

# Dataprodukt i Helseanalyseplattformens økosystem

Versjon 1.0

*Sist oppdatert: 03.06.2020*



# Innhold

<b>1</b>	<b>Innledning .....</b>	<b>3</b>
1.1	Bakgrunn.....	3
1.2	Formålet med dokumentet.....	3
1.3	Omfang og avgrensninger .....	4
<b>2</b>	<b>Dataprodukter .....</b>	<b>6</b>
2.1	Definisjon av dataprodukt.....	6
2.2	Hvorfor dataprodukt?.....	7
2.3	Innholdet i dataprodukt.....	7
2.4	Verdikjede for sekundærbruk av helsedata .....	8
2.5	Produksjonslinje for dataprodukt .....	9
2.6	Metadata om dataprodukt.....	9
2.7	Rettsgrunnlag og anbefalte bruksområder.....	10
2.8	Overføring av dataprodukt til Helseanalyseplattformen .....	10
<b>3</b>	<b>Dataprodukt som en del av Helseanalyseplattformens økosystem .....</b>	<b>12</b>
3.1	Roller i økosystemet av dataprodukter .....	12
3.2	Prosesser i dataøkosystemet .....	13
3.3	Regulering av ansvarsforholdet for dataprodukt på Helseanalyseplattformen.....	14
<b>4</b>	<b>Realiseringsstrategi for dataprodukt i Helseanalyseplattformens økosystem .....</b>	<b>16</b>
4.1	Forutsetninger og avhengigheter for dataprodukt på Helseanalyseplattformen ....	16
4.2	Prinsipper for realisering av dataprodukt på Helseanalyseplattformen .....	17
4.3	Helsedataprogrammets og dataleverandørens rolle i realiseringen av data på Helseanalyseplattformen.....	17
4.4	Prioritering av dataprodukt på Helseanalyseplattformen.....	18
4.5	Overordnet realiseringsplan .....	20
<b>5</b>	<b>Begrepsliste .....</b>	<b>22</b>

# 1 Innledning

## 1.1 Bakgrunn

Direktoratet for e-helse fikk i 2017 oppdrag fra Helse- og omsorgsdepartementet om å iverksette tiltak for å følge opp helseregisterstrategien 2010-2020. Oppdraget er organisert som et program, Helsedataprogrammet, og er en videreføring av Nasjonalt helseregisterprosjekt.

I april 2018 leverte Helsedataprogrammet Konseptvalgutredning for Helseanalyseplattformen. Her anbefales det blant annet at «Det etableres en felles analyseinfrastruktur og dataplattform der helsedata kan tilgjengeliggjøres for bruk på en sikker måte. Andre datakilder og eksterne analyseinfrastrukturer tilknyttes ved behov.»

## 1.2 Formålet med dokumentet

Formålet med dette dokumentet er å beskrive rammene for håndtering og overføring av data til Helseanalyseplattformen. Dette bidrar til å sette rammebetingelser for hvordan dataleverandører kan jobbe med data som skal eksporteres til Helseanalyseplattformen. Målgruppen for denne versjonen av dokumentet er beslutningstakere og andre medarbeidere i virksomheter som skal ha rollen som dataleverandører til Helseanalyseplattformen i løpet av 2020 og 2021. Dette vil primært være forvaltere av de sentrale helseregistrene, de nasjonale medisinske kvalitetsregistrene, utvalgte helseundersøkelser og eventuelt andre offentlige registre. Fremtidige versjoner av dette dokumentet vil potensielt adressere andre typer av dataleverandører. Dokumentet skal:

- Introdusere begrepet dataprodukt og forklare hvordan dataprodukt inngår som en del av Helseanalyseplattformens økosystem.
- Forklare hvorfor Helseanalyseplattformen bør etableres basert på dataprodukt fremfor databasedump fra helseregistrene.
- Beskrive det overordnede ansvarsforholdet mellom dataleverandører, tilgangsførvalter og sluttbrukere knyttet til data på Helseanalyseplattformen.
- Beskrive gjennomføringsstrategien for etablering av data på Helseanalyseplattformen.

Rollen som tilgangsførvalter skal ivaretas av Helsedataservice<sup>1</sup>. Helsedataservice skal også kunne inneha rollen som dataleverandør av sammensatte dataprodukter og sluttbruker av dataprodukter for å levere analysetjenester som eksempelvis en kohortutforsker. Vi benytter derfor begrepet *tilgangsførvalter* i dette dokumentet for å tydeliggjøre rollen Helsedataservice skal ha som dataansvarlig for Helseanalyseplattformen.

Begrepslisten i kapittel 0 forklarer sentrale begreper benyttet i dokumentet.

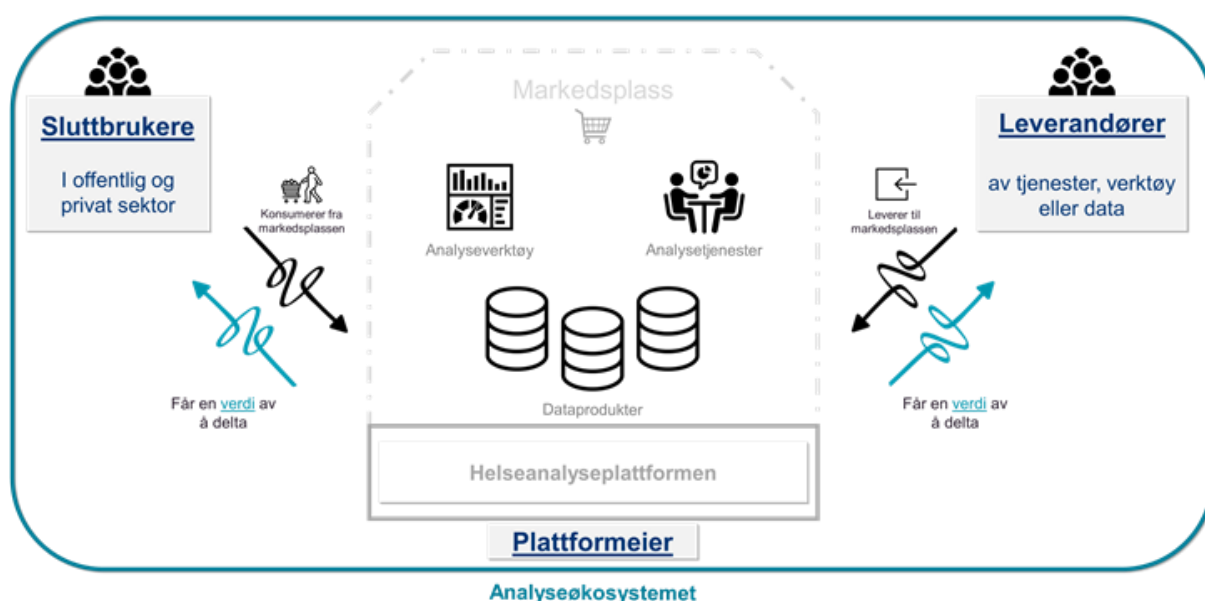
---

<sup>1</sup> Helsedataservice er organisasjonsenheten som skal være dataansvarlig for data på Helseanalyseplattformen. Helsedataservice detaljeres nærmere i rapporten *Etablering av Helsedataservice som en nasjonal enhet for å forvalte helsedataområdet* som skal overleveres til Helse- og omsorgsdepartementet 1. mai 2020.

### 1.3 Omfang og avgrensninger

Det er besluttet at Helseanalyseplattformen skal realiseres gjennom å etablere et analyseøkosystem.<sup>2</sup> Med analyseøkosystem menes et økosystem av virksomheter som leverer analysetjenester, analyseverktøy og dataprodukter gjennom en felles digital plattform og markeds plass og et felles sett med forvaltningsprosesser. For at digitale økosystem skal gi verdi, er det behov for at følgende tre aktører finnes:

- **Plattformei**er står for styringsregime (governance) for den digitale plattformen og markeds plassen
- **Leverandø**rer tilbyr tjenester, verktøy eller data gjennom plattformen og markeds plassen
- **Sluttbruk**ere konsumerer verktøy, tjenester og data som tilbys på plattformen og markeds plassen



Figur 1 Aktører og tjenester i Helseanalyseplattformens analyseøkosystem

Dette dokumentet avgrensner seg til å se på dataproduktenes rolle i økosystemet, samt grenseflatene mellom dataleverandører og tilgangsforsvalter.

Dokumentet adresserer ikke ansvarsdelingen mellom Helsedataservice som dataansvarlig for Helseanalyseplattformen og Norsk Helsenett som drifts- og forvaltningsleverandør av plattformen. Dette vil bli adressert nærmere i *Etablering av Helsedataservice som en nasjonal enhet for å forvalte helsedataområdet*.<sup>3</sup> Deler av oppgaver og ansvar beskrevet for tilgangsforsvalter i dette dokumentet kan i endelig organisering besluttes delegert til Norsk

<sup>2</sup> <https://ehelse.no/publikasjoner/konseptvalgutredning-for-helseanalyseplattformen>

<sup>3</sup> *Etablering av Helsedataservice som en nasjonal enhet for å forvalte helsedataområdet* er en rapport som skal beskrive hva Helsedataservice skal være og gå opp ansvarsforholdene mellom Helsedataservice, Norsk Helsenett og registerforvaltningen. Rapporten ble overlevert Helse- og omsorgsdepartementet 15. mai 2020.

Helsenett. Dokumentet adresserer derfor ikke hva hver av disse virksomhetene vil ha anledning til å gjøre med dataene.

Gjennom eget forskriftsarbeid vil Helse- og omsorgsdepartementet regulere ansvarsforholdene mellom Helsedataservice og registerforvaltningen. De ansvarsforholdene som ikke reguleres i forskrift vil måtte reguleres gjennom avtaleverk. Dette dokumentet inneholder overordnede forslag til ansvarsfordeling knyttet til håndtering av data som skal overføres fra registerforvaltningen til Helseanalyseplattformen. Detaljerte innspill til forskriftsarbeidet og forslag til avtaleverk for regulering av ansvarsforhold vil bli utarbeidet i egne dokumenter.

Dokumentet presenterer de identifiserte prosessene som vil være nødvendig for overføring av registerdata til Helseanalyseplattformen. Detaljering av disse prosessene er håndtert utenfor dokumentet. Dokumentet adresserer heller ikke hvordan innbyggers rett til innsyn i egne data på Helseanalyseplattformen skal håndteres.

Overføring av data til Helseanalyseplattformen defineres nærmere i *Spesifikasjon for overføring av data til Helseanalyseplattformen*<sup>4</sup> med støtte i *Nasjonal spesifisering for metadata om helsedata*<sup>5</sup>.

---

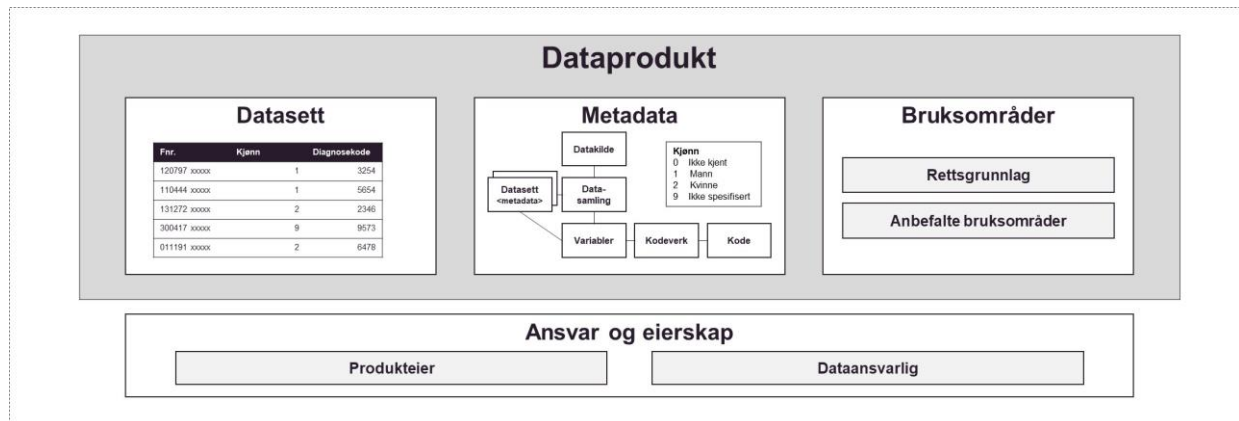
<sup>4</sup> Første versjon av denne spesifiseringen er ferdigstilt og tilgjengeliggjøres på forespørsel. En oppdatert versjon tilpasset teknologiplattform for helseanalyseplattformen vil publiseres på ehelse.no i løpet av 2020.

<sup>5</sup> <https://ehelse.no/standarder/ikke-standarder/nasjonal-spesifisering-for-metadata-om-helsedata>

## 2 Dataprodukter

### 2.1 Definisjon av dataprodukt

Dataprodukter er definert som en sammenstilling av datasett, metadata og bruksområder der en dataleverandør er produkteier. Dataprodukter skal genereres med sikte på å dekke et sett med sluttbrukerbehov. Dataprodukter kan inneholde anonyme eller personidentifiserbare data.



Figur 2: Overordnet modell av knytningen mellom metadata, datasett og dataprodukt.

Som illustrert i Figur 2 introduseres begrepene produkteier og dataansvarlig. Vi definerer begrepene som følger:

- **Produkteier:** Begrepet er hentet fra smidige utviklingsprosjekt og beskriver personen som har ansvar for omfangsstyringen av produktet.<sup>6</sup> I konteksten der data er produktet vil produkteier betegne personen eller virksomheten som er ansvarlig for dataproduktets produksjonslinje i forkant av publisering på Helseanalyseplattformen. Produkteier skal generere dataprodukt med sikte på å dekke sluttbrukeres behov, og sikre at dataproduktet har de tekniske egenskapene og den faglige kvaliteten nødvendig for publisering. Dette innebærer blant annet å gi et riktig bilde av datakvaliteten i dataproduktet. Produkteierskapet ligger hos dataleverandøren.
- **Dataansvarlig:** Begrepet betegner personen eller virksomheten som er ansvarlig for behandling av helseopplysninger, jf. personvernforordningen artikkel 4 nr. 7. Definisjonen er hentet fra helseregisterloven § 2 bokstav d, og det vil måtte tas høyde for eventuelle lov- og forskriftsendringer knyttet til forståelsen av og plasseringen av dataansvar. Produkteier er dataansvarlig for all behandling av helseopplysninger og andre personopplysninger som er nødvendig for å generere, kvalitetssikre og publisere et dataprodukt. Tilgangsforsvalter er dataansvarlig for mottak, lagring, sammenstilling, tilgjengeliggjøring og annen behandling av helseopplysninger og andre personopplysninger som overføres til Helseanalyseplattformens dataplattform. Se også om regulering av ansvar i kapittel 3.3.

Dataansvaret og produkteierrollen defineres uavhengig av hverandre. I motsetning til dataansvaret vil ikke Helsedataservice overta produkteierrollen for registerspesifikke

<sup>6</sup> <https://www.prosjektveiviseren.no/roller/produkteier>

dataprodukt når disse publiseres på Helseanalyseplattformen. Dette fordi man ønsker at dataproduktene på Helseanalyseplattformen i størst mulig grad skal være basert på kjennskapen til dataene som finnes i det enkelte register, spisskompetansen rundt data fra det enkelte register ligger hos registerforvaltningen.

## 2.2 Hvorfor dataprodukt?

Man kan i utgangspunktet se for seg to prinsipielt ulike tilnæringer til data på Helseanalyseplattformen. Data kan enten håndteres som bearbeidede dataprodukt som tilgjengeliggjøres på plattformen, eller som databasedumper fra utvalgte registre. Realisering av Helseanalyseplattformen basert på bearbeidede dataprodukt vil være den mest kostnadseffektive løsningen der man utnytter eksisterende kompetanse i registermiljøene og reduserer gjennomføringsrisikoen. Dette skyldes primært at:

- 1) Man gjennom tilnærmingen med dataprodukt begrenser dupliseringen av funksjoner og arbeidsoppgaver.
- 2) Man viderefører fagkompetansen hos eksisterende registerforvaltere innen dataforvaltning og kvalitetsforbedring ved at dataleverandørene har det faglige ansvaret for dataprodukt som skal publiseres på Helseanalyseplattformen.
- 3) Dataprodukt innebærer høy etterlevelse av valgt konsept for Helseanalyseplattformen, ved at man med denne tilnærmingen legger til rette for et økosystem av dataleverandører som kan publisere ulike dataprodukt på plattformen.
- 4) Dataprodukt med anbefalte bruksområder på et standardisert dataformat legger til rette for at Helsedataservice kan håndtere en bred portefølje av data.
- 5) Dataprodukttilnærmingen sikrer at Helseanalyseplattformen inneholder relevante og etterspurte data.

Tilnærmingen med dataprodukt vil vel å merke kunne innebære merarbeid for dataleverandørene og i første omgang de eksisterende nasjonale registermiljøene. Dette skyldes at det vil være behov for en initiell innsats for å tilpasse dagens produksjon av data for å kunne utvikle HAP-spesifikke dataprodukt. Dette vil kunne være krevende for registerforvaltere som allerede er presset på kapasitet. Helsedataprogrammet har derfor iverksatt flere finansielle tiltak for å støtte registerforvaltningen, se kapittel 4.3. Tilnærmingen innebærer også en viss økt usikkerhet for Helsedataprogrammet, siden det oppstår en økt prosjektavhengighet til aktiviteter hos registerforvalterne. Denne risikoen er detaljert nærmere i kapittel 4.5.2. For at tilnærmingen skal lykkes er det også avgjørende at dataproduktene designes slik at de lykkes med å dekke den store majoriteten av sluttbrukerbehov. Prinsipp for realisering av dataprodukt er adressert nærmere i kapittel 4.2.

## 2.3 Innholdet i dataprodukt

Dataleverandører inkluderer variabler i dataproduktet basert på sin kjennskap til egne data, etterspørsel fra brukere av data og relevant domenekunnskap. Ved generering av nye dataprodukt er det avgjørende å ha et bevisst forhold til mulige sluttbrukere av dataproduktet og hvorvidt produktet vil kunne dekke sluttbrukeres behov. Personidentifiserbare dataprodukt skal inneholde data på lavest mulige granulatetsnivå, og dataminimering skal gjennomføres av tilgangsforvalter basert på sluttbrukers søknad om tilgang til data.

Det må legges opp til et samarbeid der dataleverandører kan bruke tilgangsforvalter som sparringspartner ved utforming og produksjon av dataprodukt. Tilgangsforvalter vil også kunne etterspørre nye variabler inn i et nytt eller eksisterende dataprodukt basert på ny innsikt i sluttbrukerbehov. Det legges til grunn en arbeidshypotese om at tilgangsforvalter

ikke vil ha myndighet til å pålegge dataleverandør å inkludere enkeltvariabler i et dataprodukt på Helseanalyseplattformen. Tilgangsforvalter er likevel tenkt å kunne vedta tilgjengeliggjøring av registerdata som ikke er inkludert i dataprodukt eller ligger fysisk på Helseanalyseplattformen i forbindelse med sluttbrukers søknad om data. Dataleverandør vil på denne måten få et insentiv til å inkludere ofte etterspurte variabler i et dataprodukt for å slippe administrasjonen som følger med enkeltforespørsler.

Foruten å definere hvilke variabler som skal inngå i et dataprodukt, må dataleverandøren også dokumentere metadata for dataproduktet, vise til det aktuelle rettsgrunnlaget som ligger til grunn for behandling og utlevering av personopplysningene, samt definere anbefalte bruksområder.

## 2.4 Verdikjede for sekundærbruk av helsedata

Med verdikjede for sekundærbruk av helsedata menes alle operasjonene som gjøres på helsedata fra dataene registreres i helsetjenesten til data er tilgjengelig for bruk i kvalitetsindikatorer, statistikkbanker, forskning og andre sekundære formål. Figur 3 gir en forenklet illustrasjon av denne verdikjeden når Helseanalyseplattformen og Helsedataservice er etablert.

Det første steget i verdikjeden er at data registreres i helsetjenesten. Data overføres så til ulike helseregistre der de kvalitetssikres, lagres og brukes i henhold til registrenes formål. Eksempler på bruk her er statistikkproduksjon og styringsinformasjon som innsatsstyrt finansiering. Mange av helseregistrene har egne produksjonslinjer for denne type bruk. Hos enkelte registre vil disse produksjonslinjene være et godt utgangspunkt for produksjon av dataprodukt som skal publiseres på Helseanalyseplattformen. I andre registre vil det være behov for en større jobb knyttet til å etablere produksjonslinjer for HAP-spesifikke dataprodukt

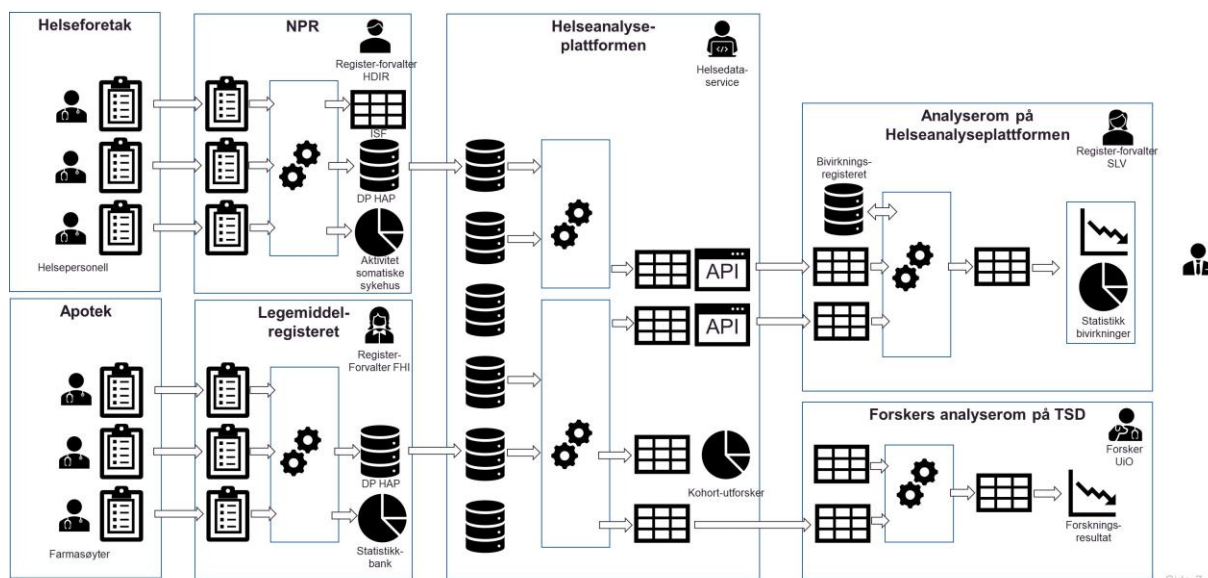
På Helseanalyseplattformen vil det også gjøres teknisk kvalitetssikring og bearbeiding av dataproduktene som mottas fra helseregistrene. Dette inkluderer blant annet å bearbeide dataene for å legge til rette for effektiv tilgjengeliggjøring etter søknad, å etablere nye dataprodukt basert på data fra flere registre og å tilrettelegge for at data kan inngå i tjenester som tilbys på Helseanalyseplattformen, som for eksempel en kohortutforsker. Denne jobben vil måtte gjøres av Helsedataservice.

Det neste steget i verdikjeden inkluderer operasjoner utført av ulike sluttbrukere. Dette kan være forskere, helseledere, næringslivsaktører eller myndighetsaktører. I denne delen av verdikjeden vil det også være behov for å gjøre både teknisk og faglig kvalitetssikring, bearbeiding og analyse av dataene for at de skal være tilpasset deres bruk.

Det er derfor viktig å understreke følgende:

1. Kvalitetssikring, bearbeiding og analyse vil gjøres av mange aktører i hele verdikjeden.
2. Mange av registerforvalterne har i dag etablerte produksjonslinjer for ulik bruk av data. For noen forvaltere vil produksjonslinjer for HAP-spesifikke dataprodukt kun innebære å gjøre justeringer i dagens produksjonslinjer, mens for andre vil det være behov for å etablere nye produksjonslinjer for dataprodukt som skal publiseres på Helseanalyseplattformen.





Figur 3 Forenklet verdikjede for sekundærbruk av helsedata

## 2.5 Produksjonslinje for dataprodukt

Med produksjonslinje for dataprodukt menes de prosessene og det tekniske miljøet som skal til for å utvikle et dataprodukt basert på data som er tilgjengelig hos en dataleverandør. På et overordnet nivå innebærer dette å motta rådata, kvalitetssikre og foredle dataene og generere dataprodukt. Dette er illustrert i Figur 4. Dataprodukt kan bygges på ferdig kvalitetssikrede årganger av helseregisterdata. Dette kan typisk være godt tilpasset forskningsformål og statistikk. Produktene kan også bygge på løpende mottatte data der kvalitetssikringsprosessen har vært minimal. Dette vil eksempelvis kunne være relevant i et dataprodukt tiltenkt beredskapsformål eller helseovervåking. Samme dataleverandør vil kunne lage flere dataprodukt med tilhørende produksjonslinjer for å dekke ulike behov.



Figur 4: Illustrasjon av generisk produksjonslinje for et dataprodukt

## 2.6 Metadata om dataprodukt

Alle dataproduktene må dokumenteres med gode metadata slik at det kan lages gode, informative løsninger for sluttbrukere av Helseanalyseplattformen. Metadata skal også legges til rette for at sluttbruker enkelt kan finne relevante dataprodukt basert på egne problemstillinger.

Helsesdataprogrammet leverte i januar 2020 *Nasjonal spesifisering for metadata om helsedata*. Her beskrives metadata som dataleverandører kan og må definere som et ledd i arbeidet med å utarbeide dataprodukt til Helseanalyseplattformen.

I tillegg til metadataene definert i *Nasjonal spesifisering for metadata om helsedata* må dataproduktene vise til rettsgrunnlaget som ligger til grunn for utlevering og behandling av personopplysningene i registeret, samt anbefalte bruksområder. Dette er beskrevet nærmere i kapittel 2.7. På sikt skal også rettsgrunnlag og anbefalte bruksområder innarbeides i *Nasjonal spesifisering for metadata om helsedata* og implementeres i Nasjonal variabelkatalog.

## 2.7 Rettsgrunnlag og anbefalte bruksområder

Som en del av jobben med å generere dataproduktet, skal rettsgrunnlag og anbefalte bruksområder for dataproduktet angis. Dette er informasjon som skal gjøre det enklere for tilgangsforsvalter å vurdere søknader om tilgang til data på plattformen, samt å synliggjøre informasjonen overfor sluttbrukere slik at de vet hvilke muligheter og begrensninger som finnes ved et dataprodukt.

- **Rettsgrunnlag:** Rettsgrunnlaget for utlevering av personopplysningene og den videre behandlingen. Rettsgrunnlaget og særlig formålsangivelsen for den enkelte datakilde er styrende for hvilke rammer som gjelder for behandlingen av helseopplysninger og andre personopplysninger som inngår i et dataprodukt. For eksempel vil formål og bruk for dataprodukt fra de lovbestemte helseregistrene bestemmes av den relevante forskriften, mens formål og bruk for dataprodukt fra befolkningsbaserte helseundersøkelser vil bestemmes av den registrertes samtykke og forskrift om befolkningsbaserte helseundersøkelser av 2018. Det må tas høyde for planlagt lov- og forskriftsarbeid.
- **Anbefalte bruksområder:** Anbefalte bruksområder skal beskrive hva dataene kan og vil være egnet til å benyttes til. Eksempelvis kan en datakilde ha samlet inn data både til forskning, statistikk og beredskap, men på grunn av egenskapene ved det aktuelle dataproduktet er produktet kun egnet til forskning og statistikk. Dette kan for eksempel være tilfellet for dataprodukt med årlig oppdateringsfrekvens. Dataprodukt skal kun ha anbefalte bruksområder som er innenfor formålet data er samlet inn til. Anbefalte bruksområder vil ikke være bindende for tilgangsforsvalter da dette ville undergravd tilgangsforsvalters selvstendige dataansvar.

Det er ønskelig å etablere en standardisert tilnærming til angivelsen av rettsgrunnlag og anbefalte bruksområder. Videre utredninger vil konkretisere hvordan denne tilnærmingen bør være. Hvordan dette praktisk skal håndteres vil dokumenteres i en revidert versjon av *Nasjonal spesifisering for metadata om helsedata*.

Det må understrekes at tilgangsforsvalteren som dataansvarlig for data på Helseanalyseplattformen og i tråd med vedtaksmyndigheten som er foreslått i nytt regelverk, vil fatte selvstendige beslutninger om tilgjengeliggjøring og sammenstilling av data til sluttbruker. Tilgangsforsvalter vil dermed stå ansvarlig for å påse at tilgjengeliggjøring og sammenstilling er i samsvar med rettsgrunnlaget. Dette innebærer at Helsedataservice vil gjøre selvstendig dataminimering av datasett som skal utleveres og ikke levere ut fullstendige dataprodukt.

## 2.8 Overføring av dataprodukt til Helseanalyseplattformen

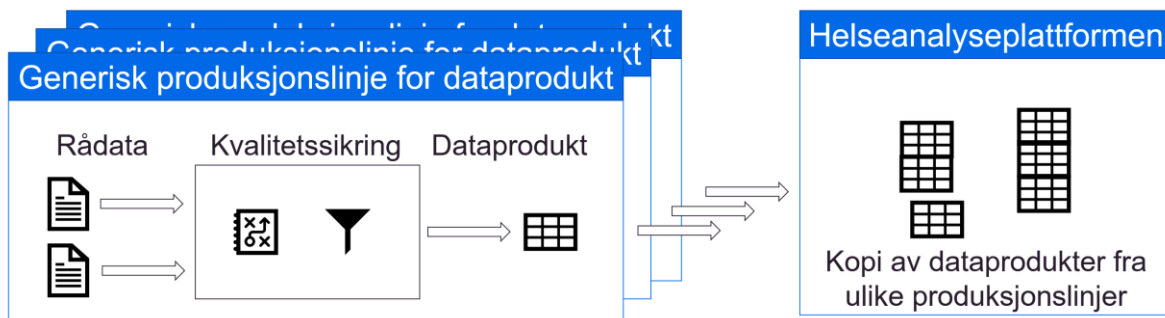
Dataproduktet må i mange tilfeller overføres til Helseanalyseplattformens dataplattform for at tilgangsforsvalteren skal kunne ta et helhetlig ansvar for å tilgjengeliggjøre data til

sluttbrukere. Samtidig vil det også kunne være dataprodukter som aldri fysisk vil ligge på Helseanalyseplattformens dataplattform. Disse vil i stedet kunne være tilgjengelig gjennom grensesnitt (API) for datadeling der Helseanalyseplattformen kan gjøre spørringer mot predefinerte datasett fra gitte dataleverandører. For hvilke dataleverandører dette vil kunne være aktuelt og hvordan dette skal realiseres i praksis må utredes nærmere.

Egenskapene rundt selve dataproduktet er førende for hvilken integrasjonsmekanisme det er hensiktsmessig å benytte for å flytte dataproduktet til Helseanalyseplattformen. Eksempelvis vil oppdateringsfrekvensen og mengden data være av stor betydning for valg av integrasjonsmønstre. Helseanalyseplattformen må derfor støtte en rekke ulike former for integrasjon – fra manuell filoverføring til automatisert datadeling.

For å legge til rette for effektiv dataforvaltning på Helseanalyseplattformen bør data overføres på et enkelt standardisert format som samtidig bevarer presisjonsgraden i registerdataene. Videre må Normens og Nasjonal sikkerhetsmyndighets krav til informasjonssikkerhet i grensesnittene ivaretas. Mer detaljert informasjon om overføring av data til Helseanalyseplattformen spesifiseres i *Spesifikasjon for overføring av data til Helseanalyseplattformen*.

Figur 5 illustrerer på et overordnet nivå hvordan dataprodukter fra produksjonslinjer hos ulike dataleverandører overfører kopier av dataproduktet til Helseanalyseplattformen. I eksempelet under ligger kopier av tre ulike dataprodukter på Helseanalyseplattformens dataplattform.



Figur 5: Illustrasjon over hvordan kopier av *dataprodukter* overføres til *Helseanalyseplattformen*

## 3 Dataprodukt som en del av Helseanalyseplattformens økosystem

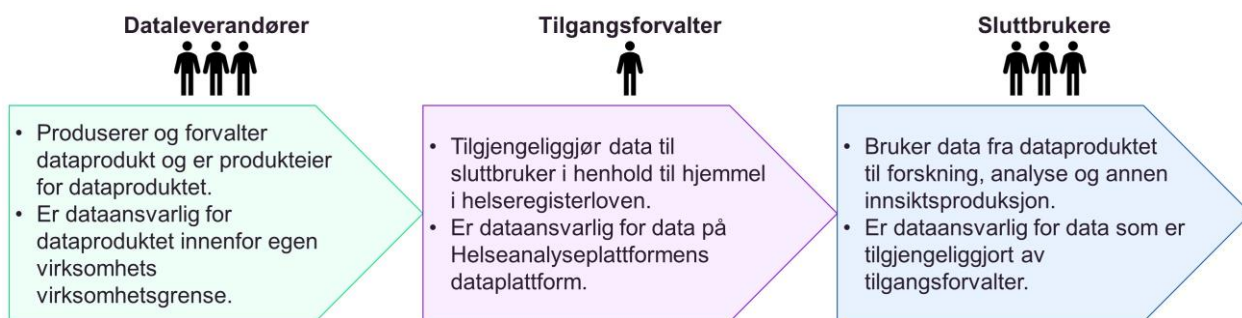
Etableringen av dataprodukter i Helseanalyseplattformens økosystem forutsetter en digital plattform der ulike dataleverandører kan publisere dataprodukter. Den digitale plattformen består i hovedsak av to sentrale komponenter:

1. En felles datakatalog (helsedata.no) der informasjon og metadata om dataproduktene publiseres. Helsedata.no fungerer dermed som markedsplassen for data der sluttbrukere kan identifisere og søke om å kunne ta i bruk helsedata og andre relevante data fra mange ulike dataleverandører. Datakatalogen på helsedata.no skal også vise variabler som ikke inngår i dataprodukt men kun ligger lokalt i det enkelte register.
2. En dataplattform der selve innholdet (dataene) i dataproduktet ligger. Gjennom dataplattformen gis den faktiske tilgangen til data fra dataproduktene. Tilgangsstyring av data fra disse dataproduktene vil håndteres gjennom enhetlige prosesser av en felles tilgangsforvalter (Helsedataservice).

Dette kapitlet tar utgangspunkt i ansvarsbegrepene introdusert i kapittel 2.1 og beskriver de viktigste prosessene og rollene knyttet til etablering og forvaltning av dataprodukt i økosystemet. Videre sees det nærmere på behovet for regulering av ansvarsforholdet gjennom standardvilkår, avtaler og forskrift.

### 3.1 Roller i økosystemet av dataprodukter

I økosystemet av dataprodukter vil det primært være tre roller; dataleverandør, tilgangsforvalter og sluttbruker. De ulike rollene vil ha ulike oppgaver og ansvar i økosystemet. Dette er illustrert på et overordnet nivå i Figur 6. Hva de ulike rollene innebærer går nærmere gjennom i dette kapitlet.



Figur 6: Overordnet illustrasjon av rollene i økosystemet av *dataprodukt*

#### 3.1.1 Dataleverandører

Dataleverandører vil som produkteier produsere og forvalte dataprodukter med tilhørende produksjonslinjer innenfor sine virksomhetsgrenser. Dataleverandøren skal sikre den nødvendige faglige kvaliteten på og nødvendige tekniske egenskaper til dataproduktet. Dette innebærer blant annet å gi et riktig bilde på forventet datakvalitet. Gjennom rollen som produkteier vil dataleverandøren generere dataproduktet basert på egen kjennskap til

dataene og eksisterende sluttbrukerbehov. Dataleverandøren skal videre overføre metadata om dataproduktet til Nasjonal variabelkatalog i henhold til *Nasjonale spesifikasjoner for metadata om helsedata* og overføre dataprodukt til Helseanalyseplattformens datamottak i henhold til *Spesifikasjon for dataoverføring til Helseanalyseplattformen*.

Ulike aktører kan inneha rollen som dataleverandør. Dette kan eksempelvis være sentrale helseregistre som leverer dataprodukt basert på ferdig kvalitetssikrede årganger av registrene, SSB som leverer dataprodukt med demografiske og sosioøkonomiske data, Statens kartverk som leverer dataprodukt basert på Matrikkelen, forskningsmiljøet som leverer dataprodukt basert på egne innsamlede data eller gendata fra ferdigstilte forskningsprosjekt, en næringslivsaktør som leverer et dataprodukt basert på innsamlede data eller Helsedataservice som etablerer og forvalter nye dataprodukt basert på dataprodukter fra andre dataleverandører.

### 3.1.2 Tilgangsforvalter

Med tilgangsforvalter menes aktøren som er dataansvarlig for dataprodukt på Helseanalyseplattformens dataplattform. Tilgangsforvalter skal veilede sluttbrukere, behandle søknad og fatte vedtak om tilgang til data og eventuell dispensasjon fra taushetsplikten, tilgjengeliggjøre data, kommunisere med dataleverandører om behovet for nye dataprodukter og forvalte rammebetingelsene for håndtering av data på Helseanalyseplattformen. Rollen som tilgangsforvalter skal ivaretas av Helsedataservice.

### 3.1.3 Sluttbrukere

Sluttbrukere i kontekst av dataprodukter er personen eller virksomheten som søker om og bruker data fra dataproduktene for å skape ny kunnskap eller innsikt. Det legges til grunn at sluttbruker vil være dataansvarlig for data som er tilgjengeliggjort av tilgangsforvalter, enten data tilgjengeliggjøres i dedikerte prosjekter på Helseanalyseplattformen eller utenfor plattformen. Eksempler på aktører som kan inneha rollen som sluttbruker kan være forskere som analyserer helsedata som en del av et forskningsprosjekt, helsepersonell som analyserer helsedata i forbindelse med kvalitetsforbedringsprosjekter i klinikken, myndighetsaktører som analyserer helsedata for å skaffe til veie et beslutningsgrunnlag for politiske beslutninger, næringslivsaktører som ønsker å teste ut hypoteser eller utvikle innsikt og algoritmer basert på helsedata og media som ønsker statistisk helseinformasjon til en nyhetssak.

## 3.2 Prosesser i dataøkosystemet

Det er identifisert en rekke prosesser som må defineres for håndtering av data på Helseanalyseplattformen i samspillet mellom dataleverandør, tilgangsforvalter, sluttbruker og innbygger. Prosessene kan grovt sett kategoriseres i etableringsprosesser, forvaltningsprosesser og avslutningsprosesser. Prosessene skal legge til rette for et hensiktsmessig samspill mellom aktørene i økosystemet. De identifiserte prosessene er på et overordnet nivå illustrert i HUKI-matrisen<sup>7</sup> i Figur 7. Ansvarsforholdene beskrevet i HUKI-matrisen baserer seg på de antagelser som er tatt rundt aktørenes ansvar i dokumentet for øvrig og må gås opp nærmere i lys av forskriftsarbeidet som skal gjennomføres av Helse- og omsorgsdepartementet i løpet av 2020 og våren 2021.

---

<sup>7</sup> En HUKI-matrise dokumenterer roller og ansvar for oppgaver gjennom å angi hvem som er Hovedansvarlig, Utførende, Konsulterende og Informert.

		H(ovedansvarlig)	U(tførende)	K(onsulterende)	I(nformert)	
Etablering	Etablere samarbeid med ny dataleverandør	Tilgangsforsvalter	Tilgangsforsvalter, Dataleverandør			
	Identifisere og implementere nytt dataprodukt	Dataleverandør	Dataleverandør	Tilgangsforsvalter		
	Publisere nytt dataprodukt til HAP	Tilgangsforsvalter	Tilgangsforsvalter, Dataleverandør			
Forvaltning	Overføring av dataprodukt til HAP	Dataleverandør	Tilgangsforsvalter, Dataleverandør	Tilgangsforsvalter	Sluttbruker	
	Tilgjengeliggjøre data fra dataprodukt til sluttbruker	Tilgangsforsvalter	Tilgangsforsvalter	Dataleverandør, Sluttbruker	Dataleverandør	
	Veilede sluttbrukere om dataprodukter på HAP	Tilgangsforsvalter	Tilgangsforsvalter	Dataleverandør	Sluttbruker	
	Behandle innbyggerforespørsel for HAP (innsyn i registrerte data)	Tilgangsforsvalter	Tilgangsforsvalter		Innbygger	
	Behandle innbyggerforespørsel for HAP (innsyn i bruk)	Tilgangsforsvalter	Tilgangsforsvalter		Innbygger	
	Behandle innbyggerforespørsel for HAP (sletting/reservasjon)	Tilgangsforsvalter	Tilgangsforsvalter		Innbygger	
	Behandle innbyggerforespørsel for HAP (retting)	Tilgangsforsvalter	Tilgangsforsvalter		Innbygger	
	Endre eksisterende dataprodukt	Dataleverandør	Dataleverandør	Tilgangsforsvalter	Tilgangsforsvalter, Sluttbruker	
	Utbedre datakvalitetsproblemer	Tilgangsforsvalter	Dataleverandør	Sluttbruker	Sluttbruker	
	Revisjon av leverandører og dataprodukter	Tilgangsforsvalter	Tilgangsforsvalter			
	Forvalte rammebetingelser for dataøkosystemet	Tilgangsforsvalter	Tilgangsforsvalter		Dataleverandør	
	Avslutning	Fjerne dataprodukt fra Helseanalyseplattformen	Tilgangsforsvalter	Tilgangsforsvalter, Dataleverandør		Sluttbruker
		Avslutte samarbeid med dataleverandør	Tilgangsforsvalter	Tilgangsforsvalter, Dataleverandør		Sluttbruker

Figur 7 Utkast til ansvarsmatrise for prosesser knyttet til forvaltning av data på Helseanalyseplattformen (under arbeid)

### 3.3 Regulering av ansvarsforholdet for dataprodukt på Helseanalyseplattformen

For at prosessene knyttet til generering og publisering av dataprodukt på Helseanalyseplattformen skal fungere som ønsket, er det nødvendig med klare reguleringer av de enkelte aktørenes oppgaver og ansvar.

Det planlegges å vedta forskrift for Helseanalyseplattformen og Helsedataservice, samt å endre forskrifter knyttet til helseregistrene, i løpet av første halvår 2021. I det videre skisseres de behov for reguleringer av ansvar, oppgaver og plikter som er identifisert så langt. Nærmere vurderinger knyttet til innholdet i reguleringene, samt av hva som er hensiktsmessig å regulere direkte i forskrift og hva som eventuelt må avtales, vil være en nødvendig del av forskriftsarbeidet.

Det vil være nødvendig å regulere en eventuell plikt for visse aktører til å være dataleverandør, samt plikt til å generere og publisere dataprodukt på Helseanalyseplattformen. En slik plikt vil være relevant for forvaltere av lovbestemte sentrale helseregistre, befolkningsbaserte helseundersøkelser, medisinske kvalitetsregistre og andre offentlige registre.

For aktører som ikke vil få en plikt til å generere og publisere dataprodukt må det nødvendigvis settes vilkår for å være dataleverandør til Helseanalyseplattformen. Slike vilkår vil kunne knytte seg til økonomisk og teknisk evne til å generere egnede dataprodukt.

Avhengig av finansieringsmodell for realiseringen av dataprodukt på Helseanalyseplattformen, vil betingelser knyttet til eventuell økonomisk kompensasjon for generering og publisering av dataprodukt på plattformen måtte reguleres.

Det må videre settes vilkår og spesifikasjoner for dataprodukt som skal publiseres på Helseanalyseplattformen. Vilkår og spesifikasjoner vil knytte seg til utformingen av dataproduktet, herunder spesifikasjoner knyttet til metadata, rettsgrunnlag og anbefalte bruksområder. Metode og frekvens for overføring, dataproduktets eventuelle gyldighetsperiode, plikter og rettigheter knyttet til mangler ved dataproduktet og veiledning, samt endringer av dataproduktet, er andre ting som må reguleres. Vilkår og spesifikasjoner vil måtte gjelde alle dataleverandører som skal overføre dataprodukt til Helseanalyseplattformen. Detaljnivået i forskrift vil avgjøre hva som må reguleres i avtale.

Man må sikre at det ikke er tvil om hvilke behandlinger av dataproduktets data som dataleverandør er dataansvarlig for, og hvilke behandlinger som tilgangsførvalter er dataansvarlig for.

Det må også vurderes hvordan egnet kontroll og håndheving av reguleringene skal gjennomføres.

Videre vurdering og konkretisering av nødvendige reguleringer vil måtte skje gjennom egne juridiske utredninger, og vil i stor grad avhenge av det kommende forskriftsarbeidet. Utforming av eventuelle avtaler vil være en videre prosess der registerforvaltere vil involveres ytterligere.

## 4 Realiseringsstrategi for dataprodukt i Helseanalyseplattformens økosystem

Realiseringsstrategien beskriver overordnet tilnærming til hvordan dataprodukter og nødvendige produksjonslinjer for dataprodukter er tenkt etablert som en del av Helseanalyseplattformens økosystem. Dette inkluderer både tilnærming til etablering av de første dataprodukter og etablering av nødvendige prosesser slik at nye dataprodukter vil kunne etableres som en del av økosystemet. Realiseringsstrategien avgrens seg til de nasjonale helsedatakildene, samt eventuelt andre offentlige registre. Sentrale punkter i realiseringsstrategien inkluderer:

- Prioritering av dataprodukter og plan for involvering av nødvendige dataforvaltere
- Rammebetingelser og grensesnitt for etablering av dataprodukter på plattformen
- Organisering av arbeidet i HDP og samarbeidsform med aktører i sektoren
- Overordnet plan og milepæler
- Tilnærming til finansiering av arbeidet
- Gjennomføringsrisiko

### 4.1 Forutsetninger og avhengigheter for dataprodukt på Helseanalyseplattformen

Forutsetningene angir de overordnede rammene som må være på plass for at data skal kunne behandles på Helseanalyseplattformen og tilgjengeliggjøres for sluttbrukere av Helsedataservice. Det er primært to overordnede forutsetninger som må være på plass.

- 1) Helsedataservice må ha hjemmel til å kunne behandle helseopplysninger på Helseanalyseplattformen. For å behandle det omfanget av opplysninger per 01.01.2021 som er skissert i kapittel 0, forutsetter dette endringer i Helseregisterloven og registerforskriftene, samt etablering av egen forskrift for Helseanalyseplattformen og Helsedataservice. Foreløpige signaler tilsier at hjemmel vil være på plass 01.07.2021. Dette dokumentet tar derfor utgangspunkt i at de første dataproduktene skal være ferdigstilt lokalt i registerforvaltningen til 01.01.2021, men at identifiserbare helse- og personopplysninger ikke skal overføres til Helseanalyseplattformen før tilstrekkelig hjemmel foreligger.
- 2) Helsedataprogrammet må ha tilstrekkelig finansiering til å videreutvikle dataplattformen i 2021, samt å kunne støtte dataleverandørene til å kunne gjennomføre aktiviteter i sine virksomheter.

Foruten forutsetningene nevnt over er det også en rekke avhengigheter som må håndteres for at dataprodukt skal kunne eksporteres til Helseanalyseplattformen:

- 1) Dataleverandørene må gjennomføre aktiviteter i sine virksomheter for å definere innhold i dataproduktene, etablere produksjonslinje og forvaltningsprosesser for dataprodukt og utvikle nødvendig infrastruktur for forvaltning og eksport av dataprodukt.



- 2) Hersedataprogrammet må ferdigstille anskaffelse av Data- og analysetjenester som planlagt i løpet av Q2 2020 og dataplattformen inkluderte datamottak må være etablert før data kan overføres.
- 3) Prosesser, roller og ansvarsfordeling mellom Hersedataservice og andre dataleverandører må være besluttet før data kan overføres.
- 4) Hersedataservice må være klar til å ivareta sine plikter som dataansvarlig for data på Helseanalyseplattformen før data kan overføres.
- 5) Prioriterte dataleverandører må være klare til å avgi data til Helseanalyseplattformen fra 01.01.2021. Skarpe data skal ikke overføres før Hersedataservice har hjemmel til å behandle disse på plattformen.

## 4.2 Prinsipper for realisering av dataprodukt på Helseanalyseplattformen

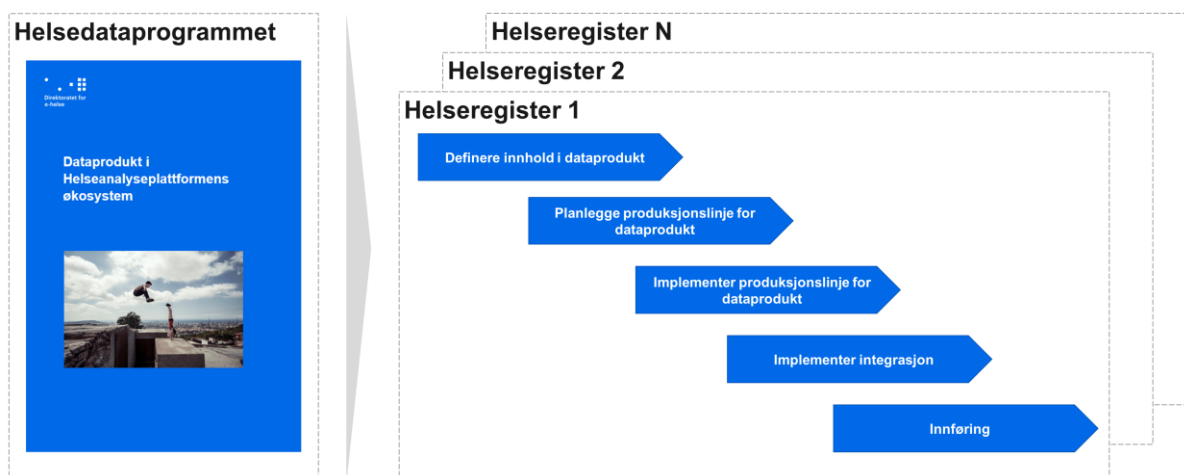
Prinsippene for realisering av dataprodukt på Helseanalyseplattformen følger de generelle realiseringsprinsippene for Helseanalyseplattformen som beskrevet i Konseptvalgutredning for Helseanalyseplattformen. Følgende prinsipper er utledet fra de generelle realiseringsprinsippene:

- 1) Økosystemet med dataleverandører og dataprodukter skal realiseres stegvis.
- 2) Ansvarsdelingen mellom aktørene i økosystemet skal være basert på lov- og forskriftsregulering og avtaler. Ansvarsfordelingen må bygge på prinsippet om mest mulig effektiv offentlig forvaltning, samt ivareta kravene i personvernforordningen.
- 3) De første dataproduktene skal være tilpasset forskningsformål. Dette innebærer at de første dataproduktene bør
  - være hentet fra de mest brukte, direkte identifiserbare helseregistrene.
  - kunne dekke minst 80 % av forskerutleveringene fra det aktuelle helseregisteret.
  - inneholde godt dokumenterte data som i begrenset grad krever manuell veiledning og forklaring.
  - inneholde kvalitetssikrede data og fortrinnsvis ha lav oppdateringsfrekvens.
- 4) Videre utvikling av dataprodukt skal også være tilpasset andre formål, eksempelvis statistikk, næringsutvikling, helseovervåking og beredskap.

## 4.3 Hersedataprogrammets og dataleverandørenes rolle i realiseringen av data på Helseanalyseplattformen

Realisering av dataprodukt på Helseanalyseplattformen forutsetter et godt samarbeid mellom Hersedataprogrammet og dataleverandørene innenfor Hersedataprogrammets programperiode. Som illustrert i Figur 8 er det til nå lagt opp til en tilnærming der Hersedataprogrammet i samarbeid med registermiljøene definerer rammebetingelsene for hvordan dataprodukt skal håndteres og overføres til Helseanalyseplattformen, mens registermiljøene selv må gjennomføre aktiviteter i sin virksomhet for å definere innhold i dataproduktene, etablere produksjonslinje for dataprodukt og utvikle nødvendig infrastruktur for forvaltning og eksport av dataprodukt til Helseanalyseplattformen.

Hersedataprogrammet vil samtidig inneha en koordinerende og støttende rolle ut mot dataleverandørene. Dette for å sikre en helhetlig tilnærming til data på plattformen, indentifisere avhengigheter mellom prosjektaktiviteter i registermiljøene og utviklingen av plattformen og å legge til rette for god informasjonsflyt på tvers av dataleverandørene.



Figur 8 Helseanalyseplattformen vil stå ansvarlig for å sette rammebetingelsene for *dataprodukt* på Helseanalyseplattformen, mens registermiljøene vil måtte gjennomføre prosjektaktiviteter for å realisere *dataprodukt* fra sine registre.

Formålet med en slik tilnærming vil være å etablere noen felles rammer alle dataproduktene skal operere under. Samtidig er det registermiljøene som kjenner sine data, sine forvaltningsprosesser og sine tekniske løsninger best. Det er derfor avgjørende at registermiljøene har hovedansvaret for gjennomføringen av aktivitetene som skal gjennomføres i registrene.

Under programperioden vil det tilrettelegges for at dataleverandøren kan søke om medfinansiering for etablering av nødvendige produksjonslinjer for sine dataprodukter. Det jobbes også med å se på ulike modeller for finansiering av drift og forvaltning av Helseanalyseplattformen og Helsedataservice. Blant annet vurderes brukerfinansiering der deler av inntektene for bruk av data tilbakeføres dataleverandørene. Endelig modell er foreløpig ikke avklart.

Tilnærmingen over er gyldig for de nasjonale helsedatakildene. Helsedataservice vil samtidig fungere som et serviceorgan som vil kunne ta imot dataprodukter fra andre potensielle dataleverandører som private virksomheter og forskningsprosjekt. Denne typen dataleverandører vil ikke kunne søke om medfinansiering til etablering av sine dataprodukt, men vil kunne inngå som en del av en fremtidig finansieringsmodell og kunne motta midler basert på bruken av dataprodukter de har tilbudt på Helseanalyseplattformen.

## 4.4 Prioritering av dataprodukt på Helseanalyseplattformen

Basert på prinsippet om stegvis realisering av økosystemet må det gjøres en prioritering av hvilke dataleverandører og datakilder som skal få bistand til å levere dataprodukter til Helseanalyseplattformen. Datakildene som i første omgang vurderes er lovbestemte helseregistre, nasjonale medisinske kvalitetsregistre, befolkningsbaserte helseundersøkelser, demografiske og sosioøkonomiske data fra SSB og eventuelt andre offentlige registre. Det er ikke gitt at dataprodukt fra alle disse kildene skal overføres til Helseanalyseplattformen og det vil hele veien være behov for kost-nytte-vurderinger av den enkelte datakilden og det enkelte dataproduktet.

I dette kapittelet introduseres kvalifikasjons- og prioriteringskriterier for datakilder på Helseanalyseplattformen. Basert på kriteriene presenteres også et forslag til prioritering av datakilder i ulike faser.

#### 4.4.1 Kvalifikasjonskriterier

For at et dataprodukt skal kunne publiseres på Helseanalyseplattformen må det oppfylle et sett med kvalifikasjonskriterier. Kriteriene er definert for å sikre at Helsedataservice har hjemmel til å behandle data fra de aktuelle kildene og at dataene kan gjøres forståelige på Helseanalyseplattformen.

- 1) **Rettsgrunnlag:** Formålet med datakilden må være i samsvar med formålet for behandling av data på Helseanalyseplattformen.
- 2) **Metadata:** Dataproduktet må ha metadata dokumentert i henhold til *Nasjonal spesifisering for metadata om helsedata*.

#### 4.4.2 Prioriteringskriterier

For å sikre transparens i prioriteringene av datakilder på Helseanalyseplattformen er det definert et sett med prioriteringskriterier. Kriteriene omfatter dagens bruk, potensialet for merbruk, politiske føringer, kostnader og risiko. For å vurdere merbrukspotensialet tas det utgangspunkt i parameterne datakvalitet og koblingsmuligheter. Kriteriene danner grunnlaget for en helhetlig og kvalitativ vurdering av når datakildene bør prioriteres inn på plattformen.

- 1) **Bruk av data:** Dagens bruk av datakilden. For eksempel antall søknader om utleveringer.
- 2) **Datakvalitet:** En vurdering av datakvaliteten i datakilden basert på kildens kompletthet, korrekthet/validitet, aktualitet, sammenlignbarhet, reliabilitet og relevans.<sup>8</sup>
- 3) **Koblingsmuligheter:** Hvorvidt det er mulig å sammenstille data fra den aktuelle kilden med andre datakilder. For eksempel om datakilden inneholder personopplysninger som gjør at den kan sammenstilles på individnivå og om det er hjemmel til å sammenstille den med andre datakilder.
- 4) **Politiske og strategiske føringer:** Direkte og indirekte føringer for hva som skal prioriteres. For eksempel utvalgte datakilder eller forskning, kvalitetsforbedring og næringsutvikling innen utvalgte helsefaglige tema.
- 5) **Kostnader:** Anslåtte totale kostnader knyttet til å etablere dataprodukt fra den aktuelle datakilden på Helseanalyseplattformen.
- 6) **Risiko:** Generell risiko for at datakilden ikke vil kunne integreres mot Helseanalyseplattformen. For eksempel risiko ved manglende kapasitet hos dataforvalter eller ved at datakilden ligger på en ukjent teknisk plattform.

#### 4.4.3 Prioritering av datakilder på Helseanalyseplattformen

Det er en målsetning å etablere de første dataproduktene på Helseanalyseplattformen for nasjonale helseregistre og eventuelt andre offentlige registre innenfor følgende faser.

---

<sup>8</sup> Konseptvalgutredning for Helseanalyseplattformen, Vedlegg F:

<https://ehelse.no/publikasjoner/konseptvalgutredning-for-helseanalyseplattformen/Konseptvalgutredning%20for%20Helseanalyseplattformen.pdf>

- **Fase 1:** 01.01.2021<sup>9</sup>
- **Fase 2:** 01.07.2021
- **Fase 3:** 01.01.2022
- **Fase 4:** 01.07.2022

Basert på prioriteringskriteriene fra kapittel 4.4.1 foreslås de ulike datakildene inn i de ulike fasene. Datakildene som foreslås inkludert i Fase 1 har de siste årene vært de mest brukte, direkte identifiserbare, sentrale helseregistrene i forskningsøyemed. Det foreligger i tillegg politiske føringer fra Helse- og omsorgsdepartementet om at Norsk Pasientregister (NPR), Kommunalt pasient og brukerregister (KPR), Dødsårsaksregisteret (DÅR), Medisinsk fødselsregister (MFR) og Kreftregisteret (Kreg) bør inngå i en tidlig versjon av Helseanalyseplattformen.

For Fase 2, 3 og 4 er det gjort et arbeid på å konkretisere hvilke registre man ser for seg vil kunne inngå når. Det er foreløpig for tidlig å kunne beslutte hvilke kvalitetsregistre og helseundersøkelser som med sikkerhet vil kunne inngå i de ulike fasene. Dette henger sammen med målsetningen om stegvis realisering, der man ønsker å detaljplanlegge for det som kommer nært i tid. Det vil samtidig være behov for å bygge mer kunnskap om det som skal skje lengre frem i tid.

Fase	Datakilder
<b>Fase 1</b>	Norsk Pasientregister (NPR), Kommunalt pasient og brukerregister (KPR), Dødsårsaksregisteret (DÅR), Medisinsk fødselsregister (MFR), Kreftregisteret (Kreg) og Folkeregisteret
<b>Fase 2</b>	SSB-data, Legemiddelregisteret, Hjerter- og karregisteret (HKR), Bivirkningsregisteret og utvalgte kvalitetsregistre
<b>Fase 3</b>	SYSVAK, MSIS, flere kvalitetsregistre og utvalgte helseundersøkelser
<b>Fase 4</b>	Flere kvalitetsregistre og helseundersøkelser

Tabell 1 Oversikt over forslag til hvilke datakilder som skal inngå i de ulike fasene av eksport av data til Helseanalyseplattformen

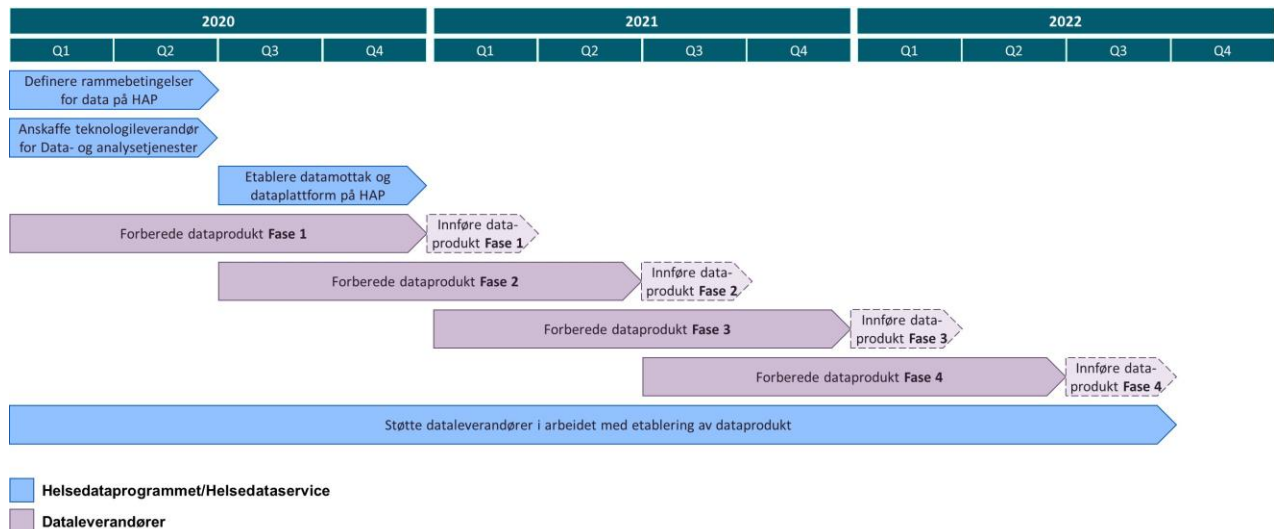
## 4.5 Overordnet realiseringsplan

Som beskrevet i 4.3 vil Helsedataprogrammet være avhengig av prosjektaktiviteter i registermiljøene for å kunne levere gode og relevante dataprodukt på Helseanalyseplattformen. Dette kapitlet beskriver den overordnede realiseringsplanen med fokus på samspillet mellom Helsedataprogrammet og registermiljøene.

### 4.5.1 Tidsplan

Figur 9 illustrerer den overordnede realiseringsplanen med prosjektaktiviteter i regi av Helsedataprogrammet og dataleverandørene. Tidsplanen tar utgangspunkt i at rammebetingelsene beskrevet i kapittel 4.1 er oppfylt.

<sup>9</sup> Helsedataservice vil ikke ha hjemmel i helseregisterloven med tilhørende forskrift for å kunne behandle data før 01.07.2021, og eventuell overføring av data før dette må skje basert på et midlertidig rettsgrunnlag.



Figur 9 Overordnet realiseringsplan for data på Helseanalyseplattformen

## 4.5.2 Gjennomføringsrisiko

Etablering av dataprodukt på Helseanalyseplattformen med tilhørende prosjektavhengigheter i flere virksomheter vil medføre flere former for risiko. De mest sentrale risikoelementene er oppsummert i punktlisten under. Risikovurderingen og tiltaksplan for hvert av risikoelementene håndteres internt i prosjektet og oppdateres løpende. Den inkluderes derfor ikke i dette dokumentet.

- 1) Dataleverandørene har ikke kapasitet til å etablere nødvendige produksjonslinjer for overføring av dataprodukter til Helseanalyseplattformen.
- 2) Forsinkelser i forskriftsarbeidet medfører at det ikke vil foreligge nødvendig hjemmel for overføring av dataprodukter til Helseanalyseplattformen i det omfanget som er skissert i kapittel 0.
- 3) Uklar og/eller ulik oppfatning av ansvarsdeling mellom Hersedataservice/ Helseanalyseplattformen og dataleverandørene hindrer overføring av dataprodukter og etablering av en fungerende forvaltningsprosess.
- 4) Kompleksitet og/eller uenighet ved sikring og kryptering av dataen hindrer at dataprodukter kan overføres til Helseanalyseplattformen.
- 5) Manglende kapasitet og kompetanse i Helsedataprogrammet forsinker tydeliggjøringen av rammebetingelsene for etablering av dataprodukt hos dataleverandører og dermed overføringen av data til Helseanalyseplattformen.
- 6) Manglende kapasitet og kompetanse i Hersedataservice/NHN etter endt programperiode forsinker overføringen av nye datakilder til Helseanalyseplattformen.
- 7) Forsinket anskaffelse og etablering av Helseanalyseplattformens datamottak og dataplattform forsinket overføring av dataprodukt til Helseanalyseplattformen.

## 5 Begrepsliste

Begrep	Forklaring
<i>Analyseøkosystem</i>	Virksomheter som leverer analysetjenester, analyseverktøy og <i>dataprodukter</i> gjennom en felles digital plattform og markeds plass og et felles sett med forvaltningsprosesser. Aktører inkluderer plattformeier, leverandører og <i>sluttbrukere</i>
<i>API</i>	Forkortelse for det engelske begrepet Application Programming Interface. Hjelpesverktøy ved programmering, et grensesnitt mot en eller flere tjenester i et operativsystem, en databasetjener eller lignende
<i>Behandlingsansvarlig</i>	Se <i>dataansvarlig</i>
<i>Dataansvarlig</i>	<p>Ansvarlig for behandling av <i>helseopplysninger</i>.</p> <p>Kilde: Helseregisterloven § 2 bokstav d</p> <p>En fysisk eller juridisk person, en offentlig myndighet, en institusjon eller ethvert annet organ som alene eller sammen med andre bestemmer formålet med behandlingen av <i>personopplysninger</i> og hvilke midler som skal benyttes; når formålet med og midlene for behandlingen er fastsatt i unionsretten eller i medlemsstatenes nasjonale rett, kan den <i>behandlingsansvarlige</i>, eller de særlige kriteriene for utpeking av vedkommende, fastsettes i unionsretten eller i medlemsstatenes nasjonale rett.</p> <p>I helse- og omsorgssektoren brukes begrepet <i>dataansvarlig</i> i stedet for <i>behandlingsansvarlig</i>, for ikke å skape usikkerhet knyttet til om det er snakk om den som er ansvarlig for pasientbehandlingen eller ansvarlig for behandling av <i>personopplysninger</i>.</p> <p>Kilde: Personvernforordningen artikkel 4 nr. 7</p>
<i>Databehandler</i>	<p>En fysisk eller juridisk person, offentlig myndighet, institusjon eller ethvert annet organ som behandler <i>personopplysninger</i> på vegne av den <i>behandlingsansvarlige</i>,</p> <p>Kilde: Personvernforordningen artikkel 4 nr. 8</p>
<i>Datakilde</i>	Samlebetegnelse av de tekniske løsningene som avgir data til <i>Helseanalyseplattformen</i> .

<i>Dataleverandør</i>	Rolle i <i>Helseanalyseplattformens økosystem</i> som forvalter og overfører <i>dataprodukter</i> til <i>Helseanalyseplattformen</i> .
<i>Dataprodukt</i>	En sammenstilling av datasett, metadata og betingelser der en dataleverandør er produkteier med faglig ansvar for innholdet i produktet. Dataproduktet eksisterer for å dekke konkrete behov, og det bør være behov som oppstår mer enn én gang. Dataprodukter kan inneholde anonyme eller personidentifiserbare data.
<i>Datasett</i>	Et datasett er en samling med data, for eksempel i form av en tabell, liste eller en database som kan gjøres tilgjengelig som en nedlastbar fil, og/eller nåes via et Web-API.  Kilde: Standard for beskrivelse av datasett og datakataloger (DCAT-AP-NO)
<i>Eksplorativ analysetjeneste</i>	Digital tjeneste der en <i>sluttbruker</i> kan analysere data uten å måtte søke om bruk av helsedata.  Eks: <i>kohortutforsker</i> .
<i>Grensesnitt</i>	Et grensesnitt er en abstraksjon for kommunikasjonen mellom to systemer eller konsepter
<i>Helseanalyseplattformen</i>	Den tekniske løsningen som skal tilby følgende digitale tjenester til ulike sluttbrukere slik at tilgang til helsedata skal bli enklere, raskere og sikrere. <ul style="list-style-type: none"> <li>• Data- og analysetjenester</li> <li>• Søknad- og veiledningstjenester</li> <li>• Forvaltningstjenester</li> </ul>
<i>Helsedataservice</i>	Den organisatoriske enheten som vil ha <i>dataansvaret</i> for data på <i>Helseanalyseplattformen</i> og være ansvarlig for alle prosessene knyttet til tilgangsforvaltning av helsedata som ligger på <i>Helseanalyseplattformen</i> .
<i>Helseopplysninger</i>	<i>Personopplysninger</i> om en fysisk persons fysiske eller psykiske helse, medregnet om ytelse av helsetjenester, som gir informasjon om vedkommendes helsetilstand.  Kilde: Helseregisterloven § 2 bokstav a, Personvernforordningen artikkel 4 nr. 15.

<i>Helseregister</i>	<p>Enhver strukturert samling av <i>personopplysninger</i> som er tilgjengelig etter særlige kriterier, og som inneholder <i>helseopplysninger</i>.</p> <p>Kilde: Helseregisterloven § 2 bokstav c, Personvernforordningen artikkel 4 nr. 6</p> <p>Registerbegrepet i loven er et logisk begrep. Registerbegrepet er ikke en datateknisk definisjon. Et helseregister er en samling opplysninger som stadig er i endring, ved at opplysningene korrigeres, suppleres, bearbeides og så videre, og hvor man etter bearbeidelsesprosessen har fremskaffet noe mer enn hva summen av de enkelte opplysningene skulle tilsi.</p> <p>Kilde: Prop. 72 L (2013-2014) punkt 10.1.3 merknad til helseregisterloven § 2.</p>
<i>Kohortutforsker</i>	<p>En digital analysetjeneste som gjør det mulig for <i>sluttbruker</i> å identifisere hvor mange individer som har bestemte karakteristika uten at sluttbruker ser <i>personopplysninger</i>.</p>
<i>Metadata</i>	<p>Data som definerer og beskriver data.</p>
<i>Personopplysning</i>	<p>Enhver opplysning om en identifisert eller identifiserbar fysisk person («den registrerte»); en identifiserbar fysisk person er en person som direkte eller indirekte kan identifiseres, særlig ved hjelp av en identifikator, f.eks. et navn, et identifikasjonsnummer, lokaliseringsopplysninger, en nettidentifikator eller ett eller flere elementer som er spesifikke for nevnte fysiske persons fysiske, fysiologiske, genetiske, psykiske, økonomiske, kulturelle eller sosiale identitet.</p> <p>Kilde: Personvernforordningen artikkel 4 nr. 1</p>
<i>Rådata</i>	<p>Data som er samlet inn, men ikke foredlet eller tilpasset bestemte formål.</p>
<i>Sluttbruker</i>	<p>I denne sammenheng rollen som søker om, og bruker data fra <i>dataproduktene</i> for å skape ny kunnskap eller innsikt.</p>
<i>Tilgangsførvalter</i>	<p>Rolle i <i>dataøkosystemet</i> med ansvar for søknadsmottak, saksbehandling, tilrettelegging, tilgangsstyring og tilgjengeliggjøring av <i>datasett</i> fra <i>dataprodukter</i> til <i>sluttbrukere</i>. Denne rollen er tenkt å ligge under <i>Helsedataservice</i>.</p>
<i>Tilgjengeliggjøring</i>	<p>Begrepet brukes vidt og omfatter all data som databehandler/dataansvarlig gjør tilgjengelig for</p>



	analyse ved utlevering (distribusjon av data) eller ved å gi bruker tilgang til data i et analysemiljø.
<i>Variabel</i>	En variabel er et informasjonsbærende element som gir et mål på en enhet. For eksempel kan 'M' og 'K' være mål på variabelen kjønn på enheten person.

