

Konseptbeskrivelse for

Open ARI – The Open Arctic Research Index for publications and datasets

The Open Arctic Research Index for publications and datasets er en overlay-tjeneste bygget på toppen av eksisterende data i nasjonale og eventuelt også internasjonale databaser med relevante data for det definerte geografiske området.

Med en **overlay-tjeneste** menes det her en tjeneste som er basert på eksisterende metadata fra ulike kilder og hvor disse data gjøres tilgjengelig for gjenfinning i et felles søkegrensesnitt. Gjenfinning skal kunne skje dels ved tekstlig søk med avgrensningsmuligheter i resultatlista, og dels ved søk fra kart i metadata som har geodata inkludert (dette vil ikke gjelde for alle metadatabeskrivelser). Selve datasettet eller publikasjonen blir liggende i kilden, slik at metadata i resultatlista alltid vil styre brukeren til kilden hvor publikasjonen eller datasettet befinner seg. *Open ARI* bidrar på den måten til økt synlighet av innholdet i de enkelte kildene.

Innholdet i *Open ARI* bygges opp og oppdateres ved regelmessig, automatisk innhenting (høsting) av metadata fra utvalgte generelle og spesialiserte dataarkiv (kilder), basert på standard protokoll for høsting av metadata (*OAI-PMH – The Open Archives Initiative Protocol for Metadata Harvesting*). Denne høstingsmuligheten er normalt tilgjengelig i nyere arkivløsninger og er teknologi og problemstillinger som UiT har lang erfaring med, blant annet når det gjelder tjenesten *High North Research Documents* (<http://highnorth.uit.no/>). I de tilfeller hvor arkivkilden ikke støtter OAI-PMH vil *Open ARI* måtte benytte spesialiserte programmer for uttrekk av data, i et samarbeid med eier av arkivet.

Kvaliteten på en overlay-tjeneste er i stor grad avhengig av kvaliteten på de metadata som det enkelte arkiv gjør tilgjengelig for høsting. Når arkivene er fagspesifikke eller tematiske, kan dette påvirke muligheten for tverrfaglig, standardisert presentasjon av metadataene på grunn av ulike fagspesifikke standarder eller egenutvikla metadatastandard. *Open ARI* vil derfor standardisere og eventuelt berike metadata fra det enkelte arkiv, i samarbeid med arkiveier, for å oppnå en best mulig funksjonalitet for brukerne.

En overordnet målsetting (og forventning) til arkiv over forskningsdata er at de oppfyller FAIR Data Principles¹ (FAIR: Findable, Accessible, Interoperable, and Re-usable). I Forskningsrådets policy *Tilgjengeliggjøring av forskningsdata*² er krav om dette tatt inn som et hovedprinsipp, og er også sterkt vektlagt i *Nasjonal strategi for tilgjengeliggjøring og*

¹ <https://www.force11.org/group/fairgroup/fairprinciples>

² Forskningsrådet 2017, Åpen tilgang til forskningsdata – <https://www.forskningsradet.no/servlet/Satellite?cid=1254032061080&pagename=VedleggPointer&target=blank>

*deling av forskningsdata*³. FAIR forutsetter internasjonale standarder for metadata og lisenser, og bruk av globale og varige adresser til selve innholdet, som f.eks. DOI-identifikator. Arkiver som oppfyller FAIR vil ha vesentlig større synlighet i det europeiske og globale nettverket for forskningsdata, og slik også enkelt kunne integreres i gjenfinningstjenester. *Open ARI* vil derfor ha som et viktige delmål å bidra med hjelp til å identifisere eventuelle hindringer for dette i de ulike kildearkivene og gi råd om hvordan de kan bli et FAIR-arkiv. Men *Open ARI* skal også ha som mål å kunne tilby standardiserte metadata fra arkiver som selv ikke oppfyller FAIR-prinsippene, til andre nasjonale og internasjonale søketjenester, for økt synlighet av innholdet i det enkelte arkiv og av de involverte forskerne.

Prosjektidéen og konseptskissen er så langt bare presentert for Norsk Polarinstitutt som har sagt seg interessert i å delta i et pilotprosjekt som skal kartlegge og vurdere muligheter og interesse for en slik overlay-tjeneste. En realisering av tjenesten vil være avhengig av et **samarbeid** med de nasjonale datasentrene innenfor norsk polarforskning. Disse kan kobles på som aktive samarbeidspartnere når *Open ARI* ser på integrasjonsmuligheten for det enkelte arkiv. Her vil det være naturlig å se for seg at arkiver som tilfredsstiller FAIR-prinsippene og dermed kan høstes automatisk, vil bli realisert først.

Open ARI-konseptet overlapper med målformuleringen til forskningsinfrastrukturprosjekt *NorDataNet* (DNMI), men representerer en enklere og rimeligere tilnærming som ikke (uten videre) forutsetter at integrerte kilder først tilpasser seg en felles mal. Som breddeuniversitet representerer også UiT en større faglig bredde og forankring enn DNMI når en slik tverrfaglig tjeneste skal realiseres og forvaltes.

Internasjonalt så ligger det gode muligheter for et samarbeid med den kanadiske tjenesten *Polar Data Catalogue – The Arctic and Antarctic metadata and data repository for Canada*⁴. Denne tjenesten bygger på mye av den samme tenkningen som ligger bak *Open ARI* og de inviterer til samarbeid og deling av metadata for økt synlighet og bruk av forskningsdata. Og tilsvarende vil det være ønskelig å etablere samarbeid med land øst for oss.

Innholdet i *Open ARI – The Open Arctic Research Index for publications and datasets* vil, som navnet tilsier, omfatte både relevante forskningsdata og publikasjoner. UB har i tolv år utviklet tjenester for publikasjoner som er åpent tilgjengelig (open access) og lanserte i januar 2012 tjenesten *High North Research Data*, – en overlay-tjeneste med metadata om «High North»-relevante publikasjoner. Denne tjenesten bygger på et samarbeid med Bielefeld University⁵ og inkluderer nå (pr. 1.4.18) metadata for over 550 000 publikasjoner og 150 000 datasett fra mer enn 2500 kilder fra hele verden. Selv om denne tjenesten bare

³ Kunnskapsdepartementet 12/2017, Nasjonal strategi for tilgjengeliggjøring og deling av forskningsdata – <https://www.regjeringen.no/contentassets/3a0c9eeaa1c9b4611a1b86fc5616abde7/no/pdf/f-4442-b-nasjonal-strategi.pdf>

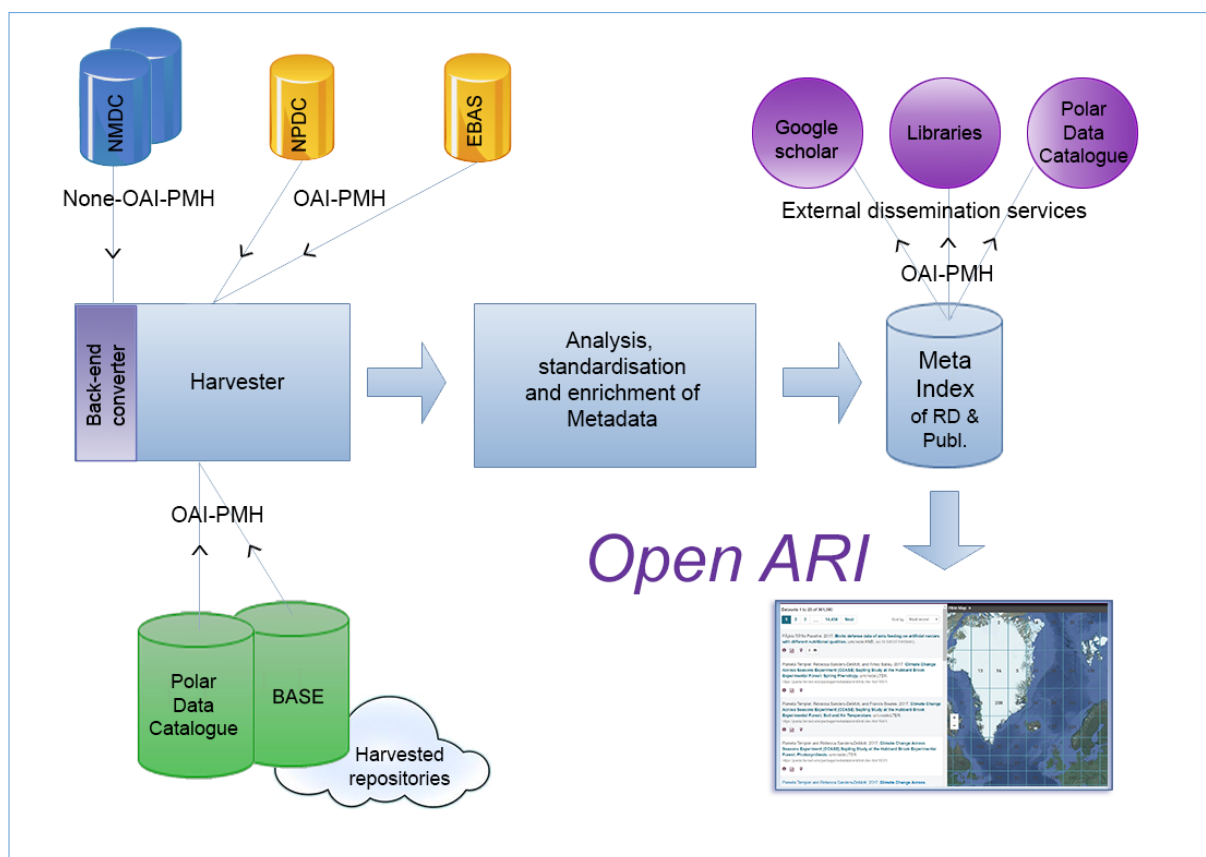
⁴ <https://www.polardata.ca>

⁵ Bielefeld Academic Search Engine – BASE <https://www.base-search.net/>

inkluderer kilder som har åpen tilgang til publikasjoner som policy, og derfor ikke dekker alle publikasjoner, vil dette kunne synliggjøre og gi tilgang til sentrale publikasjoner, sammen med forskningsdata fra Arktis. En viktig side ved dette er at andelen åpne publikasjoner nå øker betydelig år for år som konsekvens av stadig tydeligere krav fra finansiører og forskningsinstitusjoner.

Hva kan Open ARI bidra med

- for det enkelte arkiv? – Økt synlighet for arkivet – et supplement til dagens status
- for den enkelte forsker? – Enklere og tverrfaglig oversikt over relevante forskningsdata
- for den enkelte forsker? – Mulighet for flere siteringer av egne data ved økt synlighet
- for UiT? – Styrke sin posisjon som ledende innen utdanning og forskning om Arktis og som nav for åpen tilgang til nasjonal og internasjonal arktisk forskning
- for Forskningsrådet? – Mulighet for økt nytteverdi av forskningsresultater, forenkler tverrfaglig bruk av resultater
- for Norge? – Økt internasjonal synlighet av norsk polarforskning



Figur 1 Prinsippkisse for Open ARI

Bakgrunn for

Open ARI – The Open Arctic Research Index for publications and datasets

I Tildelingsbrev 2018, om utviklingsavtalen med UiT er bl.a. følgende mål gitt for avtaleperioden:

UiT skal være ledende innen utdanning og forskning om Arktis, og styrke sin posisjon som Norges arktiske universitet.

- *UiT skal styrke sin ledende posisjon i Norge på polarforskning og være en viktig premissleverandør for politiske beslutninger om polare spørsmål.*

UiTs Drivkraft i nord-strategi er klar på at UiT skal ha nasjonal forskningsinfrastruktur som særlig bygger opp under de tematiske satsingene og de forskningsområdene hvor UiT er internasjonalt ledende. Polarforskningen ved UiT er et slikt område, – noe som ble underbygget i evalueringen (juni 2017) av norsk polarforskning⁶. I denne evalueringen heter det blant annet⁷:

- *Collaboration or even communication among the larger polar data centers seems limited. We did not find much evidence for connections of the polar data centers with other environmental data centers across Norway.*
- *We identify a relative degree of insularity in Norway's polar data community.*

Med dette forslaget kan vi forenkle og bøte på noen av utfordringene dagens infrastruktur byr på, ved å lage en overbygning over disse dataene. En slik overbygning innebærer å bygge et nettsted som fungerer som en felles søketjeneste med lenker direkte til arktiske forskningsdatasett og publikasjoner, uten å gå i konkurranse med eller endre betydningen/rollen til de ulike nasjonale datasentre.

Også i OECDs gjennomgang av digitaliseringen av offentlig sektor i Norge⁸ pekes det på at systemet er for fragmentert. Det utvikles sektorinterne løsninger uten at det i tilstrekkelig grad tas hensyn til om bedre løsninger allerede er utviklet andre steder eller til behovet for at egen løsning bør fungere sammen med andres løsninger. Dette er den samme utfordringen som evalueringen av norsk polarforskning peker på innenfor det arktiske forskningsfeltet. I tillegg peker den siste evalueringen på utfordringer for de ulike

⁶ Norwegian Polar Research, An Evaluation –

https://www.forskningsradet.no/servlet/Satellite?cid=1254028931995&pagename=VedleggPointer&target=_blank

⁷ Samme sted, pkt. 4.5

⁸ <https://www.oecd.org/gov/digital-government/digital-government-review-norway-recommendations.pdf> og Nasjonal strategi for tilgjengeliggjøring og deling av forskningsdata, s. 42 – <https://www.regjeringen.no/contentassets/3a0ceeaa1c9b4611a1b86fc5616abde7/no/pdf/f-4442-b-nasjonal-strategi.pdf>

datasentrene knyttet til prioriteringer av ressurser til dette og på mangel på kompetanse (s. 23):

In common with other polar data centers and with data centers in general, many of these Norwegian polar data centers suffer the symptoms of general inattention given to data: lack of priority in science or infrastructure funding, difficulty to recruit and maintain long-term skilled staff, increasing demands for access, interoperability and services on flat or declining budgets.

UiT Norges arktiske universitet dekker de fleste forskningsfelt innenfor det polare og erfarer selv de utfordringene som evalueringen av norsk polarforskning og OECD peker på. Som en betydelig aktør innenfor polar forskning bør UiT være villig til å ta på seg rollen som en langsiktig, nasjonal tilbyder av tjenester og verktøy for tilgjengeliggjøring av forskningspublikasjoner og -datasett om Arktis.

Dette initiativet kan bidra til økt nytteverdi for de store investeringene som er gjort under Forskningsrådets *Nasjonal satsing på forskningsinfrastruktur*, gjennom større synlighet av arkiverte forskningsdata og med det bedre mulighet for gjenbruk generelt og spesielt på tvers av fagområder.