesielle behov, og både bruken og mengde informasjon varierer mye mellom standardene. Dette kan være problematisk for semantisk mapping og interoperabilitet. Metadatastandarder tillater ofte bruk av fritekst, og selv om standardene tilbyr bruk av formelle ontologier og vokabularer er disse ikke nødvendigvis utfyllt fra tilbyder, og dermed vil forståelsen av dataene mangle. Dette er en av de største utfordringene når man forsøker å finne og kombinere data fra forskjellige kilder.

Det er fortsatt mange utfordringer på metadatanivå, men det er ennå flere når man begynner å se på selve dataene. Dette skyldes blant annet mangel på klare regler for datakoding med tilhørende veiledninger. Mye data inneholder også arv fra tiden da produsentene i stor grad produserte data for å ivareta en dedikert brukergruppe, og når brukerne ofte er andre forskere eller noen velkjente tjenester.

Mange data produsenter kjenner ikke til beste praksis for dataforvaltning. De har sjeldent muligheter til å få hjelp i egen organisasjon. Det medfører som oftest dårlig dataforvaltning gjennom hele dataenes levetid. Manglende kompetanse og underestimert arbeid med å forvalte data sammen med manglende opplæring og støtte i dataforvaltningsprosesser fra dataarkivene er en utfordring.

2.1.5 Data suverenitet og lisensiering

Fra en datatilbyders synsvinkel er det også spesifikke utfordringer som hindrer aktiv deltagelse i datadelingsaktiviteter. I et ikke-styrt delingsøkosystem, kan organisasjonene være bekymret over å miste kontroll på dataene de produserer. De vet ikke om dataene blir kopiert, modifisert eller videredelt og heller ikke til hvem og hvordan delingen blir gjennomført. Denne bekymringen vil fortsatt være til stede selv om en datalisens blir lagt ved dataene. Det er en generell mangel på tillit til ukjente aktører. I dagens raskt endrede datadelingslandskap, og spesielt nå med inntoget av nye teknologier, flyter data enkelt mellom disipliner og muligheten til mistolking av data er høy. Dataprodusentene er bekymret for at mistolking av deres data medfører økt risiko for samfunnet.

2.1.6 Datakvaliteten er inkonsistent og vanskelig å håndtere

Mange av utfordringene som er nevnt ovenfor kan til en viss grad håndteres ved å sørge for standardiserte, fylldige, maskinlesbare og anvendbare metadata. Dette krever at dataprodusentene og arkivene har godt etablerte arbeidsflyter for produksjon av data, der

både standardisering og datakuratering inngår som faste elementer. Dessverre er det ikke alle som har gode prosesser for kuratering. Dette fører til at det kan være krevende, eller i verste fall umulig, å vurdere kvaliteten og innholdet i dataene som publiseres.

Utfordringene forsterkes av at prosedyrer for sikring av datakvalitet ikke er harmonisert, hverken i form av standarder eller beskrivelser på tvers av ulike fagfelt. Det er også viktig å være klar over at data med god kvalitet ikke automatisk er egnet for alle typer applikasjoner eller analyser. Alle disse detaljene er kritiske for å sikre korrekt tolkning, gjenbruk og kombinasjon av datasett. Hvordan man på en effektiv måte skal presentere denne informasjonen, er imidlertid ingen enkel oppgave. Mangelfull håndtering av denne utfordringen kan skape usikkerhet hos konsumentene. I et større perspektiv kan dette også få konsekvenser for hvordan dataprodusentene blir oppfattet. En systematisk tilnærming til beskrivelse av datakvalitet er derfor helt nødvendig for å øke tilliten til de tilgjengelige datakildene.

2.1.7 Finansiering og ressurser

De fleste datadelingsinitiativer er avhengige av nasjonal offentlig finansiering, kortvarige forskningsprosjekter eller EU-tilskudd. Mange slike ferdig etablerte delingsinitiativer er vanskelige å opprettholde i et langsiktig perspektiv. Mangel på robust finansiering av datadeling kan føre til underutnyttelse av dagens ressurser. Det kan blant annet føre til dobbeltarbeid mellom delvis overlappende tjenester levert av ulike prosjekter, og utfordringer med å utvikle merverdi for samfunnet. I tillegg kan datadelingsplattformer ha svært ulikt fokus, for eksempel avhengig av om de er offentlige eller private, og de er ofte bygget på ulike og ikke-overlappende arkitektur.

2.2 Oppsummering

De viktigste utfordringene som må løses for å muliggjøre effektiv datadeling innenfor miljø- og havdata, og som dermed vil muliggjøre mer innovasjon og kapitalisering på allerede eksisterende informasjon er:

- **Datasuverenitet og kontroll:** Behovet for å tydelig definere eierskap til data og tjenester, samt hvordan disse skal deles på en trygg måte.
- **Tillit og åpenhet:** Etablering av felles mekanismer for å verifisere datakilder (produsenter) og brukere, garantere dataintegritet, og sikre at tilgangskrav blir etterfulgt.
- **Interoperabilitet:** Å adressere barrierer på teknisk nivå (systemer/formater), semantisk nivå (terminologi/metadata) og juridisk nivå (lisensiering og reguleringer) for å muliggjøre sømløs integrasjon og bistå til å kunne sette sammen data på tvers av domener.
- **Kompetanse på dataforvaltning hos dataprodusentene:** Å sikre opplæring og støtte til dataprodusentene vil adressere utfordringene med lav kvalitet på metadata og data.

- **Støtte brukerbehov og etablering av verdiøkende tjenester:** Sikre at det eksisterer en tilstrekkelig databeskrivelse og aksesspunkt, slik at brukeren kan vurdere om dataene er egnet for sitt formål.

Ved å løse disse utfordringene vil samfunnet få tilgjengeliggjort et enormt spekter av miljø- og havdata som kan brukes til verdiskaping.

3 Effekter - gevinster og ulemper

Gevinstene i dette kapitlet er vurdert kvalitativt. Verdien av bedre datadeling avhenger av hvor mange som deltar og hvilke tjenester som blir bygget, og et tallfestet gevinstanslag vil derfor bli misvisende på dette stadiet. En kvantifisering bør gjøres når første fase av implementeringsprosjektet har gitt erfaringstall.

3.1 Hvordan løser et felles dataøkosystem problemet?

Et felles dataøkosystem er et rammeverk og en digital infrastruktur som gjør det mulig for forskning, forvaltning og næringsliv å dele, forvalte, finne, koble og bruke data trygt på tvers av organisasjonsgrenser. Norges felles økosystem for deling av miljø- og havdata skal bygge på EUs [dataspace-prinsipper](#)³³, tilpasset norske behov, og legge til rette for samhandling uten at én aktør eier eller kontrollerer hele infrastrukturen. Løsningen skal sikre norsk datasuverenitet, tillitsbasert datautveksling, tilgang til et bredere kunnskapsgrunnlag med bedre støtte for tverrfaglig datakobling som f.eks. hav-land koblinger som kan inngå i dokumenterte, reproducerbare dataflyter, samtidig som den muliggjør sømløs kobling til europeiske data- og KI-initiativer. Eksempler på brukerhistorier er beskrevet i Appendix B. Den skal adressere dagens hovedutfordringer ved å:

- Gjøre data synlige og søkbare på tvers av domener, slik at mindre tid går med på å finne riktig datasett fra en veldokumentert sikker kilde
- Etablere felles metadata-, vokabular- og provenansforståelse, så opprinnelse, prosessering og usikkerhet er kjent
- Sikre interoperabilitet og felles lisensforståelse, slik at datasett kan kombineres uten proprietære hindringer
- Gi autoritative kilder (man vet hvem som tilbyr dataene, og at de er ubehandlet fra kilden)
- Koble Norges miljø- og havdata mot europeiske initiativ og samtidig sikre norsk datasuverenitet
- Gi KI-agenter og andre tilgang til data med kjent kvalitet og opprinnelse

Hvor gode KI-modeller for klima, miljø og hav kan bli, avhenger av tilgangen til sammenkoblede data. Det finnes gode data på hvert enkelt område i dag, men de kan ikke kombineres automatisk og modellene blir dermed smalere enn kunnskapsgrunnlaget skulle tilsi.

Med et felles økosystem kan modeller enklere trenes og kjøres på observasjoner fra atmosfære, hav og miljø samtidig, med kjent kvalitet og opprinnelse. Det gir forvaltningen et mer helhetlig kunnskapsgrunnlag. Data fra flere kilder kan også kobles automatisk i sanntid, for eksempel ved skogbrann, akutt forurensning eller ekstremvær, uten at tilretteleggingen må gjøres på nytt for hver hendelse. For næringslivet betyr dokumentert kvalitet og maskinlesbare bruksvilkår at tjenester kan bygges på data uten at tilgang forhandles datasett for datasett.

³³ Common European data spaces, <https://digital-strategy.ec.europa.eu/en/policies/data-spaces>

Økosystemet skal ikke være en KI-plattform. Økosystemet skal eksponere data, metadata, kvalitetsinformasjon, proveniens og lisenser standardisert og maskinlesbart, slik at eksisterende og framtidige KI-løsninger og andre datakonsumenter kan bruke dem. Europakommisjonens [Data Union Strategy](#)³⁴ fra november 2025 peker i samme retning, med mer og bedre data for KI, forenkling av regelverket og europeisk datasuverenitet som hovedsøyler.

3.2 Fordeler og ulemper ved et felles økosystem

I de neste avsnittene gjennomgås de ulike aktørene "Offentlig forvaltning", "Forskning", "Privat næringsliv og tjenesteutvikling" og "Andre". I en oppsummerende tabell presenteres fordeler og ulemper på tvers av aktørene.

3.2.1 Offentlig forvaltning

For offentlig forvaltning gir et felles dataøkosystem et bredere og bedre dokumentert kunnskapsgrunnlag, tilpasset behovene på de ulike forvaltningsnivåene. For staten styrker det grunnlaget for nasjonal styring, regelverksutvikling og rapportering til internasjonale avtaler og EU. For fylkeskommunen blir det enklere å samordne areal- og klimatilpasningsarbeid på tvers av kommunegrenser. For kommunene gir det raskere tilgang til oppdaterte data i arealplanlegging, beredskap og håndtering av hendelser som flom og skogbrann – i mange tilfeller automatisk. Alle tre nivåer får bedre uttelling for de dataene de selv finansierer, siden dataene blir enklere å finne, kombinere og gjenbruke på tvers av forvaltningsnivåene.

Samtidig innebærer en slik satsing forutsetninger som må håndteres på alle nivåer. Etablering og drift krever investeringer og forutsigbar finansiering, og den fulle verdien nås først når løsningen har tilstrekkelig oppslutning fra stat, fylke og kommune. Felles styring og regelverk gjør beslutninger mer sammensatte og kan ta noe tid å forankre på tvers av nivåene, og økt etterspørsel etter data kan gi mer arbeid hos dem som leverer dataene lokalt.

3.2.2 Forskning

For forskningen ligger den største verdien i at data lettere kan settes sammen på tvers av fagfelt og mellom ulike dataprodusenter, samt at nye dataserier blir tilgjengelig. Det åpner for nye forskningsområder og ny kunnskap som i dag kan være vanskelig å få til, samtidig som nasjonale forskningsdata blir tatt vare på, dokumentert og gjort tilgjengelige på et varig sted. Et felles økosystem styrker også koblingen mot europeiske initiativer som de felles europeiske dataområdene – blant annet Green Deal Data Space og EOSC. For den enkelte dataprodusent fungerer det som et utstillingsvindu som gir synlighet og anerkjennelse, og som gjør det mulig å se hvordan dataene faktisk brukes. Også spesialiserte og prosjektbaserte datasett som ellers forblir lite kjent, blir lettere å finne. Dette sparer tid, og egne data kan dessuten bli en ny inntektskilde.

På den andre siden vil noen forskningsmiljøer kunne oppleve et felles økosystem som en konkurrent til egne løsninger for datadeling, og være bekymret for at det svekker både tidligere investeringer og pågående søknader på feltet. Deltakelse kan kreve at datasett

³⁴ Celexnr.: 52025DC0835, European Data Union Strategy, <https://digital-strategy.ec.europa.eu/en/policies/data-union>

tilrettelegges og beskrives etter felles minstekrav, noe som medfører noe arbeid og en viss startkostnad, og driften vil ofte forutsette egenfinansiering. Dette er reelle hensyn, men de oppveies av at dataene blir mer synlige, mer brukt og lettere å bygge videre på – noe som over tid kommer det enkelte miljøet til gode.

3.2.3 Privat næringsliv og tjenesteutvikling

For privat næringsliv og tjenesteutviklere gjør et felles dataøkosystem det enklere å finne data som kan inngå i egne produkter og tjenester. Fordi dataene har kjent opprinnelse og kvalitet, gir de et pålitelig grunnlag å bygge på. Den europeiske koblingen gjør at norske aktører lettere kan knytte seg til større markeder. På sikt kan en felles markeds plass også åpne for at data kan kjøpes og selges, slik at det oppstår nye forretningsmuligheter rundt dataene.

For mange næringslivsklynger gir et felles økosystem for deling av data en vesentlig forenkling i implementering av gjensidig datadeling for felles nytte innenfor klyngen. Et eksempel på frivillig datadeling mellom konkurrenter, andre finansaktører og det offentlige er [DSOP-samarbeidet](#)³⁵ i finansnæringen. Der deler aktørene blant annet informasjon om metoder for avdekking av økonomisk kriminalitet. Et annet eksempel på slike datadelingssamarbeid er Søkeldirektoratets [DISKOS](#)³⁶, et felles nasjonalt datalager for lete- og utvinningsrelatert informasjon fra norsk sokkel, som er delt mellom styresmakter og oljeselskap.

Som for de andre aktørene er det en viss terskel og startkostnad ved å delta, og tjenestene man utvikler vil være avhengige av at den underliggende løsningen fungerer som den skal. Felles regler for blant annet salg og bruk, og kravet om å identifisere seg som databruker, kan oppleves som ekstra hindre for datatilbydere. Enkelte kan være bekymret for å miste kontrollen over egne dataleveranser eller få konkurranse til egne plattformer. Erfaringsmessig er dette håndterbart, og for de fleste vil tilgangen til nye data og nye markeder veie tyngre enn ulempene.

3.2.4 Andre

For alle andre – fra enkeltpersoner og virksomheter til KI-løsninger – er den store gevinsten at man får én felles inngang til oversikten over hvilke miljø- og havdata som finnes i Norge. KI-agenter kan hente ut data med kjent opprinnelse og kvalitet og bygge videre på dem, dataprodusenter får ett sted å gjøre dataene sine kjent, og det sikrer at norske data, som ofte er finansiert med offentlige midler, er godt beskrevet og forståelige for dem som skal bruke dem. For at dette skal fungere godt, må økosystemet være enkelt å orientere seg i, slik at den ikke i stedet oppleves som uoversiktlig. Ingen enkelt løsning vil dekke absolutt alle behov, og det finnes en risiko for at det vokser fram for mange parallelle dataøkosystemer i Norge. Nettopp derfor er det viktig å samle kreftene om én felles satsing og å innfri det man lover, slik at brukerne får tillit til løsningen.

³⁵ DSOP-Digitalt samarbeid offentlig og privat, FinansNorge, <https://www.finansnorge.no/tema/digitalisering/dsop/>

³⁶ Diskos National Data Repository (NDR), Søkeldirektoratet, <https://www.diskos.no>

3.3 Oppsummering av fordeler og ulemper for ulike aktører

Tabellene oppsummerer hvert poeng på tvers av aktørene. Et kryss viser hvilke grupper poenget er mest relevant for. Flere kryss er mulig, men få poeng gjelder alle fire aktører.

3.3.1 Fordeler for brukere

Beskrivelse	Forskning	Forvaltning (stat/fylke/kommune)	Privat	Andre
Bredere og bedre dokumentert datagrunnlag: data fra flere autoritative kilder kan brukes sammen med maskinlesbar informasjon om kvalitet, opphav og usikkerhet, slik at brukeren kan vurdere om dataene er egnet til formålet	X	X		
Økt kvalitet på dataene som følge av flere brukere, økt tverrfaglig bruk og bredere og flere tilbakemeldinger	X	X		X
Én felles inngang gjør det lettere å finne riktige data og bruke mindre tid på å lete, i stedet for å forholde seg til mange ulike kilder med hver sin tilgangsmåte	X	X	X	X
Data kan enklere settes sammen og kombineres på tvers av fagområder og dataprodusenter, noe som åpner for nye innsikter og forskningsområder som i dag er vanskelige å få til	X		X	
Data med kjent opprinnelse og kvalitet gir et pålitelig grunnlag for å utvikle tjenester og for at KI-løsninger kan bygge videre på dem	X	X	X	X
Bedre utnyttelse av data som det offentlige og forskningen allerede finansierer	X	X		
Mer målrettede målinger, som på sikt kan redusere behovet for et stort antall målestasjoner og dermed kostnadene		X		
Et mer oppdatert og helhetlig situasjonsbilde som styrker varsling og håndtering av hendelser, i mange tilfeller automatisk		X	X	
Gjør det mulig å dele data på en kontrollert måte selv om de ikke kan gjøres helt åpent tilgjengelig	X	X		
Trygg og varig forvaltning av nasjonale data hos autoritative kilder med en felles inngang si gjør dataene søkbare på tvers	X	X		
Økt synlighet for spesialiserte og prosjektbaserte dataprodukter som ellers forblir lite kjent; fungerer som et utstillingsvindu med anerkjennelse og innsikt i hvordan dataene brukes	X			X
Innsikt i hvem som faktisk bruker dataene	X	X		

Bedre kobling mot de europeiske dataområdene (bl.a. Green Deal Data Space, Copernicus og EOSC), der norske data kan bidra	X	X	X	
Mulighet for kommersialisering: data kan kjøpes og selges, slik at egne data kan bli en ny inntektskilde	X		X	
KI-agenter kan hente ut data med kjent opprinnelse og kvalitet og bygge videre på dem			X	X
Sikrer at norske data, som ofte er finansiert med offentlige midler, er godt beskrevet og forståelige for dem som skal bruke dem	X	X		X

3.3.2 Ulemper for brukere

Beskrivelse	Forskning	Forvaltning	Privat	Andre
Etablering og drift krever betydelige og forutsigbare investeringer, og den fulle verdien nås først når løsningen har et tilstrekkelig omfang	X	X		
Det er en terskel og en startkostnad ved å bli datatilbyder, og prosessen med å bli med kan oppleves som byråkratisk og tung	X		X	
Felles styring gjør beslutningene mer sammensatte; de krever enighet mellom mange aktører og kan derfor ta tid	X	X	X	
Oppfølging av lover, regler og tekniske løsninger blir mer komplekst å ivareta over tid		X		
Når data blir lettere å kombinere, stilles det tydeligere krav til sikkerhet og til å vite hvem som leverer og bruker dem	X	X		
Kravet om å identifisere seg som datatilbyder eller databruker kan øke terskelen for å delta	X		X	
Datasett og metadata må ofte tilrettelegges og beskrives etter felles minstekrav, noe som medfører arbeid	X		X	
Tjenester og bruk avhenger av at den felles løsningen og samspillet mellom systemer fungerer som det skal	X		X	
Noen kan oppleve et felles økosystem som en konkurrent til egne løsninger for datadeling, og være bekymret for at det svekker investeringer i disse	X		X	
Usikkerhet om felles regler for kommersialisering, og om balansen mellom samfunnsnytte og økonomisk nytte	X		X	
Når flere finner dataene, kan etterspørselen gi mer arbeid og økt press lokalt hos dem som leverer dataene	X	X		
Hvis løsningen ikke er enkel å orientere seg i, kan den oppleves som uoversiktlig for sluttbrukerne				X

3.4 Brukermatrise og eksempler

Matrisen oppsummerer typiske brukere, behov, gevinster og konkrete dataeksempler per aktørgruppe.

Aktør	Typisk bruker	Behov	Gevinst	Dataeksempel
Offentlig forvaltning (stat, fylke og kommune)	Saksbehandler (stat), regional planlegger (fylke), kommunal saksbehandler/beredskapsansvarlig (kommune)	Sikkert og bredt beslutningsgrunnlag fra autoriserte kilder – til nasjonal styring og rapportering (stat), regional samordning (fylke) og lokal planlegging/beredskap (kommune)	Sikrere kunnskap og raskere beslutninger; bedre utnyttelse av finansierte data	Nasjonal rapportering på miljøindikatorer (stat); regional klima- og arealplanlegging på tvers av kommuner (fylke); sanntidsbasert flom- og skogbrannberedskap (kommune)
Forskning	Forsker eller dataprodusent ved et institutt	Finne og kombinere data på tvers; varig og dokumentert lagring	Ny kunnskap, synlighet og sitering, kobling til EU-initiativer	NORCE og HI kombinerer havforsuring- og temperaturdata med langtidsserier på nitrogenavsetning fra EBAS ³⁷ for å studere økosystemeffekter i kystsonen
Privat næringsliv og tjenestutvikling	Tjenestutvikler eller dataanalytiker i en bedrift	Pålitelige data med kjent kvalitet til egne tjenester	Raskere produktutvikling og tilgang til en marketplace	Rederier bruker METs vind- og bølgevarsler sammen med HIs strømdata til ruteoptimalisering og drivstoffbesparelse
Andre	Privatperson, journalist eller KI-agent	Oversikt over hvilke data som finnes; maskinlesbar tilgang	Åpen innsikt; KI med kjent opprinnelse og kvalitet	En KI-agent kombinerer METs varsel om ekstremvær med luftkvalitetsdata fra EBAS og Miljødirektoratets luftkvalitetsdatabase for å generere automatiske helseråd til sårbare grupper

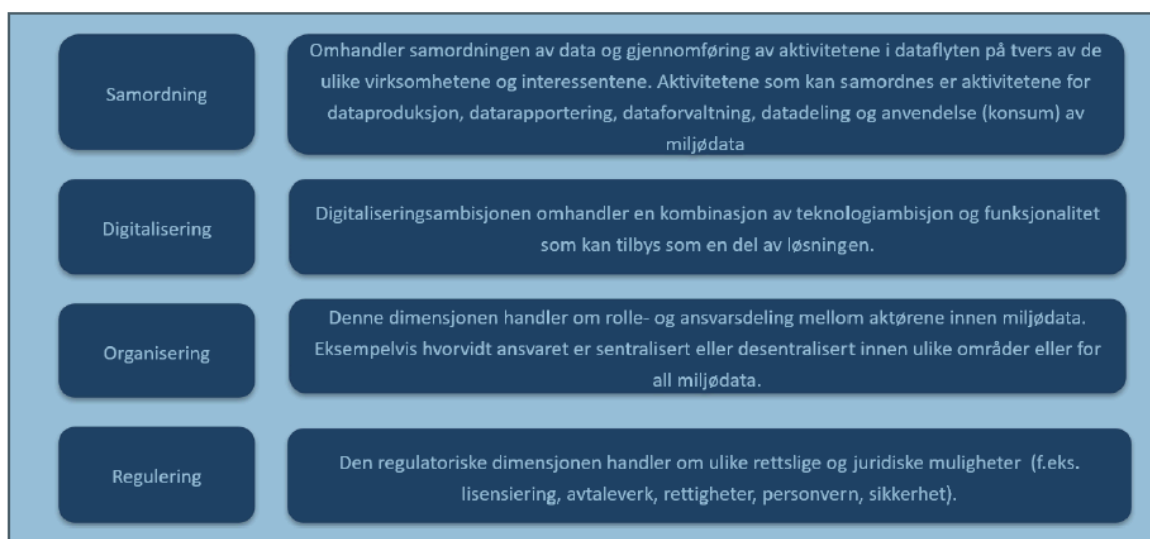
³⁷ *Environmental and Atmospheric data infrastructure - EBAS*: EBAS er NILU's database for observasjonsdata om atmosfærens sammensetning fra overvåkingsprogrammer og forskningsprosjekter, etablert i 1990: [EBAS home](#)

4 Konsepter

Utredningen *Fremtidens miljødata* (Menon) ³⁸fra 2021, heretter referert til som Menon (2021), hadde følgende mandat:

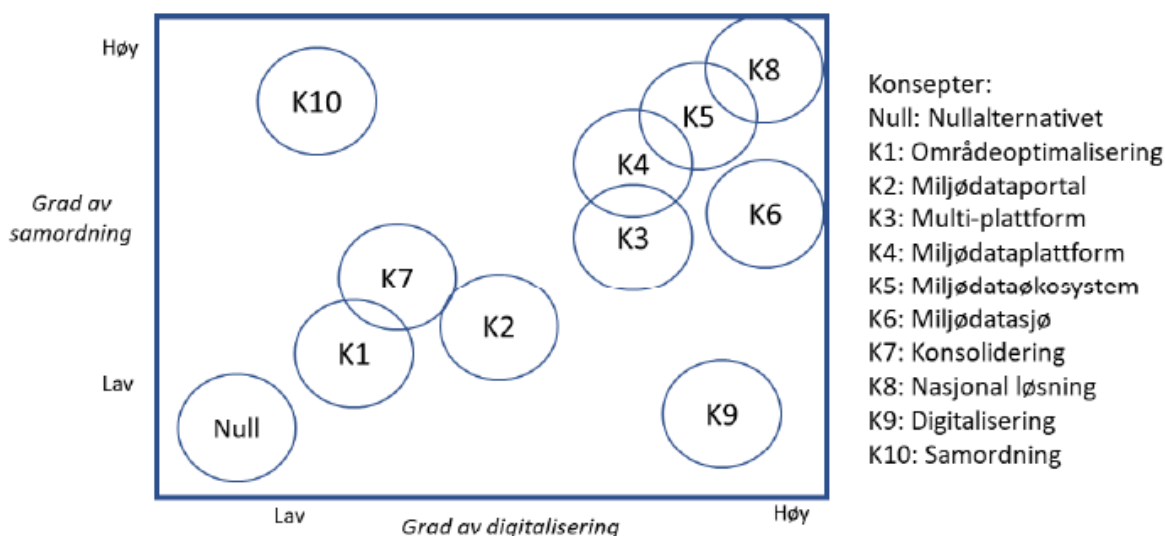
“....Klima- og miljødepartementet og Miljødirektoratet har et overordnet nasjonalt samordningsansvar for miljødata. Dette er en arena med mange aktører, og det er store utfordringer knyttet til spørsmål om framtidig datainfrastruktur, som vi per i dag ikke har god nok oversikt over. Miljødirektoratet skal på vegne av departementet gjennomføre en utredning av dagens utfordringer, status og behov i miljødataforvaltningen. Det skal i tillegg utredes alternative løsninger på gjeldende utfordringsbilde innen miljødataforvaltning i Norge....”

Utredningen dekker både en problembeskrivelse og en behovsanalyse. Mulige datainfrastrukturkonsepter er vurdert etter fire dimensjoner: Samordning, digitalisering, organisering og reguleringsperspektiv.



³⁸ Fremtidens miljødata, Menon Economics og A-2. Menon-publikasjon nr. 153/2021, <https://menon.no/prosjekter/fremtidens-miljodata>

10 ulike konsepter for datainfrastruktur er beskrevet og vurdert grundig i rapporten.



Fire konsepter hadde vesentlig høyere grad av måloppnåelse enn de andre. Samtidig var gjennomføringsrisikoen vurdert fra middels til svært høy (fra 4 - 6 på en skala fra 1-6). Utredningen anbefaler at man skal starte med en konsolidering, K7, for deretter å gå videre med å etablere en Miljødataportal, K2. Etablering av Miljødataplattform, K4, og Nasjonal løsning, K8, vurderes på det tidspunktet som mer urealistisk og krevende å få gjennomført.

Selv om Menon (2021) omhandler miljødata, er konseptene som er skissert like relevante for havdata. De økonomiske beregningene i rapporten er dog kun for miljødata og kan ikke gjenbrukes uten videre her.

I dette prosjektforslaget har vi vurdert

1. Nullalternativet
2. Konsolidering og en felles miljø- og havdataportal
 - a. presentert for miljødata i Menon (2021)
3. Nasjonal løsning - samling av miljø- og havdata på ett sted
 - a. presentert for miljødata i Menon (2021)
4. Dataområde for deling av miljø- og havdata (Data space)

For å sammenligne de ulike alternativene har vi valgt vurdere dem opp mot hvordan konseptene tar inn over seg dataenes innebygde evne til interoperabilitet (*dataplan*), og hvordan de håndterer krav til kontroll og styring på datadeling (*kontrollplan*). *Teknologiplan* omhandler hvordan konseptene ser for seg at økosystemet skal bygges opp rent teknisk. Til slutt har vi sett på *kompetanseoverføring og formidling*.

Evnen til å oppnå gevinstene skissert i kapittel 3.1 er vurdert for hvert av konseptene.

4.1. Nullalternativet

Nullalternativet innebærer at dagens praksis for datadeling videreføres uten etablering av et koordinert, nasjonalt dataøkosystem. Utviklingen av datadelingsløsninger vil dermed forbli fragmentert, drevet av enkelte institusjoners egne behov, begrensede forskningsprosjekter eller ad hoc samarbeid.

4.1.1 Dataplan

Det gjennomføres ingen felles forsøk på å harmonisere arbeidsprosesser knyttet til datainnsamling og distribuering. Hver og en datatilbyder arbeider videre med å tilnærme seg FAIR prinsippene og etablere metadata, metadatakataloger og vokabularer innenfor sine fagfelt.

4.1.2 Kontrollplan

Det er opp til hver og en datatilbyder å etablere løsninger for tilgang til data, og regler for bruk av egne datadistribusjonsløsninger. Datatilbyderene kan velge å bruke kjente standardiserte datalisenser eller lage sine egne. Det samme gjelder tilgangsstyring til dataene.

4.1.3 Teknologiplan

Det er opp til hver og en datatilbyder å benytte den teknologien den ønsker. Det vil være mange som benytter eventuelle felles løsninger eller opensource løsninger for å løse enkelte utfordringer avhengig av egen kompetanse og kapasitet på området. Det vil ikke være enhetlige metoder og formater for distribuering av data utover det som er fastslått i dagens lovverk.

4.1.4 Kompetanseoverføring og formidling

Hver og en datatilbyder må løse eventuelle kompetanseutfordringer på datadeling selv ved å oppsøke konferanser, kurser og relevant litteratur. Det vil ikke eksistere noe rådgivende organ som kan bistå datatilbyderen i arbeidet med metadata og distribusjon. For datakonsumenten blir det opp til datatilbyderne å tilby tjenester for å bistå i forståelsen av dataene som blir delt. Det kan være skriftlig brukerdokumentasjon eller mer konsultative tjenester.

4.1.5 Konsekvenser og måloppnåelse

- **Interoperabilitet:** Uten en felles infrastruktur vil miljø- og havdata forbli i institusjonelle siloer. Dette fører til vedvarende dobbeltarbeid, der ulike aktører utvikler overlappende tekniske løsninger, metadata-standarder og portal-tjenester. Ressurser som kunne vært brukt på innovasjon og verdiskaping, vil fortsette å gå til vedlikehold av ikke-interoperable systemer. Muligheten til å koble ulike data innenfor

miljø og hav vil forbli krevende. Konseptet medfører vedvarende fragmentering og ineffektivitet.

- **Manglende evne til å møte kravene fra KI:** Mens KI-modeller krever data som er strukturerte, kontekstualiserte og sikkert forvaltede, vil nullalternativet bety at en stor andel av våre miljø- og havdata forblir «ikke-KI-klare». Vi vil ikke evne å høste gevinstene fra den teknologiske utviklingen, da dataene mangler den nødvendige semantiske konsistensen og kvalitetssikringen som kreves for maskinell prosessering.
- **Manglende tverrfaglig innsikt:** Den største gevinsten med et felles økosystem for deling av miljø- og havdata ligger i å kombinere data på tvers av domener (atmosfære, hav, miljø). Ved å opprettholde dagens situasjon vil denne muligheten i stor grad forbli urealisert. Vi vil fortsatt ha store utfordringer med å generere ny, helhetlig vitenskapelig innsikt som er kritisk for forvaltning og beredskap, fordi dataene fra de ulike domene ikke kan settes sammen effektivt og maskinelt. Dermed vil det fortsatt være krevende for forvaltningen å få presis informasjon ved bl.a. konsekvensutredninger og arealplanlegging. Ved beredskap og krisehåndtering vil det fortsatt være sannsynlig at ulike aktører har ulikt datagrunnlag og dermed forskjellig situasjonsforståelse. Informasjon om nye bedre datakilder når ikke fram til datakonsumentene. Lovpålagte rapporteringskrav bla til internasjonale avtaler vil fort kunne være etablert ut fra ikke-optimale data.
- **Lav kontroll og tillit:** Uten et felles rammeverk for datasuverenitet vil bekymringen hos dataprodusenter om tap av kontroll og misbruk av data vedvare. Dette kan hemme viljen til å dele data. Vi vil fortsatt mangle felles mekanismer for å verifisere dataintegritet og opphav, noe som vanskeliggjør identifisering av autoritative kilder og reduserer tilliten til de dataene som faktisk deles.
- **Risiko for akterutseiling:** Internasjonalt, og særlig innen EU, er utviklingen av økosystemer for deling av data (Data Spaces) rask. Ved å ikke etablere en nasjonal koordinert tilnærming, risikerer Norge å falle av utviklingen i Europa (f.eks. tilknytning til Green Deal Data Space). Vi vil bli stående med løsninger som ikke er kompatible med europeiske standarder, noe som gjør det vanskelig for norske aktører å delta i det voksende europeiske datamarkedet.

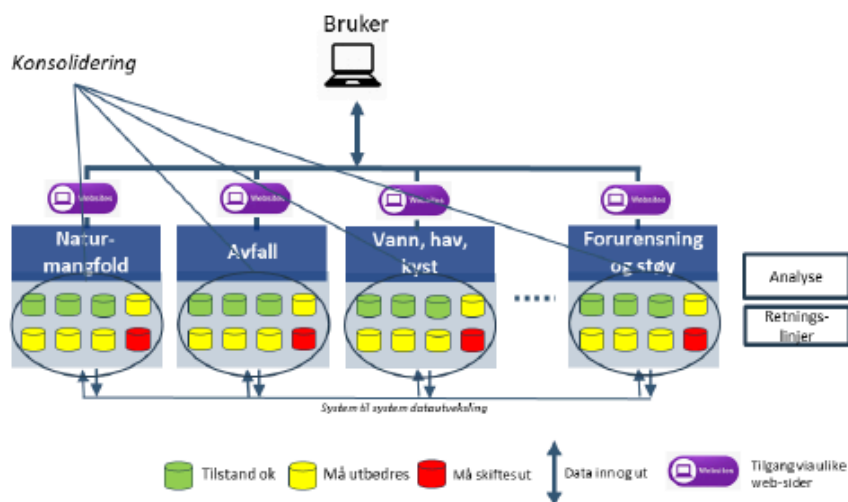
Mål	Vurdering av måloppnåelse for Nullalternativet
Tverrfaglig synlighet	Svært lav. Data forblir låst i institusjonelle siloer
Metadata & Provenans	Lav. Hver aktør ivaretar sine egne vokabularer og kataloger uten felles samordning
Interoperabilitet & Lisens	Lav. Ustandardiserte lisenser og formater krever krevende manuell ad-hoc-sammenstilling

Datasuverenitet & Tillit	Lav. Uten felles styringsrammeverk vedvarer bekymringen for tap av kontroll over egne data, noe som hemmer delingsviljen
EU-kobling	Lav. Norske data vil ikke være tilgjengelig i EU, da norske løsninger ikke er kompatible med europeiske standarder
KI-klargjøring	Lav. Dataene mangler den nødvendige semantiske konsistensen og strukturen som kreves for maskinell KI-behandling

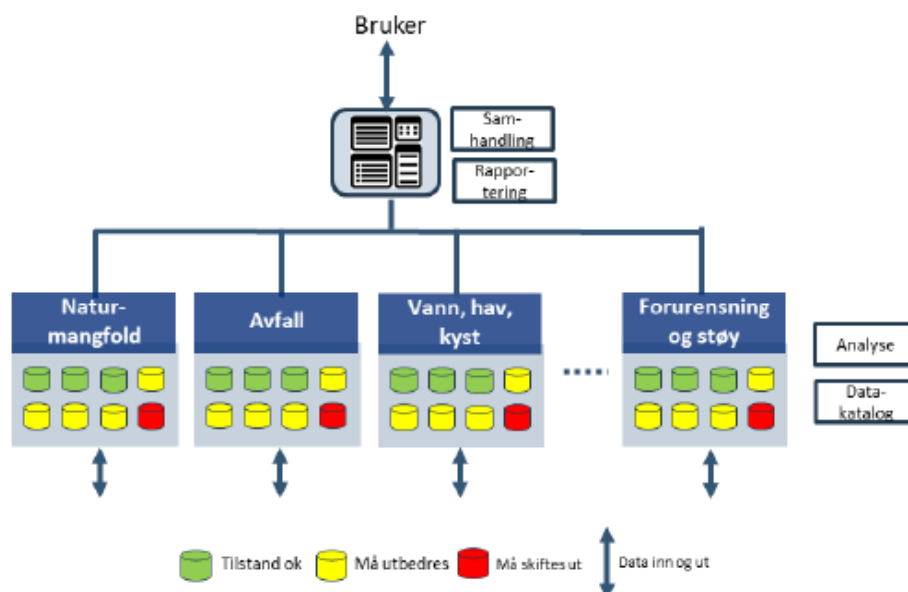
4.2 Konsolidering + Miljødataportal

Det finnes en rekke datadelingsløsninger, data-baserte tjenester og datakataloger både i Norge, EU og i verdensnivå (se Bakgrunn og Status), som fungerer i varierende grad. Landskapet er imidlertid fragmentert og heterogent, med forskjellige nivåer av tverrfaglighet og modenhet. Formålene til løsningene varierer også. Noen er designet kun for svært spesifikk bruk, mens andre er mer generiske.

Konseptet Konsolidering + Miljødataportal som er beskrevet i Menon (2021) tar utgangspunkt i dagens løsninger for deling av data. Konseptet tar sikte på at det gjennomføres mer samordning innenfor hvert fagområde og hos hver datatilbyder, pluss en felles inngangsport: Miljødataportal til miljødata, og tilsvarende en Havdataportal for havdata.



Figur Konsolideringskonsept fra *Fremtidens miljødata* rapport



Figur Portalkonsept fra Menon (2021).

4.2.1 Dataplan

Arbeidsprosesser for rapportering, dataforvaltning, formidling og tilgjengeliggjøring av data harmoniseres innenfor de enkelte domene. Dette betyr at det skal utarbeides en felles datamodell og felles retningslinjer for standardisert oppbygging og bruk av datakataloger og vokabularer innenfor hvert domene.

I enkelte tilfeller kan man også samordne med tilgrensende fagområder.

Det vil også stilles krav om rapportering av innkommende data fra næringslivet til aktører underlagt KLD (og NFD). Disse samordnes slik at ikke samme data rapporteres flere ganger på ulike måter.

4.2.2 Kontrollplan

Det etableres felles retningslinjer for god dataforvaltning. For miljødata er det på Miljødirektoratet gjennom rollen som miljødatakoordinator som har dette ansvaret. Det er ikke kjent om det finnes en tilsvarende nasjonal koordinator for Havdata.

Konsolideringsarbeidet skal intensiveres gjennom tydelige føringer gjennom blant annet tildelingsbrev, anbefalinger og annen informasjon. Blant annet skal det utarbeides

- retningslinjer for standardisering knyttet til data og formater
- god praksis for kreditering av dataprodusenter
- god bruk av avtaler og lisenser
- krav til deling av data i avtaleverk og etterlevelse av dette
- krav til bruk av datakataloger og vokabularer innen hvert fagområde

4.2.3 Teknologiplan

Det legges ikke opp til storstilt modernisering av eksisterende løsninger. De datatilbyderne som har kompetanse og midler vil fortsette å utbedre sine løsninger.

Datatilbyderne vil konsolidere eksisterende løsninger innenfor sine domener. Dette vil noen ganger kreve samarbeid mellom ulike dataeiere. I tråd med systemforvaltningsprinsipper lager man planer for å utbedre, eventuelt slå sammen og til en viss grad skifte ut eksisterende fagsystemer. Målet er å få totalt sett litt færre og bedre løsninger.

Datakonsumentene skal få tilgang til løsningene som i dag, dvs gjennom nettsider, portaler eller plattformer. Dataene vil fortsatt formidles via e-post, api-er, nedlastninger av filer og strømming.

Ved å legge på et portalkonsept utenfor de enkelte datatilbyderne vil datakonsumentene få lettere oversikt over hvilke data som finnes innenfor de forskjellige domenene. Informasjon om hvem som er datatilbydere, lisenser og lignende vil også komme tydelig fram.

4.2.4 Kompetanseoverføring og formidling

Det er lagt opp til at Miljødatakoordinatoren skal ha ansvaret for oppfølging og veiledning av datatilbyderne i deres opprydningsarbeid i tråd med Digitaliseringsdirektoratets “orden i eget hus” prinsipp.

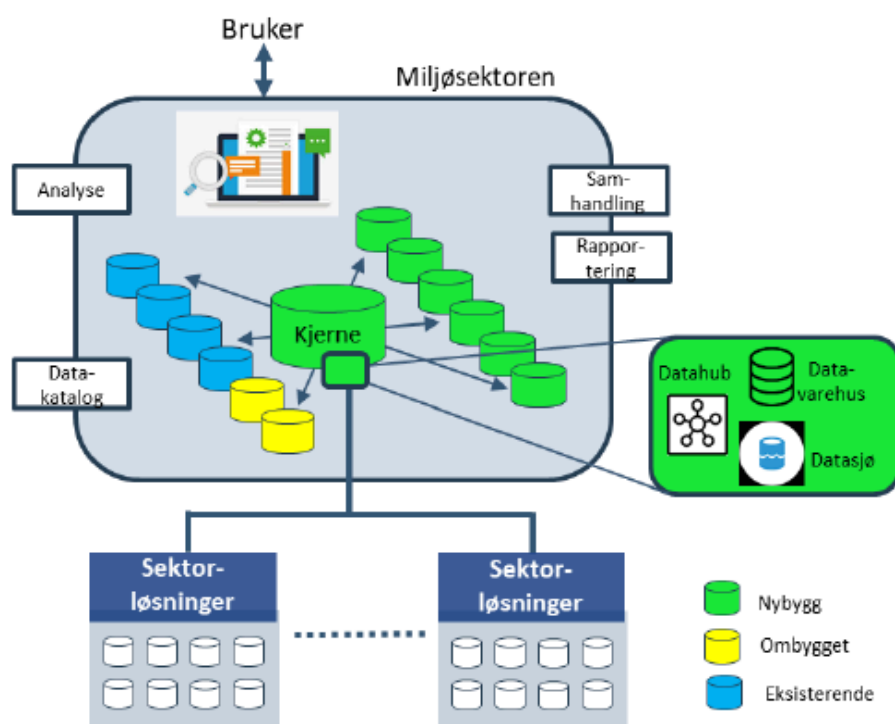
4.2.5 Konsekvenser

- **Lav interoperabilitet:** Konsolideringsløsningen arbeider mot en større grad av samordning innenfor enkeltdomener. Dette er bra. Det er også fornuftig å bygge på eksisterende løsninger, men selve gjennomføringen vil ofte kreve store løft for de eksisterende løsningene. Det vil fortsatt være mye dobbelarbeid med å etablere tildels overlappende funksjonalitet i hver enkelt løsning både innenfor et domene og mellom domener. En ny portalløsning på toppen av en konsolidering vil også fort kunne framstå som en unødvendig konkurrent til data.norge.no eller geonorge.no. Konseptet vil innebære vedvarende fragmentering og ineffektivitet.
- **Manglende evne til å møte kravene fra moderne KI:** Konsolideringsløsningen vil ikke møte kravene som KI stiller. Selv om det jobbes med regler for oppbygging av metadata og krav til datakataloger, legges det fortsatt opp til at data formidles som i dag. KI-modeller krever ikke bare data som er strukturerte, kontekstualiserte og sikkert forvaltede, men også standardiserte API-er for henting av data.
- **Manglende tverrfaglig innsikt:** Konsolideringsløsningen tar ikke innover seg behovet for tverrfaglig innsikt, men rendyrker samordning innenfor hvert enkelt domene. I og med at konsolideringsarbeidet er basert på frivillighet og uten vesentlig gevinster, men innebærer høye kostnader for datatilbyder, vil arbeidet ta svært lang tid. Det er dermed dårlige utsikter til at en konsolidering bidrar til bedre og enklere tverrfaglig innsikt.

- **Lav kontroll og tillit:** Det er ingen ting i Konsolideringsløsningen utover at det skal etableres regler for avtaler og lisens som forbedrer kontroll og tillit. Datatilbyderne må selv kontrollere tilgang og bruk for sine løsninger. Arbeidet med å få innsikt i hvordan dataene benyttes vil være som i dag. Datakonsumentene må fortsatt gjøre egne og omfattende undersøkelser på hvor de kan finne aktuelle data og selv vurdere om disse dataene kommer fra en autoritativ kilde og om dataene de har funnet er de mest presise og pålitelige.
- **Høy risiko for akterutseiling:** Konsolideringsløsningen medfører at datadeling vil fortsette omtrent som før. Muligheten til å knytte seg til europeiske økosystemer for deling av data vil måtte gjennomføres av hver enkelt datatilbyder og Norge vil få mindre muligheter blant annet til å kunne benytte europeiske digitale tvillinger for å få ut nyttig informasjon.

Mål	Vurdering av måloppnåelse for Konsolidering + Miljødataportal
Tverrfaglig synlighet	Begrenset. Rendyrker samordning <i>innenfor</i> enkeltdomener (f.eks. avfall, naturmangfold), men løser i liten grad tverrfaglig kobling på tvers av domener.
Metadata & Provenans	Delvis. Felles datamodeller utarbeides internt i sektorene, men mangler tversgående vokabulartjenester.
Interoperabilitet & Lisens	Begrenset. Etablerer interne retningslinjer, men data formidles fortsatt via tradisjonelle filoverføringer og uensartede API-er.
Datasuverenitet & Tillit	Uendret. Tilbydere må selv bygge og vedlikeholde tilgangskontroll i sine egne fagsystemer.
EU-kobling	Krevende. Hver enkelt datatilbyder må selv bygge uavhengige koblinger mot europeiske økosystemer.
KI-klargjøring	Fraværende. Legger opp til tradisjonell formidling uten krav om standardiserte maskin-til-maskin API-er.

4.3 Nasjonal løsning - samling av miljø- og havdata på ett sted



Et alternativ til et økosystem er beskrevet i utredningen som Nasjonal løsning. Konseptet går ut på å samle data fra alle dataeierne i en felles database, et datavarehus, eller i en datasjø (et sentralt lagringssted for store mengder tildels ustrukturerte rådata), og gi datakonsumentene tilgang til å hente data derfra.

4.3.1 Dataplan

En nasjonal løsning krever høy grad av standardisering. FAIR prinsippene vil ligge i bunn. Datatilbydere må samordne arbeidsprosesser innenfor både dataproduksjon og dataforvaltning innenfor samme domene, og mellom flere domener. Det innføres også felles regler for standardisert oppbygning og bruk av datakataloger og vokabularer innenfor miljø- og havsektorene. Miljødatakoordinatorrollen til Miljødirektoratet vil være førende i dette arbeidet knyttet til miljødata, likeledes vil det være en tilsvarende rolle innen havdata. I denne løsningen vil koordinatorrollen ha stort handlingsrom for å stille krav til dokumentasjon og formater og følge opp standarder.

Løsningen vil hente inn data fra de ulike sektorløsningene. En harmonisert datamodell legges til grunn for forvaltningen av disse dataene. Det legges opp til bruk av færre dataformater, men samordningen gjør at man kan håndtere en økende kompleksitet i innkommende data.

4.3.2 Kontrollplan

Løsningen krever sannsynligvis endring i lover og/eller forskrifter for å kunne sikre den tverrsektorielle datainnsamlingen, oppbevaringen og delingen.

Det kan også være personvern hensyn som må vurderes, og også eventuelle konkurransemessige hensyn overfor spesielt private aktører.

Samtidig kan det også være behov for å etablere egne lover knyttet til miljødata og havdata for å sikre ønsket deling og bruk av data til andre formål enn opprinnelig tiltenkt.

4.3.3 Teknologiplan

Plattformen som etableres vil ha kapasitet til å lagre og behandle store datamengder. Dataene vil bli lagret i datasjøer, databaser eller som datafiler. Det etableres felles funksjonalitet som datakataloger, ulike verktøy for søk, visualisering og analyse, grafisk informasjon og et datavarehus. Alt, helt ned til hvert datafelt, har innebygd tilordnede metadata.

Funksjonalitet blir flyttet fra dagens systemer og inn i denne løsningen. Ved overflytting av data legges det til grunn automatisert kontroll og kvalitetssikring. Det lages også egne moduler for rapportering av f.eks. innkommende data fra næringslivet. Data skal fortsatt formidles videre til datakonsumenter som i dag gjennom ulike medier som e-post, API-er, filoverføringer etc.

Det skal etableres avanserte søketjenester for søk på tvers av alle domener og datatyper. Det legges opp lærende brukerdialoget i søketjenesten.

Løsningen vil kreve store investeringer i infrastruktur for datalagring.

4.3.4 Kompetanseoverføring og formidling

Koordinatorrollen vil stille krav til datatilbydere, og må derfor forventes å veilede disse i avlevering av data til løsningen. Det må vurderes å legge til rette for en veiledningsfunksjon for bruk av data, utover dokumentasjon.

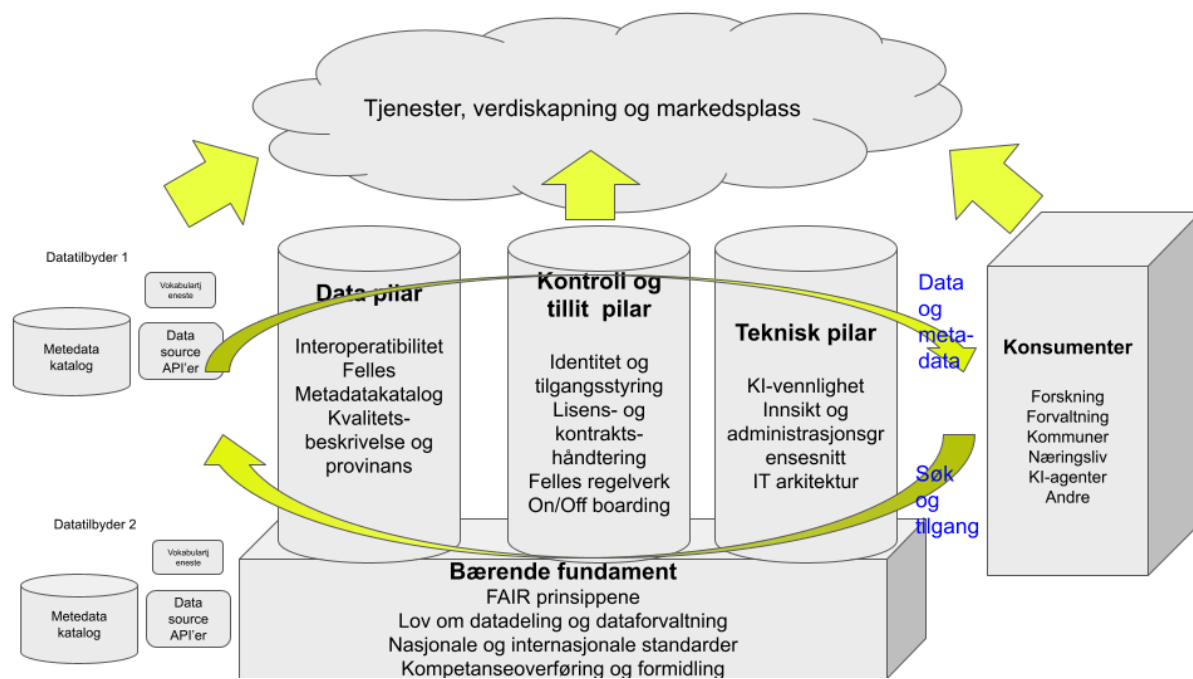
4.3.5 Konsekvenser

- **Redusert fragmentering og ineffektivitet:** Løsningen vil sikre samordning. De tidligere fragmenterte løsningene vil kunne saneres og ressursene kan settes inn på å utvikle den nasjonale løsningen og å få lagt inn data i denne. Samtidig kan løsningen framstå som veldig rigid og dermed redusere interessen for å levere data inn i løsningen. Dette kan medføre at nasjonale og internasjonale domaneløsninger fortsatt blir foretrukket for tilgjengeliggjøring av data.
- **God evne til å møte kravene fra KI:** Løsningen vil ha veldefinerte metadata- og datamodeller, og dermed vil det enkelt kunne etableres veldefinerte API-er som enkelt kan benyttes av KI.

- **God tverrfaglig innsikt:** Løsningen vil legge til rette for god tverrfaglig innsikt. Felles datamodell og gode metadatabeskrivelser og vokabulartjenester vil muliggjøre etablering av tverrfaglige sammenstillinger.
- **God kontroll og tillit:** Løsningen legger opp til stor grad av kontroll. Det vil være lett å få ut rapporter på bruk av data og lignende. Hvorvidt en slik stor løsning oppnår tillit kan være mer usikkert. Løsningen fratar datatilbyderene/dataeierne kontrollen over sine data. Enkelte dataeiere kan føle seg fremmegjort og dermed nedprioritere arbeid lokalt. Det er også en fare for at sentral datakurator ikke har tilstrekkelig kompetanse på domenet og vil komme til kort når det ettervært kommer nye formater og internasjonale standarder blir utviklet.
- **Moderat risiko for akterutseiling:** Løsningen bygger på et plattformkonsept, mens resten av Europa går mot mer federerte løsninger. Store plattformer har lett for å bli vanskelige å vedlikeholde og tilpasse over tid. Løsningen kan forholdsvis enkelt knyttes mot digitale tvillinger og også europeiske dataområder (data spaces). Løsningen vil potensielt sette sterkere begrensinger til framtidig tilgjengeliggjøring av nye typer data enn en mer federert tilnærming.

Mål	Vurdering av måloppnåelse for Nasjonal løsning
Tverrfaglig synlighet	Høy teoretisk måloppnåelse. Samler alle data i én sentral løsning med felles søketjenester.
Metadata & Provenans	Høy. Toppstyrt standardisering med felles datamodeller og integrerte metadata.
Interoperabilitet & Lisens	Høy. Krever omfattende lov- og forskriftsendringer for tverrsektoriell datainnsamling.
Datasuverenitet & Tillit	Variert. Fratar datatilbyderne eierskapet og den direkte kontrollen over dataene, noe som kan svekke motivasjonen for lokal dataforvaltning.
EU-kobling	Moderat. Kan kobles mot EU, men plattformkonseptet avviker fra EUs fødererte Data Space-strategi.
KI-klargjøring	God. Sentraliserte og harmoniserte datamodeller gjør det enkelt å eksponere strukturert data via API-er.

4.4 Dataområde (data space) for miljø- og havdata



Figur Konseptskisse Dataområde for deling av miljø- og havdata

Det fjerde konseptet som har vært vurdert er et økosystem basert på etablering av et dataområde bygd opp i henhold til EUs Data Space support center. Følgende funksjonalitet bør være inkludert:

- Dataplan
 - sikrer grunnleggende interoperabilitet mellom datasett herunder tid og rom-dimensjonene
 - datakatalog
 - sikrerer kjent kvalitet på data
- Kontrollplan
 - muliggjøre autentisering og autorisering av tilgang til dataområdet både for konsumenter og datatilbydere
 - sikrerer lisens og avtaleforvaltning mellom aktørene slik at dataeierskapet ivaretas
 - kjente regler for deltagelse (se kapittel om governance)
- Teknisk plan
 - ivaretar AI vennlighet
 - ivaretar grunnleggende IT-arkitektur inkludert maskin-til-maskin grensesitt og mulighet for enkel utvidelse av dataområdet
- Brukerengasjement
- Tjenester og verdiskapning
 - kobling mot andre dataområder, og et grunnlag for verdiøkende tjenester i senere faser

Dataområdet skal bli føderert. Autoritative data forvaltes hos datatilbyderne, og skal ikke kopieres til en sentral dataplattform med mindre et konkret bruksbehov tilsier det. Felleskomponentene kobler sammen eksisterende datakilder gjennom standardiserte og

maskinlesbare grensesnitt. Konseptet innebærer at det bygges opp en drift, forvaltning og veiledningsorganisasjon parallelt med etablering av funksjonalitet.

Ved etablering av den grunnleggende funksjonaliteten i dataområdet, bør den bygges i samarbeid med erfarne datatilbydere og erfarne datakonsumenter. FAIR-prinsippene vil være et bærende fundament for dataområdet. Dataene som skal deles i første omgang vil være et utvalg av datatilbydernes data. Datautvalget må være slik at både interoperabilitet og kvalitetsaspektet utfordres.

4.4.1 Dataplan

4.4.1.1 Sikre grunnleggende interoperabilitet mellom datasett

Metadata og datainteroperabilitet er et av kjerneprinsippene i FAIR (se Appendix A), og er ofte det vanskeligste prinsippet å oppnå i praksis, spesielt for data. Det handler om å strukturere og formatere ulike datasett slik at maskiner kan lese, tolke og koble dem sammen på en meningsfull og effektiv måte. Det er også betinget av at de ulike datatilbyderne og konsumentene deler samme forutsetninger og forventninger. Teknisk interoperabilitet (grensesnitt, protokoller), syntaktisk interoperabilitet (skjemaer, formater) og semantisk interoperabilitet (vokabularer, ontologier) må ivaretas. Når samhandlingen utvides til å gjelde flere tilbydere, innlemmes også rammeverk for organisatorisk, juridisk og forretningsmessig samhandling.

Vårt Dataområde for miljø- og havdata skal legge til rette for metadata- og datainteroperabilitet på tvers av datasett, domener og datatilbydere. Målet er å maksimere interoperabiliteten slik at dataområdet kan utvides og kobles til andre systemer med bruk av eksisterende rammeverk og standarder ved senere faser. Modenhetsgraden knyttet til interoperabilitet er varierende hos aktørene. Det må derfor utvikles et felles interoperabilitetsrammeverk i form av krav til et mellomlag (connectorer) som hver datatilbyder må etablere utenfor sine datakilder.

Data- og metadatainteroperabilitet – en avgrensning

Metadatainteroperabilitet handler om å strukturere metadatainformasjon slik at den er forståelig for andre systemer. Dette gjelder format og struktur, men aller viktigst: semantikk. Når et metadataelement er utformet med utgangspunkt i et veldefinert begrep – med både betegnelse, definisjon og eventuell kobling til andre relevante vokabularer – blir det mulig å avgjøre om to metadataoppføringer fra ulike vokabularer har samme betydning. For eksempel kan betegnelsen «ingen begrensninger for offentlig tilgang» (som beskriver tilgang til data) tilsvare «åpne data» i et annet vokabular, eller rollen «principal investigator» kan tilsvare «creator».. Vokabulartjenester kan tilby denne funksjonaliteten gjennom semantisk kobling.

Datainteroperabilitet handler om å strukturere og formatere ulike datasett slik at de kan kobles sammen eller kombineres på en meningsfull og effektiv måte. Det utgjør det tekniske og faglige fundamentet i et dataområde, og spenner over tre kjerneaspekter; teknisk, syntaktisk og semantisk. Når samhandlingen utvides til å gjelde både innenfor og på tvers av datadomener og datatilbydere, innlemmes også rammeverk for organisatorisk, juridisk og forretningsmessig samhandling. Datainteroperabilitet kan

kreve manuell og automatisk databehandling (endring i opplysning, geografiske projeksjon, omformatering, etc.), tilpassende infrastruktur og videre dokumentasjon (provenans og metadata om datatransformasjoner/behandlingsmetadata).

4.4.1.2 Metadata-innhøsting - felles metadatakatalog

Metadata-høsting danner grunnlaget for opprettelsen av ethvert metadatakatalogsystem som tilgjengeliggjør poster fra ulike kilder. Innhøstningssystemer kalles ofte for metadata-aggregatører. Høstingsprosessen baserer seg på interoperabilitetsreglene som datatilbyderne må følge for at informasjonen om deres data skal kunne finnes i en felles metadatakatalog. Interoperabilitetsreglene for metadata skal utformes slik at de sikrer maksimal fleksibilitet, bærekraftig forvaltning av infrastrukturen og full frihet for deltakende institusjoner til å velge og utvikle egen programvare, så lenge de standardiserte grensesnittene leveres i henhold til interoperabilitetsreglene.

Metadatahøstingen er en grunnkomponent i vårt Dataområde for miljø- og havdata. Aggregatøren etablerer faste tilkoblinger til maskinlesbare kataloger som tilgjengeliggjøres av de deltagende datatilbyderne. Metadatainformasjonen høstes og oversettes om nødvendig til en felles metadatatmodell. Videre lagres informasjonen i en database og gjøres tilgjengelig i kataloger, både for brukergrensesnitt og maskinell avlesning.

Datatilbyderne har ansvaret for påliteligheten til egne metadatakataloger, og for at metadataenes innhold er korrekt. Metadataenes innhold skal alltid inkludere lenker til dataene. Datatilbyderne forblir alltid den autoritative kilden til overleverte metadata.

Metadatakataloger representerer den mest hensiktsmessige og effektive metoden for å gi datakonsumenter rask og enkel tilgang til alle tilgjengelige produkter. Implementering av standarder og maskinlesbarhet (se kapittel 4.3.1.1. om interoperabilitet) i vårt dataområde sin metadatakatalog er en kritisk forutsetning for å få et brukervennlig dataområde for deling av data. For å muliggjøre tilkoblinger til andre systemer, må metadatakatalogen også kunne vise metadata på ulike standarder, og det er nødvendig å implementere ulike konverterer som kan gjenbrukes av andre.

4.4.1.3 Kvalitetsbeskrivelse og proveniens

Skal norske miljø- og havdata gi mest mulig verdi for forvaltning, forskning og næringsliv, må de faktisk kunne finnes, forstås og brukes av andre enn dem som samlet dem inn. Det er dette FAIR-prinsippene handler om. Uten dette forblir selv svært verdifulle datasett i praksis låst inne i det enkelte dataeiers systemer, og kjent bare for dem som allerede vet at de finnes, og hvordan de kan brukes.

Dataområdet må inkludere regler relatert til minimumsnivå for metadata og kvalitet. Det er ikke realistisk å kreve full proveniens for dataene som deles, men sporbarhetsinformasjon som følger metadataene til dataene bør være synlig for konsumentene. Det samme gjelder eventuelle kvalitetsbeskrivelser av dataene. Synlighet av denne informasjonen er nødvendig for å riktig bruk og tolkning av data på tvers.

4.4.2 Kontrollplan

I beskrivelsen av dataområdet bruker vi begrepene «governance» og «data governance» slik de er definert i boksen nedenfor. Dataområdet vil ha en overordnet governance-modell som fastsetter ledergruppens sammensetning, parternes beslutningsmyndighet, finansiering og medbestemmelse, samt formelt ansvar overfor Klima- og miljødepartementet (KLD) og Nærings- og fiskeridepartementet (NFD). En underliggende data governance-modell regulerer datakvalitet, metadata, lisensiering, tilgangsstyring og code of conduct for datatilbyderne.

Governance og data governance – en avgrensning

*Med **governance** menes den overordnede styringsstrukturen for en virksomhet, et samarbeid eller et system: Hvem som har beslutningsmyndighet, hvordan beslutninger fattes, hvordan ansvar er fordelt, og hvordan etterlevelse av regler og mål sikres over tid. Governance handler om strukturene og prosessene for styring, ikke om innholdet som styres.*

***Data governance** er en avgrenset del av dette: Styring rettet spesifikt mot data som ressurs – eierskap og ansvar for det enkelte datasett; krav til kvalitet, metadata og standarder; rettigheter til å produsere, dele, endre og bruke data; tilgang, lisensiering og informasjonssikkerhet; etterlevelse av regelverk som personvernforordningen (GDPR), offentlighetsloven og dataforvaltningsloven; samt håndtering av avvik og brudd på fellesreglene. Data governance er dermed en underkategori av governance, avgrenset til dataforvaltning.*

4.4.2.1 Identitet og tilgangsstyring

Adgangskontroll til data involverer autentisering (bevis på hvem man er) og autorisering (tilgangsstyring). Teknologier for autentisering er etter hvert blitt standardisert, og vi ønsker en føderert autentiseringsløsning hvor brukeren skal kunne autentisere seg mot sitt hjemorganisasjonssystem, såkalt single sign-on. Eksempler på mulige aktører her er Feide eller ID-porten, men det må også vurderes andre løsninger for å gjøre det enklere for utenlandske og private aktører å koble seg opp. I tillegg vil det legges opp til at deltagere i dataområdet skal kunne lage maskin-identiteter via f.eks Maskinporten, eller bruke sertifikater.

Autorisering er prosessen hvor man, etter at en identitet er avklart, avgjør hvilke tilganger man skal ha. Her er det et spekter fra enkle regler, som at enkelt-identiteter (personer eller maskiner) skal ha tilgang, eller at spesifikke institusjoner skal ha tilgang, til mer kompliserte regler.

Vi ser for oss at en første løsning bør støtte å dele data

- totalt åpent med mulighet for anonym tilgang, for eksempel søk i metadata og bruk av åpent lisensierte datasett
- med alle som er innlogget i dataområdet
- begrenset til enkeltpersoner og enkeltinstitusjoner.

Analysen av kombinasjoner av datasettenes metadata krever autentisering av brukeren. God tilgangskontroll krever et grensesnitt hvor dataeier kan styre hvem som har tilgang til sine data. Dette kan implementeres ved f.eks å etablere en "policy-engine" som automatisk beregner tilganger basert på identiteter og institusjoners egenskaper (f.eks hvilke sertifiseringer de har) og hvilken avtaler datakonsumentene har signert. Dette ser vi som en naturlig utvidelse i senere faser av implementeringsprosjektet. Når identiteter autoriseres for tilgang til et datasett skal dataområdet kunne logge dette, og gi dataeier en oversikt over hvem som har hatt tilgang til dataene.

4.4.2.2 Kontrakt/lisenstjeneste

Kontrakt- og lisenstjenesten er den funksjonen i dataområdet som formaliserer avtaleforholdet mellom den som leverer data (datatilbyder) og den som skal bruke dem (datakonsument). Tjenesten skal gjøre det tydelig, og maskinlesbart, hvilke vilkår som gjelder for bruk av et gitt datasett, og sørge for at disse vilkårene faktisk følger dataene når de forlater kildesystemet.

I første omgang kan dataområdet ha en kontrakt- og lisenstjeneste som kun består av faste standardiserte lisenser basert på internasjonale anerkjente lisenser for åpne data. Alle lisenser er bestemt av datatilbyder, og er en avtale mellom datatilbyder og datakonsument.

Tjenesten bør også håndtere kontraktene/terms-of-service mellom dataområdet og datatilbyderne og mellom dataområdet og datakonsumentene. Tjenesten må kunne gi svar på hva som blir gjeldende lisens ved kombinasjon av flere datasett med ulike lisenser. Tjenesten må være utbyggbar til også å kunne inneholde lisenser for ikke-åpne og kommersielle data.

4.4.2.3 Innrulling og fratreddelse av datatilbydere

Innrulling og fratreddelse er de to formelle prosessene som styrer hvordan nye partnere påkobles, integreres og avslutter sin deltagelse i dataområdet i rollen som datatilbyder, samt tilhørende avtaler. Det må fastsettes tydelige regler og krav for deltagelse slik at prosessen styres på en transparent måte. Aksept av dataområdets vilkår og betingelser, samt etterlevelse av tekniske krav og kvalitetskrav, er obligatorisk for innrullingsprosessen. Fratreddelseprosessen må omfatte både frivillig og tvungen uttreden. Det må etableres en prosess for regelmessig overvåking av etterlevelse av retningslinjer og standarder, samt oppfyllelse av inngåtte avtaler, for å ivareta dataområdets integritet, og det må iverksettes tiltak dersom det avdekkes avvik.

4.4.2.4 Fellesregler for deltagelse

Dataområdet for miljø- og havdata må inkludere et entydig sett av regler for både datatilbydere og datakonsumenter. Fellesreglene må undertegnes ved innrulling.

Reglene for datatilbyderen kan inkludere regler innenfor disse områdene:

- **Juridiske regler**

- Eksempler på mulige juridiske regler
 - **Lisensiering og bruksrettigheter:** Klar definisjon av hva dataene kan brukes til (f.eks. kun analyse, ikke viderealgoritmetrening), hvor lenge de kan brukes, og om de kan videreformidles.
 - **Eierskap og rettighetsavklaring:** Garanti for at tilbyder faktisk har rett til å dele data (inkludert immaterielle rettigheter og samtykker).
 - **Personvern og GDPR:** Krav om at personopplysninger enten er anonymisert/pseudonymisert, eller at det foreligger et gyldig behandlingsgrunnlag og databehandleravtale.
 - **Ansvarsbegrensning og SLA:** Tydelige rammer for erstatningsansvar ved feil i data, nedetid eller utilsiktet datalekkasje.
- **Tekniske og operasjonelle regler**
 - Eksempler på mulige tekniske og operasjonelle regler
 - **Standardiserte grensesnitt (API-er):** Krav om å benytte godkjente protokoller og åpne API-standarder for tilgjengeliggjøring.
 - **Metadata og datakvalitet:** Krav om beskrivelse av dataene i henhold til dataområdets felles datakatalog- og metadatastandard (f.eks. beskrivelse av struktur, oppdateringsfrekvens og historikk).
 - **Interoperabilitet:** Bruk av felles datamodeller og vokabularer (semantisk interoperabilitet) slik at data lett kan kobles med data fra andre tilbydere.
 - **Identitet og tilgangsstyring (IAM):** Krav om at tilbyder integrerer sine endepunkter mot dataområdets identitets- og tillitsrammeverk
- **Sikkerhet og etterlevelse**
 - Eksempler på mulige regler
 - **Datasikkerhetsstandards:** Krav til kryptering av data (både under lagring og overføring), samt jevnlig sikkerhetsrevisjoner.
 - **Informasjonssikkerhetsstandard:** Krav til adekvat informasjonssikkerhet hos datatilbyder
 - **Loggings- og sporbarhetsplikt:** Tilbyder må støtte eller føre ubestridelige logger over datatransaksjoner for å kunne verifisere hvem som har hentet hva og når.
 - **Tilgangsstyring (Data Sovereignty):** Regler for hvordan tilbyder setter og håndhever tilgangskontroll (policies) som datakonsumentens kobling må respekteres.
- **Organisatoriske og etiske regler**
 - Eksempler på mulige regler
 - **Retningslinjer for endringsledelse:** Plikt til å varsle i god tid ved endringer i API-er, datamodeller eller nedleggelse av datastrømmer.
 - **Code-of-conduct:** Hvordan datakonsumentene skal opptre i dataområdet

4.4.3 Teknologi plan

4.4.3.1 KI-vennlighet og maskinlesbarhet.

Dataområdet bør utformes slik at data og metadata er lett tilgjengelige for automatiserte analyser, beslutningsstøttesystemer og fremtidige KI-løsninger. Dette innebærer ikke at

dataområdet selv trenger å inneholde avanserte KI-funksjoner, men at data, metadata, lisenser og kvalitetsbeskrivelser eksponeres gjennom standardiserte og maskinlesbare grensesnitt.

KI-systemer er avhengige av å kunne finne, forstå og kombinere data uten omfattende manuell tilrettelegging. Dataområdet må sørge for at metadata beskriver innhold, kvalitet, geografisk dekning, tidsmessig dekning, opphav og bruksvilkår på en strukturert måte. Dette vil gjøre det mulig å gjennomføre automatiserte søk etter relevante datasett, bygge datadrevne arbeidsflyter og utvikle KI-løsninger som kan kombinere miljø- og havdata på tvers av institusjoner. På metadata-nivå bør det legges til rette for både tradisjonelt søk og semantisk søk, slik at brukere og KI-assistenten kan identifisere relevante datasett basert på faglige problemstillinger fremfor kun eksakte nøkkelord.

Mange brukere innen forskning, forvaltning og næringsliv benytter i økende grad språkmodeller (LLM-er) som grensesnitt mot data og informasjon. Dataområdet bør derfor tilby åpne og veldokumenterte API-er som gjør det enkelt å integrere metadata, datakataloger og autoriserte datakilder i eksterne KI-løsninger. Det kan vurderes å støtte samhandlingsmønstre mellom KI-agenter og dataområdet gjennom standardiserte protokoller. Konkrete protokoller velges ut fra gjeldende åpne standarder og europeiske anbefalinger i implementeringsfasen. Dette kan gjøre det enklere for KI-assistenten å oppdage datasett, hente metadata, forstå tilgangsbetingelser og initiere dataforespørsler på en kontrollert måte. Slike mekanismer bør imidlertid betraktes som et tillegg til, og ikke en erstatning for, åpne API-er basert på etablerte standarder.

KI-vennlighet handler ikke bare om metadata. Dataformater, semantiske beskrivelser, kvalitetsindikatorer, proveniensinformasjon og maskinlesbare lisenser er også viktige forutsetninger for at KI-systemer skal kunne tolke og anvende data korrekt. God KI-vennlighet bygger derfor direkte på FAIR-prinsippene og på arbeidet med interoperabilitet. Utviklingen innen generative KI-modeller går svært raskt, og det vil normalt være mer kostnadseffektivt å legge til rette for at eksterne og fremtidige modeller kan benytte dataområdets data enn å bygge/forvalte/drifte egne modeller.

I tillegg til å gjøre metadata tilgjengelig gjennom standardiserte API-er, bør dataområdet tilby enkle KI-baserte støttetjenester for søk og oppdagelse av data. Dette kan for eksempel være lette språkmodeller eller semantiske søketjenester som hjelper brukere med å finne relevante datasett basert på naturlig språk, faglige problemstillinger eller beskrivende spørsmål.

4.4.3.2 Innsikt og administrasjonsgrensesnitt

For å gjøre dataområdet tilgjengelig for både datatilbydere og datakonsumenter bør det etableres en innsiktsløsning som fungerer som inngangsport til dataområdet. Innsiktsløsningen skal gi tilgang til den felles metadatakatalogen og gjøre det enkelt å finne, forstå og få tilgang til tilgjengelige datasett.

Innsiktsløsningen bør støtte søk, filtrering og visualisering av metadata på tvers av deltakende organisasjoner. Brukere skal kunne søke etter datasett basert på tema, geografi, tidsperiode, kvalitet, tilgangsrestriksjoner og lisensvilkår. Innsiktsløsningen skal også gi oversikt over hvem som eier datasettene, hvordan de kan brukes, og hvordan tilgang

eventuelt kan bestilles. Innsiktsløsningen skal være en frontend til den felles metadatakatalogen, mens autoritative metadata fortsatt forvaltes hos de enkelte datatilbyderne og innhøstes av den sentrale katalogen.

Innsiktsløsningen skal også eksponere funksjonaliteten gjennom åpne og maskinlesbare API-er, slik at brukere og fremtidige KI-løsninger kan søke, hente og integrere metadata direkte i egne systemer. Videre bør innsiktsløsningen tilby veiledning om deltakelse i dataområdet, alle reglene for dataområdet, lisensmodeller og onboarding av nye datatilbydere og datakonsumenter.

Administrasjonsgrensesnittet skal gi datatilbydere mulighet til å konfigurere tilgangen til sine datasett.

4.4.3.3 Arkitekturprinsipper ut fra et IT-teknisk perspektiv

Dataområdet for miljø- og havdata skal være effektivt i gjennomføring av oppgaver. Det skal også være enkelt å etablere, forvalte og videreutvikle. Et dataområde består naturlig nok av mange komponenter og flere partnere. Det er dermed viktig at alle komponentene som inngår har veldefinerte funksjoner og veldefinerte, åpne maskingrensesnitt (API-er) basert på kjente standarder.

I all hovedsak bør kjernefunksjonaliteten i dataområdet bestå av allerede etablerte og forvaltede byggeklosser, enten nasjonale eller internasjonale. Ved utvalgelse av felleskomponentene er det viktig å se på eierskap til komponentene og planer for videre forvaltning og utvikling. Det er også naturlig å vurdere hvorvidt vi kan være med på å planlegge videreutvikling av de utvalgte komponentene. For mange av komponentene er det en stor fordel at de består av en åpen kildekode og er lisensiert for gjenbruk og videreutvikling. Ved utvalgelse og sammensetning av komponenter må vi se på mest mulig fullautomatiserte verdikjeder. Samtidig må det være brukergrensesnitt for administrering og konfigurering av tilgang til egne data for partnerne i dataområdet. Alle brukergrensesnitt skal være universelt utformet.

Alle datatilbydere må etablere en connector for å tilgjengeliggjøre sine metadata, og her må det også legges opp til et rammeverk og samarbeid mellom datatilbyderne.

Fordeler ved bruk av felleskomponenter:

1. Kostnadseffektivitet: Reduserer behovet for å utvikle og vedlikeholde egne løsninger.
2. Standardisering: Sikrer at systemer og prosesser følger nasjonale og internasjonale standarder.
3. Sikkerhet: Felleskomponenter, spesielt statlige, er ofte designet med høy sikkerhet og samsvar med personvernregelverk som GDPR.
4. Skalerbarhet: Muliggjør enklere tilpasning til endrede behov og økt brukermasse.
5. Innovasjon: Frigjør ressurser som kan brukes til å utvikle nye tjenester og funksjonalitet.

Mulige utfordringer ved bruk av felleskomponenter:

1. Kompleksitet ved integrasjon

2. Avhengighet til leverandører: hvis kildekoden ikke er åpen kan lisensbetingelsene endre eller støtten opphøre over tid, noe som skaper risiko for å bli låst til en bestemt leverandør.
3. Fleksibilitet og tilpassingsevne: det er viktig å vurdere om komponenten også vil kunne utvikles til å møte fremtidens behov

Dataområdet skal være desentralisert. Data skal ikke flyttes fra kilden unødvendig. Det er likevel nødvendig at man bygger opp en felles metadatakatalog. Felleskomponentene kan også etableres desentralt hvis det er hensiktsmessig.

I de tilfeller hvor det ikke finnes hensiktsmessige felleskomponenter, må disse utvikles. Det er da viktig at de utvikles slik at de er gjenbrukbare og skalerer godt. Disse komponentene bør utvikles under åpen kildekode slik at andre dataområder også kan benytte dem.

4.4.3.4 Markeds plass og verdiskaping

For å gjøre dataområdet mest mulig verdifullt for brukerne er det fornuftig at det etableres en markeds plass for verdiskaping. Markeds plassen vil kunne inneholde koblinger inn i andre tilsvarende dataområder, mulighet for salg av data og å tilby andre verdiøkende tjenester. Andre verdiøkende tjenester kan f.eks være digitale tvillinger, karttjenester eller andre simuleringstjenester som benytter data som er tilbudt via dataområdet.

I første omgang vil det være behov for å muliggjøre tilknytning til andre tilsvarende dataområder, og å etablere en markeds plass for å vise eksempler på anvendelse av dataområdet gjennom verdiøkende tjenester.

4.4.4 Kompetanseoverføring og formidling

Et dataområde får verdi først når flere enn de fire første partnerne kan bruke det. Prosjektet skal derfor ha kompetanseoverføring og formidling som en egen leveranse, og ikke som noe som gjøres til slutt dersom det er tid.

Etablering, drift og bruk av et nytt dataområde krever ny kunnskap og kompetanse hos både datatilbydere og konsumenter. Formidling må foregå der konsumentene er å finne: På konferanser, i faglige fora og i egnede nyhetssaker på nett.

Hos datatilbyderne er domenespesifikk kompetanse så vel som overordnet forståelse av brukerbehov essensielt. Partnerne må få mulighet til å øke kompetansen innen aktiv datakurering og dataforvaltning gjennom kurs, veiledere eller best practices. Dette kan omfatte praktisk anvendelse av FAIR-prinsippene, beherskelse av internasjonale metadastandarder, forvaltning av felles ontologier og kontrollerte vokabularer, samt implementering av sporbarhet.

Datakonsumentene vil måtte forholde seg til det nye systemet og kan få hjelp til bruk av dette i hjelpemenyer, webinarer, brukerhistorier på nett eller som videopresentasjoner, og gjennom datatilbydernes eksisterende brukerstøtteløsninger.

Kompetanseoverføring omfatter standardoppsett og veiledning for onboarding, slik at nye datatilbydere kan koble seg til uten å gjenta arbeidet de første partnerne har gjort, og opplæring og overlevering til lokale team hos den enkelte datatilbyder. Formidling omfatter en åpen faglig arena der miljø-, klima- og havinstituttene, forvaltningen og Forskningsrådet kan møtes om europeiske dataområder, datasuverenitet og erfaringene fra første fase, med både fordeler og ulemper.

Å samle fagmiljøene til en felles diskusjon har verdi i seg selv. Skissen peker i kapittel 3 på risikoen for at det vokser fram flere parallelle dataøkosystemer i Norge, og Dataområdet for miljø- og havdata er avhengig av tillit hos aktører som ikke deltar fra start. Begge håndteres best ved at sektoren møtes og diskuterer europeiske økosystemer, datasuverenitet og erfaringene fra første fase åpent, med fordeler og ulemper. Flere av miljøinstituttene planlegger allerede en nasjonal fagkonferanse om kunstig intelligens for klima og miljø. En egen bolk der vil gi formidling til hele sektoren uten at prosjektet må etablere et eget arrangement.

4.4.5 Konsekvenser

- **Lav fragmentering og ineffektivitet:** Etter hvert som nye datatilbydere innrulles i løsningen, vil det gjøre datadelingslandskapet mer homogent og effektivt. Felleskomponentene legger til rette for federering mot datatilbyderens egne løsninger. Konseptet krever ingen omskriving av eksisterende kode, men datatilbyderen etablerer mellomlag i form av en connector som sikrer at datatilbyderens metadata, vokabulaer, lisenser og tilgangsregler blir synlig på en ensartet måte for datakonsumentene.
- **God evne til å møte kravene fra KI:** Felles regler for struktur på metadata, krysskobling mellom ulike domener og etablering av api-er for både søk og maskinlesbare resultater etter forhåndsdefinerte strukturer gjør at datatilbyder kan dele data på en KI-vennlig måte.
- **God tverrfaglig innsikt:** Felles regler for struktur på metadata og krysskobling mellom ulike domener vil gjøre det mulig å finne enhetlige metoder for å tolke og sette sammen data på tvers av domener.
- **Høy kontroll og tillit:** Dataområdet legger til rette for både autentisering og autorisering av datakonsumenter. Dette gjør at datatilbyder får bedre innsikt i hvem som benytter egne data og hvordan. Overholdelse av dataområdets regelverk knyttet til datatilbyder skal følges opp med datatilbyder og vil medvirker til at datakonsumenter får tillit til dataområdet. De får en form for kvalitetsstempel fra dataområdet på at datatilbyderen er pålitelig, at de samler inn og forvalter dataene i tråd med beste praksis og at den oppgir opprinnelsessted og proveniens for dataene de deler. Det vil også være mulig for datatilbydere å skille bedre mellom brukere, da de er autentisert via f.eks MinID, virksomhetssertifikater eller tilsvarende europeiske tjenester.
- **Lav risiko for akterutseiling:** Løsningen er en federert løsning i tråd med den europeiske utviklingen. Etablering av løsningen er basert på gjenbruk og

videreutvikling av allerede utviklede europeiske open source komponenter. Da det etter all sannsynlighet vil være mange i Europa som benytter disse komponentene, vil ny teknologi og ny funksjonalitet som regel bli utviklet og tatt i bruk raskt.

Mål	Vurdering av måloppnåelse for Dataområde miljø- og havdata (Data Space)
Tverrfaglig synlighet	Høy. Felles føderert metadatakatalog og innhøstingssystem (aggregator) gjør data fra alle domener søkbare på ett sted uten at dataene flyttes fra kilden.
Metadata & Provenans	Høy. Felles interoperabilitetsrammeverk med vokabulartjenester, felles metadatamodell og krav til synlig provenans og kvalitetsbeskrivelser.
Interoperabilitet & Lisens	Høy. Standardiserte mellomlag (connectors), maskinlesbare lisenser og en kontraktstjeneste som sikrer at lisensbetingelser følger dataene.
Datasuverenitet & Tillit	Høy. Dataene forblir hos autoritative kilder. Tilbyder styrer tilgang og lisenser direkte via føderert identitetsstyring (f.eks. Maskinporten, SSO).
EU-kobling	Optimal. Bygget direkte på EUs Data Space-arkitektur og gjenbrukbare åpen kildekode-komponenter (Simpl-open), noe som gir direkte samhandling med Green Deal Data Space og EOSC.
KI-klargjøring	Høy. Eksponerer strukturert metadata, kvalitetsindikatorer og maskinlesbare lisenser over åpne API-er og protokoller (som MCP), spesielt tilrettelagt for maskin til maskin-kommunikasjon og KI-agenter.

4.5 Sammenligning av de vurderte konseptene

Konsept	Dataplan	Kontrollplan	Teknologiplan	Kompetanseoverføring og formidling	Konsekvenser & Måloppnåelse
Nullalternativet	Ingen felles harmonisering. Hver aktør ivaretar sine data, metadata og	Ustandardiserte lisenser og tilgangsstyringsregler opp til hver enkelt tilbyder.	Ad-hoc teknologivalg uten enhetlige grensesnitt utover	Ingen felles rådgivning eller støtte; konsumenter må forholde seg til ulik	Vedvarende fragmentering, data er i liten grad KI-klare, krevende tverrfaglig kobling, lav kontroll/tillit og risiko for nasjonal akterutseiling mot EU.

	kataloger internt.		gjeldende lovkrav.	dokumentasjon .	
Konsolidering + miljø- og havdataportal	Harmoniserte prosesser og felles datamodeller innenfor enkeltfagområder (siloer).	Felles retningslinjer for god dataforvaltning internt i sektoren (f.eks. ved Miljødirektoratet).	Konsolidering av eksisterende domenesystemer og en felles portal (Miljødataportal /Havdataportal) på toppen.	Veiledning fra koordinatorrolle for opprydding i egne systemer ("orden i eget hus").	Bedre orden innen enkeltdomener, men fortsatt lite tverrfaglig samhandling, begrensede maskinlesbare API-er for KI og risiko for akterutseiling mot europeiske initiativer.
Nasjonal løsning	Høy grad av standardisering med felles sentral datamodell og færre dataformater.	Krever lov/forskriften dringer for tverrsektoriell innhenting; sentralisert styring der dataeier avgir kontroll.	Sentral dataplattform/datasjø med stor lagrings- og beregningskapasitet og felles analyseverktøy.	Sentral koordinator veileder leverandører ved avlevering av data til den sentrale plattformen.	God KI-vennlighet og tverrfaglighet, men høy gjennomføringsrisiko, fjerner kontroll/motivasjon fra fagmiljøene, og rigid arkitektur som risikerer avvik fra EUs dataområdemodell.
Dataområde for miljø- og havdata (Data Space)	Felles interoperabilitetsrammeverk (teknisk, syntaktisk, semantisk) via mellomlag (connectors) og føderert metadatakatalog.	Federert autentisering/autorisering (SSO, Maskinporten), avtalespesifikke lisenser, onboarding-regler og bevart datasuverenitet hos tilbyder.	Desentralisert/føderert arkitektur basert på open-source (Simpl-open), åpne API-er, KI-vennlighet og markedsplass.	Kurs/veiledning i FAIR og datakurering for tilbydere; webinarer, brukerhistorier og brukerfora for konsumenter.	Reduserer fragmentering, sikrer full KI-klargjøring og tverrfaglig kobling, opprettholder datasuverenitet og tillit, samt sømløs kobling til EUs Green Deal Data Space.

En sammenligning av konseptene viser en ulik grad av måloppnåelse. *Nullalternativet* og *Konsolidering + Miljø- og havdataportal* gir svært lite i forhold til synlighet og interoperabilitet på tvers av domener. I *Konsolidering + Miljø- og havdataportal*-konseptene vil Miljødatakoordinatorrollen kunne styre hva som tilgjengeliggjøres i portalen, og således gi en bedre oversikt over autoritative kilder enn *Nullalternativet*. Det gir også en mer strukturert tilnærming til å etablere felles metadata-, vokabular og provinansforståelse. På teknologisisiden fortsetter konseptene som i dag, og vil ikke gi bedre tilgang til data for KI-agenter eller andre. Det blir heller ikke noen forenkling i å koble Norges miljø- og havdata mot andre europeiske initiativer.

De to andre konseptene, *Felles nasjonal løsning* og *Dataområde (data space) for miljø- og havdata*, løser hovedutfordringene beskrevet i kapittel 2, og muliggjør gevinstene beskrevet i kapittel 3. Konseptene er derimot vesentlig forskjellig når det gjelder tilnærming til løsning. *Felles nasjonal løsning* framstår som svært komplekst og krever meget store investeringer. Risikoen for å ikke lykkes framstår som svært høy. *Dataområde (data space) for miljø- og havdata* har en mer lettbenyttet tilnærming ved bruk av allerede eksisterende komponenter. Løsningen er vesentlig rimeligere å etablere, og risikoen er vesentlig lavere. For alle

løsningene er det krevende arbeid å sikre felles metadata-, vokabular- og provenansforståelse.

I det videre arbeidet ser vi på etablering av et Dataområde (data space) for miljø- og havdata.

5 Styringsmodell for ferdig løsning

Ambisjonene til Dataområde for miljø- og havdata er å etablere et økosystem som gjennom sine partnere, regelsett og felles tjenester gir tilgang til autoritative kilder hos datatilbydere av miljø- og havdata. For å oppnå denne tillitten må både dataprodusenter, datatilbydere og konsumenter vite at dataene er av kjent kvalitet, deles under klare lisenser, og at de forvaltes transparent og i tråd med klare fellesregler etablert av dataområdet. I tillegg må dataområdet eksistere over tid og forvaltes godt. Dette krever finansiering, insitamenter for å være med som datatilbyder, insitamenter for å benytte løsningen som konsument, samt medbestemmelse.

I et dataområde vil det i tillegg til rollene datakonsument og datatilbyder være tjenesteleverandører som tilbyr tjenester inn i dataområdet og en styringsmyndighet.

5.1 Skisse til organisering og styring

I kapittel 7.1 er det foreslått en enkel styringsmodell for prosjektfasen med de fire initiativtakerne, og med oppdragsgivende departementer som premissgiver. Parallelt utarbeider prosjektet et forslag til varig styringsmodell. Endelig organisering, operatørmodell og juridisk form anbefales først når pilotfasen har gitt erfaring med deltakelse, ansvar, kostnader og brukerbehov.

Den varige styringsmodellen skal utformes slik at nye offentlige, forskningsbaserte og private datatilbydere kan delta på likeverdige vilkår. Miljødirektoratet har i dag en rolle som miljødatakoordinator og Kartverket har en rolle som geodatakoordinator. Disse rollene skal også ivaretas videre.

Dataområdets styringsmyndighet er organisasjonen som ivaretar dataområdet. Styringsmyndigheten beslutter og vedlikeholder regelverk for datatilbydere og datakonsumenter.

Dataområdets styringsmyndighet godkjenner også nye datatilbydere og sikrer verifisering av datakonsumenter.

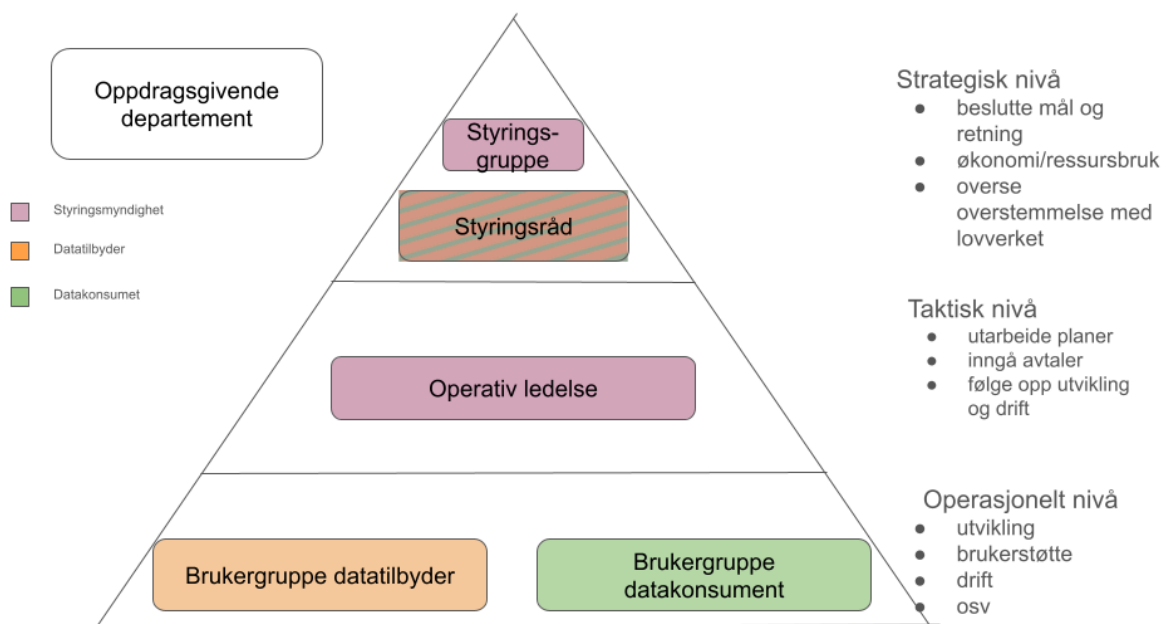
Styringsgruppen er den øverste ledelsen av styringsmyndigheten. Sammensetningen av den varige styringsgruppen avgjøres når pilotfasen er evaluert. Det bør da vurderes hvordan Miljødirektoratet og Kartverket, i sine roller som miljødatakoordinator og geodatakoordinator, best kan delta. Det kan være en fordel at styringsgruppen på sikt også inkluderer en representant for datatilbydere og en representant for datakonsumenter. Dette vil sikre god brukermedvirkning og dermed økte muligheter for suksess.

Tilordnet til styringsmyndigheten ligger en operativ enhet. Den operative enheten ivaretar dataområdet på daglig basis. Enheten ivaretar drift, videreutvikling og forvaltning av den tekniske løsningen, samt veiledning og kommunikasjon med ulike brukergrupper. De har forvaltningsmyndighet over dataområdet, det vil si at de kan håndheve dataområdets regler,

foreslå nye avtaler om deltagelse og implementere brukerhistorier, etablere ny funksjonalitet, operere dataområdet innenfor rammen av tildelt budsjett, osv.

Datatilbydere deltar i dataområdet i form av medlemskap. Dataområdet vil ha et sett klare og transparente regler for datatilbydere. Dette vil gjøre det enklere for nye deltagere å bli med, og i tillegg være en forsikring for relevansen til dataområder for både datatilbydere og datakonsumenter. Et utvalg av datatilbydere vil danne et styringsråd hvor operativ ledelse legger fram planer for videre utvikling, endringer i regelsett og forslag til implementering av nye brukerhistorier. Styringsrådet vurderer gevinster og prioritering av planer/oppgaver innenfor de økonomiske rammene som eksisterer. Planene skal prioriteres ut fra hva som gir størst mulig gevinst for hele dataområdet. Det vil alltid være dataområdets styringsmyndighet i form av styringsgruppen som tar den formelle beslutningen, mens styringsrådet har høringsrett.

På operativt nivå er det to hovedbrukergrupper: Datakonsument og datatilbyder. Det må etableres operative brukermøter med representanter fra begge brukergruppene. Brukermøtene holdes regelmessig og er både tekniske og funksjonelle. Det bør være hyppigere brukermøter i etableringsfasen. Det bør på sikt, også kunne etableres et brukerforum for tjenesteutviklere som ønsker å tilby tjenester inn i dataområdet.



Figur Styringspyramide

5.2 Anbefalt finansieringsmodell

Å etablere et Dataområde for miljø- og havdata krever finansiering. EU finansierer nå en rekke ulike europeiske dataområder. For vårt dataområde er det viktig å så langt som mulig gjenbruke eksisterende komponenter. Samtidig kreves det en initiell investering i etablering, implementering og markedsføring, samt årlige drifts- og forvaltningsmidler.

Det anbefales at staten sikrer framtidig drift, videreutvikling, veiledning og forvaltningsmidler. Dette er i tråd med anbefalingene i alternativ D fra Forskningsrådets utkast til *Anbefalinger om hvordan en ny ordning for langsiktig drift av viktige datainfrastrukturer kan innrettes*.³⁹

Utvikling av funksjonalitet bør i størst mulig grad prøves å medfinansieres gjennom Digital Europe-programmet, der medfinansiering på 50 prosent er normalordningen for utrullingstiltak. Det forutsetter at løsningen bygger på de europeiske rammeverkene, og at norsk deltakelse forankres tidlig i arbeidet med Green Deal Data Space. En annen mulig finansieringskilde kan være Forskningsrådet. Mulige finansieringskilder blir diskutert nøyere i kapittel 7.4.

Vi mener at datatilbydere bør få mulighet til å få dekket sine etableringskostnader knyttet til innrulling (onboarding). Dette gjelder først og fremst i oppstartsfasen av dataområdet, men også ved innrulling på et senere tidspunkt. Dette for å sikre at så mye data fra miljø- og havområdet som mulig blir tilgjengeliggjort i løsningen.

Det kan settes av en finansieringspott for innrulling knyttet til realisering av brukerhistorier med store gevinster for samfunnet der det ikke er andre finansieringsmuligheter.

Eventuelle betalingsmodeller for kommersielt salg av data eller tjenester må utredes særskilt, som del av en framtidig forretnings- og finansieringsmodell. Her kan det være mulig å ta en viss andel av omsetningen eller en kostnad basert på trafikk tallene data eller tjenestene genererer. Dette vil først bli aktuelt på et senere tidspunkt, og en beslutning om dette er ikke en forutsetning for å etablere den grunnleggende infrastrukturen.

³⁹ Utkast sendt ut som underlagsdokument i forkant av Forskningsrådets innspillsmøte 24.09.2026.

6 Forutsetninger

Det er en rekke forutsetninger som må være avklart eller bli håndtert underveis i etablering av et dataområde for miljø- og havdata. Disse forutsetningene er beskrevet i de påfølgende delkapittelene.

6.1 Langsiktig og forutsigbar finansiering

For å lykkes må Dataområdet for miljø- og havdata finansieres som en langsiktig nasjonal infrastruktur, ikke som et tidsavgrenset prosjekt. Prosjektmidler vil være nødvendige for etablering av den tekniske og organisatoriske grunnstrukturen, men vil alene ikke være tilstrekkelig for å sikre videre drift, forvaltning og utvikling av dataområdet. Skal dataområdet skape varig verdi, må deltakende organisasjoner og nye aktører kunne ha tillit til at løsningene, tjenestene og styringsmodellene består over tid.

Langsiktig og forutsigbar finansiering er særlig viktig for å stimulere til deltakelse fra nye datatilbydere. Tilknytning til dataområdet vil kreve investeringer i metadataforvaltning, interoperabilitet, kvalitetsarbeid og integrasjoner. Slike investeringer er vanskelig å forsvare dersom den underliggende infrastrukturen har usikker finansiering eller kort tidshorisont. En stabil finansieringsmodell vil derfor være en viktig forutsetning for å bygge opp et dataområde som gradvis vokser i omfang, datamengde og antall deltakere.

Det bør etableres en modell som skiller mellom finansiering av etablering og finansiering av drift, forvaltning og videreutvikling. Drift og forvaltning bør omfatte tekniske felleskomponenter, metadataforvaltning, bruker- og deltakerstøtte, innrulling av nye aktører, sikkerhet, governance og videreutvikling av dataområdet. Dette tilsier et behov for en langsiktig finansieringsmekanisme, for eksempel gjennom nasjonal grunnfinansiering, supplert med prosjektmidler eller inntekter fra tilleggstjenester.

Det er de oppdragsgivende departementene som har ansvaret for at det finnes midler til drift, forvaltning og videreutvikling av dataområdet. De bør også ta ansvaret for at det finnes midler til å støtte datatilbydernes mulighet til å innrulleres i dataområdet. Arbeidet utover etablering av en POC med en pilotimplementering, bør ikke påbegynnes uten at midlene til fremtidig drift og forvaltning er avklart.

6.2 Åpne data og brukerregisterering

For å sikre effektiv etablering av dataområdet forutsettes det enighet om at åpne data som hovedregel skal kunne brukes uten registrering, og at registrering eller API-nøkkel bare kreves der det er nødvendig for å styre belastningen på systemet, eller når datatilbyderen krever det.

Tilgang via innsiktsløsningen må være åpen og uten autentisering for å sikre at informasjon om eksistensen av åpne data er tilgjengelig for alle.

Åpne data med åpen lisens skal som hovedregel kunne lastes ned og brukes uten innlogging. Datasuverenitet betyr at datatilbyderen bestemmer vilkårene for egne data og kan få dem håndhevet.

Det kan kreves registrering eller API-nøkkel for nedlasting og maskin-til-maskin-tilgang, når begrunnelsen er å styre belastningen på systemet, for eksempel ved strømming av store datamengder, eller når datatilbyderen krever dette.

Styringsgruppen til Dataområdet for miljø- og havdata har ansvaret for at denne framgangsmåten følges, eller at en annen beslutning er tatt før arbeidet startes og at dette ikke diskuteres gjennom hele prosjektet.

6.3 Fortsette arbeidet med datakuratering og implementering av FAIR kompatibilitet

Dataområdet skal bygge på eksisterende løsninger og videreutvikle disse, og forutsetter at datatilbyderne fortsetter sin domenespesifikke datakuratering og kvalitetskontroll. Erfaringene fra SENDA ⁴⁰prosjektet legges tas med videre. Datatilbyder skal levere kuraterte data (jf. [Sikt: Kurateringsnivåer for forskningsdata](#)⁴¹), hvilket blant annet omfatter tillegging av metadata, versjonskontroll, aggregering eller omkoding av data, samt etablering av nye datasamlinger basert på ulike kilder.

Arbeidet med å heve nivået på metadata og datakvalitet hos datatilbyderne må ikke stanse opp i påvente av mulige finansieringsmuligheter i tilknytning til dataområdet. Likeledes må datatilbyderne sikre at metadataene inkluderer lisensvilkår som er i tråd med dataeiers krav.

Datatilbyderne må ta ansvaret for å sikre høy kvalitet på metadata, og at de er i tråd med regelverket som etableres for dataområdet. Styringsgruppen må ta ansvaret for å sikre koblingen mellom metadata fra ulike domener. Arbeidet må gjøres i samarbeid med fagressurser i domenet. Datatilbyderer må frigjøre kapasitet til å bidra til dette arbeidet.

6.4 Åpen kildekode og gjenbrukbare komponenter

Arkitekturen til dataområdet må understøtte robust drift, sikkerhet og skalerbarhet allerede fra etableringsfasen. Den må også være så enkel og rimelig å implementere hos de forskjellige aktørene som mulig. Det er krevende å utvikle dette fra grunnen av, så dataområdet skal bygge på eksisterende åpen kildekode-løsninger, som f.eks Simpl-open.

6.5 Eksisterende systemer for tilgang til data

Jo enklere dataområdet kan gjøre innrulling av nye datatilbydere, jo mer data vil bli gjort tilgjengelig, og jo større sjanse er det for suksess.

Innrulling krever en connector, som er enkelt å sette opp og som stiller få, men nødvendige krav til relevante eksisterende IT-systemer.

⁴⁰ SENDA - Satsning - enhetlig dataforvaltning(2019-2022) Meteorologisk institutt, NILU, NIVA, NINA, Kartverket, Miljødirektoratet, Sluttrapport. MET arkivreferanse 2020/74-6

⁴¹ Kurateringsnivåer for forskningsdata, Sikt, <https://sikt.no/studere-eller-forske/arkivere-data/kurateringsnivaer-forskningsdata>

Dataområdet forventer at datatilbyderne som ønsker å være med har eksisterende systemer for data og metadata, og at connectoren vil kunne samhandle med deres eksisterende systemer.

6.6 Driftsleverandør og suverenitet

Selv om data og tilgangskontroll ligger federert, vil løsningen kunne gjøres rimeligere og enklere hvis en av partnerene er vertskap for dataområdets grunnkomponenter, som API-er, vokabulartjeneste, web-portal for søk og administrasjon o.l. Å finne den riktige balansen mellom federerte og sentraliserte komponenter vil ha stor praktisk betydning. Hovedløsningen må kjøre på norsk jord.

Det er Prosjektleder/Operativ ledelse som har ansvaret for at dette blir overholdt

6.7 Maskinlesbarhet og KI-vennlighet

Maskinlesbarhet skal være en bunnplanke i dataområdets IT-arkitektur. Det er vår antagelse at gode API-er og velkjente internasjonale standarder er essensielle for dette. Det er mulig at mer KI rettede teknologier som MCP/webMCP, eller teknologier som ennå ikke finnes, blir viktige.

Det er viktig at man i implementasjonsfasen har nok frihet til å finne de riktige tekniske valgene i det aktuelle tekniske landskapet, forutsatt at dette ikke går ut over robustheten til dataområdet.

Prosjektledelse/operativ ledelse for dataområdet har ansvaret for å holde seg oppdatert i teknologilandskapet, og sikre at dataområdet utvikles og forvaltes i tråd med beste praksis.

6.8 Tjenestedesign, brukerorientering og samarbeid

Suksessen til dataområdet vises ved at dataområdet blir brukt, og at datatilbyderne ønsker å være en del av det. Dette krever en tett involvering med både datatilbydere og datakonsumenter fra tidlig i utviklingsfasen. Metoder med tjenestedesign og brukerhistorier skal benyttes. Mulighet for påvirkning av prioritering av både funksjonalitet og implementeringsrekkefølgen på brukerhistorier er essensielt for tillitten og brukernytten til dataområdet. Et eksempel på et slikt brukerstyrt prosjekt er BarentsWatch⁴². Tjenestene der er utviklet i tett kontakt med brukerne og de ansvarlige etatene. Den kontinuerlige videreutviklingen av tjenestene skjer også basert på brukerinnspill.

Departementer, Forskningsrådet og akademia må også aktivt med under både utviklings- og driftsfasen. Vi foreslår å følge [Digitaliseringsrådets råd for deling av data i kunnskapssektoren](https://www.digdir.no/digitaliseringsradet/kunnskapsdepartementet-fremtidig-delning-av-data-i-kunnskapssektoren/2007)⁴³. Her vektlegges god, åpen og løpende kommunikasjon med datakonsumenter og datatilbydere, og viktigheten av å aktivere både datakonsument- og datatilbydersiden ved å tydelig vise verdiene løsningen gir, enten det er i næringsvirksomhet

⁴² <https://www.barentswatch.no/en/>

⁴³ Digitaliseringsrådets anbefalingsbrev, Kunnskapsdepartementet: Fremtidig deling av data i kunnskapssektoren, <https://www.digdir.no/digitaliseringsradet/kunnskapsdepartementet-fremtidig-delning-av-data-i-kunnskapssektoren/2007>

eller i forvaltningen. Både datakonsument og datatilbyder er inkludert i styringsmodellen skissert i avsnitt 5.1.

Prosjektledelse/operativ ledelse har ansvar for å sikre involvering og åpen og god kommunikasjon.

6.9 Kravstille bruk av dataområdet

Offentlig forvaltning og forskningsrådet bestiller eller finansierer på annen måte oppdrag som betinger fremskaffing av datasett. Forskningsrådet har i lengre tid stilt krav om datamanagement-planer og bruk av DOler. Framover bør alle slike oppdrag også inkludere krav om at data skal være utstyrt med tilstrekkelig metadata i henhold til dataområdets krav, at dataene skal forvaltes over et bestemt antall år, og at de skal deles i dataområdet.

Departementene har ansvar for å stille krav om at Dataområdet for miljø- og havdata tas i bruk.

6.10 Kompetanse

Etablering og drift av et velfungerende dataområde krever en sammensetning av kunnskap og kompetanse på flere nivåer. Domenespesifikk kompetanse så vel som overordnet forståelse av brukerbehov er essensielt for å levere data inn i dataområdet for deling av miljø- og havdata.

Personell med relevant teknisk bakgrunn er nødvendig for å etablere og videreutvikle de påkrevde infrastrukturkomponentene, samt for drift og vedlikehold av disse. Ressursene må også bidra med veiledning/bistand til datadistributører som innrulleres.

Juridisk og organisatorisk støtte er nødvendig for å definere dataområdets styringsstruktur, etablere tydelige driftsretningslinjer og forvalte formelle kontraktsavtaler mellom de deltakende partene.

Samlet må styringsgruppen stille nødvendig kompetanse for etablering av dataområdet til rådighet, og de selv må ha en dynamisk forståelse av fagmiljøenes behov, det internasjonale datalandskapet og pågående standardiseringsarbeid, for å sikre at dataområdet kontinuerlig kan tilpasse seg teknologiske endringer, nye samfunnsbehov og europeiske rammebetingelser over tid.

6.11 Eierskap / annet

Dataområdet for miljø- og havdata må være tydelig forankret og ha en tydelig eier i departementene. Det må være entydig hvem som har etablering og driftsansvar av den tekniske løsningen ved oppstart av prosjektet. Dette kan med fordel klargjøres i Prop 1 S og tildelingsbrev.

Tillit er en essensiell suksessfaktor for attraktiviteten til dataområdet. En entydig, helhetlig og jevnlig kommunikasjon om dataområdets mål og gevinster, og prosjektets framdriftsplan bidrar til tillitten dataområdet trenger for å nå ut til datakonsumenter og datatilbydere.

Regler for å bli med i dataområdet vil bli utviklet som en del av prosjektet, men en klargjøring om hvem som kan ta beslutningen om nye aktørers deltagelse i dataområdet må være på plass ved prosjektoppstart. Dette bidrar til forutsigbarhet for aktører som ikke er med i oppstartsfasen, og kan forhindre at ressurser spres på mange forskjellige, men lignende aktiviteter eller initiativer.

6.12 Omfang av dataområdet

Før et dataområde for miljø- og havdata kan etableres, videreutvikles og brukes med suksess, kreves det en enighet om et klart definert omfang og at dette omfanget kommuniseres til aktive og fremtidige brukere.

I utgangspunktet bør de aller fleste miljø- og havdata kunne inkluderes i dataområdet. Samtidig vil det ta tid å etablere alle data på grunn av arbeidet med semantisk og syntaktisk interoperabilitet mellom datadomener.

Det må være en felles forståelse for at man i oppstartsfasen må implementere en brukerhistorie for hver kombinasjon av domener før dette kommer på plass.

6.13 Sammenstilling av forutsetninger for Dataområdet for miljø- og havdata

Forutsetning	Vilkår / Hva må være på plass før prosjektstart	Ansvarlig / Rolle	Konsekvens dersom vilkåret ikke er oppfylt
6.1 Langsiktig og forutsigbar finansiering	Sikre forutsigbar nasjonal grunnfinansiering til drift, forvaltning og brukerstøtte.	Oppdragsgivende departementer (KLD, NFD)	Manglende tillit og deltagelse fra datatilbydere, manglende vilje til integrasjonsinvesteringer, og risiko for stagnering eller nedleggelse.
6.2 Åpne data og brukerregistrering	Enighet om at åpne data som hovedregel skal kunne brukes uten registrering, og at registrering eller API-nøkkel bare kreves der det er nødvendig for å styre belastningen på systemet, eller når datatilbyderen krever det	Styringsgruppen for dataområdet	Vedvarende diskusjoner om tilgangsmodeLL som forhindrer prosjektgjennomføringen, manglende kontroll på databruk/overbelastning, samt svekket datasuverenitet
6.3 Datakuratering og FAIRkompatibilitet	Videreføring av domenespesifikk datakuratering, kvalitetskontroll og lisensiering hos datatilbyderne,	Potensielle datatilbyderne	Lavere hastighet på intern jobbing med datakuratering og metadata vil redusere etableringstiden for et dataområde med verdifullt innhold
6.4 Åpen kildekode og gjenbruk	Arkitekturen må bygge på etablerte, åpne kildekodeLøsninger (f.eks. Simpl-open) for å sikre kostnadseffektiv og enkel implementering.	Operativ ledelse / Utviklingsteam	Høye utviklingskostnader, fare for proprietær innlåsing og manglende skalerbarhet.

6.5 Eksisterende systemer for datatilgang	Datatilbyderne må ha fungerende systemer for data/metadata, og det må utvikles standardiserte connectorer med lavest mulig krav til kildesystemene.	Datatilbyderne og Prosjektledelse	Høy terskel for innrulling av nye datatilbydere og lav mengde tilgjengeliggjorte data.
6.6 Driftsleverandør og suverenitet	Gunstig balanse mellom desentraliserte og sentrale felleskomponenter, der vertskap for felleskomponenter ligger hos en partner på norsk jord.	Prosjektleder / Operativ ledelse	Svekket nasjonal datasuverenitet, økte driftskostnader og juridiske utfordringer ved datalagring.
6.7 Maskinlesbarhet og KI-vennlighet	IT-arkitekturen må sikre maskinlesbarhet via åpne API-er og internasjonale standarder, med handlingsrom til å ta i bruk nye KI-teknologier (f.eks. MCP).	Operativ ledelse	Dataene blir uegnede for automatiserte analyser og KI-agenter, og dermed redusert verdi for moderne datakonsumenter.
6.8 Tjenestedesign og brukerinvolvering	Tett brukerinvolvering, bruk av tjenestedesign/brukerhistorier, løpende kommunikasjon, og representasjon av tilbydere/konsumenter i styringsmodellen.	Prosjektledelse / Operativ ledelse	Lav bruksfrekvens og manglende oppslutning, samt risiko for etablering av parallelle/overlappende datadelingstiltak.
6.9 Kravstille bruk av dataområdet	Statlige aktører, departementer og Forskningsrådet må stille krav om at finansierte datasett skal deles i dataområdet med tilstrekkelige metadata.	Departementene, Forskningsrådet og offentlige oppdragsgivere	Frivillig og vilkårlig deling, dobbeltarbeid og manglende utnyttelse av offentlig finansierte data.
6.10 Sammensatt kompetanse	Tilgang på tverrfaglig kompetanse innen domenekuratering (FAIR, ontologier), IT-infrastruktur, samt juridisk/organisatorisk styring.	Styringsgruppen og Datatilbydernes fageksperter	Dårlig datakuratering, forsinket onboarding, juridisk usikkerhet og manglende evne til å skalere løsningen.
6.11 Eierskap og forankring	Tydelig eierskap og forankring i departementene (klargjort i Prop. 1 / tildelingsbrev), entydig plassering av driftsansvar og avklart beslutningsmyndighet for opptak av nye aktører.	Departementene (KLD, NFD)	Uklar ansvarsfordeling, styringsmessig usikkerhet, spredning av ressurser og svekket tillit fra eksterne aktører.
6.12 Omfang av dataområde	Det må være enighet om omfang og framgangsmåte for etablering av Dataområde for miljø- og havdata	Departementene (KLD, NFD)	Opplevelse av manglende gevinster, utålmodighet i realiseringen og oppstart av konkurrerende konsepter

7 Prosjektorganisering, fremdriftsplan og kostnadsoverslag

7.1 Prosjektorganisering

I etableringsfasen foreslås en enkel styringsmodell med de fire initiativtakerne, og med oppdragsgivende departementer som premissgiver. Parallelt utarbeider og prøver prosjektet ut forslaget til varig styringsmodell. Endelig organisering, operatørmodell og juridisk form anbefales først når pilotfasen har gitt erfaring med deltakelse, ansvar, kostnader og brukerbehov.

HI, MET, NILU og NORCE foreslås som initielle pilotpartnere fordi de dekker ulike miljø- og havdatadomener og har modne datadelingsmiljøer. De bør også danne den initielle operative ledelsen av dataområdet. Partnerne bør også utgjøre styringsgruppa til prosjektet. Det etableres også en referansegruppe bestående av relevante aktører. Det er ønskelig at Miljødirektoratet og Kartverket inngår aktivt i denne referansegruppen.

Styringsgruppen har ansvar for å styre etableringen av Dataområdet for miljø- og havdata. Styringsgruppen har ansvar for å vedta planer, regelsett, operativ drift osv. Det er også de som må påse at aktuelt lovverk på området overholdes - offentlighetsloven, GDPR, osv.

En av partnerne får hovedansvaret for etableringen. Denne partneren har også den operative lederen. Lederen håndterer planlegging, oppfølging av planer samt håndtering av ressurser både i utvikling og videre i forvaltning, videreutvikling og driftsfasen. Alle partnerne deltar aktivt inn i arbeidet med utvikling og drift.

7.2 Fremdriftsplan

Gjennom EUs Green Deal program har det de siste årene vært arbeidet mye med å utvikle gode rammeverk for dataområder (data space). En rekke konsortier hvor også norske organisasjoner har deltatt har etablert pilot dataområde. Det er gjort ulike erfaringer med implementasjoner. Noen har utviklet all infrastruktur selv, Geo4Water⁴⁴ (European dataspace for smart communities, DS4SSCC) og MeteoGate (EUMETNET med Meteorologisk institutt som deltager), mens andre har tatt utgangspunkt i bl.a. SIMPL-Open rammeverket og videreutviklet dette (IFE⁴⁵).

I prosjektplanen nedenfor er det planlagt gjenbruk og videreutvikling av allerede eksisterende komponenter. Blant annet SIMPL-Open rammeverket synes nå å ha god funksjonalitet som det er lett å bygge videre på. Et dataområde for miljø- og havdata lar seg best realisere ved trinnvis utvikling av funksjonalitet, og gradvis inkludere flere og flere dataområder.

⁴⁴ <https://www.ds4sscc.eu/geo4water> Norske partner: Sintef, med Osloregionen IPR, Universitetet i Oslo, BLB AS, Meteorologisk institutt

⁴⁵ IFE- Institutt for energiteknikk, oppdrag for Jernbanedirektoratet finansiert gjennom [Motional-prosjektet](https://rail-research.europa.eu/rail-projects/fp1-motional/), Europe's Rail program, <https://rail-research.europa.eu/rail-projects/fp1-motional/>

I appendix B har vi beskrevet en rekke mulige brukerhistorier som omhandler datadeling, og hvor bruk av dataområdet muliggjør sammensetning av data på tvers av domener. Appendixet viser noe bredden av konkrete forenklinger dataområdet kan realisere.

For å sikre tidlige gevinster fra prosjektet og samtidig redusere gjennomføringsrisikoen, bør man implementere en pilotbrukerhistorie tidlig i prosjektet. Det er naturlig at dataene som inngår i denne brukerhistorien kommer fra aktørene som deltar i etableringen av dataområdet. Havforskningsinstituttet, Meteorologisk institutt, Norce og NILU er alle aktører med lang erfaring på deling av data i henhold til FAIR-prinsippene. Brukerhistoriene må velges på en slik måte at dataene som inngår kommer fra ulike domener og har ulike standarder. Dataområdet vil dermed bli utviklet og kvalitetssikret ved bruk av datasett som har god diversitet i metadata og struktur, og dermed er godt egnet til å benyttes ved uttesting av interoperabilitet med moderat risiko.

Prosjektet skal besøke fagmiljøer i land som har kommet lenger med dataområder for miljø og havdata, f.eks. Spania. Hensikten er å lære av feilene de har gjort og av det som har lyktes, før prosjektet tar egne arkitekturvalg. Erfaringene brukes i piloten og legges fram på fagkonferansen.

Prosjektet deles opp i 6 arbeidspakker, hvor hver av arbeidspakkene representerer hver av av innsatsområdene beskrevet i konseptet.

Arbeidspakke 1: Prosjektledelse inkluderer generell prosjektstyring og daglig oppfølging av arbeidet, kvalitetssikring og risikostyring, oppfølging av økonomi.

Arbeidspakke 2: Data inkluderer utvikling av metodikk for interoperabilitet både innenfor og på tvers av domener, standarder for felles metadatakatalog og metadatainnhøsting, samt felles forståelse av innholdet i kvalitetsbeskrivelse og provenans-begrepene.

Arbeidspakke 3: Kontroll og tillit inkluderer omhandler alt som skal til for å sikre tilliten og suvereniteten til dataområdet. Dette omhandler regler for identitet og tilgangsstyring, lisens og kontrakthåndtering, felles regelverk som krav ved bruk av dataområdet(code-of-conduct), regler for informasjonssikkerhet, personvernregler, Terms-of-service, retningslinjer for dataforvaltning, og onboarding og offboarding av datatilbydere.

Arbeidspakke 4: Teknologi inneholder vurdering av ulike eksisterende rammeverk, IT-arkitektur og design, etableringer av innsikt og administrasjonsgrensesnitt, både som brukergrensesnitt og api'er, og konfigurering og tilpassing av komponenter for autentisering, autorisering, og implementering av mal for connectorer mellom datatilbyderne og dataområdet.

Arbeidspakke 5: Kompetanseoverføring og formidling handler om å engasjere brukere. Etablering og videreutvikle brukerhistorier, deltagelse på konferanser og fora hvor dataområder diskuteres. Opprettelse av brukerfora for datatilbydere og datakonsumenter,

etablere brukerstøtte og veiledningsfunksjon både knyttet til dataforvaltning og etablering av verdikjende tjenester.

Arbeidspakke 6: Innrulling av datatilbydere omhandler implementering av brukerhistorier som inkluderer nye datatilbydere og/eller nye domener. Samtidig skal man fortsette tilgjengeliggjøring av nye datasett fra eksisterende leverandører og domener.

Prosjektet bemannes av prosjektleder, dataforvaltere/dataspesialister, tjenstedesign, kommunikasjon og juridisk kompetanse.

Prosjektet deles opp i ulike faser.:

- En proof-of-concept-fase (POC) for å høste tidlige erfaringer med konseptet. Mot slutten av POC fasen skal det gjennomføres en pilotimplementering.
- Fase 1: Operasjonelt dataområde med noe begrenset funksjonalitet og implementering av 2-3 nye brukerhistorier.
- Fase 2: Funksjonaliteten og brukeropplevelsen perfeksjoneres ytterligere basert på brukernes behov og man begynner å systematisk innrulle flere datatilbydere.

I etterkant av hver fase, og før oppstart av ny fase, tas det beslutning om prosjektet skal fortsette eller stoppes opp basert på resultater fra sist fase og oppdaterte estimater for neste fase.

Leveranser POC:

- L1.1 Plan for engasjement av brukere (brukerforum og styringsråd)
- L1.2 Plan og kostnader for fase 1
- L2.1 Felles krav og metodikk for harmonisering av metadata
- Felles regelbok for dataområdet
 - L3.1 Regler for metadatainnhøsting
 - L3.2 Regler for innrulling og fratreddelse
 - L3.3 Regler for håndtering av sikkerhet
 - L3.4 Juridiske regler
- L3.4 Styringsmodell for operativ drift av dataområdet
- L4.1 Vurdering og beslutning om valg av rammeverk for teknisk implementering av dataområdet
- L4.2 Etablert API og enkelt innsiktsportal for tilgang til søk etter data
- L4.3 Etablert autentisering for brukere
- L5.1 Kommunikasjonsplan
- L5.2 Brukermøte for datatilbydere og datakonsumenter
- L6.1 Implementert utvalgt brukerhistorie
- L6.2 Utkast til prioriterte brukerhistorier for innrulling av datatilbydere i fase 1

Pilotbrukerhistorien for POC-en må være en reell verdikjende brukerhistorie med definert gevinst. Samtidig må den være med begrenset antall datasett og kjente datakilder. Brukerhistorien for POC'en må inkludere minimum 2 domener som skal sammenstilles. Dette for å sjekke ut at semantisk, syntaktisk og teknisk interoperabilitet fungerer.

Viktige spørsmål som dataområdet må undersøke og løse:

- Hvor finnes aktuelle datasett?
- Er det riktig geografisk område og tidsperiode?
- Er datasettene tilgjengelige?
- Kan de evt. lastes ned, eller bør man bruke et API?
- Har du lov å bruke dem, hva er lisensbetingelsene?
- Hvordan overlapper parameterbeskrivelsesinformasjonen? Kan parametere fra de ulike datasettene kombineres, og hvordan?

Output fra dataområdet skal være en maskinlesbar tilbakemelding som svarer ut spørsmålene over.

7.2.1 Faseoversikt

Arbeidspakke	POC	Fase 1	Fase 2 ...
Varighet	År 1	År 2-3	År 4
Prosjektledelse	Planlegging, oppfølging, risikostyring og økonomi		
Kostnader	10 Personmåned (PM) pr år		
Data	• Etablere felles regler for metadata-håndtering og syntaktisk og semantisk interoperabilitet • Kravstille harmonisering av datasett for bruk i teknisk interoperabilitet (connectorer) • Kravstille metadatainnhøstning for bruk i teknisk interoperabilitet (connectorer)	• Etablere felles Kvalitet og provenans vurderingsforståelse • Utvide syntaktisk og semantisk interoperabilitet til å gjelde flere domener • Revidere krav fra POC	• Utvider syntaktisk og semantisk interoperabilitet til å gjelde flere domener • Revidere regler
Kostnad	35 PM	25 PM	20 PM

Arbeidspakke	POC	Fase 1	Fase 2 ...
Kontroll og tillit	• Kun åpne data • Regler for sikkerhet og etterlevelse for datadistributør og dataområdet selv • Enkle regler for metadatainnhøstning • Ferdigsstille styringsmodell • Juridiske regler • Etske regler	• Utvide til regler for ikke åpne datasett • Kravstille logg-funksjonalitet for å ivareta datasuverenitet • Regler for identitet og tilgangsstyring • Forbedre regelverk	• Etablere regler for håndtering av ulike lisenser
Kostnader	6 PM	6 PM	5 PM
Teknologi	• Etablere IT-arkitektur • Teste ut rammeverk (Open / Simple API) • Etablere API og Enkel innsiktsportal (søke opp ihht til brukerhistorier), • Etablere enkel autentisering av brukere • Etablere metadata aggregator • Etablere mekanisme for sikkerhet mellom leverandør og dataområdet • Sikre teknisk interoperabilitet ved å etablere mal for connector mellom datatilbyder og dataområdet for de funksjonene som tas fram i POC'en	• Operasjonalisert funksjonalitet tatt fram i POC • Etablere enkel autorisering for tilgang til data og utvidet funksjonalitet for autentisering (federering mot datadistributørs egne, maskinporten eller tilgangssystemer eller MinId) • Etablere lisenstjeneste • Forbedre api og innsikt. Etablere rapport over datasøk og datanedlastinger • Utvide og forbedre mal for connector • Etablere automatisert og manuelt styrt innrulling i dataområde (mulig innenfor allerede	• Etablere KI i dataområdets søkefunksjonalite 5t (forslag til script for sammenstilling av data osv) • Abonnementsordning om nye data/oppdaterte data • Etablere kontrakts og utvide lisenstjeneste • Kobling mot andre dataområder • POC markeds plass • Etablere automatisert og manuelt styrt fratreddelse fra dataområde

Arbeidspakke	POC	Fase 1	Fase 2 ...
	Alle oppgavene inneholder prosjektplanlegging, krav-analyse, og detaljert løsning og tjenstedesign	eksisterende domener) Alle oppgavene inneholder prosjektplanlegging, krav-analyse, og detaljert løsning og tjenstedesign	Alle oppgavene inneholder prosjektplanlegging, krav-analyse, og detaljert løsning og tjenstedesign
Kostnader	30 PM	60 PM	50 PM
Kompetanseoverføring og formidling	• Informere og deltagelse på konferanse • Etablere brukerhistorier og tjenstedesign for fase 1 • Opprette brukerforum	• Informere og deltagelse på konferanse • Etablere brukerhistorier for fase 2 • Etablere veiledningsfunksjon , inkludert brukerdokumentasjon • Etablere brukerstøtte dataforvaltning og tjenstedesign ved bruk av dataområdet	• Informere og deltagelse på konferanse • Etablere markeds plass • Etablere brukerstøtte for bruk av dataområdet og forståelse av metadata • Etablere brukerhistorier for videreutvikling av dataområdets funksjonalitet
Kostnader	5 PM	12 PM	12-20 PM
Pilotering/implementering av brukerhistorier	• Åpne data som i utgangspunktet ikke krever autentisering eller autorisering • Implementering av pilotbrukerhistorie	• Innrullere flere 2-4 nye datatilbydere	• Innrullere flere datatilbydere
Kostnader	5 personmåneder pr datatilbyder (ikke allerede deltager i arbeidspakke 2)	5 personmåneder pr datatilbyder og nye domene	pr datatilbyder og nye domene

7.3 Estimater

I tabellen over prosjektfaser og prosjektoppgaver i kapittel 7.2 har vi anslått prosjektkostnadene våre. Prosjektet vil ha høyest kostnader de første 3 årene. Den største utfordringen i prosjektet er å sikre semantisk og syntaktisk sammenstilling av metadata mellom ulike domener. Det er derfor satt mye ressurser på dette området i POC og fase 1.

Etter det vil man ha mer erfaring med dette arbeidet og ressursinnsatsen fra dataområdets side vil reduseres, og flere ressurser går med til drift, vedlikehold, videreutvikling og veiledning og bistand. Det er anslått at en driftsenhet vil kreve 5-6 årsverk, men det avhenger av hvor mye veiledning som skal tilbys.

Ut fra anslagene vil gjennomføring av POC'en kreve ca 96 månedesverk. Det vil være nødvendig å gjøre nye kostnadsberegninger i løpet av POC'en da konseptet vil være mer kjent.

Det vil være noen kostnader til infrastruktur de første årene. Dette er anslått til å beløpe seg på 50 000 NOK pr år. Reisekostnader og andre administrative kostnader vil i POC fasen beløpe seg til opp mot 200 000 NOK.

Kostnadene til innrullering per datatilbyder og nye dataområder er basert på anslag. Her vil det kunne være store forskjeller, så dette må vurderes for hvert domene.

Estimerte kostnader til POC-fasen er anslått til 18 500 000 NOK i 2026 kroner. Dette inkluderer både personal-, reise- og infrastrukturkostnader. Etter POC-fasen oppdateres estimatene for fase 1 og 2 basert på erfaringene og videre drift.

Det er usikkerhet rundt kostnadsberegningen da det er liten erfaring fra bruk av rammeverkene som SIMPL-Open. SAGE prosjektet under Green Deal programmet har et totalbudsjett på 16 millioner EURO over 3 år. Prosjektet tar fram 10 brukerhistorier som demonstrerer gevinstene ved å benytte mekanismene i et dataområde. Forskjellen på SAGE prosjektet og dette prosjektet er at mye av rammeverkene begynner å bli ferdig implementert så det vil være enklere å sette sammen dataområdet rent teknisk og juridisk nå enn tidligere.

7.4 Mulige finansieringskilder

Det er mange ulike mulige finansieringskilder for etablering av et dataområde for miljø- og havdata. I tabellen nedenfor er det listet opp et utvalg kjente finansieringskilder. Vi har gjort en vurdering om hvorfor disse er aktuelle, og hvilke betingelser det er for å få midler fra den aktuelle kilden. Vi har også angitt antatt størrelsesordenen for støttebeløp.

Noen av finansieringskildene kan bidra til større etableringsoppgaver, og noen er godt egnet til å sikre en varig driftsorganisasjon. Andre finansieringskilder er bedre egnet for å

gjennomføre et eller flere brukerhistorier. Dette gjelder blant annet næringslivsklynger. I en næringslivsklynge vil deltagerne i klyngen gjerne produsere data som de andre i klyngen er avhengig av, eller som er til gjensidig nytte. Det å ha et felles etablert dataområde for deling av disse dataene, og som ivaretar sikkerhet som autentisering, autorisasjon, lisensbetingelser og interoperabilitet, vil forenkle deres jobb betraktelig.

Aktør	Hvorfor vi kan søke	Betingelser for søk	Størrelsesorden
Direkte finansiering over statsbudsjettet	Oppdraget er gitt av NFD og KLD.	Godt prosjektforslag og formål. Avhenger av budsjettprosessen.	Sjekk ut grenser for KS1 og KS2.
Digital Europe	Dataområde vil muliggjøre sikker og interoperabel deling av hav- og miljødata på tvers av organisasjoner og landegrenser, i tråd med EUs satsing på felles europeiske dataområder for grønn omstilling og digitalisering.	Må demonstrere europeisk nytteverdi, involvere et sterkt konsortium, bygge interoperable og delbare datatjenester i tråd med EUs data space-strategi og ha konkrete use-caser for næringsliv og offentlig sektor.	5-10 MEUR
EU Horizon programme Cluster 6, EU RAISE	Muliggjør KI-tilgang på data, arbeider med å forbedre interoperabilitet og dataforvaltning og gjenbruk av bl.a. forskningsdata med andre sektorer.	Må ha en tydelig europeisk samfunnsnytte, bidra til forskning og innovasjon innen hav, miljø eller bærekraft, involvere et internasjonalt konsortium og levere målbare effekter for EUs grønne omstilling.	3-10 MEUR
Nordisk ministerråd	Forutsatt at dataområdet etablerer et felles nordisk samarbeid for deling og utnyttelse av hav- og miljødata på tvers av landegrenser, kan det finansieres via Nordisk ministerråd.	For å få finansiering må minst tre nordiske land (eller relevante aktører derfra) være involvert i tiltaket.	0,5-5 MNOK
Forskningsrådet (NFR infrastruktur)	Vi kan søke da dataområdet også håndterer deling av forskningsdata, og dermed gjør resultatene fra forskning mer tilgjengelig for forvaltning, privat sektor og også for forskere.	Vi antar at vi kun kan få delvis finansiering herfra da dataområdet også vil inneholde aksess til andre typer data enn forskningsdata og at brukerne også er forvaltningen.	50-200 MNOK
Fiskeri- og havbruksnæringens forskningsfinansiering (FHF)	FHF støtter initiativ angående digitalisering og bedre bruk av data i sjømatnæringen.	Prioriteringer i gjeldende handlingsplan FHF. Finansieringen egner seg best til implementering av brukerhistorier	3-20 MNOK
Forskningsmidlene for jordbruk og matindustri (FFL/JA)	FFL støtter sammenstillinger av kunnskapsgrunnlag knyttet til miljø- og havdatas påvirkning på jordbruk og matproduksjon og vice versa	Midler lyses ut gjennom forskningsrådet og egen utlysning fra Landbruksdirektoratet Finansieringen egner seg best til implementering av brukerhistorier	Disponerer opp mot 330 MNOK

Miljødirektoratet	Miljødirektoratet f.eks FKT- Tilskudd til tiltak for å øke kunnskapsgrunnlag i forbindelse med rovvilt. Det kan også potensielt brukerhistorier innenfor forenkling av rapportering som kan søkes dekket	Miljødirektoratet har flere utlysninger og også interne forenklingsarbeide. Finansieringen egner seg best til implementering av brukerhistorier	
Innovasjon Norge	Dataområdet gjør det enklere for private aktører å bygge tjenester på data med kjent kvalitet, opphav og bruksvilkår, og å dele egne data på kontrollerte vilkår	Prosjektet må ha tydelig næringsnytte, utløse innovasjon og verdiskaping i norsk næringsliv, involvere bedrifter som aktive brukere eller partnere, og vise potensial for kommersialisering eller økt konkurransekraft.	1-20 MNOK
Norge digitalt/Kartverket	Samarbeidspartner/ medfinansierer?	Samarbeidspartner/ medfinansierer?	?
Siva -katapult ordning (selskap og industriveks)	Test og innovasjonsarenaer Ved å fokusere på bredere tilgang på data og også muligheten for å nå et bredere publikum med sine data	Kan egne seg for implementering av brukerhistorier	
Regionale forskningsfond (RFF)	Enkelte enkle brukerhistorier, men er i utgangspunktet orientert rundt sine regioner og støtter ikke nasjonale prosjekter.	Aktuelt for mindre brukerhistorier	200 000 NOK
Stiftelsen Det Norske Veritas	Støtter uavhengig forskning som har til hensikt for å trygge liv, eiendom og miljø	Aktuelt for større brukerhistorier og muligens mer funksjonalitet	opp til 100 MNOK
Næringslivsklynger som offshore Norge, Norsk forum for havdata, Aquakultur +++	Vi tilbyr en ferdig infrastruktur som forenkler sikker og trygg deling og sammenstilling av data	Aktuelt samarbeid rundt brukshistorier og kommersialisering	

8. Synergier med eksisterende dataløsninger

8.1 Integrering av eksisterende datadelingstjenester

Noen av dagens eldre og mindre interoperabilitets-tilrettelagte løsninger for deling av data overlapper med løsningen foreslått i konsept 4 Dataområde for miljø- og havdata (kap. 4.4), og kan ha potensial for å absorberes inn i det nye dataområdet. Eksempler på dette er databasegrensesnitt, eldre metadataportaler, eldre autentiseringsløsninger, samt brukeropplæring og kunnskapsformidling knyttet til disse eksisterende løsningene. Nye initiativer kan samutvikles med konseptet, gjennom en gradvis tilpasning av eksisterende utviklingsplaner (veikart) til det nye dataområdet, fremfor å bygge parallelle løsninger fra bunnen av.

8.2 Verdiøkning av eksisterende løsninger

I dag finnes det flere ulike kilder med høyverdig data, men det er ikke enkelt å sette disse sammen på tvers. Et dataområde kan bidra til å realisere gevinster fra eksisterende investeringer, gjennom verdiøkning knyttet til datasikkerhet, datakvalitet, umoden data og sammenstilling av data. Komplementært til dette kan aktører som Kartverket bidra til å bygge bro mellom ulike domener, slik at man kan finne innsikt på tvers. Fremtidig utvikling kan blant annet dreie seg om nye metadatastandarder og nye dataformatstandarder.

8.3 Økt av synlighet av norske data gjennom samarbeid

Å øke synligheten av norske data i europeiske initiativer er et strategisk mål for dataprodusenter, forskningsprosjekter og institutter. Det foreslåtte dataområdet har som ambisjon å være interoperabelt med større europeiske rammeverk, og er dermed på linje med andre dataløsninger i Europa. Flere nasjonale initiativer, som f.eks. Norsk forum for Havdata sitt [datadeling-initiativ](#)⁴⁶ og [Kartverkets nye satsning](#)⁴⁷ kan være relevant å følge opp.

Det finnes også flere relevante internasjonale prosjekter det kan være aktuelt å samarbeide med, som de fleste av initiativene nevnt i 1.2. Økt interoperabilitet – gjennom rikere og maskinlesbar metadata, inkludert prosesshistorikk og beskrivelser av bruksområder, samt krav om bruk av vokabulartjenester og lignende – vil kunne gi en betydelig verdiøkning i form av bedre datasikkerhet og høyere metadatakvalitet.

⁴⁶ Mot en havdatarevolusjon med Norge ved roret!, Nasjonalkomiteen for havforskningstiet og Norsk forum for havdata,
<https://www.forskningsradet.no/portefoljer/klima-miljo/havforskningstiet/offensivt-forslag-havdata-satsning>

⁴⁷ Norge trenger ny samhandlingsplattform for geodata, Kartverket,
<https://www.kartverket.no/geodataarbeid/felleslosningene/nasjonale-felleslosninger-for-geodata>

8.4 Gjenbruk

Potensialet for gjenbruk er allerede stort. Europeiske initiativer har utviklet komponenter, teknisk rammeverk, samt juridisk og kontraktuelt rammeverk, i tillegg til forslag til beste praksis (jf. kap. 1.2). For innlogging og autentisering finnes det blant annet etablerte løsninger som kan integreres. Dataområdet for miljø- og havdata bygger på internasjonale standarder fra fagmiljøene, noe som gir sterke synergier med dataløsninger utviklet både nasjonalt og internasjonalt. Hensikten er ikke å endre eksisterende arbeidsflyt for data – inkludert publisering og distribusjon – men å bygge videre på arbeidet som allerede er utført i fagmiljøene. Målet er å utvide dagens infrastrukturer med lette komponenter som legger til rette for enkel kobling mellom grensesnitt og deling av data. I tillegg blir trygg og sikker utveksling av digitale ressurser stadig viktigere for alle initiativer innen datadeling, og dataområdet utvikler seg i tråd med denne retningen.

9. Instituttene anbefaling

Norce ser stor nytte i et dataområde for miljø- og havdata, spesielt med tanke på evnen til å dele og bruke privat data som ellers ikke ville vært tilgjengelig for forskning. Vi ser for oss at konsept 4, Dataområde for miljø- og havdata, vil kunne realisere betydelig forskningspotensial, og la oss dele data med eksterne partnere som vi ellers ikke vil ha kapasitet til å dele. Deltagelse i et slikt dataområde vil involvere investeringer fra partnerene, så det er essensielt at dataområdet har langsiktig og trygg finansiering så en slik integrering kan rettferdiggjøres. Vi anbefaler derfor Konsept 4, med langvarig finansiering for drift.

Havforskningsinstituttet genererer og benytter en omfattende og mangfoldig mengde data fra havet, inkludert informasjon om fiskebestander, fysiske og kjemiske miljøparametere i havet. Disse dataene stammer fra flere fagområder og krever derfor en integrert tverrfaglig tilnærming for å kunne utnyttes optimalt. Dagens utfordringer knyttet til tverrfaglighet og deling av data på tvers av brukere og sektorer gjør det nødvendig med en løsning som sikrer interoperabilitet, sikkerhet og tillit i datautvekslingen. Et Dataområde for deling av miljø og havdata som muliggjør tilgjengeliggjøring og samhandling av data på tvers av fagdomener vil være avgjørende for å fremme tverrfaglig forskning, samarbeid og innovasjon. På bakgrunn av dette anbefaler Havforskningsinstituttet etableringen av konsept 4 Dataområdet for miljø- og havdata (Data Space), med sikte på å styrke den federerte, sikre og interoperable delingen av data.

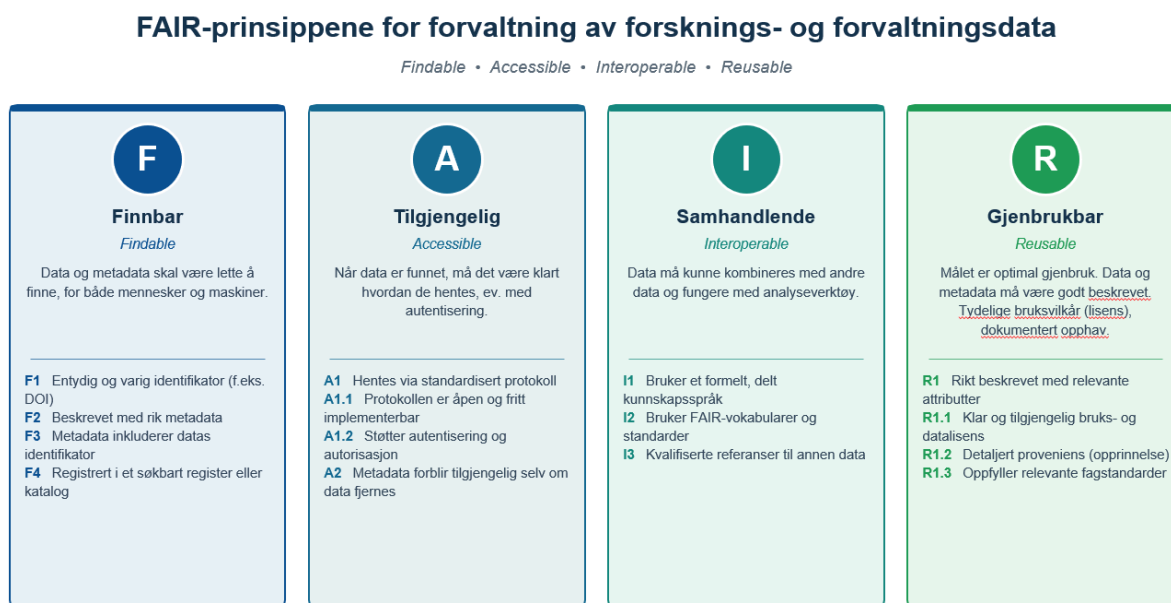
Meteorologisk institutt har som oppgave å levere data om vær, hav og klima til samfunnet gjennom pålitelige og sikre systemer. Samtidig har instituttet et langsiktig engasjement i større initiativer og prosjekter, både gjennom kjernevirksomhet og forskningsprosjekter. MET anbefaler at bruken av standarder står sentralt i dataområdet for å maksimere gjenbruk og bærekraft, også med tanke på en potensiell erstatning av eksisterende løsninger som kan fases ut. For å minimere utviklingskostnadene bør den tekniske arkitekturen bygge på etablerte europeiske åpen kildekode-komponenter, som etter hvert tilpasses dataområdets behov. MET anbefaler videre å utvikle et bredt spekter av bruksområder for å dekke behovene til ulike relevante aktører, inkludert departementene. Et fullt implementert økosystem som i konsept 4 Felles Dataområde for miljø- og havdata vil kunne utvides med flere partnere og koble norske data mot relevante europeiske rammeverk.

NILU stiller seg positive til visjonen om et dataområde for miljø- og havdata, og ser fra egen erfaring med miljø- og atmosfæredata at det fortsatt er reelle utfordringer knyttet til å finne, forstå og faktisk ta i bruk data på tvers av kilder. To forhold bør derfor tillegges vekt i det videre arbeidet. For det første bør dataområdet legge til rette for AI-drevne og LLM-støttede grensesnitt for utveksling og utforskning av metadata. Tradisjonelle metadataportaler forutsetter ofte at brukeren allerede kjenner fagterminologi og katalogstruktur, noe som utestenger mange potensielle brukere. En KI-assistent som oversetter naturlig språk til søk mot de faktiske metadatakatalogene og vokabulartjenestene vil senke denne terskelen betydelig – forutsatt at tjenesten er validert, slik at svarene forankres i reelle, kontrollerte metadata og ikke fremstår som gjettede eller hallusinerter. For det andre er det ikke nok at dataområdet kun peker videre til datakilder og API-er – brukeren må også vises hvordan dataene faktisk kan tas i bruk. Her bør konkrete bruksområder brukes som illustrasjon, ved at man tar utgangspunkt i reelle anvendelser og viser hvordan disse i praksis fører frem til spørring og uthenting av data. På denne måten bygges det bro mellom at data er tilgjengelig

og at data faktisk er i bruk – noe NILU mener er en forutsetning for at dataområdet skal oppleves som reelt verdiøkende.

Appendix A. FAIR prinsipper

FAIR data er data med veldefinerte metadata som er finnbar (Findable), tilgjengelige (Accessible), samhandlende (Interoperable) og gjenbrukbare (Reusable). Et sentralt element i FAIR er å sikre at data er maskinlesbarhet: metadata og data skal være strukturert slik at datamaskiner, ikke bare mennesker, kan finne, hente ut, tolke og kombinere dem med minst mulig manuell innsats. FAIR-prinsippene ble lansert i 2016 og Figur 1 gir en oversikt over de fire hovedprinsippene og tilhørende delprinsipper som vi refererer til i denne seksjonen.



Figur A1: FAIR-prinsippene (Findable, Accessible, Interoperable, Reusable), slik de er formulert av GO FAIR-initiativet. Kilde: Wilkinson, M.D. et al. (2016) «The FAIR Guiding Principles for scientific data management and stewardship», *Scientific Data* 3, artikkel 160018; www.go-fair.org

Delmålene i Figur A1 er essensielle for å sikre maskinlesbarhet og utveksling og tilgang på tvers av områdene klima, miljø og havforskning. FAIR er også en forutsetning for å ta i bruk kunstig intelligens på miljø- og havdata, og for å koble norske data til europeiske initiativer som Green Deal Data Space og EOSC. Uten et felles, godt forankret FAIR-grunnlag vil økosystemet verken kunne levere på ambisjonen om økt kunnskapsgrunnlag, eller på sikt legge til rette for kommersialisering, siden kjøpere betaler for dokumentert kvalitet, ikke for rå datamengde.

FAIR er ikke det samme som åpne data. Et FAIR-datasett kan være beskyttet eller åpent og kommersielt, så lenge det er godt beskrevet, har en klar og maskinlesbar lisens, og kan gjenfinnes og forstås uten manuell etterspørring hos dataeier. Dette skillet er avgjørende for økosystemet: det gjør det mulig å kombinere data på tvers av fagområder og institusjoner, redusere dobbeltarbeid, og bygge nye tjenester og innsikt som ingen enkeltinstitutt kan levere alene.

Prinsippet om åpne data står sterkt i norsk klima, miljø og havforskning og forvaltning. Generelt for et "Felles økosystem for deling av miljø- og havdata" bør datasett deles åpent som mulig, og følge (CC BY 4.0/CC0) i lisenskatalogen, uten registrering som standardvalg. Dette er i samsvar med Norsk lisens for offentlige data (NLOD) som har vært gjeldende anbefaling for offentlig norsk datadeling og innholdet i en ny "Veileder ved tilgjengeliggjøring

av data fra offentlig sektor” sendt på høring fra Dataprioriteringsrådet. Hovedtrekkene i NLOD ny veileder er:

- NLOD er utformet for å være kompatibel med internasjonale åpne lisenser, i første rekke Creative Commons Attribution 4.0 (CC BY 4.0), noe som gjør det enklere å kombinere norske data med internasjonale datasett.
- Data kan brukes, kopieres, distribueres og bearbeides fritt, også til kommersielle formål.
- Data kan kombineres med andre datasett eller inngå i egne produkter og tjenester.
- Bruk krever kreditering: brukeren må oppgi kilde, og eventuelt at data er bearbeidet eller endret.

Data leveres «som de er», uten garanti for nøyaktighet, fullstendighet eller oppdateringsfrekvens, og uten ansvar for feil eller skade som følge av bruk.

For beskyttede data, som ikke kan deles under CCBY4/CC0, må lisenskatalogen i tillegg tilby et lite sett bilaterale/restriktive avtalemaler, jf. drøftingen under Markedsplass (kapittel 4.4.3.4). FAIR og åpne data er to forskjellige ting, og denne forskjellen bør ligge til grunn for hele konseptet: FAIR handler om at data er godt beskrevet, identifiserbare og forståelige FAIR data betyr ikke at data er gratis eller uten vilkår. I FAIR kravene om gjenbruk av data (resusable data) stiller kun R1.1 i FAIR-prinsippene krav om en klar, tilgjengelig og maskinlesbar lisens, som også kan være en kommersiell lisens.

Appendix B. Brukerhistorier

I dette appendixet er det tatt fram en rekke brukerhistorier som illustrerer behovet for å sammenstille datasett på tvers av domener og datatilbydere innenfor miljø og havdata.

B.1 Climate-Driven Air Quality & Extreme Heat Risk Assessment (CLIM-AIR)

B.1.1. User Story and Core Objective

As a municipal urban planner and public health advisor,

I want to evaluate how future climate warming projections (temperature spikes and stagnant weather patterns) interact with local atmospheric air quality (NO₂, PM_{2.5}, and tropospheric ozone),

So that I can design targeted climate adaptation measures, optimize urban green infrastructure, and issue automated health alerts for vulnerable populations during compound heatwave-pollution events.

B.1.2. Required Data Sources and Domains

Domain	Data	Provider	Parameters / Attributes
Climate Projections	Downscaled regional climate model scenarios (e.g., RCP/SSP projections, DestinE/Copernicus)	MET (Meteorological Institute) / ECMWF	High-resolution daily max/min temperature, precipitation patterns, wind velocity, drought indices.
Atmospheric Data	Continuous air quality observations, chemistry transport models	EBAS hosted by NILU	NO ₂ , PM _{2.5} , PM ₁₀ , tropospheric O ₃ concentration, nitrogen/carbon deposition time series.
Terrestrial and Urban	Land cover, building density, vegetation cover	NORCE / Kartverket / Municipalities	Canopy cover, impervious surface percentage, topographically driven microclimate zones.

Socio-Demographic	Vulnerability mapping, population distribution	Public Administration / Statistics	Age distribution, hospital proximity, high-density residential areas.
--------------------------	--	------------------------------------	---

B.1.3. Data Space Interoperability and Technical Workflow

1. **Metadata Discovery & FAIR Alignment:** The user or an automated AI agent queries the federated metadata catalogue using e.g. standardized DCAT-AP profiles. The catalog harvest references both gridded netCDF climate projections and RESTful atmospheric monitoring APIs without duplicating the underlying datasets.
2. **Semantic & Spatial Harmonization:** An automated interoperability connector layer reconciles grid alignments and parameter definitions (e.g., mapping atmospheric chemistry model coordinates to climate scenario spatial grids) using shared domain ontologies.
3. **Trust & Sovereignty Control:** Data access policies are automatically verified through the Control Plane. Public open datasets (e.g., MET weather models) are governed by CC BY 4.0 / NLOD open licenses, while restricted local sensor streams could require federated single sign-on (via Feide/Maskinporten).
4. **AI & Machine Readability:** The ecosystem exposes model-ready endpoints allowing LLM agents and analytical pipelines to execute multi-variable queries and automated risk calculations.

B.1.4. Expected Output and Societal Value

- **Predictive Urban Heat & Pollution Maps:** Identifies specific urban micro-districts where future climate projections predict worsening air stagnation during high-temperature months.
- **Automated Public Health Services:** Enables AI agents to generate real-time health warnings combining meteorological extreme weather alerts with atmospheric air quality thresholds for sensitive groups.
- **Documented Decision-Making:** Provides municipal planners with reproducible, verifiable data workflows to justify investments in green spaces and emissions control zones.

B.2 Arctic Sea Ice Dynamics & Safe Marine Operations (ARCTIC-NAV)

B.2.1 User Story and Core Objective

As an Arctic vessel operator, Coast Guard search-and-rescue (SAR) coordinator, or marine climate researcher,

I want to combine real-time and projected sea ice conditions (concentration, thickness, and drift) with ocean hydrodynamics (currents, surface temperature, and salinity),

So that I can optimize safe, low-emission shipping corridors, execute rapid emergency drift modeling during polar storms or oil spills, and monitor long-term sea ice retreat under climate change scenarios

B.2.2 Required Data Sources and Domains

Domain	Data	Providers	Parameters / Attributes
Sea Ice Data	Satellite sea ice concentration, ice thickness, drift models	MET	Ice edge position, thickness (m), drift speed/direction, ridge density, sea ice age.
Oceanographic Data	Ocean physics, hydrographic observation series	HI (Institute of Marine Research)	Sea surface temperature (SST), salinity, ocean current velocity/direction, bathymetry, water column profiles.
Meteorological Data	Polar weather observations, short/medium-range forecasts	MET	Wind velocity, polar low indicators, air temperature, atmospheric pressure, fog/visibility.
Marine Biodiversity	Marine mammal migration routes, vulnerable benthic habitats	HI / NORCE / Miljødirektoratet	Protected marine areas, seasonal species presence, environmental vulnerability indices.

B.2.3. Expected Output and Societal / Economic Value

- **Optimized Arctic Route Planning:** Blending sea ice thickness, surface current vectors, and wind forecasts reduces fuel consumption and vessel transit risk in ice-infested Arctic waters.
- **Enhanced Polar Emergency Response:** Multi-vector drift modeling (combining sea ice drift, ocean currents, and wind fields) improves Search & Rescue (SAR) accuracy and oil spill containment response times.
- **Climate Impact & Ecosystem Safeguarding:** Provides documented, reproducible observation series for marine researchers and public authorities to evaluate how sea ice decline impacts Arctic marine ecosystems

B.3 High-Latitude Sea Ice Dynamics & Marine Ecosystem Health (OSI-MARINE)

B.3.1. User Story and Core Objective

As a marine ecologist, fisheries manager, or climate risk analyst,

I want to combine EUMETSAT OSI SAF high-latitude satellite products (sea ice concentration, ice edge position, surface temperature, and ice drift) with oceanographic observations (sea surface salinity, currents, and chlorophyll-a concentrations),

So that I can evaluate how rapid polar sea ice retreat interacts with ocean hydrodynamics to impact marine habitat shifts, the timing of spring phytoplankton blooms, and fish stock distribution in Arctic transition zones.

B.3.2. Required Data Sources and Domains

Domain	Data	Provider	Parameters / Attributes
High-Latitude Remote Sensing (OSI SAF)	Satellite High Latitude Ocean, Sea Ice Products	MET (OSI SAF High Latitude Processing Center) / EUMETSAT / Copernicus	Sea Ice Concentration (SIC), Sea Ice Edge, Sea Ice Surface Temperature (IST), High Latitude SST, Sea Ice Drift & Type.

Physical Oceanography	Ocean Physics, Hydrographic Profiles	HI (Institute of Marine Research) / EMODnet / Copernicus Marine	Sea Surface Salinity (SSS), ocean current velocity/direction, CTD water column profiles, mixed layer depth.
Marine Biology and Ecology	Biological Monitoring, Ecosystem Indicators	HI / NORCE / Miljødirektoratet	Chlorophyll-a concentration, primary production timing, acoustic fish biomass estimates (e.g., capelin, polar cod), marine mammal tracking.
Fisheries and Governance	Resource Management, Vessel Activity	Directorate of Fisheries / Coast Guard / Industry	Vessel Monitoring System (VMS) tracking, catch logs, protected marine area boundaries.

B.3.3. Expected Output and Societal / Economic Value

- **Early-Warning Indicators for Ecosystem Shifts:** Automates detection of anomalous sea ice retreat in polar margins and its immediate hydrographic impact on marine species distribution.
- **Sustainable Fisheries Governance:** Supplies evidence-based, reproducible ocean-ice data pipelines for setting sustainable fishing quotas in climate-sensitive Arctic transition zones.
- **European Data Space Synergy:** Directly links national Norwegian marine data with European Green Deal Data Space (GDDS/SAGE) and Digital Twin of the Ocean (EDITO) capabilities.

B.4 Oslofjorden — næringsavrenning fra jordbruk

Problem. Avrenning av næringsstoffer fra jordbruk har over tid forringet marine og kystnære økosystemer i Oslofjorden, med konsekvenser for fiskerier, fiskebestand og annen næring. Eutrofiering kan ikke analyseres isolert: den henger sammen med sedimenttransport, miljøgifter, og plastrester, og hele systemet er koblet til luft. Atmosfærisk deponisjon påvirker jordas næringsretensjon, mens arealbruk påvirker luftkvaliteten gjennom

ammoniakk, lystgass og partikler. Ambisjonen er analyse fra enkeltkilde til kumulativ regional effekt i ett og samme system.

B.4.1 Brukerhistorie og mål

Som rådgiver hos Statsforvalteren med ansvar for vannforvaltning og landbruk,

ønsker jeg å koble data om jordbruksdrift, avrenning, atmosfærisk avsetning og fjordens tilstand, fra enkeltkilde til samlet regional effekt,

slik at jeg kan målrette tiltak i jordbruket der de gir størst effekt for vannkvaliteten i Oslofjorden, uten at sensitive gårdsdata må deles.

B.4.2 Nødvendige datakilder og domener

Domene	Data	Eier	Tilgang
Luft	Deposisjonskjemi, ammoniakk, partikler, pesticider	NILU gjennom EBAS, MET	Åpne
Meteorologi/hydrologi	Nedbør, vind, avrenning, varsel	MET, NVE	HVD, åpne
Ferskvann	Vannkjemi, grunnvann, sedimenttransport	NIVA, Miljødirektoratet	Åpne, fragmenterte
Kyst og hav	Fjordoseanografi, næringssalter, økosystemtilstand	HI	Blandet
Areal og jordbruk	Arealbruk, jordsmonn, driftspraksis, gjødsling	NIBIO, gårdsnivå	Sensitivt
Fjernmåling	Sentinel, drone, proksimal sensing	Copernicus, NIVA, NORCE	Åpne
Punktkilder	Kommunale utslipp, renseanlegg	Kommuner	Ujevn kvalitet
Modeller	Prosessmodeller som referanse	Flere	Tjenester, ikke datasett

Skalaspennet er en egen utfordring: vann og jord fra subregionalt til lokalt nivå, atmosfære fra kontinentalt nivå og ned.

B.4.3 Hva dataområdet løser

- **Maskinlesbare lisenser blir teknisk nødvendig.** Casen bygger et orkestreringslag der KI setter sammen data og modeller på forespørsel. Et slikt lag kan ikke respektere vilkår som bare finnes i en fil. Lisenskatalogen må dekke begge ytterpunkter samtidig: CC BY for observasjonsdata, bilaterale maler for gårds- og utslippsdata.
- **FAIR-minimum gjør analysen mulig.** Uten felles vokabular, kvalitetsbeskrivelse og proveniens blir kobling av luft–jord–vann–kyst manuelt arbeid hver gang. Brukerreisen krever PID-er, semantiske metadata, med mer, samme inngangskrav som vi diskuterer.
- **Federering framfor sentralisering.** Dataområdet trenger ikke eie dataene. For gårdsdata går selv ikke deling. Der er svaret federert læring, som i stedet stiller krav til identitet, tilgangsstyring og logging.
- **Betalingsvilligheten ligger i vurderingen.** Brukerreisen lover vesentlig reduksjon i tid og kostnad for miljøvurderinger. Kundene finnes: Statsforvalteren, kommunene rundt fjorden, landbruksrådgivning, konsulenter.

B.4.4 Tre konsekvenser for konseptet

De fire initielle institusjonene dekker ikke casen alene. NIVA og NIBIO må med, så arkitekturen må tåle påbygg. AI-bruk skjerper FAIR-kravene fra ambisjon til forutsetning. Og hvis dataområdet får finansiering, er dette en pilot med reelle krav og frist.

B.5 Norges klimagassbalanse – fra målte konsentrasjoner til samfunnsøkonomiske konsekvenser (KLIMA-BALANSE)

Bakgrunn. Norsk klimapolitikk bygger på kunnskap om utslipp og opptak av klimagasser. Utslippene beregnes i det nasjonale utslippsregnskapet, opptaket i skog og arealer beregnes separat, og framtidige utviklingsbaner vurderes i klimamodeller og økonomiske modeller. Disse kunnskapsgrunnlagene utvikles i dag i stor grad hver for seg, ved ulike institusjoner og med ulike metoder. Målte konsentrasjoner i atmosfæren gir i tillegg en uavhengig kontroll av om beregnede utslipp og opptak er i samsvar med det som faktisk observeres. Et helhetlig bilde forutsetter at disse kildene kan kobles.

B.5.1 Brukerhistorie og mål

Som rådgiver i forvaltningen med ansvar for oppfølging av klimamålene,
ønsker jeg å sammenstille målte klimagasskonsentrasjoner, utslippsstatistikk, beregnet skogtilvekst og modellerte klima- og økonomiske utviklingsbaner,
slik at jeg kan vurdere hvor robust måloppnåelsen er, og hvilke tiltak som gir størst effekt per krone.

B.5.2. Nødvendige datakilder og domener

Domene	Data	Tilbyder	Tilgang
--------	------	----------	---------

Atmosfære	Langtidsserier for konsentrasjoner av CO ₂ , CH ₄ og andre klimagasser (bl.a. Zeppelin og Birkenes)	NILU gjennom EBAS	Åpne
Utslipp	Nasjonalt utslippsregnskap fordelt på sektorer og kilder	SSB, Miljødirektoratet	Åpne
Klimamodellerin g	Scenarioer fra jordsystemmodellen NorESM	NORCE, MET m.fl.	Åpne, store datamengder
Skog og arealbruk	Tilvekst, stående volum og karbonlager i skog (Landsskogtakseringen)	NIBIO	Aggregerte data åpne; prøveflatenes posisjoner skjermet
Økonomi	Modeller som kobler utslipp, opptak og tiltak til økonomiske konsekvenser	NTNU	Modelltjenester; vilkår må avklares

B.5.3 Hva dataområdet løser

- **Finnbarhet på tvers av institusjoner:** Den felles metadatakatalogen gjør det mulig å finne relevante datasett og modellresultater hos NILU, SSB, NORCE, MET, NIBIO og NTNU uten å kjenne hver enkelt kilde på forhånd.
- **Proveniens og etterprøvbarehet:** Proveniensinformasjon dokumenterer hvilken versjon av utslippstallene og hvilket modellscenario en analyse bygger på. Det er avgjørende for at resultatene skal være etterprøvbare i en politisk beslutningsprosess.
- **Maskinlesbare vilkår:** Lisenstjenesten skiller mellom åpne data, skjermede data og modelltjenester med egne vilkår, slik at åpne data forblir åpne, mens tilgangskontroll kun gjelder der det er nødvendig.
- **Europeisk kobling:** Økosystemet legger til rette for kobling mot europeiske forskningsinfrastrukturer og rapporteringsforpliktelser på klimaområdet.

B.5.4. Hva dataområdet ikke løser alene

Den største utfordringen er granularitet. Datakildene beskriver prosesser på svært ulike skalaer i tid og rom: målinger ved enkeltstasjoner, utslipp per sektor og år, modellresultater i rutenett, skogdata fra et utvalg prøveflater og økonomiske modeller per næring. Å koble disse på en faglig forsvarlig måte krever valg om aggregering, fordeling og usikkerhet som ikke kan automatiseres med dagens verktøy. Det finnes i hovedsak to veier videre: å vente til KI-løsninger blir i stand til å gjøre slike koblinger, trolig med lavere krav til hvordan kildene er

tilrettelagt, eller å samle eksperter i et tverrfaglig prosjekt der de bidrar med data og kompetanse og gjør koblingene manuelt.

De to veiene utelukker ikke hverandre. Et tverrfaglig pilotprosjekt kan dokumentere koblingene, som aggregeringsregler, begrepskoblinger og usikkerhetsvurderinger, som metadata og proveniens i dataområdet. Da blir ekspertarbeidet gjenbrukbart, og det gir samtidig det kvalitetssikrede grunnlaget framtidige KI-løsninger trenger. Uten et felles økosystem blir resultatet et enkeltstående prosjekt; med dataområde for deling av miljø- og havdata blir det en varig nasjonal kapasitet.

B.5.5. Forventet resultat og samfunnsverdi

- **Uavhengig kontroll av utslippsregnskapet:** Målte konsentrasjoner kan sammenholdes med beregnede utslipp og opptak, og avvik kan identifiseres tidligere.
- **Bedre grunnlag for oppfølging av klimamålene:** Forvaltningen får et samlet og dokumentert kunnskapsgrunnlag for å vurdere hvor robust måloppnåelsen er.
- **Kostnadseffektive tiltak:** Et sammenlignbart grunnlag gjør det mulig å vurdere ulike tiltak opp mot hverandre, for eksempel skogtiltak sammenlignet med utslippskutt i andre sektorer.
- **Utvidbar arkitektur:** Brukstilfellet viser at dataområdet må kunne utvides med aktører utover de fire partnerne, blant annet SSB, NIBIO og NTNU.

B.6 Oslofjorden – fra kilde på land til samlet effekt i fjorden (OSLOFJORD)

Bakgrunn. Oslofjorden er i dårlig miljøtilstand. Tilførsler av næringsstoffer og miljøgifter fra jordbruk, avløp, industri og atmosfære bidrar til overgjødning, oksygensvikt og tap av viktige leveområder, med konsekvenser for fiskebestander og annen næringsvirksomhet. Regjeringen har lagt fram en helhetlig tiltaksplan for fjorden, og fiskeriforvaltningen har innført strenge reguleringer for kysttorsk. Effekten av tiltakene kan likevel ikke vurderes isolert. Overgjødning henger sammen med sedimenttransport, miljøgifter og plastforurensning, og hele systemet er koblet til luft: atmosfærisk avsetning påvirker jordas evne til å holde på næringsstoffer, mens arealbruk påvirker luftkvaliteten gjennom utslipp av ammoniakk, lystgass og partikler. Ambisjonen er å kunne analysere sammenhengen fra enkeltkilde til samlet regional effekt i ett og samme kunnskapsgrunnlag.

B.6.1. Brukerhistorie og mål

Som saksbehandler hos Statsforvalteren eller i fiskeriforvaltningen, eller miljøansvarlig i en industribedrift ved fjorden,

ønsker jeg å følge næringsstoffer og miljøgifter fra kilde til samlet effekt i fjorden,

slik at tiltak i jordbruket, avløpssektoren, industrien og fiskeriforvaltningen kan vurderes samlet og prioriteres der effekten er størst.

B.6.2. Nødvendige datakilder og domener

Domene	Data	Tilbyder	Tilgang
Luft	Avsetningskjemi, ammoniakk, partikler, plantevernmidler	NILU gjennom EBAS, MET	Åpne
Meteorologi og hydrologi	Nedbør, vind, avrenning, varsler	MET, NVE	Åpne, inkl. datasett med høy verdi
Ferskvann	Vannkjemi, grunnvann, sedimenttransport	NIVA, Miljødirektoratet	Åpne, men fragmenterte
Kyst og hav	Fjordoseanografi, næringssalter, økosystemtilstand	HI	Delvis åpne
Areal og jordbruk	Arealbruk, jordsmonn, driftspraksis, gjødsling	NIBIO, gårdsnivå	Sensitive
Fjernmåling	Satellitt- og dronedata	Copernicus, NIVA, NORCE	Åpne
Kommunale punktkilder	Kommunale utslipp, renseanlegg	Kommuner	Ujevn kvalitet
Industri	Utslipp til vann og luft, egenkontroll og resipientovervåking	Chemring Nobel, Miljødirektoratet	Delvis åpne; bedriftsdata kan være kommersielt sensitive
Modeller	Prosessmodeller for luft, jord, vann og fjord	Flere	Modelltjenester

B.6.3. Hva dataområdet løser

- **Maskinlesbare lisenser blir en teknisk nødvendighet:** Brukstilfellet forutsetter at data og modeller kan settes sammen på forespørsel, også ved hjelp av KI. Vilkår som kun finnes i et PDF-dokument kan ikke håndheves automatisk. Lisenstjenesten må derfor håndtere begge ytterpunkter samtidig: åpne lisenser for observasjonsdata og avtalebaserte vilkår for gårds-, industri- og utslippsdata.

- **FAIR-minimum gjør analysen mulig:** Uten felles vokabular, kvalitetsbeskrivelse og proveniens blir koblingen mellom luft, jord, vann og kyst manuelt arbeid hver gang analysen skal gjøres.
- **Føderering framfor sentralisering:** Dataområdet trenger ikke eie dataene. For sensitive gårdsdata er deling i mange tilfeller ikke aktuelt; da må analysen gjøres der dataene ligger, gjennom føderert analyse, noe som stiller krav til identitetshåndtering, tilgangsstyring og logging.
- **Industrien som både tilbyder og bruker:** Med en industriaktør som Chemring Nobel ved Sætre viser brukstilfellet hvordan private virksomheter kan delta på egne vilkår. Bedriften kan dele utslipps- og overvåkingsdata med definerte bruksvilkår, og samtidig bruke dataområdet til å finne data til å dokumentere egen miljøpåvirkning i sammenheng med øvrige kilder i fjorden.

B.6.4. Forventet resultat og samfunnsverdi

- **Raskere og rimeligere miljøvurderinger:** Et samlet kunnskapsgrunnlag kan redusere tid og kostnader ved miljøvurderinger vesentlig, for Statsforvalteren, kommunene rundt fjorden, fiskeriforvaltningen, landbruksrådgivningen, industrien og rådgivende konsulenter.
- **Samlet tiltaksvurdering på tvers av sektorer:** Tiltak i jordbruk, avløp, industri og fiskeriforvaltning kan vurderes i sammenheng, slik at innsatsen rettes dit den gir størst effekt for fjorden.

Konsekvenser for konseptet: De fire partnerinstitusjonene dekker ikke behovet alene, slik at arkitekturen må kunne utvides med blant andre NIVA, NIBIO, NVE og industriaktører. Bruk av KI gjør FAIR-kravene til en forutsetning snarere enn en ambisjon. Et tydelig skille mellom åpne og skjermede data er nødvendig for at åpne data ikke underlegges unødvendig tilgangskontroll.

B.7 Arendal Kommune — Forbedring av beslutningsgrunnlaget for kommunal arealforvaltning i marine områder

Problemet. Kommuner langs kysten har et betydelig ansvar for å balansere bærekraftig bruk og vern av marine naturverdier, men mangler ofte et helhetlig og tilgjengelig kunnskapsgrunnlag. Dette svekker kvaliteten på beslutninger og gjør det vanskelig å vurdere konsekvenser av tiltak som utbygging, høringsaktiviteter, ledningsanlegg og annen aktivitet i sjø.

B.7.1. Brukerhistorie og mål

Som forvalter i Arendal kommune med beslutningsansvar i marine arealsaker

ønsker jeg å ha en helhetlig og tverrfaglig oversikt over kystnære områder i Arendal kommune,

slik at kyst- og sjøområder i Arendal forvaltes og brukes på en bærekraftig måte, med hensyn til effektene i havmiljø og biologisk mangfold.

B.7.2. Nødvendige datakilder og domener

Domene	Data	Eier	Tilgang
Dybdedata	Multistråle ekkolodd	Kartverket/HI/Arendal Kommune	Egne data, klassifisert, begrenset tilgang
Biologisk mangfold	Artsobservasjoner, modellerte habitater, forvaltningsprioritert natur (Tareskog, Blandet tangsamfunn, Ålegress, Mosaikkbunn, Oppvekstområder, Vandringsruter, Konnektivitetes-korridorer, Plankton)	HI/NINA/Miljødirektora tet/Andre kilder	Åpne, fragmenterte
Bunnskart	Kornstørrelse, Sedimenttransport	NGU	Åpent
Meteorologi/ Hydrologi	Nedbør, vind, Avrenning, Farevarsel	MET, NVE	HVD, åpne
Miljøkjemi	Vannkjemi, Grunnvann, Næringsalter, Oksygen, annet?	HI, NIVA, andre kilder	Åpne, fragmenterte
Topografiske kart	Terskler og grunner, Undersjøiske fjelltopper, Strømrrike nes, Elveutløp, Bløtbunn (NiN-type?)	Lages basert på egne bathymetridata, Elveutløp fra NVE, NiN fra HI	

Fjernmåling	Sentinel, drone, proksimal sensing	Copernicus, NIVA, NORCE	Åpne
Eksisterende infrastruktur	Kommunal infrastruktur som Renseanlegg, bygninger etc	Arendal kommune	Egne data, full intern tilgang?
Menneskelig påvirkning	Utslipp, vrak, Modifisert kyst, Kunstige rev	Ulike kilder	Åpne

B.7.3 Hva dataområdet kan bidra til å løse

Å gjøre eksisterende forvaltningsrelevante data tilgjengelig og intuitive for offentlig kommunal forvaltning.

Prosedyrer og kunnskap om dette kan også hjelpe andre norske kystkommuner til å få et bedre kunnskapsgrunnlag knyttet til sine marine områder.

- **KI fra en sikker datakilde:** Bruk av kunstig intelligens (KI) i offentlig sektor kan brukes til å klassifisere arter og bunntyper, analysere store datamengder og utvikle søkbare løsninger som kan brukes av både fagpersoner og innbyggere.
- **Maskinlesbare metadata med fylldig og maskinlesbar informasjon om egnede bruksområder og provenans.** KI setter sammen data og modeller på forespørsel, det er da en forutsetning at metadata er fylldig og maskinlesbart. Et forslag er å bruke tidligere høringssvar til å forbedre «hjelp til bruk».
- **FAIR-minimum gjør analysen mulig.** Uten felles vokabular, kvalitetsbeskrivelse og proveniens blir kobling av ulike datakilder manuelt arbeid hver gang.

To konsekvenser for konseptet:

KI-bruk skjerper FAIR-kravene fra ambisjon til forutsetning.

Arendal er én av 280 norske kystkommuner, så gevinstpotensialet er stort.

B.8 Rømt laks

Problemet: Data fra en lang rekke kilder blir samlet og presentert i rapporten «Rømt laks» fra HI. Dataene inngår videre i andre prosjekter, f.eks. i Risikorapporten, prosjekter hos Fiskeridirektoratet etc. Rapporten har vært en pdf frem til 2024, deretter en PowerBI rapport uten mulighet for å få tak i selve datagrunnlaget, dette sendes rundt som filer på etterspørsel.

B.8.1 Brukerhistorie og mål

Som forsker på rømt laks v/ HI,

ønsker jeg å hente data fra laksetellinger, anadrome strekninger for alle norske elver, fra de relevante enkeltkildene, samle disse og kunne videreformidle datautvalget til publikum.

slik at jeg kan presentere årlig status for rømt laks i norske elver, og andre kan bygge videre på samme datagrunnlag.

B.8.2. Nødvendige datakilder og domener

Domene	Data	Eier	Tilgang
Elvenettverk	Vassdragsnavn, ID-nummer, anadrome strekninger etc for norske elver	NVE	Åpent, tilgjengelig gjennom ELVIS API. Lakseregisteret benyttes også
Biologi	Drivtelling	HI, NINA, NORCE, FVB/SNA, LFI, Rådgivende Biologer AS, Naturtjenester i nord, Veterinærinstituttet	Lite tilgjengelig
Biologi	Individdata	HI, NINA, NORCE, LFI, Rådgivende Biologer AS, Naturtjenester i nord, Veterinærinstituttet	Lite tilgjengelig
Statistikk	Fangststatistikk	SSB	

B.8.3 Hva dataområdet kan bidra til å løse

Å forenkle innsamling og bruk av data på tvers av domener, og å kunne videreformidle datagrunnlaget for rapporten på en ryddig og federert måte, uten å mellomlagre og videredistribuere andre etaters data.

Datagrunnlaget brukes blant annet av Eksterne ekspertgrupper (VRL, ICES) og internt til videre rapportering i f.eks. Risikorapporten.

B.9 Fra spredte naturdata til et dokumentert kunnskapsgrunnlag for arealbeslutninger

Brukercase: Naturdata i bruk fra NINA

Kjerne i casen

En kommunal miljørådgiver skal sammenligne arealalternativer i en kystkommune. For å vurdere naturverdier, endring over tid og samlet belastning må rådgiveren kunne finne og kombinere data fra flere institusjoner - med opphav, kvalitet, skala, versjon og faglige begrensninger intakt. Verdien av et felles økosystem ligger i å gjøre dette mulig på en gjenbrukbar og etterprøvbar måte, ikke i å opprette enda en isolert naturdatabase.

B.9.1 Brukercasen i kortform

Kommunene tar arealbeslutninger som ofte får varige og vanskelig reversible konsekvenser for naturen. Samtidig ligger kunnskapen de trenger spredt på tvers av datakilder, fagområder og institusjoner. Det gjør det krevende å få et samlet, oppdatert og etterprøvbart bilde tidsnok til at kunnskapen faktisk påvirker beslutningen.

Casen tar utgangspunkt i en kommunal miljørådgiver som skal oppdatere naturgrunnlaget til kommuneplanens arealdel i et landskap med skog, myr, semi-naturlig mark, vassdrag, strandsone og tilgrensende sjøområder. Rådgiveren skal sammenligne foreslåtte arealendringer og svare på ett hovedspørsmål:

Hvilke naturverdier og økologiske sammenhenger finnes i landskapet, hvordan utvikler de seg over tid, og hva vil nye inngrep bety når de ses sammen med tidligere inngrep og andre påvirkninger?

Naturmangfoldloven § 10 krever at påvirkning på et økosystem vurderes ut fra den samlede belastningen økosystemet er eller vil bli utsatt for.⁴⁸ To arealalternativer kan derfor ha samme arealomfang, men gi svært ulike konsekvenser fordi de berører ulike naturtyper, leveområder, økologiske forbindelser eller påvirkningsveier til vassdrag og kyst.

Det ønskede sluttproduktet er ikke ett nytt kart, men et dokumentert kunnskapsgrunnlag som viser hva som er observert, hva som er modellert eller beregnet, hvor kunnskapen er god, hvor den er svak, og hvilke forskjeller mellom alternativer som kan underbygges.

B.9.1.1. Behovet er dokumentert - og aktuelt

Behovet er ikke bare et lokalt effektiviseringsproblem. KLDs kunnskapsstrategi 2025-2030 peker på varierende datakvalitet, manglende data og at eksisterende miljødata ofte er vanskelige å finne, sammenstille og gjenbruke. Strategien vektlegger bedre koordinering og deling på tvers av sektorer og institusjoner.^{49,50} OECD har tidligere pekt på behovet for bedre kobling mellom lokale arealbeslutninger og samlede konsekvenser for natur og økosystemtjenester.⁵¹

Riksrevisjonens undersøkelse fra 28. april 2026 er særlig relevant for denne casen. Den konkluderer med at kommunene ofte mangler tilstrekkelig oversikt over naturverdier, og

⁴⁸ Lov om forvaltning av naturens mangfold (naturmangfoldloven), § 10 om økosystemtilnærming og samlet belastning.

⁴⁹ Menon Economics og A-2 Norge (2021). Utredning av fremtidens miljødata. Menon-publikasjon nr. 153/2021, utarbeidet for Miljødirektoratet.

⁵⁰ Klima- og miljødepartementet (2025). Klima- og miljødepartementets kunnskapsstrategi 2025-2030, særlig kapittel 5.1 om miljødata.

⁵¹ OECD (2022). OECD Environmental Performance Reviews: Norway 2022, kapitlet om arealbruk og naturmangfold.

anbefaler blant annet systematisk og sammenlignbar kunnskap om naturtap og forringelse og kvalitetssikrede naturdata som kan brukes på tvers av sektorer.⁵²

NINAs behovsintervju med Akershus fylkeskommune konkretiserer det samme fra brukersiden: data må kunne sammenstilles på tvers av skala, alder og kartleggingsdekning, usikkerhet må være forståelig, og kunnskapshull må bli synlige i den konkrete planprosessen.⁵³

Utviklingen av kommunale naturregnskap gjør behovet enda mer konkret. Miljødirektoratet planlegger første versjon av en felles digital løsning tidlig i 2027. Naturregnskap, arealplanlegging, restaurering og naturovervåking bygger i stor grad på det samme underliggende naturfaglige kunnskapsgrunnlaget. Det gir en klar gevinst dersom data og dokumenterte koblinger kan etableres én gang og gjenbrukes på tvers av formål.^{54,55}

B.9.1.2. Fra data til beslutning - en konkret brukerreise

1. Avgrense saken Rådgiveren avgrenser planområdet og den økologisk relevante konteksten - for eksempel berørte nedbørfelt, sammenhengende skogområder, strandsone og tilgrensende sjøareal - og velger tidsperioden som skal vurderes.
2. Hente og forstå kunnskapsgrunnlaget Arbeidsflyten finner relevante data og viser samtidig dekning, alder, metode, presisjon, tilgangsvilkår og usikkerhet. Feltobservasjoner og overvåkingsserier skilles fra avledede kart og modellresultater.
3. Sammenligne alternativer og samlet belastning Naturtyper, artsdata, arealendring, skogegenskaper og økologiske forbindelser sammenholdes med eksisterende og planlagte inngrep. Der det er relevant kobles påvirkning videre til hydrologi, vannkvalitet, kyst og sjø. Alternativene vurderes med dokumenterte metoder på den skalaen datagrunnlaget støtter.
4. Synliggjøre usikkerhet og kunnskapshull Fravær av artsfunn behandles ikke som dokumentert fravær. Arbeidsflyten skiller mellom nullobservasjoner, manglende kartlegging og data som finnes, men ikke er tilgjengeliggjort. Brukeren får dermed vite hvor beslutningsgrunnlaget er robust og hvor supplerende undersøkelser kan være nødvendige.
5. Levere og gjenbruke Resultatet er en etterprøvbar kunnskapspakke med kart, tidsserier, indikatorer, kilder og faglige forbehold. Data forblir hos dataeierne, mens felles metadata, tilkoblinger og dokumenterte arbeidsprosesser gjør grunnlaget oppdaterbart og gjenbrukbart.

Et viktig skille

Observasjoner, bearbejdede kartprodukter, forskningsbaserte modeller og operative beslutningsstøttetjenester er ikke det samme. Et felles økosystem må bevare denne forskjellen. Versjon, metode, modenhet, skala og bruksbegrensninger må følge data og produkter hele veien til brukeren.

B.9.2 Hvilke data må kobles - og hva tilfører de?

Casen trenger både biologiske naturdata og data som beskriver landskapet og påvirkningene. Poenget er ikke å samle alt i én database, men å gjøre relevante kilder

⁵² Riksrevisjonen (2026). Kartlegging av naturverdier og klimaeffekter som grunnlag for beslutninger i arealforvaltningen. Dokument 3:9 (2025-2026), offentliggjort 28. april 2026.

⁵³ NORA (2026). Oppsummering av behovsintervju med Akershus fylkeskommune, 15. juli 2026. Internt arbeidsnotat.

⁵⁴ Miljødirektoratet (2026). Kommunale naturregnskap - hva kommer i første versjon? Publisert 22. september 2026.

⁵⁵ United Nations mfl. (2021). System of Environmental-Economic Accounting - Ecosystem Accounting (SEEA EA).

søkbare, forståelige og kombinerbare gjennom felles metadata og grensesnitt. Geonorge er allerede den nasjonale geoportalen for kartdata og annen stedfestet informasjon; et naturdata-økosystem bør bygge videre på slike eksisterende infrastrukturer fremfor å duplisere dem.⁵⁶

Tabell: Tabellen er kun et eksempel og ingen uttømmende oversikt over datakilder

Datatrinn	Eksempler på kilder	Hva dette tilfører brukerens oppgave
Observasjoner og feltdata	Artskart/GBIF, Naturbase, ANO og historiske TOV-data, TOV-E, NINAs insektovervåking	Dokumenterer registrerte arter, naturtyper og biologisk tilstand over tid. Må brukes med informasjon om kartleggingsinnsats, metode, skala og dekning.
Areal- og endringsdata	Norge i rødt, hvitt og grått; plan-, basis- og andre geodata via Kartverket/Geonorge	Viser arealbruk, nedbygging, terreng og geografisk kontekst. Gir grunnlag for å koble naturdata til faktiske og planlagte inngrep.
Skog og landskapsstrukturer	SR16, naturskogkart, NINAs konnektivitetskart og GreenPlan	Belyser skogegenskaper, fragmentering og sammenheng mellom leveområder. Modellresultater og prototyper må presenteres med metodeversjon og bruksbegrensninger.
Kyst og hav	SEAPOPOP, SEATRACK og SeaBee	Tilfører tidsserier for sjøfugl, sporingsbasert arealbruk gjennom året og detaljert kartlegging av kysthabitater der dekning finnes. Relevans avhenger av art, livsstadium, skala og kartleggingsområde.
Påvirkningsdata	Hydrologi, vannkvalitet, forurensning, klima og menneskelig aktivitet	Gjør det mulig å undersøke påvirkningsveier og mulige forklaringer på naturendringer. Konkrete datasett og dataeiere må velges ut fra beslutningen og området som analyseres.

B.9.3 Land, ferskvann, kyst og hav må kunne ses i sammenheng

Den terrestriske delen er bærende i casen: skog, myr og semi-naturlig mark må beskrives med både areal, naturtype, biologisk tilstand og landskapsstruktur. For myr er hydrologi og inngrep sentralt. For skog må blant annet arealtap, skogegenskaper, kanteffekter og fragmentering kunne sees i sammenheng. For semi-naturlig mark er endringer i leveområder for arter og pollinatorer særlig relevante.

⁵⁶ Kartverket / Geonorge. Nasjonal geoportal for kartdata og annen stedfestet informasjon, med metadata, visnings- og nedlastings tjenester.

Samtidig stopper ikke påvirkning ved administrative eller økologiske grenseflater. I en kystkommune kan arealendringer på land påvirke vassdrag, strandsone og sjø. Når saken krever det, må derfor naturgrunnlaget suppleres med relevante biologiske overvåkingsdata fra ferskvann og marine områder fra blant andre NINA, NIVA og Havforskningsinstituttet. Hvilke konkrete datasett som skal inngå, må avklares for den enkelte brukerreise.

Dette er også det som gjør casen egnet for andre arealbeslutninger. Den samme grunnlogikken kan brukes når påvirkningen kommer fra eksempelvis havvind, oppdrett, bryggeanlegg eller annen utbygging i sjø: økologisk relevante data kombineres på riktig skala, med tydelig sporbarhet og metode.

B.9.4 Hva et dataområde må tilføre

Mye av dataene finnes allerede. Merverdien ligger derfor først og fremst i å gjøre dem mulig å bruke sammen uten at faglig mening og usikkerhet går tapt. For denne casen betyr det at økosystemet må gjøre følgende mulig:

- Finne relevante data og tjenester på tvers av institusjoner uten å kjenne alle kildene på forhånd.
- Forstå hva dataene representerer: metode, tidspunkt, geografisk dekning, oppløsning, kvalitet, lisens og egnet bruk.
- Koble observasjoner, tidsserier, kart og modeller med felles sted- og tidsreferanser.
- Beholde sporbarhet til dataeier, kilde, versjon og metode når data inngår i analyser og beslutningsgrunnlag.
- Skille tydelig mellom observerte data, avledede produkter, modellresultater og faglige vurderinger.
- Oppdatere sammenstillinger når kildedata eller kunnskap endres, uten å bygge hele arbeidsflyten på nytt.
- Gjøre data maskinlesbare og dokumenterte nok til videre analyse, modeller og på sikt KI-baserte tjenester.

Casen forutsetter dermed en føderert logikk: dataeierne beholder ansvaret for kildedataene, mens felles standarder, metadata, tilkoblinger og arbeidsprosesser gjør dataene sammenhengende fra brukerens ståsted. Dette reduserer også risikoen for at forskningsbaserte produkter blir løsrevet fra forskerne, metodene eller versjonene som gir dem mening.

B.9.5. Samme naturgrunnlag - flere anvendelser

Arealplanlegging og samlet belastning: Sammenligne alternativer og vise hvordan nye og eksisterende inngrep berører naturtyper, leveområder, økologiske forbindelser og påvirkningsveier.

Naturregnskap og arealregnskap: Gjenbruke areal-, naturtype- og endringsdata og supplere med tilstandsinformasjon der utvalgsdesign og lokal dekning gir grunnlag for det. Observerte endringer må holdes adskilt fra modellerte scenarioer.

Restaurering og naturovervåking: Identifisere områder der restaurering kan være relevant, og følge utvikling etter tiltak med vegetasjon, hydrologi, artsdata og landskapsstruktur.

Konsekvensutredning og natur-/arealrisiko: Gi et tydeligere utgangspunkt for hvilke naturverdier, påvirkninger og kunnskapshull som må undersøkes videre - både for offentlig forvaltning og næringsliv.

Forskning og kunnskapsutvikling: Koble observasjoner og tidsserier fra ulike økosystemer og institusjoner, dokumentere kunnskapshull og gjøre analyser etterprøvbare og gjenbrukbare.

B.9.6 Forventet merverdi

For forvaltningen er den viktigste gevinsten et tidligere og mer etterprøvbart bilde av naturverdier, arealendring, fragmentering, kunnskapshull og samlet påvirkning. For næringslivet kan samme grunnlag tydeliggjøre naturhensyn og dokumentasjonsbehov tidligere i prosessen. For forskningen gir det bedre mulighet til å gjenbruke data, kombinere tidsserier og gjøre metode og usikkerhet synlig i videre anvendelse.

Den langsiktige samfunnsverdien ligger i at eksisterende investeringer i kartlegging, overvåking, forskning og geodata kan brukes mer effektivt i reelle beslutninger. Effekten bør derfor måles i faktisk gjenbruk, spart tid, kvalitet, sporbarhet og anvendelse i beslutningsprosesser - ikke bare i hvor mange datasett som er koblet til.

Kjernebudskap

Naturdata bør være et tydelig domene i Felles økosystem for miljø- og havdata fordi areal- og naturbeslutninger krever mer enn tilgang til enkelt kart. Brukeren må kunne kombinere data om arter, naturtyper, tilstand, arealendring og økologiske sammenhenger med data om påvirkning - på tvers av land, ferskvann, kyst og hav - og samtidig vite hva kunnskapen bygger på og hvor grensene går.

B.10 Kunnskapsgrunnlag for næringsutvikling i kystsonen – kommunal saksbehandling på tvers av land, vassdrag og fjord

Utviklet av NIVA

Kort scenario: En kystkommune mottar et planinitiativ og forslag til detaljregulering for et nytt nærings- og industriområde i strandsonen – f.eks. et landbasert havbruksanlegg med kai og utfylling i sjø, sjøvannsinntak og utslipp av rensset avløpsvann til fjorden. Tiltaket berører et myr- og skogområde, et lakseførende vassdrag og en terskelfjord med ålegrasenger og gyteområder for kysttorsk. Saken utløser krav om konsekvensutredning (KU-forskriften), vurderinger etter naturmangfoldsloven §§ 8–12 (kunnskapsgrunnlag, føre-var og samlet belastning) og vannforskriften § 12, og må samordnes med Statsforvalteren (utslippstillatelse), fylkeskommunen (akvakulturtillatelse), NVE og Fiskeridirektoratet.

Utfordringer i dag: Kunnskapsgrunnlaget må i dag settes sammen manuelt fra en lang rekke kilder (Vann-Nett, Vannmiljø, Naturbase, Artskart, Fiskeridirektoratets kart, MET, NVE, Geonorge, instituttenes egne portaler og rapporter). Dataene har ulike formater, koordinatsystemer, tids- og romlig oppløsning og lisenser, kvalitet og usikkerhet er sjelden maskinlesbart beskrevet og mye relevant kunnskap ligger i PDF-rapporter fra overvåkingsprogrammer, forskningsprosjekter og tidligere konsulentutredninger som ikke gjenbrukes. Kommunene har ofte begrenset kapasitet og fagkompetanse til å vurdere dette.

Resultatet er dobbeltarbeid, lange saksbehandlingstider, ulike aktører som vurderer saken på ulikt datagrunnlag – og usikkerhet om kunnskapsgrunnlaget faktisk er «godt nok».

B.10.1 Brukerhistorie og hovedformål

Som en saksbehandler/arealplanlegger i en kystkommune som skal behandle et reguleringsplanforslag med tilhørende søknader for nærings- og industrietablering i et område som berører både natur på land, ferskvann og kyst/fjord,

ønsker jeg å finne, hente og sammenstille et tilstrekkelig godt og kvalitetssikret kunnskapsgrunnlag på tvers av domener fra ett inngangspunkt (med kjent opphav, kvalitet, usikkerhet, aktualitet og bruksvilkår) slik at jeg, som tiltakshaver/konsulent og sektormyndighetene, kan:

- avgrense påvirkningsområdet (nedbørfelt, vassdrag og fjordbasseng) og få oversikt over registrerte naturverdier, miljøtilstand og eksisterende påvirkninger (samlet belastning),
- vurdere om kunnskapsgrunnlaget er tilstrekkelig (nml. § 8) og tidlig identifisere kunnskapshull som krever målrettede nye undersøkelser (i stedet for å bestille alt på nytt),
- kombinere data i analyser, f.eks. spredning av utslipp i resipienten (strøm, lagdeling, vannutskifting), tilførsler av næringssalter fra nedbørfeltet, fremtidig klima (havnivå, stormflo, ekstremnedbør) og vurdere om tiltaket kan forringe tilstanden i berørte vannforekomster,
- få tilgang (med riktig tilgangsstyring) til beskyttede data (f.eks. sensitive artsdata for rødlistede arter) som ikke kan deles åpent,
- dokumentere kunnskapsgrunnlaget sporbart i saken (hvilke datasett, versjon, kvalitet og kilde) og dele nøyaktig samme grunnlag med Statsforvalter, fylkeskommune, NVE, Fiskeridirektoratet og tiltakshaver,
- på sikt la en KI-assistent hente og sammenstille data via maskinlesbare grensesnitt (der kilde, kvalitet og lisens alltid følger med svaret),

slik at saken kan avgjøres raskere og på et bedre, likeverdig og etterprøvbart grunnlag (med lavere risiko for mangelfulle utredninger, klager og omgjøring) og slik at både næringsutvikling og miljømålene (god økologisk og kjemisk tilstand, ivaretatt naturmangfold, etc.) blir ivaretatt.

B.10.2 Eksempel på datakilder og domener

Domene	Datasett/produkt	Datatilbyder	Parametere/attributter
Vannforvaltning (ferskvann og kyst)	Vann-Nett: vannforekomster, tilstandsklassifisering, påvirkninger og miljømål	Miljødirektoratet / vannregionmyndighetene	Økologisk og kjemisk tilstand, påvirkningstype og -grad, miljømål, risiko for ikke å nå miljømål
Vannkvalitet – måledata	Vannmiljø (nasjonal database), inkl. data fra ØKOKYST og	Miljødirektoratet NIVA m.fl. som dataprodusent	Næringssalter (tot-N, tot-P, nitrat, fosfat), klorofyll a, oksygen i bunnvann, siktedyp,

Domene	Datasett/produkt	Datatilbyder	Parametere/attributter
	Elveovervåkingsprogrammet		bløtbunnsfauna, makroalger, miljøgifter tidsserier per stasjon
Tilførsler og utslipp	TEOTIL (tilførsler fra nedbørfelt): Norske utslipp (virksomheter med tillatelse)	NIVA Miljødirektoratet	N- og P-tilførsler per kilde (landbruk, avløp, industri, akvakultur), utslipp per virksomhet, nedbørfelt
Sediment og forurensset grunn	Sedimentdata i Vannmiljø, Grunnforurensningsdata basen	Miljødirektoratet NIVA, NGI m.fl.	Metaller, PAH, PCB, TBT, tilstandsklasse sediment, lokalitet og prøvetakingsår
Oseanografi og fjordsirkulasjon	Havmodellen NORKYST (varsel og historiske kjøring): fjordmodeller og spredningsanalyser	MET og HI NIVA	Strøm, temperatur, salinitet, lagdeling, vannutskifting, oppholdstid, spredning av utslipp
Marint naturmangfold	Naturbase – marine naturtyper fra nasjonalt kartleggings-program (ålegras, bløtbunn, tareskog m.m.)	Miljødirektoratet HI, NIVA m.fl. som kartleggere	Naturtype, utbredelse, verdi, kartleggingsår, stedfestingsnøyaktighet
Fiskeri og akvakultur	Gyteområder (bl.a. kysttorsk), fiskeplasser, akvakulturregister og lokaliteter	Fiskeridirektoratet HI	Område, art, viktighet, periode lokalitet, arter, kapasitet
Anadrom fisk	Lakseregisteret, nasjonale laksevassdrag, bestandsovervåking	Miljødirektoratet NINA HI	Bestandsstatus, gytebestandsmåloppnåelse, påvirkningsfaktorer
Naturmangfold på land og i ferskvann	Artskart, Norsk rødliste, fremmede arter, Naturbase (naturtyper etter NiN, verneområder), økologiske grunnkart	Artsdatabanken Miljødirektoratet NINA	Art, observasjonstid, stedfestingsusikkerhet, rødlistekategori, naturtype, lokalitetskvalitet sensitive data med tilgangsstyring
Sjøfugl	SEAPOP – kolonier og bestandsdata	NINA Norsk Polarinstitut	Koloni, art, bestand, trend, sårbarhet
Areal, jord og myr	AR5 arealressurskart, SR16, myr og karbonrike arealer	NIBIO	Arealtype, skogbonitet, myrareal, jordbruksareal
Geologi, flom og skred	Løsmassekart, marin grense; flom- og	NGU NVE	Løsmassetype, faresone, gjentakintervall

Domene	Datasett/produkt	Datatilbyder	Parametere/attributter
	skredfaresoner, aktsomhetskart		
Vær og klima – observasjoner	Frost API (observasjoner), seNorge (grid data)	MET	Nedbør, temperatur, vind, ekstremverdier, tidsserier
Klima – framskrivninger	Klimaprofiler og klimaframskrivninger; havnivå og stormflo («Se havnivå»)	Norsk klimaservice-senter (MET, NVE, NORCE m.fl.) Kartverket	Fremtidig ekstremnedbør og avrenning, havnivåstigning, stormflonivå (f.eks. 200-årsnivå)
Luftkvalitet og avsetning	Luftkvalitetsmålinger og modellert avsetning	NILU gjennom EBAS Miljødirektoratet	NO ₂ , PM, nitrogenavsetning
Kartgrunnlag og planer	DOK (Det offentlige kartgrunnlaget), FKB, dybde-/sjøkartdata, gjeldende arealplaner	Kartverket (Geonorge / Norge digitalt) kommunen	Dybde, kystkontur, planformål, eiendom, infrastruktur
Utredningsdata fra tiltakshaver	Resipientundersøkelser, feltkartlegging og tidligere KU-er	Tiltakshaver/konsulent kommunen	Rådata fra feltundersøkelser med metode, dato og kvalitet – deles tilbake til økosystemet for gjenbruk

Listen er en illustrasjon og viser bredden av datatilbydere (myndigheter, forvaltning, forvaltningsinstitutter, forskningsinstitutter og private). Brukerhistorien forutsetter at data som hovedregel blir liggende hos autoritativ kilde og at økosystemet gir felles katalog, vokabular, kvalitets- og proveniensbeskrivelse, samt tilgangs- og lisensstyring.

B.10.3 Forventet resultat og samfunnsnytte

Forventet resultat: Kommunen kan på kort tid etablere et felles, dokumentert kunnskapsgrunnlag («datapakke») for planområdet og påvirkningsområdet, der hvert datasett har kjent opphav, kvalitet, usikkerhet og bruksvilkår (og som kan deles direkte med sektormyndigheter og tiltakshaver).

- **Raskere og bedre saksbehandling:** mindre tid på å lete etter, tolke og tilrettelegge data og mer tid på faglige vurderinger. Kunnskapshull blir synlige tidlig, slik at nye undersøkelser kan målrettes.
- **Etterprøvbare og likeverdige vedtak:** sporbar dokumentasjon av kunnskapsgrunnlaget styrker oppfyllelsen av naturmangfoldsloven § 8, gir færre klager og omgjøringer og sikrer at kommune, Statsforvalter, fylkeskommune og tiltakshaver vurderer saken på samme datagrunnlag.
- **Forutsigbarhet for næringslivet:** lavere utredningskostnader, færre runder med tilleggsutredninger og raskere avklaring gir bedre vilkår for grønn og blå næringsutvikling i kystsonen.

- **Bedre miljøforvaltning:** samlet belastning på tvers av land, vassdrag og fjord kan vurderes helhetlig, noe som styrker arbeidet med vannforskriftens miljømål, naturavtalen og fremtidig naturregnskap (både terrestrisk, akvatisk og marint).
- **Økt verdi av offentlige investeringer i data:** data fra nasjonale overvåkingsprogrammer, forskning og tidligere utredninger blir gjenbrukt i stedet for å bli liggende i rapporter. Nye utredningsdata føres tilbake til økosystemet.
- **KI-klare data:** maskinlesbare data med kjent kvalitet og proveniens gjør det mulig å ta i bruk KI-baserte assistenter i saksbehandlingen på en ansvarlig og etterprøvbar måte.
- **Skalerbarhet:** samme mønster gjelder for et stort antall kystkommuner og sakstyper (havbruk, havner og kaier, industri, energi, fritidsbebyggelse og klimatilpasning) og for Statsforvalterens behandling av utslippssaker.

Tilleggs kommentar:

Brukerhistorien demonstrerer kanskje særlig økosystemets verdi der data fra mange tilbydere og domener (land – ferskvann – hav – klima) må kombineres, der kvalitet og usikkerhet må være kjent for å kunne brukes i myndighetsutøvelse og der både åpne og beskyttede data inngår. Den egner seg godt som pilot i et avgrenset fjordområde med gode eksisterende tidsserier (f.eks. i Oslofjord-regionen?), med målbare indikatorer som tidsbruk til sammenstilling av kunnskapsgrunnlag, antall kilder og andel datasett med maskinlesbar kvalitetsbeskrivelse.