

NTNU Open Data

NTNUs politikk for åpne forskningsdata 2018-2025

Bakgrunn

Tilgjengeliggjøring og gjenbruk av forskningsdata bidrar til økt etterprøvbarehet og transparens i vitenskapen, og kan forhindre at samme data samles inn mange ganger. Åpne data skaper også grunnlag for nye og innovative digitale tjenester med potensielt stor verdi for samfunnet. Det er en rimelig forventning at offentlig finansiert forskning skal komme samfunnet til nytte. Derfor er det stor nasjonal og internasjonal oppmerksomhet omkring åpne forskningsdata.

Kunnskapsdepartementet (KD) har ansvaret for å utforme en strategi for tilgjengeliggjøring av forskningsdata. EU-kommisjonen og Norges forskningsråd stiller allerede i dag krav om en plan for hvordan data fra forskningsprosjekter skal gjøres tilgjengelig. I årene som kommer vil kravene om åpne data fra offentlig finansiert forskning bli skjerpet.

NTNUs *Publiseringspolitikk 2014–2020* legger til grunn at «kunnskap skal gjøres tilgjengelig». *Handlingsplan for åpen tilgang til NTNUs forskningspublikasjoner 2016-2020* er et ledd i oppfølgingen av publiseringspolitikken og forutsetter at NTNU utreder en modell for forsvarlig lagring og tilgjengeliggjøring av åpne forskningsdata. I 2016–2017 ble det med forankring i NTNUs forskningsutvalg gjennomført et forprosjekt som så på internasjonale og nasjonale utviklingstrekk og dokumenterte behov og praksis ved NTNU.¹

Oppfølging av politikken

Det må utarbeides sentrale og lokale handlingsplaner med mål og tiltak for å følge opp denne politikken.

Kostnader knyttet til grunnleggende håndtering, lagring og publisering av forskningsdata må i utgangspunktet dekkes gjennom de enkelte forskningsprosjektene, og regnes vanligvis som legitime kostnader i søknader om finansiering. Åpen tilgang til forskningsdata skal normalt innebære gratis bruk for eksterne. Dekning av reelle kostnader ved spesiell tilrettelegging av datasett og lignende må likevel være mulig.

¹ [Lenke til forprosjektrapport]

NTNUs politikk for åpne forskningsdata

Definisjon av forskningsdata

Forskningsdata kan være både digitale og analoge. Det kan for eksempel være dokumentasjon av prosedyrer, innsamlede intervjudata, store responsdatasett, kontinuerlige datastrømmer fra sensorer, arkivmateriale og sensitive persondata, men også fysiske objekter eller digitale representasjoner av disse. Norges forskningsråd skiller mellom *kildedata* og *resultatdata*.² Kildedata er informasjon som finnes allerede i ulike formater og som danner inngang til eller grunnlag for ny forskning. Resultatdata er data som er generert gjennom forskning og som i neste omgang kan bli kildedata for annen forskning.

NTNUs politikk omfatter ikke automatisk alle former for data som samles inn eller produseres i forbindelse med de ansattes forskningsvirksomhet, men i første rekke resultatdata fra forskningsprosjekter som er helt eller delvis offentlig finansiert, eksternt eller internt. Slike forskningsdata skal lagres og forvaltes på en sikker måte og gjøres tilgjengelig for andre.

NTNUs politikk for åpne forskningsdata legger Forskningsrådets definisjon til grunn:

Med forskningsdata mener vi i denne politikken registreringer/nedtegnelser/rapporteringer i form av tall, tekster, bilder og lyd som genereres eller oppstår underveis i forskningsprosjekter.

Overordnede prinsipper

NTNUs politikk for åpne forskningsdata bygger på disse prinsippene:

1. NTNUs forskningsdata skal i utgangspunktet være åpent tilgjengelig. Data er unntatt når hensyn til sikkerhet, personvern, juridiske eller kommersielle forhold krever det.
2. Forskningsdata skal gjøres åpent tilgjengelig så tidlig som mulig i forskningsprosessen, uten at det er i konflikt med forskerens bruk, kvalitetssikring og eventuell kommersiell bruk av dataene.
3. Forskningsdata som publiseres fra NTNU skal ha tydelig merking av opphav og eierskap.
4. Forskningsdata skal kunne finnes, hentes ut og brukes uten systembegrensninger og kunne gjenbrukes av andre. NTNU slutter seg til FAIR-prinsippene for forvaltning av forskningsdata og skal benytte lisenser for forskningsdata som disse forutsetter.³
5. Forskningsdata skal lagres med sikte på å være tilgjengelige i lang tid. Klassifisering og metadata skal følge internasjonale standarder så langt som mulig, og formater må om nødvendig oppdateres over tid. Håndteringen av dataene i hele livsløpet skal skje i samsvar med gjeldende lovverk og krav fra finansieringskilde, myndigheter og relevante interessenter.
6. NTNU skal sørge for at forskerne har tilgang til gode tjenester og infrastruktur for trygg lagring og deling av forskningsdata. NTNU skal bidra til å utvikle hensiktsmessige nasjonale og internasjonale løsninger på feltet.
7. Den enkelte forsker har ansvar for å håndtere forskningsdata i henhold til de regler, prinsipper og krav som gjelder. Prosedyrene skal dokumenteres i en datahåndteringsplan.

² http://www.forskningsradet.no/no/Artikkel/Apen_tilgang_til_forskningsdata/1254001013535

³ FAIR: Findable, Accessible, Interoperable, Re-usable; <https://www.force11.org/group/fairgroup/fairprinciples>

Eierskap til forskningsdata

Som ved vitenskapelig publisering er hovedprinsippet for forskningsdata at institusjonen beholder de intellektuelle rettighetene og forfatteren/forskeren opphavsrett. NTNU eier normalt ikke forskningsdata som studenter og gjesteforskere produserer, unntatt når dette er avtalt for eksempel ved eksternt finansierte prosjekter.

Som hovedregel eier NTNU alle forskningsdata som er innsamlet og bearbeidet av personell som er tilsatt ved eller innleid av institusjonen. Det gir NTNU rett til å publisere materialet åpent, men er ikke til hinder for at

- forskeren kan kreve embargo (utsatt åpen publisering) for innholdet, men normalt ikke for metadataene.
- NTNU kan avstå fra sin rett til åpen publisering av materialet. Dette påvirker imidlertid ikke eierskapet til dataene. Hvis forskeren likevel publiserer datamaterialet åpent via et forlag, må forlaget også gi NTNU rett til å legge dette materialet åpent.

Spørsmål om rettigheter og eierskap må avklares og avtales i tilfeller der data brukes i kommersialiserings- og patenteringsøyemed. Avtaler må også inngås i situasjoner der NTNUs egne forskere gjør bruk av andres data.

Datahåndtering

A. Datasikkerhet og tilgang

Alle forskningsprosjekter skal ha en datahåndteringsplan (*Data Management Plan – DMP*) som tilfredsstiller kravene fra finansieringskildene.

Forskningsdata skal lagres i sikre arkiver, enten sentralt ved NTNU eller i nasjonale eller internasjonale arkiver som NTNU har godkjent.

Forskningsdata skal lagres/arkiveres så lenge de er av verdi for forskeren og et bredere forskningsmiljø, og så lenge som angitt av finansieringskilde, patentbestemmelser, embargokrav, lovgivning og andre myndighetskrav. Den minste lagringsperioden for forskningsdata er tre (3) år etter publisering, men i de fleste tilfeller vil data forbli lagret lenger enn minstekravet på tre år. Hvis forskningsdata skal slettes eller tilintetgjøres, enten fordi den avtalte perioden for oppbevaring er utløpt eller på grunn av juridiske forhold, må dette gjøres i samsvar med gjeldende regler og forskningsetiske prinsipper samt finansieringskildens og samarbeidspartneres krav, og med særlig hensyn til konfidensialitet og sikkerhet.

Det bør utarbeides en plan for hvordan forskningsdata som er vurdert til å ha verdi på lang sikt, skal forvaltes. Det må også finnes planer for hvordan data som er vurdert til ikke å ha langsiktig verdi, skal forvaltes og eventuelt destrueres etter en viss tid

Forskningsdata skal være korrekt klassifisert med hensyn til krav om beskyttelse og utstyres med standardiserte, kategoriserte metadata. Metadata skal følge internasjonale standarder og inkludere en beskrivelse av datakvaliteten. Ofte vil et datasett være klart avgrenset, men noen ganger vil det være behov for å spesifisere hva som inkluderes i en forekomst av et datasett.

Generelt skal forskningsdata gjøres tilgjengelig på et tidligst mulig tidspunkt, etter en periode med førstebruksrett for forskerne selv. Data som ligger til grunn for vitenskapelige artikler bør fortrinnsvis gjøres tilgjengelig ved publiseringstidspunktet. Andre data som kan være av interesse for annen forskning, bør gjøres tilgjengelig innen rimelig tid og aldri senere enn tre år etter endt prosjekt.

Forskningsdata bør være tilgjengelig for et internasjonalt forskerkollektiv. NTNU er forpliktet til universell utforming av nettbaserte tjenester. Bruk av åpne standarder og tilrettelegging av forskningsdata i tråd med kravene til universell utforming er en forutsetning.

Når forskning er støttet gjennom en kontrakt eller avtale som inneholder spesifikke bestemmelser om eierskap, oppbevaring av og tilgang til forskningsdata, vil bestemmelsene i denne ha forrang.

B. Sensitive data

Sensitive forskningsdata er normalt unntatt fra prinsippene om åpen tilgang når hensyn til sikkerhet, personvern, juridiske eller kommersielle forhold krever det. Slike forskningsdata skal likevel utstyres med metadata som kan inngå i systemer for gjenfinning, slik at det synliggjøres både at et lukket datasett finnes og hvordan en kan komme i kontakt med dem som forvalter dette datasettet.

Enkelte typer data skal ikke oppbevares, men destrueres i tråd med gjeldende avtaler og regler etter at bruksperioden er utløpt.

NTNU gir nærmere retningslinjer for hvordan forskningsdata som inneholder sensitiv og fortrolig informasjon skal håndteres og hvilke arkiver som kan benyttes for lagring. Også ikke-sensitive data som i utgangspunktet skal være åpne, må spres gjennom kontrollerte kanaler og ikke være utsatt for tyveri (brudd på IPR), utilbørlig kopiering (lisensbrudd) eller endring (integritetsbrudd).

Roller og ansvar

Forsker

Den enkelte forsker har ansvar for å håndtere forskningsdata i tråd med NTNUs prinsipper. I dette ligger også en plikt til å sette opp en datahåndteringsplan (DMP) tidlig i forskningsprosjektet og ikke senere enn seks (6) måneder etter oppstart, samt å gjøre data tilgjengelig for institusjonen etter avslutning.

I prosjekter der flere forskere deltar og felles forskningsdata blir generert, har prosjektleder ansvar for at en DMP blir satt opp og at data fra prosjektet blir arkivert og gjort tilgjengelig i tråd med NTNUs prinsipper og gjeldende retningslinjer.

Fakultet

Fakultetet skal gjøre forskerne kjent med NTNUs politikk for åpne forskningsdata og de retningslinjene som følger av denne samt gjeldende regelverk omkring datahåndtering. Fakultetet skal også sørge for at studenter og stipendiater får nødvendig opplæring.

Fakultetet har ansvar for å lage utfyllende retningslinjer for datahåndtering innenfor de fagområdene som har behov for det.

Fakultetet skal ha rutiner for arkivering av forskningsdata når ansatte slutter eller skifter jobb, slik at materialet blir tilgjengelig på egnet måte.

NTNU

NTNU skal sørge for at forskerne har tilgang til sikre og gode løsninger for arkivering og deling av forskningsdata, og skal tilby informasjon og veiledning om disse løsningene.

NTNU har ansvar for å lage institusjonelle regler og retningslinjer for håndtering av forskningsdata, og for å iverksette nødvendige fellestiltak innenfor området.

NTNU Universitetsbiblioteket (UB) gis det operative ansvaret for den langsiktige forvaltningen av institusjonens forskningsdata. Vårt institusjonelle arkiv, NTNU Open, bør brukes der det er hensiktsmessig, men UB kan også godkjenne, anbefale og bruke andre nasjonale og internasjonale løsninger for arkivering og deling.