



## Saksdokument

### Porteføljestyre for forskning og kunnskapsressurser

|                     |                                                     |
|---------------------|-----------------------------------------------------|
| <b>Til</b>          | Porteføljestyre for forskning og kunnskapsressurser |
| <b>Dato</b>         | 05.06.2024                                          |
| <b>Saksnummer</b>   | 24/24                                               |
| <b>Saksarkivnr.</b> | 24/00000                                            |
| <b>Sakstype</b>     | D-sak                                               |

|                      |                          |
|----------------------|--------------------------|
| <b>Saksansvarlig</b> | Vigdis Namtvedt Kvalheim |
| <b>Saksbehandler</b> | Bodil Agasøster          |

## Modernisering av Data om det norske samfunn (MODANSA)-satsingen

### Forslag til vedtak

Porteføljestyret anbefaler at Sikt jobber videre med satsingen basert på kommentarene og innspillene som framkom i møtet og ber om at styret holdes orientert om videre framdrift i saken.

### Bakgrunn

Porteføljestyret har i flere omganger blitt orientert om Sikts planer om å videreutvikle sin forskningsdataplattform til å kunne formidle data i databaser og tekstdata på en brukervennlig måte, (jf. Sak 16/23 MODANSA – En modernisering av Sikt infrastruktur for forskningsdata, Sak 25/23 Langtidsplan for fellestjenester 2024-2026 og Sak 37/23 Langtidsplan for fellestjenester 2024-2026). I forbindelse med behandlingen av langtidsplanen for divisjonen (til sak 37/23) vedtok porteføljestyret planen med forbehold om «...at MODANSA vurderes av Porteføljestyret når detaljert plan, inkludert budsjett for MODANSA legges fram for Porteføljestyret.»

I denne strategiperioden er MODANSA hovedprioritet i produktområdets kontinuerlige utvikling av infrastruktur tilknyttet Sikts forskningsdataplattform. Sikt har i innværende år har avsatt 5,5 MNOK til satsingen fra risikokapitalen. I tillegg er det lagt inn 1,3 MNOK fra årets bevilgning og tilskudd. I tråd med bestillingen fra styret legges saken nå fram med fremdriftsplan med leveransebeskrivelser og kostnadsestimat med flere alternativer til gjennomføring. Siden MODANSA var sak for Porteføljestyret har Produktrådet for Dataforvaltning behandlet satsingen, (se Vedlegg 1). Her ble følgende punkter fremhevet:

- Produktrådet uttrykker at Sikt bør prioritere data som antas å ha størst potensial for bruk. Dette for å sikre at ressursene allokeres effektivt og at innsatsen rettes mot de mest verdifulle dataene. Det er Forvaltningsdatabasen som har flest brukere, dette gjenspeiles i både nettrafikk og siteringer.
- Det er positivt at MODANSA vil utvikle datainfrastruktur som tilrettelegger for mer deling av åpne data.
- MODANSA-satsingen kan spille en nøkkelrolle i å fremme FAIR-prinsippene innen forskningsmiljøer som arbeider med kvalitative tekstdata. Disse dataene har til nå ikke oppnådd samme nivå av FAIR-implementering som surveydata, og en forbedring på dette området er nødvendig. Å gjøre dette vil forutsette fokus på tekstdata.



## Saksdokument

### Porteføljestyre for forskning og kunnskapsressurser

- Det ble stilt spørsmål om muligheten for å overføre teknologi og lærdom fra MODANSA til andre områder innen Sikt. Dette ble bekreftet. I MODANSA vil det utvikles komponenter, som potensielt kan gjenbrukes blant annet i formidling av tekstdata innsamlet som intervju og åpne svar i surveyundersøkelser. I tillegg vil MODANSA medføre et løft i kompetanse knyttet til bruk av kunstig intelligens, noe som vil være verdifullt for utviklingen av ytterligere tjenester.
- For å effektivt kommunisere fordelene med MODANSA til eksterne parter, anbefales det at Sikt illustrerer satsingen med konkrete eksempler og "use cases". Dette vil gjøre det lettere for interessenter å forstå hvilke gevinster MODANSA kan tilby.

### Økonomi og alternativer for gjennomføring

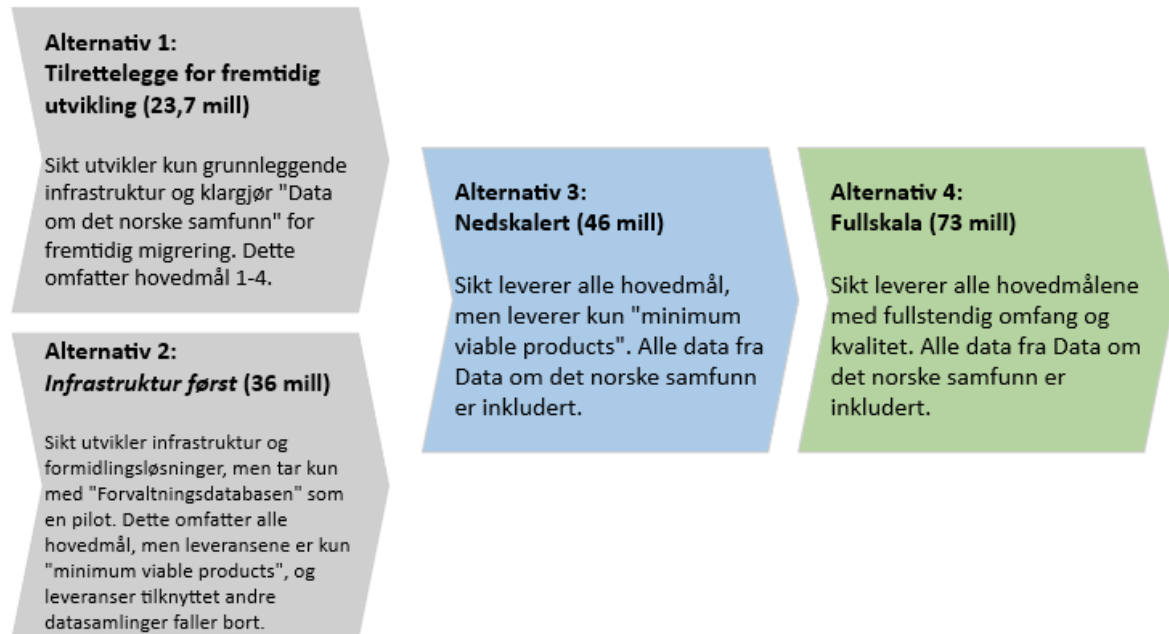
Detaljert kostnadsvurdering for satsingene er gitt i som separat vedlegg med detaljerte leveransebeskrivelser og kostnadsestimat. Vurderingen inkluderer to scenarier - en for et nedskalert initiativ, og en for fullskala realisering. Den nedskalerte modellen inkluderer etablering av dataplattform. En slik plattform danner grunnlaget for videre utvikling av datainfrastrukturen og formidlingskanaler. Fullskalamodellen omfatter en fullstendig realisering av satsingens ambisjoner relatert til omfang og kvalitet. I tillegg presenteres to alternativer til som baserer på disse i vedlegg 2. Alle fire er kort presentert under. Alternativene bygger på hverandre og kan gjennomføres suksessivt, se figur 1.

Finansieringen av MODANSA kommer fra ulike kilder, hvor en betydelig del av midlene kommer fra Sikts risikokapital. I tillegg vil vi benytte noe bevilgningsmidler og tilskudd. Det vil være behov for å knytte til seg samarbeidspartnere som kan bidra, både økonomisk og faglig. Dette kan være institusjoner eller andre aktører som ønsker å delta aktivt i satsingen, for å bevare formidlingen av data, eller fordi de besitter data som må formidles via en solid datainfrastruktur.

I fullskalaformat har satsingen større finansieringsbehov enn tidligere estimert, og som ligger til grunn i satsingsforslag, langtidsplan og intern godkjenning. Det opprinnelige kostnadsoverslaget på 35 MNOK var basert på tidligere anslag. De to siste alternativene har vesentlig større kostnadsramme, og er derfor mindre realistisk å få gjennomført i denne omgang gitt de finansielle rammene som foreligger nå. Vi finner det imidlertid riktig å vise til disse fordi de vil danne grunnlag for videre utvikling eller utvidelse av satsingen dersom økonomiske forutsetningen er til stede.

## Saksdokument

### Porteføljestyre for forskning og kunnskapsressurser



Figur 1: Alternativ 1 og 2 kan utvides til alternativ 3 og deretter til alternativ 4.

#### Alternativ 1: Tilrettelegge for fremtidig utvikling (23,7 MNOK)

Dette alternativet innebærer at MODANSA *kun* leverer de første fire hovedmålene; kompetanseheving, standardisering av metadata, omstrukturering av data og videreutvikling av forskningsdataplattformen. Disse målene er avgjørende for hva det vil være mulig å tilby i fremtidige tjenester. Ved kun å gjennomføre hovedmål 1-4, vil metadata, data og datainfrastruktur være forberedt for fremtidig utvikling av formidlingstjenester og drift. Alternativ 1 vil danne grunnlaget for å senere få på plass funksjonalitet og grensesnitt ut mot forskerne og andre sluttbrukere, som er beskrevet i alternativ 3 og 4.

#### Alternativ 2: Infrastruktur først (36 MNOK)

Vi har en visjon om at Sikts forskningsdataplattform skal videreutvikles til å håndtere kvalitative tekstdata og kvantitative matrisedata som ikke er generert fra surveyforskning. Denne type data har behov for en plattform som bearbeider og formidler dataene i samsvar med FAIR-prinsippene og oppfylle krav til universell utforming. Dette alternativet erkjenner at det vil ta tid å realisere visjonen uten god finansiering. Satsingen gjennomføres derfor i sekvenser. Det betyr at Sikt først vil utvikle infrastrukturen, og *kun* migrere data fra Forvaltningsdatabasen. Forvaltningsdatabasen er mest brukt, og den består av ulike type data – som gjør den særlig egnet som pilot. Kostnadsestimatet for dette alternativet baseres på estimatene fra nedskalert satsing (alternativ 3). Rensing og migrering av de andre datasamlingene vil komme senere i sekvenser og vil gjennomføres under andre faner og med annen finansiering, eventuelt senere i MODANSA om finansiering tillater realisering av alternativ 3 eller 4. Det betyr at alternativ 2 vil gjennomføre alle hovedmålene, som nedskalert satsing, men leveransene tilknyttet Data om det politiske system, Kommunedatabasen og Kirkedatabasen vil ikke gjennomføres. En ulempe med alternativ 2 er at datakildene ikke vil kunne kobles sammen før alle har blitt rensket og migrert over til forskningsdataplattformen.



## Saksdokument

### Porteføljestyre for forskning og kunnskapsressurser

#### Alternativ 3: Nedskalert satsing (46 MNOK)

Den nedskalerte satsingen fokuserer på de mest grunnleggende behovene for å etablere en plattform for kontinuerlig utvikling av datainfrastrukturen og formidlingskanalene. Dette innebærer at ressursene i all hovedsak anvendes til hovedmål 1-5, da disse i liten grad kan reduseres. Dette skyldes at de legger grunnlaget for alt som vil være mulig i MODANSA og i fremtidige satsinger. Utover dette vil kostnadsestimatene ta sikte på etablering av «minimum viable products», dvs. systemer og tjenester som er kokt ned til det mest essensielle for å kunne tas i bruk.

Dette innebærer at MODANSA vil realisere de grunnleggende målene i satsingen, men det vil være behov for kontinuerlig produktutvikling for å heve kvaliteten på tjenester og systemer i etterkant. Alternativet vil altså ha noe redusert funksjonalitet i forhold til en full satsing (alternativ 4), men vil gi god effekt for sluttbrukerne og på interne prosesser. Den vil ikke oppfylle alle muligheter som ligger i satsingen, men vil legge til rette for videre utvikling i framtiden. Dersom alternativet blir valgt, anbefaler vi å forlenge prosjektperioden med ytterligere to år, fram til 2030. På den måten blir belastningen per år noe lavere, både finansielt og kapasitetsmessig.

#### Alternativ 4: Full satsing (73 MNOK)

I denne satsingen skal MODANSA realisere hele omfanget beskrevet i hovedmålene. Dette innebærer en dyp standardisering og omstrukturering av metadata og data, slik at det kan bygges tilpassete tjenester – både internt og eksternt – med stor grad av automatiserte prosesser. For brukeren vil det bli enklere å finne, koble og bruke data og – i et universelt utformet og brukervennlig grensesnitt. Videre vil det effektivisere arbeidet med kuratering, prosessering og publisering.

Alternativet omfatter alle elementer og alle kostnader produktområdet har vurdert. Dette er trolig urealistisk å få finansiert og gjennomført nå. Dersom det åpner seg nye finansieringsmuligheter i framtiden, f.eks. gjennom prosjektmidler gjennom Forskningsrådets infrastrukturutlysninger, tildeling fra Kunnskapsdepartementet som svar på satsingsforslaget fra Sikt eller andre muligheter, vil man kunne gjøre en ny vurdering i forhold til dette alternativet.

#### Tidsplan

I framdriftsplanen har MODANSA oppstart andre kvartal 2024 og er planlagt å vare ut 2028, evt. 2030, dels avhengig av årlig finansiering og tilgang på utviklingskapasitet. Så langt arbeider teamet med å bygge kunnskap. I tråd med Sikts modell for produktutvikling, vil satsingen gjennomføres som kontinuerlig utvikling. Nye løsninger vil derfor ruller ut og tas i bruk suksessivt gjennom utviklingsperioden. Dette har mange fordeler, bl.a. at det interne arbeidet i Sikt og brukerne raskt kan ta løsningene i satsingen.

I tabellen på neste side vises planlagt arbeid med hovedmålene for satsingen.

## Saksdokument

### Porteføljestyre for forskning og kunnskapsressurser

Tabell 1: Fremdriftsoversikt for hovedmål.

|                   |                                                      | 2024 | 2025 | 2026 | 2027 | 2028 |
|-------------------|------------------------------------------------------|------|------|------|------|------|
| <b>Hovedmål 1</b> | <i>Kompetanseheving og opplæring</i>                 | X    | X    |      |      |      |
| <b>Hovedmål 2</b> | <i>Standardisering av metadata</i>                   | X    | X    | X    |      |      |
| <b>Hovedmål 3</b> | <i>Omstrukturering av data</i>                       | X    | X    | X    |      |      |
| <b>Hovedmål 4</b> | <i>Videreutvikling av forskningsdataplattformen</i>  |      |      | X    | X    |      |
| <b>Hovedmål 5</b> | <i>Migrering av data og metadata</i>                 |      |      | X    | X    |      |
| <b>Hovedmål 6</b> | <i>Utvikle formidlingskanal for åpne databaser</i>   |      |      |      | X    | X    |
| <b>Hovedmål 7</b> | <i>Utvikle formidlingskanal for åpne tekstdata</i>   |      |      |      | X    | X    |
| <b>Hovedmål 8</b> | <i>Utvikle tilgangssystem</i>                        |      |      |      |      | X    |
| <b>Hovedmål 9</b> | <i>Etablere ny arbeidsflyt og driftsorganisering</i> |      |      |      |      | X    |

## Konklusjon

Basert på nåværende rammebetingelser og oppdatert kostnadsoversikt framkommet gjennom arbeidet med leveranseplanen, foreslår Sikt å jobbe videre med finansiering av satsingen parallelt med oppstart av alternativ 2. Etter hvert som finansiering kommer på plass, vil vi fortsette med leveransene i henhold til alternativ 3 og 4.

Sikt har en klar plan for hvordan datainfrastrukturen skal videreutvikles for å sikre at alle data tilgjengeliggjort hos Sikt formidles i samsvar med FAIR-prinsippene. Hvorvidt ambisjonen vil realiseres innenfor et fem eller ti års perspektiv er avhengig av hvilken finansielle støtte satsingen får. Dersom satsingen ikke får innvilget støtte fra Kunnskapsdepartementet, vil Sikt søke midler fra Norges forskningsråd. Samtidig leter vi etter samarbeidspartnere som kan bidra med medfinansiering. I mellomtiden vil Sikt finansiere satsingen med risikokapital opp til de tre første årene. Hvis ekstern finansiering ikke oppnås i denne perioden, vil utviklingsarbeidet bli betydelig bremset. Videre arbeid må da finansieres gjennom kontinuerlig utvikling under normal drift, noe som vil forlenge tidsaspektet og medføre alternative kostnader og risikoer for både kunnskapssektoren og Sikt.

Alternativ 2 er utgangspunktet for arbeidet fremover. Dette alternativet har det laveste kostnadsestimatet som vil gi kunnskapssektoren infrastruktur som kan benyttes. Alternativ 1 har et lavere kostnadsestimat, men gir ingen umiddelbar gevinst, kun fremtidig tilrettelegging. En fordel med alternativ 2 er at det bygger opp innsikt og kunnskap nødvendig for å realisere ambisjonene beskrevet i alternativ 3 og 4. Kostnadsestimatene for disse alternativene vil også bli mer presise, som følge av en lavere grad av usikkerhet.

Fordelen med å starte satsingen til tross for uavklart finansiering, er at det gir Sikt muligheten til å demonstrere hva vi kan oppnå. Leveransene i første fase vil gi konsepter og prototyper som kan tiltrekke seg andre aktører, med tanke på finansiering og samarbeid. Aktivitetene som gjennomføres i oppstarten, som inkluderer heving av kunnskap og kompetanse på AI og datatyper, rensing og standardisering av metadata og data, samt etablering av samarbeid med andre aktører (Histreg, Arkivverket, DFØ (Kudos)), vil uansett ha stor verdi for Sikt som dataforvalter.

## Vedlegg

PS-FK-24-24 Vedlegg 1: MODANSA Referat Møte i produktrådet for dataforvaltning 03.04.2024

PS-FK-24-24 Vedlegg 2: MODANSA Kostnadsestimat

Separat vedlegg: MODANSA Fremdriftsplan med budsjettering