

UiO : Universitetet i Oslo

Sluttrapport Implementasjon av forskningsdataarkiv ved UiO

1. Bakgrunn for prosjektet

Med prosjektet *implementasjon av forskningsdataarkiv ved UiO* ønsket Universitetet i Oslo, ved UB, å gå inn som partner i det nasjonale forskningsdataarkiv-konsortiet DataverseNO og med dette få på plass et tilbud for åpen arkivering og kuratering av data fra forskningsprosjekt ved UiO som ikke har fagspesifikke arkiv.

2. Prosjektets viktigste leveranser

Følgende leveranser var skissert i prosjektet:

- Implementasjon av DataverseNO-samling.
- Tilrettelegge for forskerstøttetjenester for arkivering.
- Pilotfase.
- Permanente identifikatorer (PID), hvis tid.
- Kommunikasjon.
- Prosjektledelse.

Av disse er fem leveranser oppnådd, mens arbeidet med permanente identifikatorer som innebar en kartlegging av behov for PID og vurdering av behov for retningslinjer, er prioritert ned. Vi ser det som hensiktsmessig at arbeidet med PID-strategi gjøres nasjonalt eller i BOTT for å kunne ha en felles strategi for slike på tvers av institusjoner.

Prosjektet var planlagt å gå over 6 måneder, men har gått ut over dette i tid. En av grunnene var et ønske om en felles lansering med spesialsamlingprosjektet sine tilbud; Alvin og Digitaliseringsverkstedet. Videre har en glidende overgang fra prosjekt til drift vært hensiktsmessig for å kunne bygge opp erfaring og dialog med ulike miljøer gradvis. Ved å strekke pilotfasen over 10 måneder har alle kuratorer gjort flere kurateringer, vi har også fått testet deling til dobbel-blind fagfelleevaluering. Ut over dette har vi fått anledning til å utforske noen av mulighetene som ligger i å lage undersamlinger.

Det var også ønsket å få på plass en integrasjon mot leveransen fra FAIR@UiO-prosjektet som ble utviklet av IT-avdelingen. FAIR@UiO-produktet «metadatageneratoren» og bruken av dette har ikke vært på et slikt nivå at det har vært mulig eller hensiktsmessig å utvikle eksport til DataverseNO. UB sitt bidrag gjennom tilpasninger av Dataverse metadatastandard til FAIR@UiO har imidlertid blitt gjort tilgjengelig for DataverseNO-miljøet med tanke på å kunne dra nytte av tanker og tilpasninger som er gjort her.

Etter inngått avtale med DataverseNO og fullført kuratoropplæring med representanter fra UiT i august 2022, arbeidet prosjektgruppen med utforming av mandat og rutiner.

Etablering og opplæring av en fast kuratorgruppe var viktig for å komme i gang med en pilotfase. Som en del av implementering ble mandat for Drift- og kuratorgruppe for forskningsdataarkiv på UB (kuratorgruppen) vedtatt av UBs ledergruppe 20.09.2022.

Prosjektnavn: implementasjon av forskningsdataarkiv ved UiO	Sluttrapporten er utarbeidet av : Live Kvale
Prosjekteier: UB, Grete Christina Lingjærde	Prosjektleder: Live Kvale
Versjon: 1.0	Vedtatt dato:

Det ble også etablert faste møter to ganger i måneden. På disse møtene har vi diskutert og orientert om en rekke ulike tema inkludert kuraterings erfaringer og -utfordringer, videre planer og gjennomgang av opplæringsmateriale og webinarer, som er delt i konsortiet.

En intern nettside med mandat, rutiner og referat fra gruppens møter ble opprettet¹.

En ekstern nettside om dataarkivering ble opprettet under «Skrive og publisere» på UB sine nettsider, med informasjon om valg av arkiv, FAIR-prinsippene, DataverseNO og lenker til UiOs overordnede nettsider om forskningsdatahåndtering².

Etterfulgt av implementasjon og tilrettelegging av DataverseNO høsten 2022, fulgte en pilot- og kommunikasjonsfase der vi presenterte tilbudet mot ulike fagmiljø, og mottok datasett til kuratering.

Følgende kurs, workshoper, presentasjoner og informasjonskampanjer ble gjennomført i pilotfasen:

- DataverseNO ble bakt inn i eksisterende kurs i dataarkivering. Åpne kurs ble holdt tre ganger, i tillegg kommer kurs og presentasjoner i master- og PhD-kurs.
- Det har vært holdt to workshoper i arkivering i DataverseNO. På grunn av lav oppslutning på disse, ser vi det som mer fornuftig å være tilgjengelige for veiledning, enn videre workshoper i dataarkivering.
- DataverseNO har vært presentert for Open Science Lunch, Livsvitenskapsmiljøet ved UiO, forskerseminar ved RITMO og IBV, forskerstøtte på UiO, Datarøkternettverket ved UiO, forskningsrådgivere og -leder ved HF, UB-ansatte, og Bibliotekenes nettverk for ph.d.-støtte.
- Informasjon om DataverseNO har blitt sendt på epost til alle som jobber med forskerstøtte på UiO (FANE-nettverket).
- Utformet en PowerPoint-mal for bruk i DataverseNO-presentasjoner.
- Til synliggjøring ble det trykt opp og delt ut 250 DataverseNO-klistermerker.

I løpet av pilotfasen har vi blitt kontaktet av ulike miljø som ønsker å kunne bruke DataverseNO til ulike behov, dette gjelder LINK, KHM og Ibsen-senteret. I dialog med disse tre miljøene og DataverseNO-miljøet ved UiT har vi testet ut muligheten for egne undersøkelser i DataverseNO sin testinstans. Her har vi møtt på ulike utfordringer knyttet til personvern, API og import. Dette er ikke avsluttede dialoger, men arbeid og erfaringer vi tar med oss videre i drift.

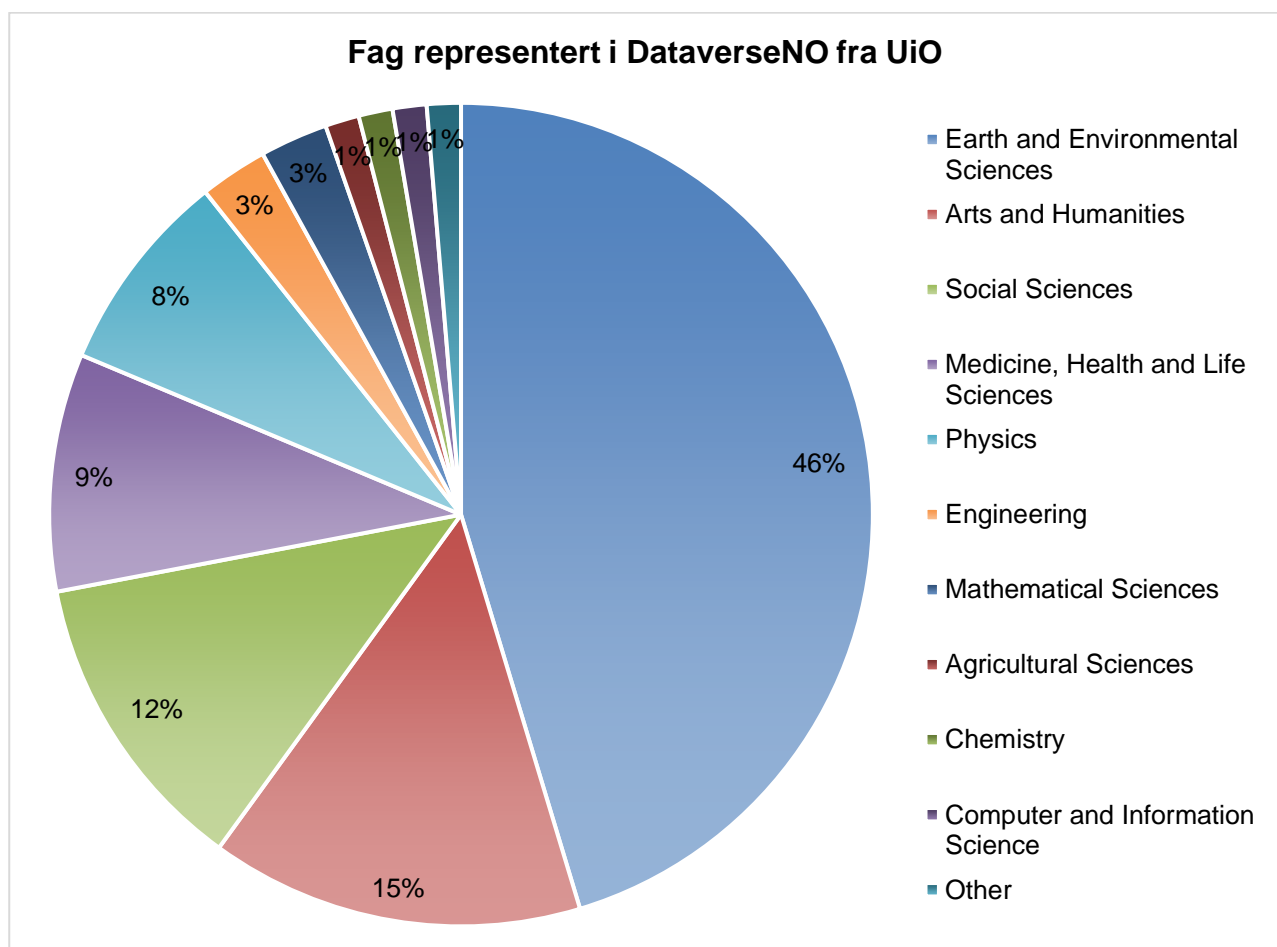
Lansering

¹ <https://www-int.uio.no/for-ansatte/enhetssider/ub/avdelinger-grupper/faste-grupper/datakurator/index.html>

² <https://www.ub.uio.no/skrive-publisere/dataarkivering/index.html>

- Som avslutning på prosjektet regner vi lanseringen av DataverseNO 29. november 2023.
- Kuratorgruppen bidro i ulike deler av planlegging og gjennomføring av samlansering for DataverseNO, Alvin og digitaliseringsverkstedet.

Ved lansering i november 2023 var 42 ulike datasett, med til sammen 1232 filer, publisert i UiO sin samling på DataverseNO. Figur 1. viser den faglige spredningen i datasettene som ligger i samlingen. I tillegg til 42 publiserte datasett, ligger 16 datasett på ulike steder i arbeidsflyten fram mot publisering.



Figur 1; Faglig bredde på datasett i DataverseNO ved lansering 27. november 2023

Ut over leveranser som ble skissert i prosjektsøknaden har det vært gjort andre bidrag som er med på å styrke tilbudet på arkivering av forskningsdata gjennom DataverseNO for forskere på UiO

- Endring av styringsmodell i DataverseNO-konsortiet, som har gått fra å være UiT-ledet til å ha et styre med representanter fra BOTT. UiO ved Grete (Tina) Lingjærde, UB, er nå er del av DataverseNO-styret.

- REASON, søknad til forskningsrådet: Vi har bidratt til innsending av en støtteinfrastruktursøknad til Forskningsrådet for å videreutvikle DataverseNO sammen med BOTT, NMBU og en rekke internasjonale partnere. Søknaden tar sikte på å utvikle infrastruktur for forskningsdata etter GORC modell, der hele forskningsdatakretsløpet inngår, DataverseNO og Dataverse-programvaren utgjør kjernen i søknaden. Fra UiO står UB og IT-avdelingen nå med ansvaret for to arbeidspakker, hvor blant annet støtte til arkivering av sensitive data inngår.
- Deltagelse i prosjektet med etablering av nasjonalt kuratornettverk: Fra kuratorgruppen på UB deltar vi i det NB-finansierte prosjektet med etablering av et nasjonalt kuratornettverk.

Innledningsvis var det viktig å skape en forankring i UB på at arkivansvaret ikke er av prosjektform, men en langsiktig forpliktelse. Gruppens mandat og prosjekteiers involvering i styret for konsortiet, har begge deler vært viktige grep for å synliggjøre en forpliktelse.

Prioriteringer og retning i prosjektet har blitt gjort av prosjektleder i tett dialog med kuratorgruppen og prosjekteier.

Prosjektleder har ikke brukt den nedsatte styringsgruppen tilstrekkelig. Det kunne vært hensiktsmessig å involvere styringsgruppen mer i prioriteringer underveis i prosjektet og sikre bredere forståelse og støtte til søknaden om videreutvikling i REASON.

3. Oppnåelse av prosjektets mål

Ved å gå inn som partner i DataverseNO har prosjektet bidratt til oppnåelsen av følgende overordnede formål:

- Arbeide systematisk for utvikling og deling av forskningsinfrastruktur og dataressurser.
- Intensivere arbeidet med forskningsetikk og forskningsintegritet.
- Effektiv utnyttelse av digitale ressurser.
- Gjøre UiO-miljøer konkurransedyktige i søknadsprosesser om eksterne forskningsmidler fra for eksempel EU og NFR.

Selv om et arkiv som infrastruktur tilsynelatende står langt fra arbeidet med forskningsetikk og integritet, spiller både muligheten til etterprøvnbarhet og bedre utnyttelse av resurser og investeringer en rolle i dette. Arkiv, tillit og integritet henger tett sammen.

Etableringen av en kuratorgruppe sammensatt med en bred representasjon av UBs kompetanse gjør at forskere ved UiO nå har fått et godt støttetilbud for åpen, tilgjengelig og gjenbrukbar arkivering. Dette er i tråd med finansørers krav om FAIR deling av data i et «Trusted repository». DataverseNO møter dermed både krav fra EU og Forskningsrådet til arkiv for forskningsdata. Ved å legge til rette for sitering av data støttes det også opp om forskeres impact. Følgene effektmål regnes dermed som oppnådd:

- UiO-forskere som trenger å arkivere data får et bedre støttetilbud
- Forskningsdata fra UiO vil bli gjort tilgjengelige og gjenbrukbare

- Økt impact for UiO-forskere gjennom deling av data
- Krav fra forskningsfinansiører relatert til deling av data kan innfris

Samtidig ser vi at DataverseNO er et generisk arkiv med begrensninger både på domenespesifikk støtte, arkivering med tilgangsbegrensninger og fulltekstsøking i innhold. Dette gjør at vi i søknaden til NFR ønsket å forplikte oss til videre arbeid med utvikling av arkivet. En innvilgelse av denne søknaden vil kunne bidra til å levere FAIR deling av data i DataverseNO på et helt annet nivå enn det som er mulig i dag.

4. Forutsetninger for gevinstrealisering

Utvidelsen av prosjektperioden, og den gradvise overgangen fra prosjekt til drift, har vært et grep for å sikre en sømløs kontinuitet og gjør at vi allerede nå kan se gevinsten ved å gi forskere et åpent arkivtilbud for forskningsdata av høy kvalitet.

5. Evaluering av prosjektet

a. Suksessfaktorer

Målet med implementeringsprosjektet var etableringen av et varig tilbud om arkivering av forskningsdata for UiOs forskere. Dette er nå på plass og vi regner UiO som godt rustet til å bidra videre inn i utviklingen av DataverseNO, gjennom deltagelse i DataverseNO-styret, samarbeid på tvers med andre institusjoner i konsortiet og med en stabil kuratorgruppe som kan sørge for at UiOs forskere får et godt tilbud om dataarkivering. Å gå inn i DataverseNO-samarbeidet på et tidspunkt hvor kurs, opplæring og informasjon om datahåndtering allerede var etablert har også vært en fordel, da vi har kunnet flette tilbudet inn i eksisterende innhold.

b. utfordringer og usikkerhet

Deling av forskningsdata er en viktig brikke i å bygge en kultur for åpen forskning. Denne kulturendringen er noe som tar tid. Vårt arbeid med å nå ut til fagmiljøene på UiO med DataverseNO er med på å bidra til kulturendring gjennom å gi forskere positive erfaringer med åpen deling av forskningsdata. Samtidig er dette en liten snøball som det tar tid å bygge større. Kontinuerlig arbeid med kvalitet kombinert med en god og åpen dialog med forskere om muligheter og begrensninger for deling er vårt bidrag til en kulturendring, der fremtidig gjenbruk av data og tillitt til forskning vil være suksessindikatorer.

c. Viktige læringspunkter

Tidsrammen for prosjektet var i overkant optimistisk. I etablering, opplæring og særlig i pilotering har vi vært avhengige av at forskere deponerer data for å kunne få erfaringer. Prosjektleder kunne også hatt fordel av å anvende styringsgruppa, særlig hva gjaldt involvering i søknaden til Forskningsrådet.

6. Prosjektregnskap

Som det kommer frem av regnskapet under, har prosjektet hatt totale kostnader på 1 668 000. Driftskostnader på 146 000 kr og personalkostnader på 1 522 000, over to år. Av dette er 970 000 dekket av midler (1 000 000) som ble bevilget til forskningsdata arkiv av SK-IT-F 22 mai 2022³. Resterende 698 000 er dekket av UBs driftsmidler.

År 2022/ 34 Universitetsbiblioteket

Prosjekt 103807 IT satsings-og porteføljemidler

Delprosjekt 103807100 UB-Arkivløsning for forskningsdata

Tall i 1000 kr

				Regnskap valgt periode
	Isolert mer- /mindreforbruk uten nettobidrag og prosjektavslutning	Inntekter	Inntekter	-911
		Kostnader	Personalkostnader	826
			Driftskostnader	85
		Sum		0
Isolert resultat			0	
Akkumulert resultat			0	

<i>Totale kostnader</i>	911 116,35
<i>hvorav dekket av IT satsings- og porteføljemidler</i>	-911 116,35
<i>hvorav dekket av UB</i>	0,00

Konto	
6751 Databehandlingstjenester	83 253,75
6861 Servering ved møter	2 199,60
9112 Lønn belastet	825 663,00
9147 Interne tjenester godskrevet	-911 116,35
Grand Total	0,00

År 2023/ 34 Universitetsbiblioteket

Prosjekt 103807 IT satsings-og porteføljemidler

Delprosjekt 103807100 UB-Arkivløsning for forskningsdata

Tall i 1000 kr

				Regnskap valgt periode

³ <https://www-int.usit.uio.no/om/it-dir/strategi/masterplan/sk/sk-itf/moter/2022/220512/index.html#toc15>

	Isolert mer- /mindreforbruk uten nettobidrag og prosjektavslutning	Inntekter	Inntekter	-59
		Kostnader	Personalkostnader	696
			Driftskostnader	61
		Sum		698
Isolert resultat			698	
Akkumulert resultat			698	

<i>Totale kostnader</i>	<i>756 970,96</i>
<i>hvorav dekket av IT satsings- og porteføljemidler</i>	<i>-59 348,75</i>
<i>hvorav dekket av UB</i>	<i>697 622,21</i>

Konto	
6751 Databehandlingstjenester	59 348,75
6861 Servering ved møter	942,66
7132 Tjenestereiser ikke oppgavepliktige	945,00
9112 Lønn belastet	695 734,55
9147 Interne tjenester godskrevet	-59 348,75
Grand Total	697 622,21