

Søknad om midler til kompetanse-hub for IT i forskning

Life science data management hub

Søker:	Senter for bioinformatikk
Vertsinstitutt (hub):	Institutt for informatikk
Samarbeidsparter (inkl. noder):	USIT, Kompetansegruppen for forskningsdata, CEES, Norment, Oslo Center of Biostatistics and epidemiology, Big Insight, CanCell, Life science, ELIXR er de som inngår ved oppstart. Ytterligere noder/samarbeidspartnere vil kunne bli inkludert.
Søknadsbeløp:	Samlet 2,42 MNOK
Søknadsperiode:	2019-2023

1: Formål og gevinster

1.1: Hva er formålet?

Formålet er å sørge for etablering og fasilitering av at forskere gjennom all sin forskning planlegger for og implementerer gode rutiner for datahåndtering innen livsvitenskapene. Datahåndtering innebærer å kontrollere den informasjon som genereres gjennom et forskningsprosjekts fulle levetid. Alle forskningsprosjekter har noen grad av behov for datahåndtering. I livsvitenskapen er dette særskilt viktig ikke minst innenfor genomikk, med sine store datamengder.

Etablering av gode rutiner for håndtering av data vil kunne sørge for at data og modeller blir "FAIR". Dette skjer vet at de:

1. kan gjøres tilgjengelige for søk for andre forskere etter publikasjon ("Findable")
2. kan gjøres lesbare og nedlastbare for andre forskere ("Accessible")
3. kan bli klart forstått i den opprinnelige ekseperimentsammenhengen dataene ble produsert i ("Interoperable"), og
4. kan brukes av andre forskere ("Re-usable").

Det foreligger en rekke tekniske løsninger forskere kan benytte, både standarder og infrastrukturelementer, for å få maksimal effekt av egne data i forskning, men det er i dag ingen som har dette som særskilt kompetanseområde, eller som kan rådggi innen dette. I stigende grad krever både finansiører og at søkere har gode datahåndteringsplaner fra start, ikke minst for å sikre maksimal verdiutnyttelse av forskningen, gjennom gjenbruk og krysskoplinger. Etablering av slike krever god forståelse av hvilke lagringsløsninger som foreligger, hvilke standarder som eksisterer for håndtering av ulike typer data, hvilke metadata som bør inkluderes for å sikre gjenbruk, hvilke internasjonale utvekslingsløsninger som foreligger, hvilke datamengder som vil genereres, hvilken livssyklus data kan forventes å ha, i hvilken grad data er personsensitive og krever løsninger for slike, og en rekke andre spørsmål det ikke er lett for den enkelte



forsker å kunne håndtere. Det er derfor behov for å etablere såkalte "datastuerter" ("data stewards"), for å utvikle kulturen for datahåndtering. Vi søker derfor spesifikt om midler til etablering av en slik stilling, og til noe kursaktivitetsmidler knyttet til dette.

1.2: Hvilket kompetanseområde skal dekkes?

Vi ser for oss å dekke kompetanse innen hele livsløpet til molekylærbiologiske data innen alle sider av biologisk og biomedisinsk forskning. Dette vil inkludere kunnskap om tekniske delingsløsninger, om standarder innen dataformater og transaksjonsprotokoller, infrastrukturer, metadatahåndtering, former for publisering av forskningsdata. Også håndtering av sensitive data innen biomedisin vil være av betydning, både i tilknytning til kravene som stilles av GDPR, og i forbindelse med generelle forskningsetiske krav, og hvordan data kan deles på ulike måter under disse beskrankninger.

1.3: Hvilke ressurs-/eInfrastrukturbehov skal dekkes?

Det vil dels dreie seg om utnyttelse av eksisterende ulike infrastrukturer (som Norwegian eInfrastructure for Life Science), dels om utnyttelse av internasjonale ressurser, som ENA, GEO, ArrayExpress og andre, men også standardiseringsarbeid og protokoller under utvikling knyttet til Global alliance for Genomics and Health og andre.

2: Organisering

2.1: Hvem skal være ansvarlig leder?

Eivind Hovig

2.2: Hvem skal være datahåndteringsansvarlig?

Eivind Hovig inntil person er ansatt.

2.3: Hvor skal hub-en være lokalisert?

Hub skal svære lokalisert ved Senter for bioinformatikk, tett tilknyttet ELIXIR-prosjektet. ELIXIR-prosjektet inkluderer både en helpdesk-funksjon og samarbeider også med Kjernefasilitet for bioinformatikk, og vil på denne måten også kunne formidle til aktuelt servicepersonale og dermed også bidra til lokal koordinering. ELIXIR-prosjektet ved UiO er også koordinert nasjonalt og internasjonalt, og vi vil på denne måten være i stand til å utnytte de mange ressurser som eksisterer og er under utvikling.

2.4: Hvilke forskningsgrupper, fagmiljøer og lignende skal inngå som noder?

Aktuelle fagmiljøer inkluderer større datagenererende enheter, som Norsk sekvenseringssenter, analyseenheter (som OCBE), større forskningsprosjekter der storskala data analyseres, i tillegg til at det vil være et stort behov for direkte samarbeid med både den eksisterende kompetansen

innen datahåndtering, og også USIT spesielt. Life science som overgripende enhet finner også at dette er viktig. Det er derfor innhentet støttebrev som demonstrerer den store interessen som eksisterer rundt etablering av en slik kompetansehøb.

2.5: Hvordan skal arbeidet i organiseres?

Stillingen som søkes vil være lagt til Senter for bioinformatikk, og vil trekke på veksler både på eksisterende datahåndteringskompetanse, og den kompetanse som befinner seg i ELIXIR-systemet, og via USIT. Det vil bli etablert en styringsgruppe som vil være representanter for de noder som er tilknyttet (se støttebrev). Det vil bli lagt opp til personlig rådgivning, til kursvirksomhet, og dokumentasjonsmateriale og instruksjonsmateriale vil bli laget.

2.6: Hvordan skal gevinstene videreføres etter avvikling?

<Kort om gevinstrealisering inkludert plan for datahåndtering>

3: Budsjett

Inntekter:	2019:	2020:	2021:	Sum:
Fra fagrådet	200000 kr	1110000 kr	1110000	2420000
Egenfinansiering	kr	kr	kr	kr
Annen finansiering	kr	kr	kr	kr
Sum inntekter:	200000 kr	1110000 kr	1110000 kr	2420000 kr
Utgifter:				
Lønnsmidler, frikjøp	200000 kr	960000 kr	960000 kr	2120000 kr
Anskaffelser	kr	kr	kr	kr
Kompetanseutvikling	kr	100000 kr	100000 kr	200000 kr
Andre utgifter	kr	50000 kr	50000	100000 kr
Sum utgifter:	200000 kr	1110000 kr	1110000 kr	2420000kr

4: Vedlegg

Vedlagt følger støttebrev fra IFI, samt fra noder/fagmiljøer/samarbeidspartnere.

Dato: 15 August 2019

STØTTEBREV

IFI stiller svært gjerne som vertsintitusjon for søknaden "Life science data management hub", og er av den oppfatning at tilgjengeliggjøring og gjenbruk av forskningsdata er svært viktig. Foreliggende søknad vil bidra sterkt til at UiO i større grad vil kunne få større nytte av sine forskningsdata, og at aktiviteten vil bidra til å bedre forskningspraksis på UiO innen dette området.

Vennlig hilsen

Ellen Munthe-Kaas

Ellen Munthe-Kaas
Instituttleder
Institutt for informatikk
Universitetet i Oslo



Postadresse: Postboks 1080 Blindern, 0316 OSLO
E-post: ifiadm@ifi.uio.no
www.mn.uio.no/ifi/

Eivind Hovig
Senter for bioinformatikk
Gaustadalléen 23 B
N-0373 Oslo

Dato: 19. august 2019

Støttebrev Life science data management hub

Kompetansegruppen for forskningsdata støtter søknaden fra Eivind Hovig og Senter for bioinformatikk om midler for å etablere en kompetanse-hub for datahåndtering innenfor livsvitenskapene.

Kompetansegruppen for forskningsdatahåndtering er et samarbeid mellom FADM, USIT og UB, og består av Margaret Fotland (FADM), Anne Bergsaker (USIT), Live Håndlykken Kvale og Elin Stangeland (UB). Gruppen arbeider i samråd med Fagråd for eInfrastruktur for å bedre praksis for forskningsdatahåndtering på UiO, samt sikre at UiOs forskningsmiljøer er i stand til å følge opp krav og målsetninger fra finansierer og myndigheter når det gjelder forvaltning av forskningsdata.

Søknaden om en Life science datahåndteringshub komplementerer Kompetansegruppens arbeid med sitt fokus på å etablere et bedre tilbud for forskere med livsvitenskap som fokusområde. Hvis huben etableres vil denne kunne bli en viktig samarbeidspartner for gruppen.

Med hilsen



Elin Stangeland
På vegne av UiOs Kompetansegruppe for forskningsdata



Til: Fagrådet for eInfrastruktur / Søker på "Life science data management hub"

Dato: 16. august 2019

Støttebrev

USIT v/Underavdeling for IT i Forskning vil gjerne støtte søknaden "Life science data management hub". Dette siden vi har erfaring for at arbeidet med tilgjengeliggjøring av data er svært viktig, både for å sikre lang levetid, utnytte ressursene som er lagt inn i dataproduksjonen, og fordi god håndtering vil kunne gi både gjenbruk og kritisk dokumentasjon på utført forskning. USIT er allerede involvert på flere nivåer med å utvikle og/eller drifte løsninger av relevans for søknaden, som for eksempel TSD, og vi ser også her helt klart nytten i å bedre planlegging og utnyttelse av systemet for den enkelte forsker. Vi vil derfor svært gjerne bidra i form av koordinering og planlegging, samt samarbeid der dette er naturlig. USIT mener videre det er fornuftig at en slik Life Science Data management HUB er organisert rundt et sterkt fagmiljø.

Med vennlig hilsen

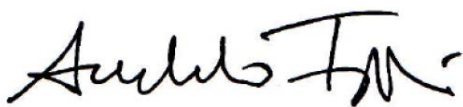


Gard Thomassen
Underavdelingsdirektør – IT i Forskning
USIT, UiO
gardot@usit.uio.no/93674926



Oslo 15 august 2019

BigInsight supports open data and our research follows the FAIR principles. This is not easy, and we are enthusiastic to see the emergence of data stewards to facilitate FAIR compliance and so in maximizing the impact of our research. We therefore foresee very useful collaboration and interaction towards this goal between the Life science data management hub and BigInsight.

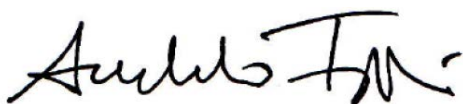
A handwritten signature in black ink, appearing to read "Arnaldo Frigessi".

Arnaldo Frigessi
Professor
BigInsight
Centre for research based Innovation
Director

<http://www.biginsight.no>
frigessi@medisin.uio.no
Tel: +4795735574

Oslo 15 august 2019

OCBE, the Oslo Centre for Biostatistics and Epidemiology, is very happy to support the application for a " Life science data management hub" as a node. We believe that it is very important to foster good data management practices to maximize the value of the research generated, and we are confident that the presented plans will significantly enable us in our support and advising work for OUS/UiO scientists. OCBE will be a collaborator and contribute towards development of data management practices.



Arnaldo Frigessi
Professor

Oslo Centre for Biostatistics and
Epidemiology (OCBE)
University of Oslo
Oslo University Hospital
Director

frigessi@medisin.uio.no
Tel: +4795735574

UNIVERSITETET I BERGEN

INSTITUTT FOR INFORMATIKK
HØYTEKNOLOGISENTERET, 5020 BERGEN
Tlf. 55 58 42 00
Telefax 55 58 41 99



UNIVERSITY OF BERGEN

DEPARTMENT OF INFORMATICS
HØYTEKNOLOGISENTERET,
N-5020 BERGEN, NORWAY
Ph. +47-55 58 42 00
Fax +47 55 58 41 99

Bergen, 15. august 2019

Støtte til søknaden “Life Science Data Management Hub”

Undertegnede er Head of Node for den norske noden av ELIXIR (den europeiske infrastrukturen for bioinformatikk – et ESFRI landmark), og er også leder av arbeidsgruppe 4 (Kompetanse og Infrastruktur) i Digitalt liv Norge. I begge egenskaper ser jeg nytten av å utvikle datahåndteringsperspektivet for forskere, og dette er et felt som bare vil bli viktigere i tiden framover, med nye krav til Open science, FAIR (Findable, Accessible, Interoperable, Reusable) deling av data på standard formater, datadelingsmekanismer og med krav til god forvaltning av data over den samlede levetid data vil ha. ELIXIR Norge, og Digitalt liv Norge vil derfor være en entusiastisk støtte dersom søknaden innvilges, og vi tror vi kan skape viktige synergieffekter, ikke minst med de internasjonale standarder som er under utvikling.

Yours sincerely

Inge Jonassen, PhD, Professor
Head of ELIXIR Norway
Head of Computational Biology Unit
Department of Informatics
University of Bergen
Norway

University of Oslo
Division of Mental Health and Addiction

Til den det måtte angå

Dato: 15. aug, 2019

Deres ref.:

Vår ref.: OAA

Støttebrev

Kjære Eivind,

Her på senter for fremragende forskning NORMENT har vi i flere år bygget mye av vår forskningsaktivitet omkring stordata analyser. Det er helt sentralt element for å forstå årsaker til komplekse menneskelige sykdommer, slik som psykiske lidelser. Alle våre viktigste publikasjoner siste år (i tidsskrift som Science, Nature, Nature Genetics og Nature Neuroscience) er basert på stordata-analyser. Vi har nå multidimensjonelle data (genotyper – fenotyper) fra over 630 000 individer, som ilt neste år vil øke til 780 000. Dette vil kreve ikke bare sikker lagring, men også sikker dataanalyse med oppdaterte verktøy og kapasitet, slik at vi kan utnytte data og skaffe mer innsikt i sykdomsårsaker, og ikke minst bidra til persontilpasset medisin. Som du vet, har vi derfor et tett samarbeid med Senter for bioinformatikk.

Vi støtter søknaden " Life science data management hub".

Vennlig hilsen,

A handwritten signature in blue ink, appearing to read "Ole A. Andreassen".

Ole A. Andreassen

Professor
Director NORMENT Centre



Klinikk Psykisk Helse og Avhengighet
Postadr.: Postboks 1039 Blindern, 0316 Oslo
postmottak@medisin.uio.no
Org.nr.: 971 035 854





Oslo 19.08.2019

Fagråd for eInfrastruktur
Universitetet i Oslo

STØTTE TIL SØKNADEN «LIFE SCIENCE DATA MANAGEMENT HUB»

CanCell er et SFF under MedFak, UiO, som har som mål å identifisere kreftcellers svake punkter og utnytte disse til omprogrammeringsbasert kreftterapi. Senteret består av 6 forskningsgrupper ved UiO/OUS og 7 assosierte medlemmer, inkludert ledende norske bioinformatikere (Eivind Hovig) og biostatistikere (Arnoldo Frigessi).

Vi kommer i kontakt med ulike datatyper og dataformater i forbindelse med senteret. Vi er opptatt av å gjøre våre data synlige, og å utnytte muligheter for krysskoplinger mellom egne data og offentlig tilgjengelige forskningsdata. Vi vil derfor ha behov for å bli kjent med eksisterende formater og delingsmekanismer, få bedre kunnskap om god dataforvaltning for våre data, herunder sekvenseringsbaserte data, men også avbildingsdata. Vi støtter derfor søknaden "Life science data management hub" og vil gjerne delta som miljø fordi vi tror vi både kan bidra til å utvikle bedre mekanismer, og selv utnytte mulighetene som kommer med et finansiert prosjekt.

Vennlig hilsen

Harald Stenmark
Professor, senterleder



UNIVERSITY
OF OSLO

CEES

Centre for Ecological and Evolutionary Synthesis

Department of Biosciences
P.O.Box 1066, Blindern
N-0316 Oslo

Kristine Bonnevie's hus
Telephone (office): +47 22 85 46 02
CEES-office telephone: +47 22 85 44 00
Telefax: +47 22 85 40 01
E-mail: k.s.jakobsen@ibv.uio.no
Web: <http://www.cees.uio.no/>

Blindern: 19 August 2019

**FACULTY OF MATHEMATICS AND
NATURAL SCIENCES**

Støtte til søknad - Life Science data management hub

Samarbeidet med Senter for Bioinformatikk (SBI) i form av en hub-node struktur er viktig for Institutt for biovitenskap (IBV) og Centre for Ecological and Evolutionary Synthesis (CEES) i særdeleshett. Dette har for eksempel gitt seg utslag i utlysning av en førsteamanuensis-stilling innen bioinformatikk knyttet både til IBV og SBI og vi har etablert et nytt endringsmiljø (CELS 2¹).

Datahåndtering er særdeles viktig innen livsvitenskap og det krever ressurser og ikke minst innsikt. CEES og IBV støtter derfor søknaden om å søke midler til «data stewardship» noe som vil kunne bidra til å gjøre data FAIR² for brukere på UiO, samt å bistå genomikk-brukere med data-delning på ulike relevante måter. Dette er en svært viktig aktivitet for senteret og for hub-node strukturen.

Hilsen

Kjetill S Jakobsen
Professor, leder CEES og NSC

¹ CELS 2: <https://www.mn.uio.no/cees/english/research/groups/cels/index.html>

² FAIR-data: <https://www.uio.no/for-ansatte/arbeidssstotte/fa/forskningsdata/fair-data/>