

Innspill til nytt norsk veikart for forskningsinfrastruktur

Uttalelser – høring

Alle institusjoner som har søkt Nasjonal satsing på forskningsinfrastruktur om midler siden oppstart i 2009 har hatt mulighet til å gi innspill til nytt norsk veikart for forskningsinfrastruktur. Hver institusjon har blitt bedt om å levere ett samlet innspill for hvert fag-/temaområde som er relevant for institusjonen som er godt forankret i egen organisasjon.

DATAINFRASTRUKTUR OG IKT

Dato: 03.10.2022

DATAINFRASTRUKTUR OG IKT

Institusjoner som har sendt inn uttalelser

- Høgskulen på Vestlandet
- NTNU
- OsloMet - storbyuniversitetet
- UiT Norges arktiske universitet
- Universitetet i Agder
- Universitetet i Bergen
- Universitetet i Oslo
- Universitetet i Sørøst-Norge
- Universitetet i Stavanger
- Havforskningsinstituttet
- Meteorologisk institutt
- NIBIO
- Simula Research Laboratory
- NORCE - Norwegian Research Centre AS
- NORSAR
- SINTEF
- Stiftelsen Nansen senter for miljø og fjernmåling (NERSC)
- Akershus universitetssykehus HF
- Oslo universitetssykehus
- Direktoratet for e-helse
- Sikt - kunnskapssektorens tjenesteleverandør
- Arkivverket
- NVE, Hydrologisk avdeling

Datainfrastruktur og IKT

Universitet / Høgskole (Offentlig)	Svar
Høgskulen på Vestlandet	Faggruppe innen Datainfrastruktur og IKT uttrykker følgende behov for forskningsinfrastruktur og e-infrastruktur: <ul style="list-style-type: none">- Spesifikt for IKT, det meste av infrastrukturen kan distribueres gjennom virtualisering (f.eks. i sky eller lignende teknologier). Som eksempler, se nrec.no for virtualisering, og det franske distribuerte laboratoriet for til og med fysisk IoT.- Vi forventer at det, spesielt for AI-trening og simulering, vil være en stor etterspørsel etter distribuerte og sentraliserte ressurser (f.eks. superdatamaskiner) som må administreres ansvarlig og enda viktigere bærekraftig og på en sikker måte.- IT-forskeres enkle tilgang til sky- eller virtualiseringssystemer er avgjørende for å oppnå de ovennevnte poengene.- I tillegg til virtualisering bør det være forskningslaboratorier som inneholder fysiske enheter (f.eks. AR/VR/datagrafikk, IoT, innebygde systemer, robotikk osv.), det vil da være nødvendig å ha både personell og rom som kan være vert for nettverk, vedlikeholde, og fornye disse enhetene over tid.
NTNU	<p>* Forskningens behov for forskningsinfrastruktur for å løse utfordringer innenfor et strategisk prioritert område</p> <p>E-infrastrukturer skal levere tjenester til forskningsprosjekter som sikker lagring, beregningsressurser og tilgjengeliggjøring av data i tråd med FAIR-prinsippene. De er viktige for mange fagfelt innen naturvitenskap og teknologi, og er også blitt vesentlig i medisin og helsevitenskap, samfunnsvitenskap, humaniora og økonomi. De er derfor viktige innen alle våre strategisk viktige områder og for kvalitet i forskning generelt.</p> <p>NTNU mener e-infrastruktur skal bidra til kvalitet i forskningen ved</p> <ul style="list-style-type: none">• å understøtte nasjonale og internasjonale målsettinger for åpen vitenskap, FAIR forskning og sikker datahåndtering, og bidra til kvalitet og relevans i forskningen• å være attraktiv og legge til rette for samarbeid på tvers av organisasjoner og landegrenser• å skape en trygghet for forskere og ledere om at lover og forskrifter enkelt kan bli fulgt• å være en varig tjeneste som er bærekraftig, godt organisert, finansiert og effektivt driftet• å legge til rette for langsiktig faglig utvikling gjennom utvikling av databaser, programvare og systemer av programvare• å ha et støtteapparat som sørger for god støtte for beslutninger og operasjonalisering knyttet til e-infrastruktur til forskere for samhandling, lagring, prosessering og analyse av ulike typer forskningsdata <p>Fagfeltene innen IKT har behov for infrastruktur som kan støtte kjerneområder i IKT-forskning. Mobile og trådløse nettverk utgjør en av de viktigste infrastrukturene for kommunikasjon og tilgang til informasjon og tjenester i samfunnet. Drevet fram av globale</p>

standarder og et omfattende økosystem av mobile terminaler, nettverks-utstyrsleverandører, mobiloperatører og applikasjonsutviklere, har mobile nettverk blitt en nødvendig del av hverdagen for milliarder av mennesker. Ellers se svar under mangler.

*** *Hvilke eksisterende nasjonale infrastrukturer er av stor verdi å opprettholde og videreutvikle***

E-infrastruktur: Fra bl.a. Sigma2 tilbys ressurser for lagring og beregninger, spesielt for CPU-baserte beregninger. Men det har vært kapasitetsutfordringer, spesielt innenfor GPU, lagring og manglende generelle løsninger for lagring, langtidslagring og prosessering av sensitive data (intern, fortrolig, strengt fortrolig). Løsninger gjennom NSD og andre dekker ikke alle behov. Vi ser det som svært viktig at nasjonale og internasjonale beregnings- og lagringsressurser blir opprettholdt og videreutviklet.

Bærekraft i egen organisasjon er prioritert ved NTNU og dermed også energieffektive anlegg som reduserer påvirkningen av miljø og kostnader. F.eks. har vi et tungregneanlegg som gjenbraker spillvarme og som er en del av den nasjonale infrastrukturen. Nasjonale infrastrukturer bør ha fokus på bærekraft og energieffektivitet.

Innen helseforskning er det viktig at det legges til rette for forskning på pseudonyme helsedata på en sikker måte som også muliggjør deling av data på tvers av institusjoner og landegrensener. Kjernefasilitet for Bioinformatikk (Biocore), en node i ELIXIR er en viktig infrastruktur å opprettholde.

Infrastruktur for IKT-forskning: Infrastrukturene NorFab/Nanolab, NORTEM og MiMac er viktige å opprettholde for forskning og utvikling av materialer på nanoskala: Effektiv og alternativ minneteknologi, nye substrat for beregninger, energieffektiv elektronikk, nye materialer innen spintronikk, kvantecomputere m.m.

*** *Innenfor hvilke områder blir det spesielt viktig å etablere nye nasjonale infrastrukturer eller samarbeide om internasjonalt***

E-infrastruktur: Se svar over. Det vil også være viktig med nasjonalt og internasjonalt samarbeid for å legge til rette for langsiktig faglig utvikling gjennom utvikling av databaser, programvare og systemer av programvare.

Infrastrukturer for IKT-forskning: På systemnivå mangler det infrastruktur for å kunne eksperimentere med system som støtter veksten og som kan svare på forventningene i et avansert, datarikt samfunn. En slik infrastruktur må være modulbasert, skalerbar og inkludere alt fra datainnsamling til bærekraftige og energieffektive beregnings- og lagringsressurser. Den må inneholde IoT (Internet of Things), enheter (typisk sensorer og aktuatorer med dataprosesseringskapasitet), overføringsnettverk (radio og kabel), distribuerte enheter for kommunikasjon med IoT-nivået, samt nettverk for kommunikasjon med beregningsressurser. Infrastrukturen skal bidra til å svare på forsknings spørsmål innen nettverk av sensorer og aktuatorer

(IoT), distribuerte beregninger, og om ressursforbruk innen «Edge computing» og tungregning.

Infrastrukturen kan bygge på eksisterende infrastruktur, typisk eksisterende lagrings- og beregningsressurser (f.eks. IDUN ved NTNU) og som suppleres med nettverk til distribuerte eksperimentelle noder på IoT-/Edge-nivået.

*** Hvilke hull kan dekkes gjennom utvikling av eksisterende nasjonale og /eller samarbeid om/tilgang til internasjonale forskningsinfrastrukturer**

E-infrastruktur: Kapasitetsproblemer og generelle og spesielle løsninger for lagring, langtidslagring og prosessering av sensitive data kan dekkes av utvidelse av eksisterende infrastruktur. NTNU HUNT Cloud tilbyr digitale laboratorier spesielt tilpasset sensitive data og store data. De har brukere fra 12 land og holder data for alle større universitet, forskningsinstitusjoner og befolkningsundersøkelser i Norge. Gjennom ISO-sertifisering på kvalitetsstyring og datasikkerhet og samsvar med Normen plasserer de NTNU blant de beste i Europa på sikker håndtering av forskningsdata. HUNT Cloud gir mulighet for nasjonal tilpasning inn mot European Health Data Space, hvilket vil gi et stort konkurransefortrinn i fremtiden. Utvidelse av de nasjonale infrastrukturen til også å omfatte HUNT Cloud ser vi derfor på som viktig.

Infrastruktur for IKT-forskning: Se svar over.

OsloMet -
storbyuniversitetet

Ordningen med forskningsinfrastruktur har vært en viktig ordning gjennom Forskningsrådet som har hjulpet med å tydeliggjøre behov og nødvendighet for forskningsinfrastruktur som en særskilt viktig rammebetingelse for fremdragende forskning og utvikling. I tillegg til at forskningsinfrastruktur er viktig for forskere i sitt virke, har det i tillegg hatt en viktig strukturerende rolle nasjonalt og internasjonalt og fått fram prioriteringer blant mange viktige forskningsområder. EU har med ESFRI vært en betydningsfull pådriver og foregangsorganisasjon i hvordan en planlegger, kvalifiserer og finansierer etablering av forskningsinfrastrukturer i Europa.

OsloMet vil anbefale Forskningsrådet å følge utviklingen av forskningsinfrastrukturprosesser i ESFRI og finansieringsformer for støtte til bl.a. «trans-national access» og andre virkemidler for å sikre optimal utnyttelse av nasjonale forskningsinfrastrukturer og samarbeid om utvikling og bruk av internasjonale infrastrukturen. Vi anbefaler også Forskningsrådet å vurdere sin ensidige tilnærming til søknader til Forskningsinfrastrukturordningen hvor det kun legges vekt på eksellense. Utvikling av forskningsinfrastrukturer i Europa som er viktige for Norge bør også vurderes i en tidligere fase enn det gjøres nå. Vi mener det er viktig å vurdere strategiske samarbeid med europeiske land om (videre-)utvikling av forskningsinfrastrukturer slik at Norge kommer tidlig nok med i utviklingen. Dette er avgjørende for å bidra substansielt i (videre-)utvikling av forskningsinfrastrukturer enn heller å komme sent med når hovedlinjene og de viktigste valgene vedrørende infrastruktur allerede er foretatt.

Det er helt avgjørende for teknologiske fag å ha adekvat IKT og datainfrastruktur. Dette innebærer kompetente folk, solide nasjonale støttefunksjoner, programvare, riktige lagringsressurser og nok beregningskraft, både klassisk og eksperimentell. Akkurat hva riktige lagringsressurser og nok beregningskraft er, vil være individuelt mellom fag og forskningsgrupper, og dette må kartlegges med jevne mellomrom for å sikre at gode rammebetingelser er på plass for forskningen. En slik kartlegging bør danne utgangspunkt i hvilke ressurser en må ha, basert på strategiske prioriteringer for fremdragende forskning i Norge. Vi anbefaler for eksempel øvelsen som ble gjennomført i regi av The Nordic e-Infrastructure Collaboration (NeIC), «Cognitus: A Science Case for High Performance Computing in the Nordic region». Tilgang til ESFRI forskningsinfrastrukturer eller andre internasjonale forskningsinfrastrukturer fordrer modenhet hos de forskningsgruppene som ønsker tilgang til disse ressursene. Det er viktig at de nasjonale rammebetingelsene er på plass slik at slik forskningsmodenhet kan utvikles. Tilgang til forskningsinfrastrukturer er ofte flerdelt og de fleste forskere vil bli vurdert om forskningen deres er fremdragende i konkurranse med andre forskere som også ønsker tilgang på infrastrukturen. Moderne forskningsinfrastrukturer er helt avhengig av velfungerende IKT og datainfrastruktur, de fleste forskningsinfrastrukturer er distribuerte og tilgang er kun mulig gjennom adekvat IKT og datainfrastruktur. I motsetning til de fleste forskningsinfrastrukturer er IKT og datainfrastruktur kortlevd, og må moderniseres etter ca. 3 år. Dette kan utfordre infrastrukturordningen ved Forskningsrådet siden behovet for modernisering av IKT og datainfrastruktur også er i utvikling. En nasjonal koordinering av IKT og datainfrastruktur med SIGMA2 støttes.

OsloMets oppfatning er at UH-sektoren strever med å få godt grep om FAIR-prinsippene og hvordan disse skal etterleves på en god måte. OsloMet vil anbefale en nasjonal tilnærming for arbeidsdeling, prioritering og operasjonalisering. Dette bør skje i tett kontakt med forskningsmiljøene og gjerne med utgangspunkt i de allerede finansierte nasjonale forskningsinfrastrukturene.

1. Forskningens behov for forskningsinfrastruktur for å løse utfordringer innenfor et strategisk prioritert område

Generelt ser UiT behovet for mer **langsiktig finansiering** av støttetjenester for håndtering av alle typer forskningsdata gjennom hele livsløpet.

Det nåværende vegkartet har et sterkt fokus på at infrastrukturer av nasjonal interesse støtter opp om fremragende forskning og svært datadrevne fag. Men hva med annen type forskning? Selv om det er gode argumenter for at ansvaret for å ta vare på de mange små og mellomstore datasett som genereres i all forskning skal ligge på institusjonene, er det behov for å intensivere samarbeid om etablering og videreutvikling av infrastrukturer for **data fra den lange halen**, i tråd med FAIR-prinsippene og andre krav og anbefalinger.

UiT ønsker spesielt å peke på viktigheten av at **urfolksperspektiver og -rettigheter** ivaretas på en god måte ved utvikling og drift av infrastrukturer, for eks. ved å implementere *The CARE Principles for Indigenous Data Governance*.

2. Hvilke eksisterende nasjonale infrastrukturer er av stor verdi å opprettholde og videreutvikle

DataverseNO har på under fem år utviklet seg til å bli et av de mest brukte nasjonale arkivene for forskningsdata fra forskere som ikke har tilgang til bedre egnede, fagspesifikke tjenester. DataverseNO er blitt utviklet av UiT og driftes gjennom det ordinære budsjettet til UiT og partnerinstitusjonene. For å holde tritt med raskt utviklende krav og anbefalinger vil det være nødvendig med nasjonalt bidrag til oppgradering og videreutvikling av infrastrukturen.

Sigma2s **HPC-infrastruktur** er en tjeneste som er viktig å videreføre for å tilfredsstille behovet innen beregningsorientert forskning. Man kan se for seg at man dekker de største nasjonale regnebehovene ved å inngå (kjøpe seg inn) i europeiske konsortier som f.eks. LUMI. Det vil minske behovet for innkjøp av HW-infrastruktur nasjonalt.

Datasett innen alle disipliner blir stadig større. Det er derfor essensielt at det eksisterer og videreutvikles et tilbud for

storskala lagring, arkivering og tilgjengeliggjøring av store datasett; jf. Sigma2 NIRD.

Det er pr. i dag få nasjonale tjenester for lagring og deling av **strengt fortrolige data**. De få tjenestene som finnes, bør derfor videreføres og videreutvikles.

3. Innenfor hvilke områder blir det spesielt viktig å etablere nye nasjonale infrastrukturer eller samarbeide om internasjonalt

UiT ser behovet for en **nasjonal skyinfrastruktur for forskningsdata**. En slik infrastruktur bør

- legge til rette for lagring, arkivering og deling av forskningsdata på tvers av institusjoner og også med internasjonale partnere;
- kunne brukes til ikke bare fillagring, men også objektlagring for å drifte skybaserte tjenester/verktøy som dekker behovet gjennom hele livsløpet til forskningsdata, fra planlegging til publisering, og dermed kunne integreres med slike tjenester;
- være EOSC-kompatibelt;
- kunne brukes for data med ulike nivå av sensitivitet;
- ha gode backup-løsninger og ellers tilby funksjonalitet på linje med kommersielle skytjenester;
- gjøre det mulig å flytte/endre status på data/filer fra aktiv til arkivert fase så enkelt som mulig.

AI/ML er et stort forskningsfelt som har store udekkede behov for regnekraft. Etablering av en nasjonal infrastruktur for AI/ML vil kunne bidra stort for at flere fagfeltet vil kunne gjøre ny typer forskning, eksempelvis basert på stordata.

4. Hvilke hull kan dekkes gjennom utvikling av eksisterende nasjonale og /eller samarbeid om/tilgang til internasjonale forskningsinfrastrukturer

Behandling og deling av **fortrolige data** er kostnadskreven, i dagens samfunn blir mediedata (video, bilde og lyd) stadig viktigere. Det bør derfor (videre)utvikles infrastruktur tjenester som kan ivareta slike data på en god måte. UiT ser også et behov for tjenester som tilbyr flere typer sikkerhetstiltak for deling av fortrolige data enn kontrollert tilgang, f.eks. Differential Privacy. Utvikling og drift av verktøy og tilknyttede støttetjenester for utarbeiding og oppfølging av **datahåndteringsplaner** bør i større grad koordineres nasjonalt og bør skje etter de føringer vi foreslår nedenfor under «Andre innspill».

UiT ser også behovet for en nasjonal instans av **Open Science Framework (OSF)**, som er en plattform bestående av tjenester for prosjektsamarbeid og for å gjøre forskningsprosesser og -resultat mer åpne og transparente. OSF er mye brukt på en rekke fagområder, men for å tilfredsstille europeiske og nasjonale krav til håndtering av forskningsdata, bør vi få på plass et samarbeid om en nasjonal instans av OSF.

5. Andre innspill

NFR bør i større grad legge **føringer** på hvordan forskningsinfrastrukturer de finansierer skal **utvikles, driftes og styres**. UiT ser det som spesielt viktig at man tar i bruk internasjonale standarder og beste-praksis-anbefalinger og styrker arbeidet med å gjøre infrastrukturene åpne. Vi viser til internasjonale initiativ og rammeverk på dette området, som f.eks. *Invest in Open Infrastructure*, *Global Sustainability Coalition for Open Science Services* og *TRUST Principles for digital repositories*.

Hovedprinsippet for nasjonale forskningsinfrastrukturer bør være at de er heftet på **internasjonalt samarbeid**. Det vil kunne friggi nasjonale midler og ressurser til å støtte opp om

- nasjonale tilpassinger av internasjonale infrastrukturer der det er nødvendig;
- lokale, forskernære støttetjenester hos institusjonene;
- kompetanseløft, særlig på juridiske og forskningsetiske aspekt ved forskningsdatahåndtering;
- nettverk og andre ressurser som støtter opp om nasjonalt samarbeid;
- klare nasjonale retningslinjer og praktiske guider og maler som hjelper forskere og støttepersonell med å navigere i det juridiske og forskningsetiske landskapet og handtere data i tråd med regelverk og anbefalinger og gjøre dem så åpne som mulig;
- verktøy som gjør det enkelt for forskere å velge tjenester/verktøy og metoder som er egnede for data av ulik sensitivitet;
- nasjonal skyinfrastruktur for forskingsdata.

Utvikling og drift av forskningsinfrastruktur krever **kompetanse**. Her snakker vi om IKT-kompetanse som ikke bare går på

maskinvare, men også på (vitenskapelig) programvare, fag-/forskningskompetanse og kompetanse på datahåndtering. Kompetanse bør være en del av infrastrukturbegrepet som er brukt i vegkartet. Dette aspektet er underkommunisert i nåværende vegkart. UiT etterlyser økt fokus på opplæring av forskere og støttepersonell til å kunne og ta i bruk og nyttiggjøre seg ny IKT-teknologi. Etablering av en kunnskapsinfrastruktur basert på Research Software Engineering vil kunne øke kunnskapsproduksjonen ved at forskere får hjelp til utvikling og effektivisering av vitenskapelig programvare de bruker, samt å gjøre kode mer reproduserbar og gjenbrukbar.

Vern av integritet er essensielt for forskningsdata.

Cybersikkerhet er derfor et krav som må gjelde drift og utvikling av all digital infrastruktur.

Universitetet i Agder

Evalueringen av INFRASTRUKTUR programmet viser at den nasjonale e-infrastrukturen (Sigma2) er godt anerkjent hos de større forskningsinstitusjonene når det gjelder HPC og tjenester for forskningsdata. Men, som rapporten også peker på, er det krevende å finansiere dette med en stadig større andel av INFRASTRUKTUR-programmets midler som konkurranseutsatt finansiering. Denne type infrastruktur har behov for regelmessige oppgraderinger (det søkes om relativt store beløp ved hver utlysning), og sammen med generelt økende behov i forskningen, kan en anbefaling være å gjøre en større andel av dette sentralt finansiert (utenom INFRASTRUKTUR programmet/direkte over statsbudsjettet).

Frem til nå er Sigma2 et resultat av samarbeidet mellom BOTT og NFR på området. Selv om BOTT fremdeles står for den desidert største bruken er det i økende grad andre forskningsinstitusjoner som benytter den nasjonale e-infrastrukturen. Det medfører at BOTT (som institusjoner) er alene om å levere tjenester til og å tilføre samarbeidet finansiering (i tillegg til NFR), og gir dem i tillegg innflytelse på styring og prioriteringer (Sigma2 styre og samarbeidsutvalg). Et nærliggende spørsmål er å se på om denne organiseringen gir øvrige forskningsinstitusjoner best utnyttelse av (og bidrag til) den nasjonale e-infrastrukturen. Dersom denne infrastrukturen skal finansieres utenfor INFRASTRUKTUR-programmet/direkte over statsbudsjettet, må det ses på endringer i organisering som gir flere institusjoner eierskap til infrastrukturen.

UiA og vårt forskningssenter CAIR er en viktig partner sammen med Sigma2 i prosjektet Norwegian AI Cloud (NAIC), hvor finansiering fra INFRA-programmet understøtter en satsing for etablering av nasjonal e-infrastruktur for AI/ML (inkl. kompetanse og brukerstøtte), som vil bli en del av den nasjonale e-infrastrukturen og tjenestetilbudet koordinert av Sigma2. Dette konsortiet består av flere enn BOTT (UiO, UiB, UiT og NTNU). Også dette reflekterer et behov for å se nærmere på organiseringen av Sigma2.

Innledning

- UiB mener at forskningsinfrastruktur av høy kvalitet i bredden av fag er helt avgjørende for god forskning. Det er derfor viktig å beholde en ambisiøs opptrappingsplan for forskningsinfrastruktur.
- UiB mener at store, permanente satsinger, som Sigma2 og Helseanalyseplattformen, bør finansieres løpende over statsbudsjettet. Det lite formålstjenlig at investering og drift av slike infrastrukturer er konkurranseutsatt på samme måte som prosjekter som anvender denne infrastrukturen. Det bør løpende vurderes om andre infrastrukturer som også betraktes som nasjonale og almene bør løftes ut av konkurransearenaen og at fagdepartementene også tar ansvar for slik nasjonal permanent forskningsinfrastruktur.
- Infrastruktur som tilrettelegger for bruk av kunstig intelligens innen hele bredden av akademiske disipliner vil være avgjørende for at Norge skal kunne hevde seg i forskningsfronten innen hele bredden av fag og the Norwegian Artificial Intelligence Cloud (NAIC) er en god modell som bør søkes å gjøre permanent.
- Nasjonalt er det viktig å ha et økosystem av tjenlige sikre forskningsservere med tilstrekkelig kapasitet. UiB bidrar her med Sikker Adgang til Forskningsdata og E-infrastruktur (SAFE) som har sin styrke i sikker adgang til forskningsdata og utgjør en nasjonal e-infrastruktur.

E-infrastrukturen Sigma2 som brukes for tungregning og lagring av store data, er en viktig fasilitet for tunge analyser og lagring av blant annet helse, biodiversitetsdata og klimadata. Det er viktig med en bred tilgang til tungregningsfasiliteter for hele bredden av fag som muliggjør bruk av kunstig intelligens. Samtidig er det viktig at disse i større grad igjen blir rene infrastrukturleverandører og ikke bygger opp organisasjoner som blir selvstendige og til dels konkurrerende projektaktører.

Innen flere områder av e-Infrastruktur er ideen bak digitale økosystemer en modell som i større grad bør utvikles. Her vektlegges samspill med flere uavhengige men samspillende parter som til sammen skaper sammenhengende verdikjeder for forskning. Samspill legger til rette for innovasjon, reduserer risiko og øker dynamikken over tid. Økonomiske modeller for slike må utvikles.

Helseanalyseplattformen (HAP) og Helsedataservice (HDS) ble etablert for å møte et stort og udekket behov for behandling av data fra helseregistre, befolkningsbaserte helseundersøkelser og biobanker. Det er avgjørende at arbeidet med HAP og HDS fullføres med fleksible og dynamiske løsninger som møter forskningens behov. For at Norge skal kunne utnytte de unike dataressursene vi har for helseforskning og innovasjon, er det avgjørende at kompetanse og kapasitet innen biostatistikk og bioinformatikk kommer på plass. Uten snarlig tilgang vil

forskere i Norge mister lederrollen i det internasjonale samarbeidet. Samtidig bør samarbeidet mellom Biobank Norge (se UiBs innspill på Helse), Elixir Norge og HAP, samt -omics data generert gjennom den samlede agendaen for persontilpasset medisin (se UiBs innspill på Helse), videreutvikles for å styrke utnyttelsen av norske helsedata for samfunnets beste. Økosystemet med sikre forskningsservere ved universitetene må også utvikles som en del av HAP.

Kunstnerisk utviklingsarbeid

Det er stort behov for en tydelig satsing på (inter)nasjonal forskningsinfrastruktur innen kunstnerisk utviklingsarbeid (KU).

Det er et økende behov for 1) å etablere og utvikle digitale og fysiske plattformer som støtter varig lagring av kunstneriske media og oppfyller krav til både åpen forskning og lagring av sensitive data og 2) å bygge ut nettverk for deling og interaksjon som muliggjør samarbeid i kunstnerisk utviklingsarbeid over større avstander og på tvers av landegrenser. Dette er ikke mulig med dagens infrastruktur. Etablering av en unik nasjonal forskningsinfrastruktur vil tilrettelegge for mer samarbeid mellom nasjonale institusjoner, slik at disse samlet sett styrkes og vil kunne gjøre seg mer gjeldende i europeisk sammenheng. En slik infrastruktur tilrettelegger for digitale og hybride løsninger, som er i tråd med bærekraftsmål (miljøbesparende; mindre transport og mobilitet), samtidig som det åpner for innovativ kunstnerisk forskning og undervisning. Deling og samarbeid gjennom høyhastighetsnettverk bidrar også til økt internasjonalisering av KU-feltet, legger til rette for økt tverrfaglighet og samfunnsrelevans. Videre vil utvikling av en robust digital infrastruktur for KU gjøre de institusjoner som baserer sin virksomhet på utøvende og skapende kunstnerisk praksis bedre rustet til å 1) søke innovasjonsrettede program og 2) initiere samarbeid med kultur- og teknologisektoren og 3) bidra til å løfte de kreative næringene og kunnskapsklyngene på feltet. En nasjonal infrastruktur kan utvikles i etapper og eksempelvis bygge videre på etableringen av ARIIS (Artistic Research and Innovation Infrastructure ved UiB). Det vil åpne opp og samordne kostbare enkeltelementer i en stor slagkraftig nasjonal forskningsinfrastruktur som kan løfte KU-feltet og som ikke har noe motstykke i Europa.

Universitetet i Oslo

Forskningen begynner å se veksten i den lange halen mtp hvem som vil bruke mer data og mer regnekraft. Dette medfører ikke at vi får mange flere og totalt sett mye mer bruk av tungregning, men det betyr at Norge får langt flere brukere med forholdsvis små ressursbehov og store forskjeller i behov. Dette krever fagnær støtte, kompetanse og enkel tilgang til ressurser. I dag preges de nasjonale HPC-leveransene av de største brukernes behov, outsourcing får å utnytte muligheter for lavere kostnader og vanskelig tilgjengelige ressurser (blant annet kompliserte søknadsprosesser). Det er for liten oppmerksomhet

knyttet til nye brukermasser og nye brukerbehov og bruksmønstre. De største brukerne kan gjerne outsources. De nye brukerne med mindre ressursbehov og større variasjon i bruksmønstre og verktøy må vies langt større oppmerksomhet. For disse må tjenestene produseres og leveres så nært brukeren som mulig for å kunne fange opp endringer i behov, støtte nye, mer interaktivt orienterte ressurser og for å senke terskelen for å ta tjenestene og ressursene i bruk. Og ikke minst, – dette er brukergrupper som har stort behov for nærhet til kompetanse og både generell og avansert brukerstøtte.

For data og lagring er verktøy for å enkelt gjøre data FAIR (metadata ++) og kobling av data + person + publikasjon + organisasjon + metadata av stor viktighet. Ruud-rapporten påpeker dette som en vesentlig utfordring fremover, og at 'one-size-does-not-fit-all'. Anbefalingen er gjøre dette så nært fagmiljøene som mulig og at Norge ved NFR følger opp de initiativene som har suksess og bidrar til å finansiere disse som en slags 'venture-kapitalist'. Dette for å la mange blomster blomstre, men dyrke de beste. Ikke satse på et fåtall hester, som fort kan feile. Dette er vanskelige oppgaver der det bør testes og evalueres før det investeres stort. Det er også lite trolig at dette kan løses langt fra forsker/institusjon og lite trolig at man vil ha en superinfrastruktur, noe som er i tråd med rapportens vurderinger og anbefalinger.

Hvilke eksisterende nasjonale forskningsinfrastrukturer er av stor verdi å opprettholde og videreutvikle

Kjernefasilitetfunksjonalitetene er gode og ser ut til å virke godt.

Sigma2 leverer viktige tjenester, men finansieringsnøkkel og fordeling må revurderes totalt for å muliggjøre framtidig bærekraft. I dag får BOTT svært lite igjen av midlene NFR deler ut til HPC/Lagring over eInfra og dette skaper urettferdighet og konkurransefortrinn. Tjenestetilbudet må opprettholdes, men leveransemetodikken og organisasjonen må endres radikalt for å sikre videre liv.

Tjenester for sensitive data (TSD) fungerer som en nasjonal infrastruktur både i form av samarbeidet med Sigma2 om HPC og lagring (se pkt over), men også de-facto med 7600 brukere, 1760 prosjekt og 72 organisasjoner representert i brukermassen. Infrastrukturer som TSD, SAFE og HUNT bør vurderes av NFR når nye ting skal på plass. I stedet for at disse må søke finansiering for å gjøre ting, kan disse i stedet bli bedt om det. Da unngår en også at det dukker opp mer døgnfluepregede løsninger for enkeltmiljø løsninger som i mange tilfeller ikke er gjenbrukbare. Her kan også metoden med 'venture-kapital' fra NFR for å styrke gode initiativ som allerede har bevist bærekraften. Det skal ikke være for mange slike infrastrukturer og de som er etablert og fungerer godt bør utnyttes bedre og dyrkes. Eksempelvis bør arbeidet med å gjøre sensitive data FAIR legges til disse infrastrukturene.

Educloud – UiOs mindre sikre versjon av TSD – er på vei til å bli en defacto nasjonal ressurs med over 100 prosjekt etter 6 måneders drift. Samarbeid på tvers av institusjons- og landegrenser er en av hovedideene bak Educloud, og den bør vurderes for rettede satsninger fra Forskningsrådet som diskutert over.

Hvilke tematiske områder det blir særlig viktig å etablere nye nasjonale forskningsinfrastrukturer på og/eller samarbeide om internasjonale forskningsinfrastrukturer

For alle initiativ som handler om data vil samarbeid og integrasjoner med internasjonale tiltak/initiativ som EOSC, European Health Data Space med mer være av stor relevans. Det er lite trolig at man bør ha et nasjonalt mellomnivå, men heller satse på at data og/eller metadata leveres direkte til de internasjonale overbyggene. Her er samarbeid essensielt. Det gjelder for standarder (men ikke la oss bli stoppet av å vente på standarder), det gjelder datatransport og det gjelder lisenser og tilganger.

Videre er det en del ting som kanskje ikke burde gjøres innen forskning og IT:

- 1: NVA bør ikke lages for å holde faktiske forskningsdata, dette burde leveres av andre arkiv som NVA kan peke til
- 2: Sigma2 NIRD-lagring bør ha tydelig policy for hvor lenge data kan ligge der uten at det får metadata og legges i NIRD-arkivet etter at data har blitt 'kald'.
- 3: Sigma2 NIRD-arkiv bør samordnes med Dataverse, og det er ikke sikkert Sigma2 burde levere et arkiv (det krever arkiv -og bibliotekskompetanse), men kanskje heller levere lagring til ulike arkivløsninger (som f.eks dataverse.no). En tung satsning på at Sigma2 skal lage arkiv (ved å tilpasse en eller annen Open Source-løsning) vil også skape et 'scope-creep' for kjernevirksomheten i Sigma2.

Hvilke hull som kan dekkes gjennom utvikling av eksisterende nasjonale og/eller samarbeid om eller tilgang til internasjonale forskningsinfrastrukturer

FAIR@UIO-prosjektet er et prosjekt som nå er i utprøvningsfasen av verktøy for å gjøre data FAIR, dette prosjektet bør monitoreres og vurderes for midler til å muliggjøre bruk av tjenestene nasjonalt.

Hullene på infrastrukturensiden når det gjelder **kunstig intelligens (AI)** vil langt på vei dekkes av NAIC-prosjektet når dette kommer i gang. Dette vil være en viktig satsning der data og tilgangen til data vil være av stor viktighet.

Når det gjelder **Helseanalyseplattformen (HAP)** er det ekstremt viktig at de er et tydelig scope og mål for alternativprosessen for denne. HAP i distribuert løsning må framover prioritere å løse levering av skarpe data raskest mulig til forsker. Det må være en edruelig pengebruk, ikke minst for å

sikre at man jobber i raske iterasjoner i motsetning til det originale prosjektet. Microdata sammen med helsedata må komme i andre rekke.

CRIS/NVA må ferdigstilles og CRIS systemet må være globalt tilgjengelig via API'er.

Innspill fra tematisk forskningsinfrastrukturområde ved UiO - MNT

1) forskningens behov for forskningsinfrastruktur for å løse utfordringer innenfor et strategisk prioritert område

Astrofysikkforskere er storbrukere av infrastruktur i tungregning og datalagring for analyse av data innhentet fra satellitter og fra internasjonale observatorier og for numeriske simuleringer av sola, galakser og storskala kosmologiske strukturer. Forskerne bruker tungregningsinfrastruktur på nasjonalt nivå gjennom Sigma2 og på europeisk nivå gjennom PRACE (Partnership for Advanced Computing in Europe) og EuroHPC. Innen kjemi vil det satses videre på UiOs internasjonalt ledende fagmiljø innen teoretisk kjemi etter utløpet av de to påfølgende SFF-sentrene, og støtte opp kjernevirksomheten innen metodeutvikling og anvendelsen av slike verktøy for nøyaktig modellering av kjemiske strukturer og prosesser. Tilgangen til nødvendig infrastruktur i form av regnekraft (i mindre grad lagring) er i dag tilfredsstillende gjennom Sigma2-konsortiet. Behovet for regnekraft er imidlertid sterkt økende ved at kjemikere på snart alle områder benytter kvantekjemiske metoder for modellering til å understøtte fortolkningen av eksperimentelle data, noe som i tiltakende grad kreves for internasjonal publisering. Av denne grunn er det presserende med hyppig utskifting av infrastrukturen for maksimal ytelse. Dette kan bare skje gjennom vedvarende og forutsigbar satsing på Sigma2 og støtte til bruk av de mest avanserte regnearbeidene utenfor landets grenser.

2) hvilke eksisterende nasjonale forskningsinfrastrukturer dere mener er av stor verdi å opprettholde og videreutvikle

De nasjonale tungregningsinfrastrukturene i Sigma2 er helt essensielle for alle forskningsgrupper innen astrofysikk

3) hvilke tematiske områder det blir spesielt viktig å etablere nye nasjonale forskningsinfrastrukturer på og/eller samarbeide om internasjonale forskningsinfrastrukturer

Det er viktig for astrofysisk forskning at Sigma2 fortsetter å utbygge og oppgradere eksisterende tungregningsinfrastruktur og datalagring i Norge, samt deltar i PRACE / EuroHPC på europeisk nivå.

Storskala dataanalyse og såkalt maskinlæring holder på å bli sentrale metoder i mange disipliner. Tungregning og 'scientific computing' er dermed relevante for flere brukergrupper. Dette krever fokus på tilpassete nasjonale beregningsressurser, særlig videre økning av GPU-kapasitet. Sigma2-samarbeidet er svært verdifullt og forutsetter løpende tilpasning og fornyelse av tilgjengelige lagrings- og beregningsressurser. Per i dag mangler

det et 'trappetrinn' mellom de svært begrensede nasjonale GPU-systemene og EuroHPC-anlegget LUMI. Videre mangler det tilstrekkelig GPU-kapasitet tilkoblet omgivelser for behandling av sensitive eller private data.

4) hvilke hull som kan dekkes gjennom utvikling av eksisterende nasjonale og/eller samarbeid om eller tilgang til internasjonale forskningsinfrastrukturer

Fra Medisin

De strategisk prioriterte områdene «persontilpasset medisin» og «helsenæringen» krever investeringer innen beregningsorientert IKT, muligheter for å overføre store mengder data og lagringskapasitet. Det er av stor verdi å opprettholde og videreutvikle ELIXIR og SIGMA2. Blant områdene som er viktige å utvikle er storskala beregnings- og lagringsmuligheter for modelleringer av både humane data og eksperimentelle data fra dyreforsøk. Vi må delta aktivt i European Open Science Cloud (EOSC) og andre løsninger for sensitive data.

Norge (v/ Helse- og omsorgsdepartementet) har signert en deklarasjon for å arbeide mot **utveksling av genominformasjon på tvers av Europa (1+MG-initiativet)**. Formålet er å utvikle mekanismer for internasjonalt samarbeid og dataforvaltning som støtter et felles mål: gjøre en million genomer tilgjengelig i EU innen 2022, utnytte investeringer som allerede er gjort på nasjonalt og europeisk nivå, spesielt innen sekvensering, biobanking og datainfrastruktur, samt å oppnå en kohort i tilstrekkelig stor skala for å understøtte ny biomedisinsk og klinisk forskning og helsehjelp. For å oppfylle forpliktelsene må det sikres at de norske miljøene settes i stand til å levere data, ved at etiske, rettslige, samfunnsmessige og infrastrukturmessige hensyn er ivaretatt. Eksisterende datakilder bør velges der det er mulig, men det må sikres en prosess for generering av supplerende data for å sikre nasjonal representativitet på befolkningsnivå, mellom ulike diagnosegrupper og kliniske datasett. Dette vil gjøre at vi bedre kan knytte oss opp mot europeisk infrastruktur, slik som eksisterende samarbeid på genetisk epidemiologi, utvide kohortgrunnlaget for klinisk- og translasjonsforskning, bl.a. innen presisjonsmedisin. Initiativet må sees i sammenheng med arbeidet med forslaget om European Health Data Space (EHDS). Innen mikroskopi og annen avbildningsteknologi er det et raskt økende behov for kvantitative analyser av kvantitative billeddata. I dag gjøres slike analyser hos den enkelte infrastruktur, men i fremtiden vil dette kreve for store ressurser og bør derfor legges til en nasjonal infrastruktur.

Det er nødvendig med bedre tjenester for sensitive data. Skal vi få til presisjonsmedisin i Norge, må det finnes IKT løsninger for innsamling og analyse av hjerneavbildningsdata, biobankdata og helseregisterdata. I dag er forsknings IKT løsninger i hovedsak fokusert på ikke-sensitive data.

TSD (Tjenester for Sensitive Data) på UiO er en allerede etablert infrastruktur for sikker lagring og analyse. Det bør bevilges midler til videreutvikling av denne tjenesten, med

integreerte databaseløsninger og enkel kobling mot helseregistre og journalsystemer.

Innen helseforskning på komplekse lidelser er det gjort store framskritt pga internasjonalt samarbeid, spesielt i Norden.

Tryggve, under NeIC, er en sikker tjeneste for forskningssamarbeid i Norden, rettet mot sensitive data innen helseforskning, som bør utvides og gi mulighet til økt kapasitet og bedre service.

Et viktig tema for enklere datadeling nasjonalt og internasjonalt er interoperabilitet. Det bør settes av ressurser til etablering av felles normer og dataharmonisering.

PraksisNett er en eksisterende infrastruktur som legger til rette for at forskere skal kunne gjennomføre gode og kvalitetssikrede kliniske studier i norsk allmennpraksis. Satsningen på kliniske studier er sentralt i Avdeling for allmennmedisins strategiske plan. For et medisinsk fakultet vil det alltid ha høy prioritet å styrke klinisk forskning, uavhengig av hvor i helsetjenesten pasientene befinner seg.

PraksisNett er en forskningsinfrastruktur som til nå har vært finansiert av Norges forskningsråd (finansiering fra 2019 til og med 2022) og partnerne i prosjektet er:

Universitetet i Bergen (UiB), UiT - Norges arktiske universitet, Norges teknisk-naturvitenskapelige universitet (NTNU), NORCE - Norwegian Research Centre AS, Universitetet i Oslo (UiO) og Nasjonalt senter for e-helseforskning (NSE) ved Universitetssykehuset Nord-Norge (UNN).

PraksisNett ledes fra Universitetet i Bergen, ved Professor Guri Rørtveit.

PraksisNett består av en tjenestebasert del og en digital del. Det tjenestebaserte nettverket i PraksisNett består av rundt 90 fastlegepraksiser med nesten 500 fastleger som har ansvar for rundt 520 000 pasienter. I tillegg inngår administrativ støtte lokalisert ved de fire universitetene. Den digitale delen består av en liten datamaskin (Snow Health Appliance Box) som installeres lokalt på fastlegens kontor og som automatisk trekker ut data fra pasientjournalen basert på en avtalt variabel-liste. Alle data som blir lagret på Snow-boksen lokalt er pseudonymisert. Aggregerte data blir overført til PraksisNett og inngår i en database som vil beskrive praksis- og pasientgrunnlaget i nettverket. PraksisNett kan benyttes for innhenting av data fra primærhelsetjenesten i form av datauttrekk, rekruttering av pasienter til ulike eksterne studier og kliniske intervensjonsstudier hos fastlegene i PraksisNett. PraksisNett avgjør hvilke prosjekter som kan bruke infrastrukturen basert på kriterier som gjennomførbarhet, etiske vurderinger og relevans (www.praksisnett.no).

PraksisNett Øst er den største regionale enheten i PraksisNett, og administreres fra Allmennmedisinsk forskningsenhet (AFE) ved Avdeling for allmennmedisin, Institutt for helse og samfunn.

Prosjektansatte i PraksisNett Øst deltar lokalt i infrastrukturensamarbeid med Regional forskningsstøtte for Helse-Sørøst ved Oslo universitetssykehus (OUS), representert ved Clinical trials unit, CTU. PraksisNett har også et samarbeid med

NorCRIN (infrastruktur for kliniske studier i sykehus), lokalt ved OUS, omkring samkjøring av kvalitetssystemene for kliniske studier.

'omics data

Forskning innen livsvitenskap baserer seg i stadig større grad på storskala 'omics data med økende omfang og kompleksitet, innen medisin og helse så vel som akvakultur og klima og miljø. For å kunne utnytte disse datamengdene best mulig, i tråd med nasjonale og internasjonale policyer for åpen forskning, trengs robust datainfrastruktur som legger til rette for kapasitetskrevene beregninger og harmonisering med internasjonale standarder, samt løsninger som utvikles for EOSC. Flere utredninger har i løpet av den siste tiden pekt på datahåndtering og -stewardship som en viktig flaskehals i å oppnå FAIR forskningsdata, og det er derfor svært viktig at det fortsatt investeres i pågående initiativ som understøtter dette. ESFRI-landemerket ELIXIR er en viktig aktør i denne sammenhengen, og bygger pan-europeisk e-infrastruktur for analyse, håndtering og publisering av livsvitenskapelige data innen de fleste domener, som mat og helse, biomangfold, klima og miljø, tilpasset til norske data og -brukere gjennom den norske noden i nettverket. Løsninger for FAIR arkivering av sensitive humane data inngår i dette. Det er derfor av stor verdi for norsk livsvitenskapelig forskning å opprettholde og videreutvikle ELIXIR Norge.

Universitetet i Sørøst-Norge / på veien av
Viserektor for
forskning, innovasjon
og internasjonalisering

Tingenes internett utgjør allerede flere milliarder enheter som er koblet til nettet og er sterkt voksende på grunn av de store fordelene ved å gjøre transport, produksjon og nesten all drift smartere. Mikro- og nanofabrikasjon muliggjør små og, ved store produksjonstall, rimelige energikilder, energilager, sensorer og aktuatorer for dette formålet. Videre er det mikroteknologiens metoder for sammenstilling og innkapsling som sørger for at systemene i seg selv kan bygges små, og ikke bare komponentene. Betydningen av dette siste understrekes av at European Chips Act har som et av delmålene å bygge og forsterke evnene til design, fremstillinger og *pakking* av avanserte brikker. For å kunne forske på denne type teknologi er det nødvendig med tilgjengelig renromsfasiliteter for mikro- og nanofabrikasjon av komponenter og tilhørende innkapsling og sammenstilling. Norfab er en nasjonal forskningsinfrastruktur for mikro- og nanofabrikasjon og består av MST-Lab i Horten, UiO og Sintef MiNaLab i Oslo og NTNU Nanolab i Trondheim. Den tilbyr *state-of-the-art* laboratorier for norske forskere uavhengig av institusjons- eller firmatilhørighet. Som sentral infrastruktur med høy kapasitetsutnyttelse sikrer NorFab open access, forutsigbarhet, konkurransedyktige priser og samlet fagkompetanse for både akademisk forskning, innovasjon og kommersiell aktivitet. Det er viktig at Norfab opprettholdes og videreutvikles som sentral infrastruktur for å støtte både akademisk forskning, innovasjon og kommersiell aktivitet innen området. Det er nødvendig med kontinuerlig oppgradering for å opprettholde kapasiteten etter

hvert som maskinparken eldes og nyinvesteringer for å holde tritt med utviklingen og skape nye muligheter innen området. Senter for helse og teknologi ved USN ble etablert i 2011 og var tidlig etablert i landskapet for forskning på eHelse, eller det som tidligere ble kalt Velferdsteknologi. Fokus i satsingen er utfordringene helsesektoren i hele den vestlige verden står overfor når det gjelder økt antall hjelpetrengende (både eldre og andre), mangel på helsepersonell og synkende skatteinntekter. Disse utfordringene gjør at dagens modell for helsetjenester ikke er bærekraftig, og eHelse kan i noen grad bidra til å møte denne utfordringen. I forbindelse med samhandlingsreformen har oppgavene i både spesialisthelsetjenesten og kommunehelsetjenesten vært i rask endring og behovet for forskning som kan komme sektoren til gode har økt tilsvarende. Tilsvarende initiativ har vært etablert ved en rekke forsknings- og utdanningsinstitusjoner i Norge, og Senter for helse og teknologi samarbeider med en rekke av dem. Imidlertid er det behov for en samordning av forskningskrefter for å gjøre forskningen mer robust, noe rapporter fra for eksempel KS og Menon Economics dokumenterer. Det vil styrke forskningen om fagområdet E-helse får en egen områdestrategi som en del av veikart for forskningsinfrastruktur.

Det er en utfordring at forskningen på e-helse framstår som fragmentert og småskala. Dette er delvis en speiling av fragmentert praksis, men synliggjør samtidig behovet for tettere samordning og for møteplasser. Som en respons på disse utfordringene ønsker Senter for helse og teknologi å etablere en plattform for å forene forskningsinitiativ og samarbeid mellom eksisterende infrastruktur, myndigheter, forskningsinstitusjoner, kommuner/KS, ansatte i helsesektoren og helseteknologi industrien. Eksisterende infrastruktur er i høy grad fokusert på helsedata (for eksempel Helsedataprogrammet og Biobank Norge), mens det er behov for nyetablering som adresserer for eksempel teknologi og teknologiimplementering som bidrag til en fremtidig bærekraftig helsetjeneste.

Internasjonalt er dette et tema som er høyt oppe på forskningsdagsorden, for eksempel i England. Imidlertid ligger de nordiske land langt fremme både når det gjelder implementering av eHelse i praksis og når det gjelder forskningsinnsats. Her har Norge mulighet til å ta en ledende rolle og Senter for helse og teknologi har utstrakt samarbeid med aktører i de skandinaviske landene, spesielt i Sverige. Utvikling av forskning rundt eHelse vil ikke bare styrke kunnskapsgrunnet for offentlig sektor, men vil også bidra til å styrke norsk næringsliv, som om talt i Stortingsmelding 18 (2019). Breivik, E., et al.

"Kunnskapsoppsummering om effekter av trygghets- og mestrings-teknologi." Tromsø: Nasjonalt senter for e-helseforskning (2021): 2021-02.; Jacobsen, E.W. et al. «Strategier for økt produksjon og eksport av norsk helseindustri». Menon publikasjon nr 45/2021; Stortingsmelding18. (2019). Helsenæringen — Sammen om verdiskaping og bedre tjenester. Oslo: Nærings- og fiskeridepartementet

Fakultet / Institutt / Svar
Senter ved
universitet eller
høgskole (Offentlig)

Flere fakultet ved UiS: Tek.nat. fakultet, Det samfunnsvitenskapelige fakultet, Arkeologisk museum.

Forskningens behov for forskningsinfrastruktur for å løse utfordringer innenfor et strategisk prioritert område
Flere fagområder ved UiS har et økende behov for datakraft til å løse forskningsutfordringene. Det gjelder spesielt innen teknisk-naturvitenskapelige fag, men også innen andre områder. "Energi" og "helse" er to prioriterte fagområder i UiS sin strategi. Innen begge disse er det mye tverrfaglig samarbeid der kunstig intelligens, og da spesielt maskinlæring er et viktig verktøy. Stordatalagring og tungregnekraft (HPC) er essensielt i trening og kjøring av maskinlæringsalgoritmene. Til dette benyttes i første rekke GPU-kraft. Men, vi ser også fagområder som fremdeles er avhengige av CPU-kraft. Tilgang på e-infrastruktur er essensielt, både på lokalt, nasjonalt og internasjonalt nivå.

Hvilke eksisterende nasjonale infrastrukturer er av stor verdi å opprettholde og videreutvikle

Det er viktig å fortsette Sigma2, den nasjonale e-infrastrukturplattformen for tjenester innen tungregning. Parallelt med nasjonal infrastruktur er det også viktig å styrke den lokale/regionale infrastrukturen, som typisk benyttes mer i undervisningsformål og i mellom- og småskala forskningsprosjekter. Lokal/regional infrastruktur kan følgelig benyttes som pilotering i forkant av bruk av nasjonal/internasjonalt infrastruktur, hvilket kan effektivisere bruken av denne. Ressursene lokalt må tilpasses forventet behov.

Innenfor hvilke områder blir det spesielt viktig å etablere nye nasjonale infrastrukturer eller samarbeide om internasjonalt
Maskinlæring og dyplæring benyttes nå i mange ulike fagområder. For å få utført denne forskningen må kapasiteten av GPU-ressurser nasjonalt økes kraftig. De aller største prosjektene bør søke mot superdatamaskiner i Europa, som forskere i Norge har tilgang til. Man bør også styrke forskningen på selve utnyttelsen av datakraften, ut fra flere hensyn: Utnytte maskinvaren bedre, sørge for distribuerte løsninger og sørge for bærekraftige løsninger som også tar hensyn til miljøet.

Hvilke hull kan dekkes gjennom utvikling av eksisterende nasjonale og /eller samarbeid om/tilgang til internasjonale forskningsinfrastrukturer»

Det er viktig å sørge for godt vedlikehold av eksisterende e-infrastruktur, både lokalt, regionalt og nasjonalt. Man må sikre at ressursene blir benyttet på en god måte, noe som fordrer godt samarbeid mellom de som drifter datasentrene både nasjonalt og lokalt, og forskerne i de ulike fagmiljøer. Det er viktig at det tilbys opplæring for riktig og god bruk av tungregneressursene.

UiS ønsker i tillegg å formidle innspill fra Kognitiv lab ved universitetet:

Forskningsinfrastruktur ved UiS i et langtidsperspektiv, innspill fra Kognitiv Lab

ved faglig leder for Kognitiv Lab, Kolbjørn Kallesten Brønnick

Tverrfaglighet, et sentralt moment i nasjonale strategier og i UiS' strategi

Tverrfaglighet er et nasjonalt prioritert område (jfr stortingsmelding 4 (2018-2019), "Langtidsplan for forskning og høyere utdanning 2019–2028"). For å bidra til å løse store faglige- og samfunnsutfordringer, er tverrfaglig samarbeid viktig. Innen livsvitenskap er det for eksempel avgjørende viktig at forskere innen medisinsk informatikk, kunstig intelligens, biostatistikk, molekylærbiologi, genetikk osv, samarbeider med f.eks. leger som har oversikt over klinisk medisin. Mange ulike viktige utfordringer nevnes spesielt og ett eksempel er "det grønne skiftet", der klimaforskere, samfunnsforskere, atferdsforskere og livsvitere alle har viktige bidrag relatert til ulike aspekter av problemstillingen. Tverrfaglighet er ikke enkelt å få til i praksis og det trengs forskningsinfrastruktur som grunnlag for datainnsamlinger, datahåndtering, dataanalyser, studentveiledning og undervisning og sosiale møteplasser hvis en skal få dette til. Ved UiS har en opprettet Kognitiv Lab som et konkret tiltak for å møte disse utfordringene, et tverrfakultært initiativ og arena for tverrfaglig samarbeid innen kognisjons- og atferdsforskning. Gjennom Kognitiv lab får realfag og forskning på kognisjon og kunstig intelligens en naturlig møteplass med helsefag, psykologi og ulike fagområder som studerer samfunnsforhold, menneskelig atferd og læring.

Kognitiv Lab kombinerer atferds- eller andre eksperimenter med bruk av måleutstyr som kan måle hjernens struktur og funksjon (elektroencefalografi (EEG), magnetresonanstomografi (MRI og fMRI), fysiologiske uttrykk for emosjoner og stress (Electrodermal activity (EDA), øyebevegelser (eye-tracking), hjertets (elektrokardiografi (EKG)) eller musklens (elektromyografi (EMG)) elektriske aktivitet, søvnmønstre og andre kroppslige signaler. Signal- og bildebehandling spiller en viktig rolle i forskningen.

For å lykkes med å møte utfordringene innen tverrfaglighet, en tematikk som møter kriteriet om nasjonal viktighet ved at dette er et gjennomgående tema i stortingsmelding 4 (2018-2019), trenger Kognitiv Lab langsiktig opprusting og nyervervelser av utstyr som vil gjøre oss i stand til å møte strategien om tverrfaglighet innen følgende områder. Inndelingen er kunstig, nettopp fordi mange av elementene vil overlappe.

Livsvitenskaper og medisin

Kognitiv lab trenger utstyr for å innhente og håndtere biologisk materiale før det videresendes til biobanker eller analyselaboratorier, som:

- Fryserer som kan nå lave temperaturer som <-40.
- Utstyr for innhenting av blodprøver og spyttprøver.
- Utstyr for oppbevaring av hårprøver osv.

Medisin og psykologi

- Målinger av blodtrykk, ekg, elektroencefalografi med 64+ elektroder, elektromyografi, galvanisk hudrespons
- Portabel magnettomografi. Slikt utstyr representerer løpende kostnader gjennom vedlikehold og software-abonnement
- Tredemøller og målinger av fysiologiske parametre under fysisk anstrengelse
- Aktigrafer og andre wearables
- Applikasjoner for ambulatorisk måling av stemmebruk/vokalisering
- Applikasjoner for "experience sampling"
- Applikasjoner for eksperimentdesign, som e-prime (<https://pstnet.com/products/e-prime/>)
- Virtual reality, for psykometriske formål og for behandling av f.eks. angstlidelser og stressreduksjon

Humaniora og pedagogikk

- Sporing av øyebevegelser og pupilldiametre, for leseforskning og studier av sosiale samspill.
- Virtual reality utstyr, for f.eks. konstruksjon av kontrollerte læringsmiljøer.

Samfunnsvitenskap

- Laben har lokaler for studier av gruppeprosesser. Her trengs det portabelt utstyr for filming og lydopptak og for analyse av videomateriale
- For analyse av gruppeprosesser trengs også sensorer som trådløst monitorerer bevegelser og fysiologiske parametre, som hjerterate, øyebevegelser osv.

Utøvende kunstfag

Også her er utstyr nevnt ovenfor sentralt- øyebevegelser er blitt brukt til forskning på notelesning og er anvendbart også i studier av musikalsk samspill, teater og dans.

Naturvitenskap

- Signalbehandling er avgjørende viktig for analyse av sensordata og de andre typene data som Kognitiv Lab gir og det trengs infrastruktur for maskinlæring og dyplæring med mye datakraft
- Det samme gjelder biostatistikk, f.eks. knyttet til bildeanalyser, EEG-analyser og annet.

Til sammen gir Kognitiv lab muligheter for tverrfaglighet ved at en har et fysisk møtested, utstyr og kompetanse som skaper en arena der dette lar seg implementere. En vil trenge en langsiktig satsning, siden det må etableres en kultur og "literacy" for tverrfaglighet over flere år. Utstyr må fornyes, det må finansieres rekvisita og abonnementer, datahåndteringskostnader for sikker databehandling etc.

Havforsknings-
instituttet**Eksisterende verdt å opprettholde/videreutvikle:****NMDC 2**

Havforskningsinstituttet vil med dette bekrefte interesse for å videreføre og oppgradere forskningsinfrastruktur for marine data Norwegian Marine Data Centre NMDC gjennom prosjektet – **Norwegian Marine Data Centre - Interoperable Infrastructure NMDC-II**. Denne forskningsinfrastrukturen er på veikartet. Søknad om NMDC-II har vært innsendt to ganger tidligere uten å få tilslag. Vi mener den bør finansieres fordi den oppfyller de tre kriterier som kreves. Faglig sett er det flere viktige prosesser som er kritisk avhengig av oppdatert datainfrastruktur. Vi vil nevne fem viktige tema her. 1) Styrket forventning og om at forskningsdata skal være FAIR. 2) Bruk av marine data for å informere digitale tvillinger. 3) Assimilering av marine data i operasjonelle modeller for forskning, varsling og beredskap. 4) Økt deling av data fra havnæringene. 5) Økt behov og mulighet for tilgjengeliggjøring av rådata. **Dette fordrer at man tar i bruk nye teknologier og styrker arbeidet med forvaltning og tilgjengeliggjøring av marine data for forskning i Norge.**

Nasjonal viktighet: Norge finansierer og samler inn store mengder forskningsdata fra det marine og har store ambisjoner for videre utvikling av havnæringene med målsetning om størst mulig samlet verdiskapning innenfor bærekraftige rammer. Gjennom den nasjonale forskningsinfrastrukturen NMDC har en i en 10-års periode samlet og tilgjengeliggjort forskningsdata. Infrastrukturen har bidratt til utvikling av åpen datapolitikk med tilhørende brukerlisensiering i de 16 deltagende partnerne. Noe som vanskelig ville latt seg gjøre uten samarbeidet. Utviklingen innen sensorer og digital kommunikasjon gjør kompliserte data tilgjengelig i nær sanntid samtidig som datamengden er eksplosiv. Uten nasjonale infrastrukturer som NMDC er det stor fare for at forskningsdata går tapt. Mer deling og gjenbruk av forskningsdata av god kvalitet åpner for helt nye muligheter, og evnen til å utnytte disse mulighetene vil være en kritisk faktor for deltakelse i forskningsfronten i årene som kommer. Det krever internasjonal standardisering, nasjonal koordinering og prioritering og endringer på institusjonelt nivå. Internasjonalt samarbeid er viktigere enn noen gang og nødvendig for at forvaltning av forskningsdata skal skape merverdi for forskning og samfunn. Forskningen blir både mer dataintensiv og datadrevet som resultat av digitalisering og ny teknologi, mens det samtidig er en sterk dreining mot åpen forskning, herunder deling og gjenbruk av data. Dette fører til at mulighetene for tverrfaglighet blir større og det blir mulig å oppnå mer effektiv og kvalitetssikret flyt av data på tvers av forskning, forvaltning og næringsliv.

Gode vurderinger, faglig og strategisk:

NMDC fikk gode evalueringer både faglig og strategisk ved etableringen. Evalueringen av de 2 søknadene om oppgradering fikk noe lavere karakterer. Internasjonalt har det etablert seg krav om sertifiserte datasentre/datalagre for å sikre forskningsdataene. I NMDC er det etablert slike sertifiserte «data repositories» som omtales som «trusted repositories» blant annet ved Havforskningsinstituttet. Norge arbeider for at forskningsdata skal følge internasjonalt etablerte prosedyrer og standarder. FAIR prinsipper er en viktig og akseptert term å arbeide mot. For å gjøre forskningsdataene FAIR trenger en infrastruktur som setter dette i fokus og arbeider for implementasjon på nasjonalt nivå.

Norge har ambisjoner om at alle fagområder innen 2030 skal ha tilgang til datainfrastruktur som kuraterer forskningsdata.

NMDC-II er en slik datainfrastruktur for marine data.

Gode infrastrukturer for tilrettelegging, tilgang til og gjenbruk av forskningsdata vil bli avgjørende for forskning på nær sagt alle områder i framtiden. For å sikre dette vil det bli mer behov for spesialistkompetanse innenfor data og dataforvaltning i og rundt forskningsmiljøene.

Utvalget mener at Norge må ha et ambisjonsnivå som svarer til de overordnede målene i gjeldende Langtidsplan for forskning. Det vil si å styrke Norges konkurransekraft og innovasjonsevne, møte store samfunnsutfordringer og utvikle fagmiljøer av fremragende kvalitet. Muligheten til å samarbeide med og måle seg med de beste miljøene internasjonalt er avgjørende for å utvikle fremragende fagmiljøer. Uten tilgang til data og muligheten til å koble data begrenses mulighetene for nye gjennombrudd i forskningen, og med det mulighetene til å utvikle forskning som hevder seg internasjonalt

For at Norge skal få mest mulig ut av investeringer i datainfrastruktur, må løsningene være langsiktige og bærekraftige. Det betyr at man må tydeliggjøre ansvar, foreta prioriteringer, samordne og koordinere overlappende datainfrastrukturer og etablere gode modeller for finansiering av etablering og drift. I

I denne sammenheng bruker vi betegnelsen nasjonal datainfrastruktur om infrastruktur som ikke bare er tilknyttet én institusjon eller ett miljø, men som tilbys for bruk til alle relevante miljøer i Norge. Slik datainfrastruktur finansieres i dag enten direkte eller gjennom konkurranseutsatte midler. I begge tilfeller er det viktig at investeringene sees i sammenheng med eksisterende datainfrastrukturer, både nasjonalt og internasjonalt.

Stor og omfattende infrastruktur:

Det har blitt lagt stor vekt på å øke grad av FAIR forskningsdata i Norge og det ble våren 2022 ferdigstilt to FAIR utredninger av Norges forskningsråd. FAIR-data krever datainfrastrukturer som legger til rette for god datahåndtering. Det er behov for datainfrastrukturer som gir rask og enkel tilgang til data som er utstyrt med gode beskrivende metadata i tråd med internasjonale

standarder, metoder for behandling av dataene, systemer for håndtering av personvern og opphavsrett og lisenser for videre bruk, slik at det er enkelt å sammenstille datasett fra mange ulike kilder. Datainfrastrukturene må fungere godt hver for seg, men ulike infrastrukturer leverer funksjonalitet for ulike faser i livsløpet til dataene. Deltakelse i internasjonalt samarbeid sikrer kvalitet og fornyelse av forskningen og er spesielt viktig for å oppnå FAIR-data. Slikt samarbeid gjør at Norge kan påvirke utviklingen av standarder for datahåndtering, metadata, metoder og teknologi, men gir også norske forskningsmiljøer tilgang til relevante datainfrastrukturer i Norge og i utlandet.

I Europa vil slikt samarbeid blant annet skje gjennom deltakelse i European Open Science Cloud (EOSC). Norge er nå medlem av partnerskapet, noe som medfører nye muligheter, men også forpliktelser. Gjennom deltakelsen i EOSC får europeiske forskere tilgang til FAIR-data og relevante tjenester i virtuelle miljøer på tvers av vitenskapelige disipliner og landegrenser. EOSC bygger i stor grad på eksisterende data- og e-infrastrukturer (også norske) for gjenfinning, tilgang til og gjenbruk av data.

NMDC samler 16 aktører i Norge som har store og diverse marine data og sikrer god tilgang til relevante forskningsdata.

Meteorologisk institutt

I løpet av de siste tiårene har forskning blitt mer dataintensiv og økende sammenkobling av data fra ulike disipliner. Dette har ført til strengere kvalitetskontroll, dokumentasjon og evne til samvirke (interoperabilitet) både på data og systemnivå (infrastruktur). En betydelig innsats er lagt ned gjennom institusjonelle og nasjonale ressurser utvikling av rutiner og systemer som skal sikre FAIR (søkbare, tilgjengelige, interoperable og gjenbrukbare) data og tilhørende tjenester. FAIR prinsippene sikrer kostnadseffektiv utvikling av systemer for effektiv forvaltning og bruk av data i nedstrøms beslutningsstøttesystemer, både innenfor forskning og forvaltning. Prinsippene legger til rette for etterprøvbare beslutninger gjennom kobling av beslutninger (vitenskaplige og forvaltningsmessige) mot underliggende datagrunnlag. Nasjonalt er innføringen av FAIR-prinsippene på et nivå hvor det kan høstes betydelige fordeler i form av utnyttelse av data og observasjoner innenfor og på tvers av fagfelt, disipliner og miljøer. Full realisering av potensialet forutsetter en kulturell endring både blant dataprodusenter, datasentre og datakonsumenter (inkludert deres systemer). Slike endringer tar tid og krever investeringer i folk, systemer og infrastruktur. En grunnleggende forutsetning er at FAIR-prinsippene kommer inn i planlegging og gjennomføring av dataproduksjon/-innsamling og ikke er noe som legges på dataene i det de skal publiseres. Dette krever fortsatt fokus på opplæring, (videre)utvikling av verktøy/støttesystemer og forretningsmodeller/måltallsforvaltning. Måltallsregimet for forskning forhindrer i dag effektiv utveksling av data mellom forskning og forvaltning, f.eks. innenfor tilgang til sann- eller nær sanntidsdata. Det krever også en sterkere grad av samordning

mellom eksisterende nasjonale infrastrukturer for å unngå isolerte siloer uten kontakt med hverandre.

Hovedmålet må være å sikre brukerne FAIR tilgang til data og tjenester; kostnadseffektiv implementasjon og drift av dataforvaltningssystemer som sikrer effektiv gjenbruk av data i analyse og beslutningsstøttesystemer (f.eks. digitale tvillinger); forenkle leveranse av FAIR data for dataleverandører; samt å adressere manglende FAIRness gjennom opplæring og utvikling av felles verktøy og tjenester. Gjennom dette kan samhandling mellom forskning og forvaltning forbedres og nasjonale ressurser posisjoneres på regionalt europeisk og globalt nivå.

NIBIO

NIBIO støtter nåværende prioriteringer i de to områdestrategiene 'E-infrastrukturer' og 'IKT'.

Simula Research Laboratory

En raskt økende grad av digitalisering akselererer behovet for enkel og sikker tilgang til ressurser for datalagring og tungregning (HPC). I løpet av de siste årene har bruken av kunstig intelligens (AI) og maskinlæring (ML) ført til en voldsom etterspørsel etter slike ressurser, og på mange måter åpnet helt nye teknologimarkeder. Det har aldri tidligere blitt investert så mye i teknologier for datalagring og beregninger som i dag, og syklusen med å bringe ny teknologi ut til anvendelser har aldri gått raskere. Denne utviklingen er en viktig ramme rundt EUs satsing på tungregning og kvanteberegninger, som også omfatter utvikling av europeisk maskinvare og programvare som et ledd i å redusere Europas avhengighet av teknologier utviklet i andre deler av verden. Etableringen av den autonome enheten EuroHPC JU, hvor Norge er medlem, er et kraftfullt uttrykk for disse ambisjonene og for etableringen av europeiske HPC-infrastrukturer i verdensklasse.

Gitt det enorme momentet i den teknologiske utviklingen, og i fagområder som drar nytte av dette, er det fundamentalt viktig at Norge styrker utdanning, forskning, innovasjon og et kunnskapsbasert næringsliv gjennom tilgang på tidsriktige infrastrukturer for HPC og datalagring. Slike infrastrukturer må også tilby grunnkompetansen som trengs for at de ulike aktørene skal kunne omsette teknologi-tilgang til konkrete, samfunnsnyttige verdier. Det er følgelig behov for å videreutvikle, og til dels etablere, flere IKT-baserte forskningsinfrastrukturer både nasjonalt og gjennom internasjonalt samarbeid.

HPC. Sigma2 gjør en svært god og viktig jobb med å tilby produksjonsorienterte tungeregnerressurser. Fortsatt satsning på jevnlig fornying av nasjonale HPC-systemer for slike produksjonsbehov, og ikke minst opprettholdelse av et nasjonalt kompetansetilbud innen HPC, er ekstremt viktig. Samtidig er det ikke realistisk at Norge *alene* skal etablere nasjonale HPC-systemer som er helt i verdenstoppen ytelsesmessig. Dette gjøres best gjennom internasjonale samarbeid, og i første rekke gjennom deltagelse i EuroHPC. Dette gir Norge allerede tilgang til Europas kraftigste datamaskin per i dag, LUMI. Det er viktig at dette internasjonale engasjementet fortsetter og inkluderer Europas

første exascale-maskin, JUPITER (2023), samt kommende post-exascale systemer.

Samspeillet mellom nasjonale HPC-infrastrukturer og norsk deltagelse i EuroHPC er særdeles viktig. Det er nødvendig å ivareta en nasjonal kunnskapsbase og enkel, lokal tilgang til middels kraftige systemer, samtidig som vi kan benytte noen av de kraftigste regneressursene i Europa i et tett samarbeid med ledende fagmiljøer internasjonalt – ikke minst gjennom deltagelse i EU-finansierte forsknings- og innovasjonsprosjekter. Dette vil være grunnleggende viktig for at flere norske fagmiljøer av høy internasjonal klasse skal kunne fortsette å hevde seg internasjonalt, for eksempel innen IKT, beregningsorientert biomedisin, klimamodellering og astrofysikk.

Tradisjonelt har regnekraft blitt oppskalert ved å stadig øke antall transistorer i prosessorene, men denne teknikken har nådd sine fysiske begrensninger. Konsekvensen er at moderne HPC-systemer får en stadig mer heterogen og kompleks arkitektur, både gjennom samspill mellom ulike prosessortyper og akseleratorer, og gjennom hierarkiske strukturer for minne, datalagring og kommunikasjon. Dette betyr at det er stadig nye og mer kompliserte hensyn å ta for at både eksisterende og nyetablert programvare skal kunne ta ut potensialet i den underliggende maskinvaren. Den nasjonale forskningsinfrastrukturen *Experimental infrastructure for Exploration of Exascale computing (eX3)*, som allerede er en del av det nasjonale veikartet, har siden 2018 spesielt adressert denne økende kompleksiteten i maskin- og programvare. Dette blir gjort gjennom å tilby tidlig tilgang til et bredt utvalg av relevante teknologier i et format som muliggjør eksperimentell HPC-forskning, noe som er umulig i et produksjonssystem dominert av veletablerte teknologivalg og strenge krav til systemstabilitet. Denne eksperimentelle tankegangen tas nå videre i SLICES, en europeisk forskningsinfrastruktur på ESFRI Roadmap 2021 som vil kombinere HPC, distribuert prosessering og Internet of Things, og hvor eX3 er tiltenkt rollen som norsk node. Forskningen basert på eX3 har vakt stor interesse og muliggjort internasjonalt samarbeid, blant annet gjennom flere EuroHPC-finansierte prosjekter, og har vært en kompetanse- og erfaringskilde for Sigma2 sin planlegging av kommende produksjonsorienterte HPC-anskaffelser. Erfaringene demonstrerer behovet og verdien av å tilby eksperimentelle teknologi-infrastrukturer i tillegg til de tradisjonelle produksjonsomgivelsene, både innenfor HPC og andre teknologitunge forskningsområder.

AI/ML. Beregningsbehovet i krevende AI/ML-anvendelser er en sterk pådriver for utviklingen av HPC-teknologier og -infrastrukturer. Dette har også vært tydelig gjennom driften av eX3, der Norges mest framoverlente fagmiljøer innen AI/ML har hatt muligheten til å bruke prosessorer og systemer som er optimaliserte for denne typen oppgaver – typisk 1-2 generasjoner lenger framme enn hva som har vært tilgjengelig i produksjonsorienterte omgivelser.

Et viktig steg for AI/ML-infrastrukturer er også den nylig vedtatte etableringen av Norwegian AI Cloud (NAIC) som kombinerer oppbyggingen av GPU-baserte ressurser dedikert for AI/ML med et omfattende kompetansetilbud for forskere som har behov å ta i bruk denne typen beregninger, på tvers av ulike fagfelt. NAIC vil være produksjonsorientert og basere seg på veletablert HPC-teknologi, og over tid er det naturlig at denne typen HPC-funksjonalitet bli mer framtrødende i den infrastrukturen som forvaltes av Sigma2. En slik utvikling vil være i samsvar med den modulære utbyggingen av EuroHPC-systemer som LUMI og JUPITER. Det er imidlertid svært viktig at den kompetanseoppbyggingen som NAIC representerer forvaltes videre på en god måte, uansett om man velger å beholde den tekniske infrastrukturen i NAIC eller smelter den sammen med annen HPC-infrastruktur under Sigma2.

Quantum Computing. Internasjonalt, og ikke minst innen rammen av EuroHPC, forberedes et potensielt paradigmeskifte relatert til tungregning – Quantum Computing (QC). Gjennom de siste tiårene har internasjonale grunnforskingsmiljøer i kvantemekanikk bygd opp basisen for kvantedatamaskiner. I løpet av de aller siste årene har denne utviklingen akselerert og man begynner nå å se konturene av at kommersielle systemer etableres og av at slike maskiner kan få en nytteverdi både i forskning (utenfor kvantefysikken) og i utvalgte industrielle anvendelser. QC-maskiner kan potensielt tilby regnekraft på en helt annen skala enn tradisjonelle HPC-systemer, og muliggjøre digitale løsninger som i dag er utenkelige. EuroHPC har nylig utlyst prosjektmidler for å etablere de første QC-maskinene integrert i noen av sine eksisterende HPC-anlegg. Dette er også den første måten QC-teknologi av funksjonell verdi vil bli tilbudt forskning og industri – som en ny type akselerator for bestemte klasser av beregningsoppgaver, integrert i et mer tradisjonelt HPC-system. Dette betyr at heterogeniteten og kompleksiteten i moderne HPC-systemer vil øke ytterligere. I dag finnes det ikke veletablerte metodikker, standarder eller verktøy som kan effektivt bringe QC inn i brede utviklingsmiljøer. Dette krever grunnleggende IKT- og HPC-forskning, som igjen forutsetter tilgang til relevante forskningsinfrastrukturer for QC.

I dag er dette temaet knapt nok på den politiske agendaen i Norge, noe som er foruroligende siden det allerede investeres enorme summer internasjonalt for å etablere QC som en tilgjengelig beregningsressurs. I Europa skjer denne utviklingen ikke bare gjennom EuroHPC, men også på nasjonalt nivå – blant annet i sammenlignbare land som Danmark, Sverige og Finland. Norge bør definitivt være aktive i forhold til å ta del i en felles europeiske satsing på QC-teknologi, men man bør også i nær framtid søke å etablere en nasjonal forskningsinfrastruktur for QC, om enn med mindre kapasitet enn systemene som utvikles gjennom internasjonale samarbeid. Som for den eksperimentelle siden av HPC-forskning, handler dette om nærhet til og styring av teknologivalg som vil bli svært viktige i løpet av de neste 10-20

årene, og om oppbygging av en nasjonal kunnskaps- og erfaringsbase.

Sikker lagring og forvaltning av data. Det samles inn og prosesseres stadig flere og større datasett på tvers av fagområder. Dette fordrer tilgang til datainfrastrukturer med stor lagringskapasitet, høy hastighet på datatilgang og nødvendige mekanismer for backup. Disse infrastrukturene må være organisert, teknisk og kompetansemessig, slik at man møter forventningene om tilgjengelighet og gjenbrukbarhet (FAIR-prinsippene), samtidig som det er enkelt for brukerne å sørge for at gjeldende lover og regler blir respektert. Dette gjelder spesielt GDPR og håndtering av sensitive data. UiOs TSD-tjeneste er et viktig og godt steg i retning av trygg håndtering av sensitive data, og disse erfaringene bør oppskaleres til bruk i et nasjonalt tilbud. Samtidig må Norge følge nøye med på de data/sky-infrastrukturene som er under utviklingen internasjonalt, og spesielt i Europa, f.eks. GAIA-X og EOSC. Det bør vurderes om Norge skal ta en aktiv rolle i slike løsninger, og hvordan disse skal balanseres mot nasjonale tilbud. Uansett lokasjon og samhandlingsmodell må de aktuelle datainfrastrukturene kunne tilby adaptive konfigurasjoner som gjør det mulig å balansere åpenhet og reproduserbarhet i forskning i forhold til påkrevd datasikkerhet. Datainfrastrukturer må også ha et fokus på god dataforvaltning – ikke bare gjennom tilgjengeliggjøring, men også ved bidra til at innsamlede data renses og kurteres slik at de kan brukes i ulike sammenhenger. Dette er spesielt viktig sett i forhold til økningen av AI/ML-basert prosessering.

Datakommunikasjon og samhandlende systemer. Vi er i dag fullstendig avhengige av at ulike systemer og tjenester er tilgjengelige via internett, og i økende grad at disse systemene kan samhandle. Dette fordrer at den underliggende teknologien for datakommunikasjon effektivt møter stadig økende krav til høy overføringshastighet og minimale forsinkelser, samt garantier med hensyn til kvalitet, stabilitet og sikkerhet. Norge har i mange tiår vært langt framme i dette feltet, både i forskning og næringsliv. For å videreføre en slik ledende posisjon vil det være behov for tilgang til forskningsinfrastrukturer som tilbyr både dagens og fremtidens kommunikasjonsteknologi. F.eks. er 5G en teknologi som nylig er tatt bruk i det norske mobilnettet og som vil være avhengig av forskningsinfrastrukturer for å videreutvikles med ny funksjonalitet, nye tjenester og ny kompetanse i årene som kommer. Tilsvarende gjelder for 6G, men med en lengre horisont. Forskningsinfrastrukturer for datakommunikasjon vil også ha naturlige grenseflater mot datainfrastrukturer og HPC-ressurser, og inkludere områder som «edge computing», «fog computing» og andre typer distribuerte beregninger.

En spesiell klasse av samhandlende systemer er såkalte «cyber-physical systems», som knytter sammen mange ulike maskinvare-enheter (f.eks. kameraer og sensorer) og mer eller mindre spesialiserte datamaskiner til et «system av systemer». Slike samhandlende systemer brukes bredt i industrielle anvendelser, gjerne som grunnlag for digitale tvillinger av industrielle

prosesser. Dette medfører mange spørsmål og problemstillinger knyttet til det totale systemets robusthet, integritet og ytelse. Videre, lar man slike system få lov til å fatte enkelte beslutninger, typisk basert på AI/ML-algoritmer, er man på rask vei til autonomi. Alle disse fasettene av avanserte, samhandlende systemer må baseres på grunnleggende forskning, og tilgang til relevante forskningsinfrastrukturer, nasjonalt og internasjonalt, vil spille en avgjørende rolle for denne utviklingen.

Sikker, robust og pålitelig datakommunikasjon er grunnlaget for digital sikkerhet og en vellykket digitalisering av samfunnet. Dette krever at forskningsmiljøene innen cybersikkerhet har tilgang på eksperimentell kommunikasjonsinfrastruktur. Slik infrastruktur er nødvendig for å forstå svakhetene i nåværende og fremtidige kommunikasjonssystemer og forhindre at slike svakheter truer kritisk infrastruktur og samfunnsikkerhet. På sikt vil dette også innebære infrastruktur for Quantum Communication som er tett knyttet til Quantum Computing, Quantum Internet og de krav til sikker kommunikasjon som disse teknologiene fører med seg.

Forskningsinstitutt (Privat)

Svar

NORCE - Norwegian
Research Centre AS

Først og fremst er det viktig å få etablert en grunnleggende **nasjonal vitenskapelige FAIR datainfrastruktur**, som dekker alle forskningsfeltet og som bør **finansieres utenfor den konkurransearenaen**. Den infrastrukturen er en avgjørende del av implementeringen av den nasjonale forskningspolitikken og kobling av nasjonale forskningsvirksomheter til internasjonal (ikke minst EU) forskning. I den sammenhengen er det viktig å fokusere på teknologi og standarder som utvikles ifm **EOSC** for å sikre verdiskapingen og synligehten av nasjonale forskningsdata i det internasjonale forskningskonteksten.

Veikartet som definerer rammen og prioriteringen for konkurransearena som styres av forskingsrådet, bør fokusere på **behov for dedikerte data tjenester** som kan bygges på som overlag til den nasjonale datainfrastrukturen. Tjenestene kan målrettes mot spesifikke fagområder, men også fasilitere **forskning "uten grense"** mellom forskningsfag, og særlig tjenester som lager **bro mellom natur- og samfunnsvitenskap**. For den **digitale grønne omstillingen** er det essensielt at det er tilgjengelige data og tjenesteplattformer for forskning og innovasjon.

Samfunnsvitenskap: SIKT og SSB er sentrale institusjonene for åpen tilgang til forskningsdata, og også viktige komparative surveys som European Social Survey (ESS) og International Social Survey Program (ISSP). Disse bør vi opprettholde, videreutvikle og samordne. ESS er nå iferd med å gå over til online intervjumodus. Men det trengs å etablere ytterligere infrastruktur som genererer data på nye måter, og som benytter digital teknologi for behandling av store datamengder.

Energi: Simuleringsmiljøet OpenLab Drilling er ferdigstilt med støtte fra Forskningsrådet **og driftes og vedlikeholdes av NORCE med brukermedvirkning.** Denne infrastrukturen er basert på beregningsmodeller for simulering av realistiske data fra boreoperasjoner og er gjort lett tilgjengelig på internett. I OpenLab kan simuleringene gjøres både i et internett-grensesnitt, i en fysisk simulator hos NORCE og fra Ullrigg brønnenanlegg. Infrastrukturen er spesielt godt egnet til utvikling og testing innen automatisert boring, og spesielt i teknologi som anvender kunstig intelligens og maskinlæring der det trengs store mengder høykvalitets data. **Den e-infrastrukturen har støttet kunnskapsbygging og innovasjon for mer miljøvennlig og effektiv boring. Den har potensialet til å videreutvikles for å støtte oppbyggingen av geotermisk energi i Norge.**

Digitalisering: Behovet for digitalisering innenfor petroleumsindustrien er stort og er ventet å medføre store besparelser for næringen, men også å gi et mindre avtrykk på natur og miljø. Det er derfor et økt behov for forskning og utvikling på teknologier som utnytter økte datamengder fra mange ulike leverandører. Behovet ligger innenfor alle teknologidisipliner i hele verdikjeden i petroleumsnæringen. Dette omfatter datainnsamling, databehandling, datakvalitet, dataintegrasjon, beslutningsstøtte og datasikkerhet for muliggjørende automatiserings-, autonomi- og IKT-teknologier. Nasjonal digital infrastruktur vil tilfredsstillere behovet for en **styrket nasjonalt samarbeid mellom de fremste sterke samfunnsvitenskapelige miljøer** som innenfor statsvitenskap, media og kommunikasjon, sosiologi, rettsvitenskap, økonomi, psykologi og folkehelse. Digital infrastruktur kan legge bedre tilrettet for **nyskapende og tverrfaglig forskning mellom** for eksempel ulike **samfunnsvitenskapelige disipliner og teknologer.**

NORSAR

Det er et stort potensial å utnytte eksisterende og planlagte kabler til kommunikasjonsteknologi for forskning og teknologiutvikling for samfunnsnyttige formål. Ved å koble fiberoptisk teknologi i samspill med seismisk analyse vil man kunne benytte eksisterende kabler for kommunikasjonsteknologi til å overvåke naturlige og menneskeopvakte vibrasjoner slik som ras, snøskred, eksplosjoner, vannlekkasjer mv. Vi foreslår derfor at SIKT tilgjengeliggjør forskningsnettet til slike formål. I tillegg er det mange undersjøiske fiberkabler som vil kunne brukes til overvåking av offshore infrastruktur, livet i havet og havstrømmninger.

SINTEF

Digitalisering er en sterk driver for økt konkurransekraft og viktig for å løse vår tids samfunnsutfordringer. Infrastruktur for datadeling og IKT er derfor kritisk for å realisere potensialet som ligger i digitalisering. De fleste forskningsfelt skaper data eller analyserer data i stadig økende omfang. Dette stiller stadig større krav til samarbeid knyttet til datalagring og tungregnerressurser, samt sikker håndtering av forsknings- og forvaltningsdata.

Datainfrastruktur

Norske forskere må ha tilgang til datainfrastrukturer som muliggjør forskning og utdanning i verdensklasse. Dette bør skje som lokale fellestjenester, nasjonale fellestjenester og gjennom internasjonalt samarbeid.

Det er ikke naturlig å se for seg en felles plattform for all forskningsdata. Standardisering og interoperabilitet derfor viktig for å få et velfungerende system for datadeling. Organisasjoner som International Data Spaces Association arbeider med dette og deres referansearkitektur er integrert i EUs strategi for datadeling. Nasjonale infrastrukturer for datadeling bør også bygges på samme prinsipper.

Nasjonale datainfrastrukturer er viktig for å gi forskere tilgang til større mengde data, delta i den datadrevne forskningen, og bidra til at norsk forskning ligger i forskningsfronten. Det er derfor nødvendig med friske midler til nasjonale datainfrastrukturer.

Det eksisterer en stor mengde forvaltningsdata hos bl.a. offentlige aktører. Disse kan være av stor interesse og nytte for forskning, men det kreves en tilrettelegging og delvis varsom handling, pga. nasjonale sikkerhet eller konkurranse hensyn. Tilrettelegging av eksisterende forvaltningsdata kan bli en viktig del av framtidig (virtuell infrastruktur) og krever egne midler.

Ekspertgruppen for deling av industridata leverte sin rapport til regjeringen 31.5.2022. I rapporten vises det til flere scenarier for datadeling. Selv om rapporten fokuserer på næringslivet, er flere av problemstillingene like relevant for deling av forskningsdata.

Flere algoritmer innen kunstig intelligens gjøres åpent tilgjengelig. Disse algoritmene er trent opp basert på enorme datamengder. Et eksempel er AlphaFold som ble utviklet av firmaet DeepMind. Algoritmen er installert på norsk infrastruktur etter initiativ fra NORA, UiO og Uninett Sigma2. Programmet er tilgjengeliggjort for alle norske forskningsmiljøer og kan blant annet benyttes til å utvikle medisiner. Det er mer hensiktsmessig at denne typen algoritmer installeres på en felles infrastruktur i Norge, enn at forskningsinstitusjonene skal sette dette opp selv. Dette er bare ett eksempel på hvorfor man trenger nasjonale datainfrastrukturer.

Tungregning

Nasjonal infrastruktur for tungregning er svært viktig for norske forskningsmiljøer. Det er viktig at norske forskningsmiljøer har tilgang til superdatamaskiner for å utføre krevende beregninger som for eksempel numerisk modellering og simulering. Behovet for stadig mer datakraft gjør det nødvendig å kontinuerlig investere i nasjonale supercomputere. Samtidig er det viktig at norske forskere har tilgang til internasjonale supercomputere for å få enda mer datakraft. Norsk deltagelse i EuroHPC er derfor avgjørende.

Tungregning er viktig for trening av algoritmer basert på machine learning. Tradisjonelle tungregningsressurser basert på CPUer er derimot ikke godt egnet for denne typen beregninger. Dataparallele arkitekturer som GPUer, eller andre spesialarkitekturer, er nødvendig for effektiv trening av systemene. Nasjonal infrastruktur som inkluderer denne typen maskinvare er derfor viktig for norske forskningsmiljøer.

Quantum computing

Kvantedatamaskiner kan potensielt løse problemer på sekunder, som verdens kraftigste supercomputer bruker flere år på å løse. Teknologien er dermed svært disruptiv, og kan endre måten man utfører beregninger på i fremtiden. Kvantedatamaskiner utfører beregninger på qubits, ikke bits som tradisjonelle datamaskiner. Dette betyr at det kreves mye forskning for å utvikle algoritmer som er egnet for denne arkitekturen.

Norge har ingen kvantedatamaskiner som det er mulig å kjøre reelle problemer på. Norge må derfor delta i internasjonalt forskningssamarbeid som sikrer norske forskere tilgang til kvantedatamaskiner. Gjennom EUs forskningsprogrammer satses det tungt på dette området.

Oppsummering basert på Forskningsrådets spørsmål

Forskningens behov for forskningsinfrastruktur for å løse utfordringer innenfor et strategisk prioritert område:

- Infrastrukturer for datadeling er viktige innen de fleste forskningsfelt. Dette gjelder både for økt tverrfaglig forskning, men også for at forskningen skal tilfredsstillende prinsippene for deling av vitenskapelige data - FAIR.
- Datainfrastruktur bidrar ikke mot et strategisk prioritert område i seg selv – men er med på å muliggjøre verdensledende forskning og utdanning innen de fleste strategisk prioriterte områdene.

Hvilke eksisterende nasjonale forskningsinfrastrukturer dere mener er av stor verdi å opprettholde og videreutvikle

- Uninett Sigma2: Utvikling av infrastruktur for tungregning og midler til drift av denne må opprettholdes og videreutvikles.
- Videreutvikling av infrastruktur for tungregning basert på spesielle krav til maskinvare – dataparallele prosessorer er ofte nødvendig for å trene store algoritmer innen kunstig intelligens.

Hvilke tematiske områder det blir spesielt viktig å etablere nye nasjonale forskningsinfrastrukturer på og/eller samarbeide om internasjonale forskningsinfrastrukturer

- Quantum computing: Tilgang til kvantedatamaskiner, enten gjennom etablering av nasjonal infrastruktur eller internasjonalt samarbeid, er viktig for norsk forskning innen fagområdet.

Hvilke hull som kan dekkes gjennom utvikling av eksisterende nasjonale og/eller samarbeid om eller tilgang til internasjonale forskningsinfrastrukturer

- Det er viktig at Norge deltar i internasjonalt samarbeid innen datainfrastruktur og IKT. Det arbeides mye med standardisering for datadeling og interoperabilitet i EU.
- Noen ganger trenger norske forskningsmiljøer tilgang til kraftigere tungregnerressurser enn det man kan bygge i Norge. Deltagelse i EuroHPC er et eksempel på viktig internasjonalt engasjement.

Stiftelsen Nansen senter for miljø og fjernmåling (NERSC)

Sigma2 og HPC infrastrukturer

Norsk klima-, polar- og hav modellering, samt AI anvendelser er 100% avhengig av Sigma2 infrastrukturer for tilstrekkelig regnekapasitet. Disse anvendelsene bør nevnes eksplisitt i nytt veikartet for Datainfrastruktur og IKT (noe det ikke gjør i gjeldende veikart).

Norsk og Forskningsrådet investeringer i HPC forskningsinfrastruktur er ikke konkurransedyktig på Europeisk nivå. Norsk fagmiljøer innen denne type anvendelser vil derfor bli mindre attraktive både som partner i større modell prosjekter og for rekruttering av eksperter for denne type modellering.

Helsesektor (Offentlig)

Svar

Akershus universitetssykehus HF

Følgende er også spilt inn under punktet om Mat og Helse og Andre infrastrukturbehov innenfor naturvitenskap og teknologi. Et overgripende tema er det forskningsetiske perspektivet som skapes innenfor ny teknologi og bruk av kunstig intelligens (AI). For å generere ny kunnskap og bidra til utvikling i forskning og innovasjon, er det nødvendig at eksisterende teknologi som benyttes i diagnostikk og pasientbehandling er tilgjengelig. En satsing på kunstig intelligens og persontilpasset medisin, er helt avhengig av infrastruktur for uttrekk av forskningsdata fra kilde-systemene i sykehusene. En parallell og tung satsing på uttrekk, lagring og analyse av forskningsdata er derfor helt nødvendig. For persontilpasset medisin er målet at så mange pasienter som mulig, bør få tilgang til utvidet molekylær diagnostikk i det sykehuset som har ansvar for utredning og behandling. Kunstig intelligens er et fremtidig viktig verktøy blant annet i vurderingen av avansert billedteknologi, men det forutsetter tilgjengelighet av stort volum av kildedata. Tilgangen til billeddiagnostikk er pr. i dag en begrensende faktor på sykehusene.

Viktig å etablere pasientnær forskningsinfrastruktur, som omfatter blant annet medisinsk billeddiagnostikk, nestegenerasjons-sequenseringsteknologi, mikroskopi, celleterapi og genredigering. Teknologi/forskningsinfrastruktur må være plassert i nærhet av pasienter og må være sikker og robust nok til å kunne benyttes i diagnostikk parallelt med forskning og innovasjon. Det bør derfor legges til rette for søknader som omhandler etablering av infrastruktur i nærhet til pasientbehandling. Dette vil muliggjøre at det enkelte foretak kan bedre utnytte allerede etablerte nasjonale

forskningsinfrastrukturer, slik som tilbudet til de ulike sentre for fremragende forskning.

Viktig å sette fokus på gjenbruk av data registrert i pasientjournal i forskning, og se dette opp mot ulike design med eksempelvis en pragmatisk tilnærming. Pragmatiske studier er basert på kliniske data fra uselekterte pasienter i vanlig praksis. På den måten bygger studiene bro mellom evidensbasert medisin og klinisk praksis. Pragmatiske studier defineres av tre kjennetegn, de skal (1) gi svar på hva som er optimal klinisk behandling med direkte betydning for pasientene, (2) inkludere pasienter og undersøke problemstillinger som er relevante for den kliniske behandlingen, og (3) benytte oppfølging og data fra etablerte kliniske systemer med klinisk relevante endepunkter. Pragmatiske studier har mange fellestrekk med tradisjonelle randomiserte kontrollerte studier, men kjennetegnes av et design der fokus er å benytte ressurser som allerede er tilgjengelige i helsetjenesten og dermed inkludere et stort antall pasienter på kort tid. De er også mindre kostnadskrevende enn studier med tradisjonelt design.

Utfordringen er tilgang på strukturerte data fra pasientjournaler, samt regulerende lovverk. Ulike lover og forskrifter regulerer klinisk praksis, forskning og kvalitetssikring. Dette gir en juridisk utfordring, for eksempel når man i pragmatiske studier samler forskningsdata som ledd i kliniske behandlingsløp. Selv om behandling, forskning og kvalitetssikring er atskilt juridisk, synliggjør pragmatiske studier hvor tett disse punktene er knyttet sammen. Dette påvirker både søknadsprosedyrer rundt etikk og personvern, samtykkekrav for bruk av data fra elektronisk pasientjournal og nasjonale registre. Det er behov for et integrert forskningsdatasystem som muliggjør generering, uthenting og tilgjengeliggjøring av helsedata fra helseforetakene som er tilpasset lovverkets krav til personvern og informasjonssikkerhet og krav til forskningens etterprøvbarehet.

Det mangler et nasjonalt forskningsadministrativt system (FAS) som gjør det mulig for forskningsinstitusjoner å ivareta institusjonens behov og lovkrav til oversikt og oppfølging av forskningsprosjekter. Helsetilsynet gjennomførte i 2020 et tilsyn der 20 virksomheter ble undersøkt om hvordan de styrte og organiserte sin medisinske og helsefaglig forskning. Hovedfunnet var blant annet at flere virksomheter hadde mangelfulle systemer (internkontroll) mht oppfølging av oppstart, gjennomføring og avslutning av forskningsprosjekter. Medisinsk og helsefaglig forskning bidrar til å frembringe viktig kunnskap for samfunnet som helhet, men slik forskning kan være inngripende for den enkelte forskningsdeltaker. For å sikre nødvendig deltakelse i slik forskning er det avgjørende at befolkningen har tillitt til forskerne og virksomhetene som utøver den. Etterlevelse av regelverket om helseforskning er viktig for å etablere og opprettholde nødvendig tillit i befolkningen. På bakgrunn av at flere virksomheter hadde mangelfulle systemer for å følge med på at lovverket ble fulgt, mener Statens helsetilsyn at det foreligger risiko for svikt i medisinsk og helsefaglig forskning og at tillitten kan svekkes. Institusjonene har et utvidet behov med parametere som skal

rapporteres til både Norges Forskningsråd og eksterne finansieringskilder, samt at det er viktig å ha en oversikt over alle behandlinger av personopplysninger i forskning. Enkelte institusjoner har utarbeidet egne løsninger, men som da er tilpasset kun den enkelte institusjon. Vi anbefaler at det utredes om etablerte in-house løsninger kan være et utgangspunkt for etablering av et nasjonalt forskningsadministrativt system. Dette for å utvikle en løsning som kan dokumentere behandling av personopplysninger i forskningsprosjekter i hele prosjektets livsløp. Et slikt system bør også sees i sammenheng med det nasjonale arbeidet med FAIR prinsippene.

Oslo universitetssykehus

- forskningens behov for forskningsinfrastruktur:
Det mangler et integrert forskningsdatasystem som muliggjør generering, uthenting og tilgjengeliggjøring av kliniske data, spesielt håndtering av sensitive helsedata, i tilpasset lovverkets krav til datasikkerhet og krav til forskningens etterprøvnbarhet. Det må etableres sikre soner og systemer for datautveksling der forskere kan generere forskningsresultater basert på individdata, både i universitets- og sykehussektoren. OUS stiller seg bak rapporten og anbefalingene fra datainfrastrukturutvalget (Investering i infrastrukturer for FAIR forskningsdata og særlig relevante forvaltningsdata for forskning, Kenneth Ruud, mai 2022). OUS imøteser opprettelse av nasjonalt genomsenter mhp organisering og bruk av genomikkdata, også sekundærbruk.
- hvilke eksisterende nasjonale infrastrukturer er av stor verdi å opprettholde og videreutvikle:
BioMedData er en eksisterende infrastruktur under ELIXIR Norway som har som mål å koble sammen data fra en rekke ulike nasjonale infrastrukturer, inkludert NALMIN, NorSeq, Biobank Norge, NorOpenscreen. Det er avgjørende at denne infrastrukturen videreføres. Etter det vi kjenner til, er denne planlagt tatt opp i ELIXIR.
ELIXIR
ELIXIR er en helt sentral aktør både europeisk og gjennom sin norske node. ELIXIR arbeider med både dataforvaltning og datadeling, samt etablerer og drifter infrastruktur for å understøtte dette. FAIR håndtering av kombinasjonen av storskala genomiske og kliniske data er en helt sentral utfordring der ELIXIR i tiden framover vil være av stor betydning for OUS. I dette inngår også stigende behov for internasjonal datadeling drevet av internasjonale kliniske protokoller. For utnyttelse av genomiske stordata er det avgjørende at disse gjøres maskinlesbare for utnyttelse i AI-sammenhenger, i kombinasjon med trygge kommunikasjonsprotokoller. Sett også i sammenheng med pågående arbeid for etablering av et nasjonalt genomsenter vil ELIXIR forbli en sentral samarbeidspartner, også sett i lys av pågående internasjonalt arbeid der Norge deltar (e.g. 1+MG). Norwegian Artificial Intelligence Cloud (se under Mat og helse)
- hvilke hull kan dekkes gjennom utvikling av forskningsinfrastrukturer

En videre satsning på nasjonale datainfrastrukturer vil styrke muligheten for gjenbruk av data og bedre koordinering av nasjonale studier.

Forvaltning (Offentlig)	Svar
Direktoratet for e-helse	<p>En tenker ofte på infrastruktur som faste anlegg og systemer som er nødvendige for at en virksomhet eller et samfunn skal fungere. Men det kan også brukes i utvidet betydning om de strukturelle forutsetningene for innholdet i en aktivitet eller et produkt. Helsedataservice (HDS) ved Direktoratet for e-helse representerer en slik infrastruktur med tanke på materielle og prosessuelle rammebetingelser for god og effektiv utnyttelse av helsedata fra ulike kilder til forskningsformål og ny kunnskap for bedre folkehelse.</p> <p>Helsedataservice (HDS) ble etablert i 2021 som en virksomhet med ansatte som jobber spesielt rettet mot søkere av helsedata, og mot forvalterne av helsedata (blant annet Folkehelseinstituttet, Helsedirektoratet, Kreftregisteret og Norsk helsearkiv). Hovedformålet til HDS er å etablere «en vei inn» for blant annet forskere, slik at det skal være enklere, raskere og sikrere å få tilgang til helsedata til forskningsformål. Det er allerede etablert nettstedet helsedata.no som inneholder informasjon, felles søknadsportal, variabelutforsker, samt veiledning om søknadsprosessen.</p> <p>Når ny forskrift om sentralt vedtak er på plass er det rimelig å tro at HDS vil få anledning til å overta saksbehandlingsoppgaver fra de sentrale helseregisterforvalterne. Dette vil ytterligere bidra til «en vei inn» til helsedata for forskningsprosjekter, ved at saksbehandlingsprosessen forenkles og legges til én aktør (HDS), istedenfor som i dag å utføres i parallelle løp hos mange aktører. Dette vil kreve en videreutvikling av kompetanse og gode samarbeidsrutiner mellom HDS og forvaltere av helsedata. Tjenestene Helsedataservice nå tilbyr har stort potensiale. Både ved at stadig flere helsedataleverandører tar tjenestene i bruk for å vise frem innholdet sitt, men også ved en videreutvikling og forbedring av felles søknadsskjema, variabelutforsker og metadata, samt veiledningsmateriell og veiledningstjeneste. Det er et mål at HDS vil bidra til en harmonisering av saksbehandlingsprosesser og regelverkforståelse, som sørger for lik behandling av søknader om data, også der hvor HDS ikke har vedtaksmyndighet. Videre kan HDS bli en nasjonal kompetanseenhet som jobber med veiledere knyttet til forståelse av regelverk, personvern innen helsedata, tilgangsstyring og verktøy. Dette er et mål som krever langsiktig planlegging, bygging av kompetanse og videreutvikling av tjenestetilbudet til HDS, men som åpenbart vil bidra til stadig raskere, enklere og sikrere tilgang til helsedata.</p> <p>Videre er det identifisert behov for enklere tilgang til sammenstilt statistikk. Infrastrukturen HDS kan bidra til et kompetansesenter for å utarbeide slik statistikk, basert på helsedata fra blant annet</p>

de sentrale helseregistrene. HDS kan, gjennom veiledere, analysekompetanse, og metodeutviklingsmiljø, tilby en slik tjeneste til forskningsprosjekter.

HDS er nylig etablert og bygges trinnvis. HDS jobber strategisk for å bli nasjonal kompetanseenhet som utvikler og tilbyr tjenester og veiledning som gjør helsedata lettere tilgjengelig for forskning og utvikling av ny kunnskap på en trygg måte. HDS har som mål å være en forskningsinfrastruktur som bygger bro mellom

- 1) aktører som samler inn og tilgjengeliggjør helsedata,
- 2) forskningsmiljøer som ønsker tilgang til data,
- 3) kunnskapsmiljøer innen ulike sektorer (helse, utdanning, forskning, myndigheter og andre), og
- 4) kompetansemiljøer som jobber for bedre løsninger for deling av helsedata både nasjonalt og internasjonalt.

For å få til dette er det behov for videre investering og finansiering, både i personell og i tekniske løsninger.

Bruk av eksisterende analyserom

I samarbeid mellom Direktoratet for e-helse ved Helsedataservice og universitets- og høyskolesektoren, vil prosessen for å tilgjengeliggjøre data til analyserom kunne forenkles og optimaliseres ytterligere. De fleste forskningsprosjekt benytter mer enn bare helseregisterdata, som data fra andre typer registre, eksempelvis SSB, prosjektets egne data, personers selvrapporterte data m.m. Videre er det ikke bare forskere i Norge som er interessert i å bruke norske data. Farmasøytisk industri og internasjonale forskningskonsortier er også aktuelle brukere. I tillegg kan helsedata kan også benyttes i større grad enn det gjøres i dag til styring og planlegging av ressursbruk i helsevesenet og til bedre utnyttelse av mangfoldet av kvalitetsregistre som finnes.

Microdata

Microdataløsningen hos SIKT/SSB er en viktig satsning som bør ses i større sammenheng og det er et ønske og behov for samarbeide tverrsektorielt om videre realisering av gevinster for denne løsningen.

Sikt -
kunnskapssektorens
tjenesteleverandør

Fra Sikt – kunnskapssektorens tjenesteleverandør (Sikt) innen området Datainfrastruktur og IKT

Videreutviklingen av forskningsinfrastruktur må løse forskningens behov på en måte som utnytter og bygger videre på tidligere investeringer i infrastrukturkompetanse, -løsninger og -tjenester. Sikt har gjennom bred kontakt med forskningsmiljøene identifisert omfattende behov for arkiveringsløsninger for sensitive forskningsdata. Norge mangler i dag datainfrastrukturer som har tilstrekkelig kapasitet til å håndtere deponering, kuratering, dokumentasjon og langtidssikring av sensitive forskningsdata i stort volum/bredde, og som forenkler tilgjengeliggjøring/deling av slike data i dertil egnede plattformer og infrastrukturer for bearbeiding, kobling, deling og analyse. Som en konsekvens blir verdifulle data underutnyttet. Muligheter for kobling og interoperabilitet som kunne gitt en eksponensiell verdiøkning for

forskningen eller støtte opp om rask sammenstilling av relevante data i krisesituasjoner, uteblir helt.

Norge har etter hvert fått frem flere forskningsinfrastrukturer og plattformer som forenkler innsamling, bruk og håndtering av sensitive data i ulike deler av forskningsdatalivssyklusen. F.eks. microdata.no som er en plattform for tilgjengeliggjøring, analyse og utforskning av sensitive registerdata, og sikre analyserom som TSD, SAFE, HUNT Cloud, Sikts datahåndteringsplanverktøy og meldeskjema, UiOs Nettskjema, etc.

KD/NFR har over flere tiår investert i NSDs (nå Sikts) forskningsdataarkiv som har løsninger for deponering, maskinlesbar og standardisert datadokumentasjon, langtidskuratering og formidlings- og delingskanaler utviklet i samarbeid med ulike forskningsmiljøer. Forvaltning, foredling og formidling av store spørreundersøkelser fra Statistisk sentralbyrå (f.eks. Arbeidskraft- og levekårsundersøkelsene), fra UH-sektoren (f.eks. Ungdata, Medborgerpanel, NorLAG, m.fl.) og fra Europa (f.eks. European Social Survey) baserer seg på denne infrastrukturen som dekker en del former for ikke-sensitive data godt.

Etter vår oppfatning bør veikartet adressere det store behovet for arkiveringsløsninger for alle typer sensitive data gjennom å bygge videre på de investeringer som er gjort på datainfrastruktururområdet nasjonalt og se satsingen i sammenheng med nasjonale prioriteringer opp mot ESFRI-prosessen og etableringen av ERICer innen ulike fagområder. Det er viktig å stimulere til integrasjoner, samarbeid og vekselvirkninger mellom disse på en måte som styrker norsk forskning og kostnadseffektivt utnytter potensialet i trygg bruk og gjenbruk av sensitive data. Ofte undervurderes krav til kompetanse om data, metadata, teknologiutvikling og om forskningsprosessen som stilles for å utvikle effektive og forenklede løsninger for håndtering av både ikke-sensitive og sensitive data i ulike faser av livssyklusen. Eksisterende infrastrukturaktører kan sannsynligvis bekrefte dette. Vi vil oppfordre veikartutvalget til nysgjerrighet på hva som kreves for at Norge skal lykkes med reell FAIRness i forskning også for sensitive data, inkludert de krevende men svært viktige Interoperability- og Reusability-dimensjonene som ligger i en slik ambisjon. FAIR-utredningen (Infrastruktur og tjenester for FAIR forskningsdata) er blant flere som påpeker at viktige funksjoner som kan støtte opp om disse dimensjonene mangler i det nåværende økosystemet for forskningsdata.

Større infrastrukturesatsinger bør legge til rette for tverrfaglig dataintegrasjon og utveksling gjennom strukturering av metadata fra ulike domener. Dette blir viktig både i EOSC-sammenheng og i det arbeidet som foregår i ESS ERIC og CESSDA ERIC, et arbeid som Sikt er tett på som dataarkiv for ESS og nasjonal tjenesteleverandør i CESSDA.

Sikt ser for seg en fremtid hvor forskere enkelt kan planlegge hele livsløpet for innsamling, bruk/analyse, arkivering, deling og gjenbruk via datahåndteringsplanleggingsverktøy som både ivaretar forskningsinstitusjonenes behov for oversikt og lokale

retningslinjer, og som samtidig gjør det enkelt for forskere å velge og ta i bruk trygge og kostnadseffektive nasjonale løsninger for datahåndtering der det er ønskelig. For planlegging, innsamling, bruk, langtidslagring, tilgjengeliggjøring og deling av forskningsdata.

Et robust nasjonalt forskningsdataarkiv for sensitive data vil kunne bidra aktivt til å støtte forskningsmiljøene og komplementere de ulike institusjonelle løsningene, som f.eks. DataVerse.no, slik at optimale gjenbruks- og koblingsmuligheter for dataene som samles inn, sikres så tidlig som mulig i prosessen (via adekvate samtykker eller på annen måte). Disse lange linjene fra planlegging til gjenbruk kan være kompliserte og ressurskrevende for forskningsprosjekter å forholde seg til uten infrastrukturer med arkivering for gjenbruk høyt på agendaen, og som har kompetanse på både dataforvaltning, forskning og personvern. Uten faktisk gjenbruk har et dataarkiv liten verdi. Sikt er derfor opptatt av en fortsatt styrking av de nasjonale infrastrukturene vi har for langtidslagring, kobling, bearbeiding og bruk av data, herunder sensitive data, som allerede er samlet inn og arkivert. Vi vet at det er behov for et spenn av løsninger for å dekke behovene både for langtidslagring og for rask tilgang (inkl. tilgang i kriser) og for avansert funksjonalitet i analyser og metodeutvikling. På tilgangssiden ser vi f.eks. at infrastrukturen microdata.no dekker en del behov for rask tilgang og rike koblingsmuligheter, mens tjenester som TSD, Safe, HUNT Cloud o.l. dekker andre behov rundt funksjonalitet, metodevalg og -utvikling, alle innenfor trygge rammer.

Vi ser viktigheten av samspill mellom disse to typene infrastruktur, og har derfor lagt til rette for at microdata.no kan brukes til ikke bare selvbetjent kobling, bearbeiding og analyse, men også som tilrettelegging av skreddersydde data for senere eksport av SSB-data til f.eks. TSD (med tilhørende saksbehandling og godgjøring av behandlingsgrunnlag i dialog med SSB). Dette er et godt eksempel på hvordan forskningsinfrastrukturer kan utfylle hverandre når det legges ressurser i koordinering.

Det finnes en rekke muligheter for integrasjon og samkjøring mellom løsninger for langtidslagring og for tilgangsløsninger som gir rask tilgang og bredere/kraftigere tilgang til sensitive data, respektive.

Men det finnes også utfordringer. Kobling av data fra mer enn én kilde kan ha høy forskningsverdi, og være uproblematisk å gi rask og konfidensialitetssikret tilgang til. Ved overføring til en ikke-konfidensiell setting, slår kompliserte prosesser for avklaring av behandlingsgrunnlag inn, og veien til sikre analyserom blir ofte lang og tidkrevende. Vi ser at Helsedataservice er etablert for å strømlinjeforme disse prosessene innenfor helseforskning, men vil peke på at lignende utfordringer innenfor andre sektorer og ikke minst på tvers av sektorer også bør løses.

**Statlige etater
(Offentlig)****Svar**

Arkivverket

Arkiv og arkivdata er dokumentasjon og informasjon som er viktig å ta vare på for ettertiden fordi det bidrar til etterprøvnbarhet og demokrati, og også med formål for fremtidig forskning. Arkivdata kan komme fra både statlig, kommunal og privat sektor, og omfatter mange ulike fagområder og mange ulike formater. Digitaliseringen av samfunnet skaper nye muligheter og utfordringer, både på arkivfeltet og forskningsfelt generelt. Forskere må forholde seg til store mengder data, nye formater og nye forskningsmetoder i en digital tid. Arkivverket har et ansvar for å utvikle datainfrastrukturer og IKT-verktøy som sikrer og tilgjengeliggjør data for forskning, både nå og til ettertid. Arkivverket er en av landets største dataforvaltere, og vi ser at våre data blir brukt til forskning på flere områder enn de blir skapt for. Arkivverket har i samarbeid med Forsvarets Sanitet og Sykehuset i Vestfold begynt å digitalisere Forsvarets personalmapper, med fokus på mannskap som deltok i internasjonale operasjoner mellom 1947- 1966. Selv om de opprinnelige arkivene ble skapt for Forsvarets egen bruk, utgjør de digitaliserte mappene en database som kan bli brukt til livsløpsstudier og forskning innenfor medisinsk epidemiologi, samfunnsmedisin og populasjonsgenetikk. IKT-teknologier som kunstig intelligens og håndskriftsgjenkjenning gjør det lettere for forskere å håndtere store datasett ved å gjøre de strukturerte, søkbare og dermed bedre egnet til analyse. En del data som er nyttige i forskning, vil også være nyttige som referanser til anvendt forskning og utviklingsarbeid, slik at infrastrukturen ofte kan være en del av en større helhet. Gjenbruk av data til nye formål kan også sees som en form for datadeling, som har ringvirkninger utover enkelte infrastrukturer. De internasjonale FAIR-prinsippene har dannet et grunnlag for e-infrastruktur i tidligere versjoner av veikartet. Arkivverket ser disse som viktig å opprettholde i videreutviklingen av datainfrastruktur og IKT-løsninger.

I tillegg har Arkivverket utviklet den nasjonale fellesløsningen Digitalarkivet som en nasjonal infrastruktur for mottak, bevaring og tilgjengeliggjøring av historiske data. Som datainfrastruktur tilbyr Digitalarkivet sikker langtidslagring og bevaring av historiske data for nå- og ettertid, for flere samfunnsaktører. Dette har relevans for flere fagfelt, som utvides i vårt innspill til temaområde Samfunnsvitenskap. Digitalarkivet vil også bli en viktig forskningsinfrastruktur for forskere i mange ulike fagområder, og det jobbes med å utvikle tjenester som er tilpasset forskernes behov.

Tidligere versjoner av veikartet har hatt som mål å levere områdespesifikk e-infrastruktur. Vi tror at det fremover vil bli mer tverrfaglig forskning, der data fra ulike forskningsinfrastrukturer vil bli stadig mer aktuelt. Flere forskningspolitiske dokumenter (e.g. Humaniora i Norge, Oppfølging av evaluering av humanistisk forskning i Norge, and Norsk veikart for forskningsinfrastruktur 2020) identifiserer et behov for en felles infrastruktur for søk,

tilgang og analyse på tvers av institusjoner og fagområder. Forskere har i dag mange forskningsinfrastrukturer å forholde seg til, og det blir viktig å etablere nye infrastrukturer som kobler sammen ulike infrastrukturer, både nasjonalt og internasjonalt. FAIR-prinsippet *interoperabilitet* illustrerer denne helhetlige tankegangen, der datainfrastrukturer som API-er og felles standarder for datadeling og kan være viktige verktøy. Det kan også være verdifullt å fokusere på tjenstedesign og arkitektur i etablering av forskningsinfrastrukturer. Forskningsinfrastrukturer kan i mange tilfeller bygges på toppen av andre datadelingsløsninger, med grensesnitt og datadelingstjenester som er egnet for forskere.

Forskningsdata som er benyttet i tidligere forskning utgjør viktig infrastruktur for ny forskning, og bevaring av forskningsdata fra et prosjekt kan være viktig dokumentasjon og underlag for å forstå og tolke resultatet av forskningen - også i ettertiden. Arkivverket har siden 2014 hatt en avtale med NSD, nå SIKT (Kunnskapssektorens tjenesteleverandør), om at SIKT skal ta imot forskningsdata fra forskere og forskningsmiljøer på vegne av Arkivverket. Avtalen med SIKT legger til rette for at forskningsdata skal lagres og tilgjengeliggjøres for tilpasset analytisk bruk basert på arkivloven og annet relevant regelverk. Forskningsdata bør i mange tilfeller bevares i et langtidsperspektiv. Forskningsdataarkivenes bevaringsperspektiv strekker seg ofte ikke lenger enn 10 år, mens Arkivverkets bevaringsperspektiv ikke har en tidsbegrensning. Avtalen med SIKT innebærer at Arkivverket skal motta dataarkivene som skal bevares i et langtidsperspektiv. Arkivverket arbeider derfor med å etablere et samarbeid med SIKT om overføring og langtidslagring av bevaringsverdige datasett. Arkivverket ønsker gjennom dette å bidra til utvikling av denne infrastrukturen gjennom å ta et større ansvar for langtidsbevaring av forskningsdata. Avtalen med SIKT legger til rette for at forskningsdata skal lagres og tilgjengeliggjøres for tilpasset analytisk bruk basert på arkivloven og annet relevant regelverk.

NVE, Hydrologisk avdeling

Innenfor det tematisk prioriterte området 'Data og tjenester overallt' og sett i forhold til store samfunnsutfordringer ligger det et stort forskningspotensial inn mot samfunnssikkerhet/naturfareforskning, dersom man får til en mer **samordnet og helhetlig innsats for repetert innsamling av romlige grunnlagsdata** (topografi, arealbruk, vegetasjonsendringer, jordart, geologi, jordsmonn) som muliggjør endringsdeteksjon, med kobling mot målinger og modeller for vann- og klimaparametre.

Selv om disse ikke er en formell nasjonal forskningsinfrastruktur, er det grunn til å løfte den innsatsen som er gjort av forvaltningsinstituttene met.no, nve og ngu for innsamling av helt essensielle langsiktige tidsserier og andre data innenfor geofagene, og for å lagre og tilgjengeliggjøre åpne data for brukere. Det er etablert databaser som samler data fra flere produsenter for nedlasting. Vi har imidlertid et målenettverk med langt lavere romlig tetthet enn land vi kan sammenligne oss med,

og behovet for oppdatert informasjon og mer detaljert romlig oppløsning for slike geodata er høyt. Generelt bør datainfrastruktur/databaser legges hos de store forvaltningsinstitusjonene som har lang erfaring med lagring og tilgjengeliggjøring.