



UNIT - Dir. for IKT og fellestjenester i høyere utdanning og forskning

Abels gate 5A
7030 TRONDHEIM

Deres ref
22/00009

Vår ref
2022/377-BJUB

Dato
02.02.2022

Hørings svar - konseptvalgutredning for Tilgjengeliggjøring og deling av forskningsadministrativ informasjon (TILDE)

Universitetet i Bergen (UiB) takker for anledningen til å gi innspill til Sikts planer for utvikling av system for tilgjengeliggjøring og deling av forskningsadministrativ informasjon. Hovedmålet med høringen er tilslutning til Sikts forslag om konsept for system for masterdata. Vi oppfatter at konseptene bør utredes nærmere før vi kan gi klar tilslutning til modell.

UiB anbefaler

- Å gå videre med arbeid for å etablere felles tjenester for forvaltning av data om personer, institusjonsenheter og prosjekter
- Å utvikle et bedre beslutningsgrunnlag for valg av modell, herunder tydeligere beskrivelser av økonomiske konsekvenser. Framstillingen av konsept 2 og 3 bør balanseres, og kritisk vurdering bør konsentreres til oppsummeringen i kapittel 8
- Å synliggjøre grensene for systemet, særlig mot beslektede administrative systemer som i liten grad er beskrevet, slik som personal, regnskap og prosjektøkonomi
- Å tydeliggjøre og konkretisere investerings- og driftskostnader, kostnader for konsumenter/institusjoner, tidshorisont for utvikling og generell lønnsomhet
- Å tydeliggjøre hvordan krav til eierskap og integritet i egne virksomhetsdata sikres i et system med felles masterdataforvaltning på tvers av institusjoner
- Å ha en god plan for involvering av brukere og brukerinstitusjoner ved et eventuelt utviklingsprosjekt

Disse momentene er utredet og begrunnet i vedlagte høringsuttalelse.

Vennlig hilsen

Margareth Hagen
rektor

Robert Rastad
universitetsdirektør

Dokumentet er elektronisk godkjent og har derfor ingen håndskrevne signaturer.

Vedlegg: Høringsuttalelse

Bakgrunn

Digitaliseringsstyret for høyere utdanning og forskning ga i mai 2021 oppdrag til Unit (Sikt fra 1.1.22) om å gjennomføre en konseptvalgutredning for løsning som skal ivareta tilgjengeliggjøring og deling av forskningsadministrativ informasjon i kunnskapssektoren. Institusjoner som bidrar i offentlig finansiert forskning bruker informasjon om forskning til en mengde formål, og har et klart behov og eget ansvar for høy kvalitet på grunnlagsdata for alt fra publikasjonslister, finansieringssøknader, godkjenninger og prosjektoversikter, til statistikk, regnskap, styringsdata og virksomhetsrapportering. Til formålene brukes en rekke uavhengige IT-systemer, og mens hver institusjon har ansvar for egne data, innebærer en rekke av formålene samhandling med enheter eller referanser til forhold utenfor egen institusjon. For å skape god datakvalitet er det nyttig å dele data mellom it-systemer og institusjoner. I dette tilfellet er formålet å skape entydige kilder til data om forskningsprosjekt, institusjoner og personer. Dette er metadata som går igjen i en rekke formål der ansvar for forvaltning er delt mellom institusjoner og it-systemer, og der det i dag ikke er tilfredsstillende felles ordninger.

Hovedmålet for prosjektet er å legge til rette for at relevant og nødvendig informasjon om sentrale entiteter i ulike typer forskningsprosjekter registreres én gang og deretter gjenbrukes som unike referanser på tvers av alle aktører og tjenester som er involvert i forskningsprosessen. På denne måten er meningen at forskernes og andre aktørers tidsbruk reduseres, datakvaliteten øker, og det legges til rette for at institusjonene på en enklere måte får nødvendig lovpålagt oversikt over egen forskning.

Formålet med konseptutredningen er å gi en retning for etablering av autoritative datakilder for deling og tilgjengeliggjøring av forskningsadministrativ informasjon. Utredningen redegjør for dagens situasjon, belyser behov og stiller krav til anbefalt konsept. Det tas utgangspunkt i konseptfasens mandat og det er tre konsepter som er behandlet:

- Konsept 1 («nullalternativet»): ingen tiltak ut over eksisterende systemer og planer
- Konsept 2: Etablere TILDE, et felles system for metadataforvaltning (MDM) for personer, institusjoner og organisasjoner, i regi av Sikt
- Konsept 3: Etablere separate system for metadataforvaltning av personer, institusjoner og organisasjoner

Konseptutredningen søker å avklare om det er hensiktsmessig å utvide etablerte tjenester, etablere relevante nye registre og ta registrene i bruk i de tjenestene som vil gi institusjonene størst gevinst og eventuelt foreslå å endre etablerte prosesser. I tillegg skal konseptutredningen avklare hvordan prosesser og tjenester best mulig samspiller med etablerte tjenester, tjenester under utvikling og øvrige planlagte tjenester og applikasjoner innenfor det forskningsadministrative domenet.

Kommentarer til utredningen

Universitetet i Bergen (UiB) støtter Sikts mål om høyere kvalitet og enklere informasjonsflyt for forskningsadministrative data. UiB har ingen prinsipielle innvendinger mot at egne data om personer, finansierte prosjekter og institusjonens enheter deles og tilgjengeliggjøres for andre. Konseptutredningen beskriver i liten grad plan for organisering av eventuelt utviklingsprosjekt. Vi vil påpeke viktigheten av å involvere brukergrupper og involverte institusjoner og tjenesteleverandører tett hvis utredningen følges opp i utviklingsprosjekt.

Produsenter og konsumenter

Generelt gir utredningen en grei beskrivelse av viktige produsenter og konsumenter av forskningsadministrative metadata, med noen svakheter. Først og fremst savner vi roller knyttet til grunnleggende administrative tjenester og systemer der data og metadata om forskning og forskere behandles: personalsystemer, regnskap, budsjett med mer. Prosjektøkonomer, forskningsrådgivere og andre administrative roller er helt sentrale i den daglige forvaltningen av data knyttet til forskning. Hvis disse rollene og systemene ikke skal omfattes av TILDE, bør det sies eksplisitt.

Som produsenter kan det vurderes om også kommersielle foretak kunne nevnes. Mange bruker eksterne tjenester som Clarivate, Scopus og Google til ulike formål, og slike data inngår også som grunnlag for Cristin-rapportering. Det er også et spørsmål om kommersielle aktører kunne hatt en synligere rolle som konsumenter av data, da samlinger av åpne data med høy kvalitet også kan åpne muligheter for næringsutvikling og utvikling av nye tjenester rettet mot privat og offentlig sektor, så vel som akademien.

Dagens bruk, behov og mål

Utredningen gir et godt bilde av hvordan dagens system skaper utfordringer knyttet til datakvalitet og samsvar i informasjon mellom en rekke relevante it-løsninger for ulike formål. Erfaringer fra lokalt utviklede IT-systemer som integreres mot Cristin og andre tjenester, bekrefter beskrivelsene av utfordringer i datakvalitet, knyttet til duplikater, manglende data i grunnlagssystemene, feil i organisasjonstilknytninger og navn og manglende muligheter til å spille inn korrigeringer til datakildene. Det er vanskelig å se for seg at slike utfordringer kan løses helt uten et system for metadataforvaltning på tvers av IT-systemer og institusjoner.

For slike formål gir utredningen en god oversikt over behov og mål. Beskrivelsen av behov indikerer imidlertid at TILDE-prosjektet primært er rettet mot systemer knyttet til tradisjonell forskningsdokumentasjon (bibliotekssystemer, Cristin/NVA, prosjektdokumentasjon) og ikke til bredden av administrative system og oppgaver der data om forskning forvaltes. Økonomi- og personalsystemer og handlinger knyttet til disse er eksempelvis fraværende i figur 4 om handlinger som gir dobbeltregistreringer, og i stor grad også i beskrivelsen av behov knyttet til forskningsadministrative data. Det kan være gode grunner til å avgrense TILDE-tjenestene mot økonomisk-administrative systemer og oppgaver, og dette bør i tilfelle sies eksplisitt. Det bør da også dokumenteres at parallell registrering og forvaltning av eksempelvis prosjektdata, personreferanser mm vil videreføres for viktige funksjoner, også ved etablering av TILDE.

Krav

I kapittel 6 åpnes det for å differensiere krav mellom ulike typer metadata. For person- og organisasjonsdata stilles det sterkere krav («MDM avansert») enn til prosjektdata. Utredningen argumenterer godt for at det kan være gode grunner til slik differensiering, både med hensyn til kompleksitet i eierskap til data, og til egenskaper ved selve datatypene, der «forskningsprosjekt» er svakere og mer utydelig definert enn de andre elementene. Modellen i figur 6 indikerer at data om forskningsprosjekt ikke skal være «live»-oppdatert, men baseres på filoverføringer. Dette kan være lite hensiktsmessig, særlig i forvaltning av prosjektinformasjon i forbindelse med søknader om finansiering, godkjenninger m.m. Her ville det være en fordel å kunne få tildelt en unik identifikator når informasjon om (potensielt) prosjekt først oppstår, slik at det kan refereres likt på tvers av ulike systemer fra rådgiving og søknadsbehandling (PVO, REK, dokumentarkiv og interne godkjenninger og søknadsoversikter, behandling hos finansieringskilder) til igangsatte prosjekt (dokumentert i NVA, finansieringskilder, regnskapssystemer mm). Vi er enige i at bare et minimum av opplysninger bør foreligge i masterdatakilden, det må ikke publiseres informasjon om

Tabell 2 — Må-krav til TILDE-konsept	
A	Personvern og sikkerhet må ivaretas
B	Gi økt samhandling innad i kunnskapssektoren
C	Infrastrukturen skal gi oversikt over forskningsadministrativ informasjon i kunnskapssektoren
D	Nødvendig datakvalitet
E	Forvaltning av masterdata

prosjekter som ikke er igangsatt, og det må være mulig å unnta prosjekter fra publisering i metadatakilden.

Listen over må-krav er det viktigste grunnlaget for å evaluere ulike konsepter.

Beskrivelsene av de ulike kravene er tilfredsstillende, men enkelte av overskriftene kan med fordel spisses for å vise ambisjoner om at tiltakene skal føre til forbedring, eksempelvis

- Personvern og sikkerhet må **styrkes**
- Datakvalitet må **økes**
- Forvaltning av masterdata må **effektiviseres**

I tillegg vil vi foreslå tre nye krav:

- Relevant eierskap til data må være utvetydig
- Integriteten i institusjonenes virksomhetsdata må sikres
- Sluttbrukeres tidsbruk må reduseres

Konseptalternativene

Utredningen beskriver tre alternative konsept for framtidig metadataforvaltning. Valg av konsept er hovedformålet med utredningen.

- Konsept 1 («nullalternativet»): ingen tiltak ut over eksisterende systemer og planer
- Konsept 2: Etablere TILDE, et felles system for metadataforvaltning (MDM) for personer, institusjoner og organisasjoner, i regi av Sikt
- Konsept 3: Etablere separate system for metadataforvaltning av personer, institusjoner og organisasjoner

Konsept 1: Null-alternativet:

Nullalternativet skal normalt beskrive en sannsynlig utvikling dersom tiltak ikke blir iverksatt. Innledningsvis slås det fast at Cristin-databasen skal fases ut og at det ikke kan legges til grunn at Cristin ivaretar masterdata om person, organisasjon eller forskningsprosjekt. Dette synes å være en unødvendig presisering. Dersom Tilde ikke blir iverksatt, må vi kunne forutsette at NVA/CRIS inneholder funksjonalitet for forvaltning av metadata minst på samme nivå som Cristin i dag. Under punktet om funksjonalitet står det at tjenestene skal forvaltes videre i regi av Tilde. Vi går ut fra at det her siktes til Sikt, ikke Tilde. Det framgår også at det er tatt en beslutning om å fase ut dagens løsning. Det framgår ikke av selve utredningen hvilken beslutning dette refererer til. Drøftingen av null-alternativet er i stor grad knyttet til NVA og ikke til andre potensielle konsumenter. Språkbruken er til dels tendensiøs ved å

bruke begrep som «stor og tung jobb» om ordinære forvaltningsoppgaver. Klart eierskap til data er en fordel som ikke synliggjøres i særlig grad.

Konsept 2: Felles MDM-plattform

Konseptet innebærer at det etableres en felles plattform for masterdataforvaltning, utviklet og forvaltet av Sikt. Dette er alternativet som anbefales i utredningen. Det argumenteres for at dette alternativet gir beste forutsetninger for å bygge digitalt tilgjengelige tjenester for metadataforvaltning, og at modellen er enklest for brukerne av tjenestene fordi de får en aktør å forholde seg til og konsistente grensesnitt mot ulike tjenester.

Konsept 3: Frittstående MDM-leverandører

Konseptet innebærer at det ikke er en enkelt aktør (Sikt) som forvalter alle masterdata, men at hver masterdatakilde utvikles separat. Det er i utgangpunktet fastslått at masterdatakildene skal fungere uavhengig av hverandre, slik at dette ikke burde føre til et samlet svakere tilbud enn ved å samle alle tjenester hos Sikt. Vi tolker konseptet slik at det kan åpnes for konkurranse mellom tjenestetilbydere. Det er ikke drøftet om denne organiseringen vil kunne gi fordeler i form av tjenesteinnovasjon, kostnader m.m.

Utredningen lister en rekke risikofaktorer med organisasjonsmodellen, og mener at den kan føre til et mer kostnadskrevenende og mindre konsistent tjenestetilbud, med potensielt lavere kvalitet til sluttbrukere.

Vi oppfatter at beskrivelsen av konsept 3 er til dels uklar, og til dels tendensiøs. Det kommer tydelig fram i beskrivelsen av konseptet at utrederne fra Sikt ikke ønsker dette alternativet. Det er uheldig at det legges sterke vurderinger inn i selve konseptbeskrivelsen. Vi anbefaler at beskrivelsen gjennomgås for balanse, og at kritiske vurderinger konsentreres til drøftingene i kapittel 8.

Utredningens kap 8 drøfter konseptene og anbefaler konsept 2. Det framstår fornuftig å etablere nasjonale tjenester for masterdataforvaltning gitt erfaringene med dagens system. Vi finner det ikke absolutt godt gjort at konsept 3 nødvendigvis vil være mer kostnadskrevenende eller resultere i lavere datakvalitet enn konsept 2, men for brukere av tjenesten kan det være formålstjenlig å forholde seg til en enkelt aktør.

En grundigere økonomisk analyse ville gjøre det enklere å forholde seg til konseptene. Kapittel 10 drøfter virkninger av tiltakene. Utredningen lister imidlertid kun ikke-prissatte virkninger. Utredningen peker på flere gunstige utfall, men i noen grad framstår virkningene som ønsketenkning heller enn sannsynlige konsekvenser. Det er særlig ikke godt gjort at tiltakene vil føre til «mer tid til forskning», slik det hevdes i punkt 10.1.3. Her ser det ut til at utrederne ikke har klart for seg rollefordeling i forskningsadministrative prosesser, og i særdeleshet ikke fordeling av tid mellom manuell input og faglig virksomhet.

Det er en stor svakhet at ikke utredningen angir selv grove anslag for investeringskostnader og omstillingskostnader. Selv om det er vist at dagens system for forvaltning av forskningsadministrative data har kostnader, er det ikke godt gjort at de skisserte tiltakene samlet gir lavere kostnader eller høyere samfunnsøkonomisk lønnsomhet. Særlig er dette usikkert siden vi ikke kan se at viktige deler av forskningsadministrative data, slik de forvaltes i økonomi, personalsystem og annet, ikke i denne omgang berøres av planene. Vi savner en vurdering av økonomiske konsekvenser for institusjonene. Vil institusjonene belastes for bruk av MDM-tjenestene? Eller de enkelte konsumentssystemene?

Universitetet i Bergen støtter intensjonen om å opprette felles tjenester for data om forskningsprosjekter, forskere og organisasjonsenheter, men finner ikke grunnlag for å velge mellom de foreslåtte konseptene.

UiB oppfordrer Sikt til å

- Balansere framstillingen av konsept 2 og 3
- Gi anslag for økonomiske konsekvenser av de ulike alternativene, både totalt og for den enkelte institusjon
- Tydeliggjøre hvordan eierskap til data og sikkerhet for integritet i egne virksomhetsdata skal ivaretas gjennom systemet