



Retningslinjer for forvaltning av data fra Norsk Polarinstitut

1. Formål

Formålet med disse retningslinjene er å sikre et best mulig datagrunnlag for forskning, overvåking, forvaltningsråd og miljøinformasjon. En åpen datapolitikk skal sikre størst mulig samfunnsnytte av instituttets investeringer i datainnsamling, allmennhetens rett til åpen miljøinformasjon, samt troverdighet, transparens og etterprøvbarehet for instituttets forskning og miljørådgiving. Retningslinjene fastlegger prinsippene for forvaltning av eierskap og opphavsrett til data, og skal sikre en god praksis for dokumentasjon, formidling og langsiktig forvaltning av selve datasettene.

Datapolitikken skal sikre at Norsk Polarinstitut kan oppfylle ICSU World Data Systems kriterier for permanente dataarkiver, samt fylle datapolitiske betingelser for å delta fullt ut i alle nasjonale og internasjonale datasamarbeid og datanettverk som er viktige innenfor instituttets fagområde.

2. Definisjon

«Data» defineres som enhver samling av observasjoner eller målinger som er registrert på skriftlig eller digital form, som er ment for bruk i vitenskap, miljøovervåking eller kartlegging, og som kan beskrives av metadata. Begrepet kan også innbefatte vitenskapelige prøver og eksemplarer.

Referansedata, overvåkingsdata, miljødata og kartdata er data som blir samlet inn med dokumenterte metoder til formål som er helt eller delvis knyttet til forvaltning, rådgiving, informasjon, dokumentasjon, kartlegging eller annet som hører under instituttets samfunnsoppdrag.

Forskningsdata er data som blir samlet inn for et avgrenset, vitenskapelig prosjekt, og som er unike for prosjektet. Topografiske kartdata berøres ikke av disse retningslinjene.

3. Eierskap og ansvar

Norsk Polarinstitut eier data og vitenskapelig materiale samlet inn av ansatte i tjeneste. Dette gjelder også eksternfinansierte prosjekter, dersom det ikke er forbehold om eierrettigheter tilknyttet finansieringen.

Avdelingsdirektørene er ansvarlige for å gjennomføre disse retningslinjene. Prosjektledere er ansvarlige for at data fra et prosjekt blir forsvarlig dokumentert og permanent lagret.

4. Hovedprinsipp: åpne data

Alle data som eies av Norsk Polarinstitutt skal være åpent og kostnadsfritt tilgjengelige for videre bruk, i samsvar med statens politikk for frigivelse av offentlige og offentlig finansierte data¹. Ved ekstern viderebruk skal data siteres i samsvar med instituttets til enhver tid publiserte retningslinjer.

Bestemte datasett kan likevel beskyttes mot tilgang hvis dette er påkrevd av hensyn til personvern eller hvis offentlig tilgang til datasettet kan medføre en vesentlig risiko for skade på sårbare arter eller naturverdier. Beskyttelsen bør begrenses til de sensitive delene av datasettet².

Tilgangen til forskningsdata kan begrenses i en angitt og avgrenset tidsperiode for å ivareta opphavspersonenes første rett til vitenskapelig publisering av datasettet.

5. Lisenser

For å sikre instituttets og opphavspersonenes rett til å bli navngitt ved viderebruk av data, og for å sikre ansvarsfrihet ved viderebruk, bør data fra Norsk Polarinstitutt publiseres under åpne datalicenser som Norsk lisens for offentlige data (NLOD) eller kompatible, internasjonale lisenser som Creative Commons navngivelse (CC BY)³.

6. Eksterne samarbeid

Når instituttet inngår eksterne prosjektsamarbeid som forutsetter utveksling av data, bør samarbeidet bygge på de samme lisensene for å sikre adgang til og rettigheter ved åpen datautveksling. Når datainnsamling ved instituttet inngår i eksterne programmer med egen datapolitikk⁴, skal denne følges, men på en slik måte at instituttets åpne publisering av egne data ikke begrenses.

7. Publiseringsrettigheter (rettighetsforvaltning)

Ansatte ved NP har første rett til vitenskapelig publisering av forskningsdata som vedkommende har samlet inn. For å ivareta denne retten, kan forskningsdata beskyttes inntil datasettet er blitt publisert første gang, innenfor en nærmere angitt tidsperiode. Beskyttelsesperioden skal datofestes, og bør normalt ikke settes til mer enn 2 år etter at datainnsamlingen er avsluttet. Innenfor rammene av de kontraktsvilkår og eksterne betingelser som gjelder for prosjektet, kan beskyttelsesperioden gjøre lengre hvis det er nødvendig for å ivareta første retten til publisering. Seksjonsleder er ansvarlig for å sette frigivelsesdato og kan utvide beskyttelsesperioden inntil 4 år etter at datainnsamlingen er avsluttet. Utvidelse av beskyttelsesperioden ut over 4 år skal kun gis unntaksvis og må godkjennes av avdelingsdirektør.

Når beskyttelsesperioden er utløpt, eller når ingen beskyttelsesperiode er datofestet, vil data som er ferdig prosessert, dokumentert og lagret ved instituttet være åpent tilgjengelige fra data.npolar.no.

¹ Herunder offentlighetsloven, miljøinformasjonsloven, geodataloven, åndsverkloven, nasjonal implementering av OECD-prinsippene og tildelingsbrevet. Offentlighetsloven og miljøinformasjonsloven må antas å medføre et forsterket åpenhetskrav for overvåkingsdata og andre data som danner grunnlag for saksbehandling og offentlig informasjon vedrørende miljøets tilstand.

² Eksempelvis posisjonene til hi- eller hekkelokaliteter for arter som er sårbare for menneskelig aktivitet.

³ CC BY 4.0 NO (norsk versjon), når den er klar.

⁴ Som SCAR-samarbeidet, CliC, Det internasjonale polaråret, GBIF, IODE og andre

8. Livsløpsforvaltning

Referansedata, overvåkingsdata og miljødata forvaltes av instituttets datasenter. Forskningsdata forvaltes samme sted når de er ferdig prosessert, dokumentert og avgitt til arkivering. For nye prosjekter som innbefatter innsamling av data, skal prosjektbeskrivelsen inneholde en plan for langsiktig forvaltning av innsamlete data. Planen utarbeides i samråd med instituttets datasenter, og bør innbefatte en kort beskrivelse av de datasettene som skal samles inn (metadata); planlagt prosessering, analyse og kvalitetskontroll; transformering til ferdig dataprodukt, lagringssted og lagringsformat, protokoller og tjenester for tilgjengeliggjøring, dokumentasjon for viderebruk samt publiseringstidspunkt.

Dataintegriteten⁵ må sikres over hele datasettets livsløp, gjennom alminnelig datasikkerhet, versjonskontroll og, om nødvendig, tilgangskontroll. Instituttets datasenter har ansvaret for å migrere data til nye lagringsmedier når dette blir nødvendig av hensyn til dataintegritet eller teknologisk utvikling. Prosjekter med datainnsamling anses ikke som avsluttet før alle data er arkivert og dokumentert. Dette sjekkes av avdelingsdirektør ved godkjenning av sluttrapport.

For å sikre framtidig etterprøvnbarhet, må data som danner grunnlag for vitenskapelige publikasjoner, rapporter og analyser kunne identifiseres entydig. Data bør derfor arkiveres med unike og uforanderlige, digitale identifikatorer som kan knyttes til en datofestet versjon av datasettet⁶.

9. Lagring

Alle datasett skal lagres i original ved instituttet, eller ved eksterne datasentre som instituttet gjør avtaler med. Lagringen bør skje i samsvar med beste praksis for langsiktig dataforvaltning, i godt dokumenterte standardformater og med sikkerhetskopier lagret atskilt fra originaldata. Digitale data skal så langt mulig lagres i systemuavhengig form, i åpne standardformater og – hvis relevant – være lesbare med åpen programvare.

Der det er aktuelt, skal data samles inn og dokumenteres etter de standarder som gjelder for vedkommende datatype eller fagfelt.

10. Dokumentasjon

Alle datasett skal registreres i instituttets metadatabase. Metadatabasen skal gi grunnlag for gjenfinning, vedlikehold og publisering av datasett. Når et datasett oppdateres eller utfylles med nye data, må også metadata oppdateres, eller eventuelt publiseres i ny og separat utgave for en ny versjon av datasettet.

Alle datasett som blir arkivert ved NP skal være tilstrekkelig dokumentert for uavhengig, framtidig bruk. Metadata skal være søkbare på data.npolar.no og kunne utveksles med eksterne datakataloger. Oppdatert lagringsinformasjon skal til enhver tid være registrert i metadatabasen.

Ofte vil seinere bruk av datasettet kreve mer dokumentasjon enn den som registreres i metadatabasen. Denne informasjonen skal i så fall samles i en datarapport fra feltarbeidet/datainnsamlingen. Datarapporten skal inneholde all informasjon som er nødvendig for å bruke datasettet, inkludert:

* Feltlogg

⁵ Dette betyr at datasettet holdes intakt, uten noen form for forvitring eller utilsiktet/uautorisert endring.

⁶ De mest aktuelle typene er UUID og DOI. Begge blir koplet til en URI til selve datasettet.

- * Dokumentasjon av innsamlingsmetode, kvalitetssikring og validering
- * Dokumentasjon av programvare som har vært brukt og/eller kreves
- * Beskrivelse av instrumenter, feilkilder, måleoppløsning m.m.

Datarapporten skal være tilgjengelig elektronisk, og URI registreres i metadatabasen.

11. Tilgjengeliggjøring

Metadata og arkiverte datasett vil bli gjort søkbare og tilgjengelige over internett via digitale tjenester bygd på standard protokoller⁷, med unike og uforanderlige identifikatorer, søketjenester, nedlastingstjenester og transformasjonstjenester.

Dataformidling over internett forutsetter at metadata og alle betingelser for formidling av datasettene er maskinlesbare. Metadata, og hvis mulig datasettet selv, skal derfor inneholde formatert informasjon om eier, opphavsrett, datalicens, sitering, publiseringsdato, versjon, eventuelle bruksbegrensinger og andre rettigheter. Metadata skal knyttes til data gjennom unike og uforanderlige, digitale identifikatorer.

12. Personvern

Data skal i utgangspunktet være fri for andre personopplysninger enn tittel, navn og arbeidssted med adresse og telefonnummer. Dersom det forekommer datasett med andre personopplysninger skal NP følge Datatilsynets anvisninger og personopplysingsloven.

⁷ HTTP, OPeNDAP, WMS, WFS og andre