



UiO : **Universitetet i Oslo**

IT i forskning ved Universitetet i Oslo

Rapport fra arbeidsgruppe for IT i forskning ved Universitetet i Oslo (UiO)

30. januar 2015

INNHOLDSFORTEGNELSE

1. SAMMENDRAG OG ANBEFALINGER	3
2. BAKGRUNN FOR ARBEIDET OG MANDAT	7
2.1 Forankring	7
2.2 Arbeidsgruppens medlemmer	7
2.3 Arbeidsgruppens mandat	8
2.4 Arbeidsgruppens arbeid	8
3. VISJON FOR IT I FORSKNING VED UIO	9
4. OPERASJONALISERING AV VISJONEN	11
4.1 IT-infrastruktur: tjenester og teknologi for forskning	11
4.2 Kompetanse-hub'er og andre samarbeidsformer	16
4.3 Styrking av USITs kapasitet, bredde og beredskap for prosjektstøtte	18
5. STATUS IT I FORSKNING VED UIO ANNO 2014	20
5.1 Fellestjenester	20
5.2 Nasjonal eInfrastruktur	24
5.3 Internasjonal eInfrastruktur	26
6. KOMPETANSEBEHOV	26
6.1 Kompetanseutvikling og kompetansedeling	27
6.2 Forholdet til "IT i utdanning"	28
7. FORSLAG TIL TILTAK	29
7.1 Utnyttelse av eksisterende eInfrastruktur	29
7.2 Forskningsdata	30
7.3 Kompetanse og kompetanseutvikling	30
7.4 Støttebehov for IT i forskning og nærhetsmodellen	31
8. ORGANISERING AV IT TIL FORSKNING VED UIO	32
8.1 Roller og ansvar	33
8.2 Hub'er og noder	34
8.3 Koordinering og prioritering	35
9. RESSURSBEOH OG ØKONOMISKE FORHOLD	36

1. Sammendrag og anbefalinger

Forskningsvirksomhet er iboende dynamisk. Det søkes etter ny viten i ytterkant av den etablerte, på stadig nye områder og med det siste av metoder og verktøy. Dette er metoder og verktøy som i seg selv er resultat av forskning, og derfor i stadig utvikling. De som først og best tar i bruk nye metoder og verktøy i forskningsprosessen, har de beste forutsetninger for å produsere grensesprengende resultater. Arbeidsgruppens har i sitt arbeid med forslag til organisering og finansiering av IT i forskning lagt særlig vekt på forskningens iboende dynamiske karakter.

eInfrastruktur

"IT i forskning" handler essensielt om tilgang til og bruk av regneressurser, lagringsressurser, nettverksressurser, dataressurser, programvareressurser og menneskelige ressurser i forskning. Lagringsressurser er maskinvare for lagring og arkivering av data, mens dataressurser her referer til forskningsdata, samlinger og databaser som skal forvaltes for ulike forskningsformål. Programvare fremstår i all hovedsak i to former; i) IT-løsninger og tjenester som gir brukere tilgang til nettverk, regnemaskiner og lagringssystemer og ii) en serie med beregningsverktøy som forskere kan bruke til å studere eller analysere sine data. Menneskelige ressurser er alt fra teknisk personale som drifter nettverk, datamaskiner og lagringsressurser, via generell brukerstøtte og kurs, til avansert brukerstøtte for utvikling av programvare og verktøy for beregninger rettet mot spesifikke vitenskapelige problemstillinger. Disse ressursene samlet kalles gjerne *eInfrastruktur*. I stadig større grad er det også nødvendig med forskningspreget metodeutvikling i informatikk for å lage de rette verktøyene. Vi inkluderer denne aktiviteten i definisjonen av *eInfrastruktur*.

I det store bildet leveres og brukes *eInfrastruktur* på fire nivåer; internasjonalt, nasjonalt, institusjonelt og lokalt ute i fagmiljøene. Våre forskere deltar på alle nivåer, men vi skal her primært forholde oss til hvordan IT i forskning bør organiseres og finansieres ved UiO. I dette sammendraget vil vi i fortsettelsen bruke begrepet *eInfrastruktur* i stedet for "IT i forskning". Selv om vi her primært omtaler *eInfrastruktur* i forhold til forskning er også *eInfrastrukturen* i ferd med å bli svært viktig for store deler av utdanningen ved Universitetet i Oslo.

Behov – det store bildet

En velfungerende *eInfrastruktur* er helt nødvendig for vitenskapelig fremskritt innen stadig flere fagfelt. Grensesprengende forskning baseres i stadig større utstrekning på beregninger og avanserte analyser av store datamengder. Dette er godt dokumentert både nasjonalt og internasjonalt¹. Frem til 2005 var investeringene i *eInfrastruktur*, både lokalt, nasjonalt og internasjonalt, konsentrert rundt innkjøp av datamaskiner og utbygging av nettverk, mens vi i løpet av de siste årene har opplevd et stadig sterkere behov for investeringer i kapasitet for lagring, håndtering og tilgang til vitenskapelige data. Stadig større datamengder og økende krav til lagring og sporbarhet av beregningsresultater og tidligere analyser bidrar sterkt til dette, selv om utviklingen i forskningsmetodikken er den drivende faktor.

¹ Rapportene "The Scientific Case for eInfrastructure in Norway" (2010) og "Norwegian eInfrastructure Roadmap" (2012) er utarbeidet av en bredt panel av brukere i Norge og viser det økende behovet for *eInfrastruktur* i hele den vitenskapelige breidde.

Lagring av vitenskapelig data krever en annen form for langsiktighet knyttet til investeringer enn det vi kjenner fra tradisjonell tungregning, noe som er en viktig motivasjon for å se på nye modeller for finansiering og organisering av eInfrastrukturen ved UiO.

Datadrevet forskning

Forskning drives i stadig større omfang av tilgangen på store datamengder, og *datadrevet forskning* har blitt en vesentlig del av hverdagen i tilnærmet alle fagmiljøer i hele den vitenskapelige bredde. Det skjer over hele verden.

Begreper som “big data”, “information driven society” og “open data access” går igjen, og de henviser til hva store datamengder i ulike former betyr og kan bety for næringsutvikling spesielt og samfunnsutvikling generelt. Innen forskning (og utdanning) blir data på digital form stadig viktigere og det skjer i tilnærmet hele den vitenskapelige bredde. Innen naturvitenskap, medisin og teknologi har det vært slik lenge, men de siste årene har dette slått gjennom med en kraft det var vanskelig å forutsi for kun få år siden. Bedre og billigere målemetoder (sensorer) som fanger data i stor hastighet fra ulike kilder og beregninger i stor skala generer en dataflom som gir både muligheter og utfordringer. Det samme gjelder menneskelig aktivitet på internett.

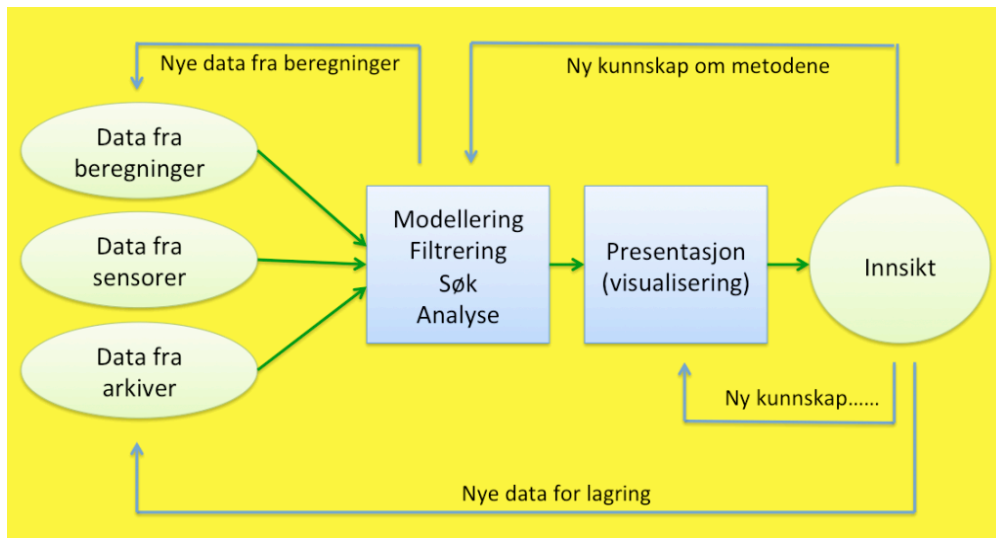
Ved UiO har vi i dag en rekke prosjekter og aktiviteter som er innrettet både mot det å kunne håndtere store datamengder og det å avdekke ny kunnskap og erkjennelse gjemt i store datamengder. Ekstra gledelig er det derfor at vi høsten 2014 fikk tildelt to nye Sentre for Forskningsdrevet Innovasjon på området – Centre for Scalable Data Access (SIRIUS) og Big Insight, statistics for the knowledge economy.

Hva er datadrevet forskning? Det er forsøkt illustrert i Figur 1. Bak denne figuren finnes den nødvendige eInfrastrukturen bestående av små og store regnemaskiner, systemer for innsamling, lagring og håndtering av store datamengder i ulike former for arkiver og samlinger, nettverk som transporterer data til og fra brukere, programvaresystemer som gir oss tilgang til disse ressursene, samt mennesker med teknisk og vitenskapelig innsikt som hjelper forskere (og studenter) i bruk av infrastrukturen.

Hva må gjøres for å få noe ut av de store datamengdene? Innen forskning må det utvikles modeller for representasjon og integrasjon av data slik at disse kan analyseres ved hjelp av ulike verktøy og beregningsmetoder. Søkemetoder og metoder for filtrering av data er også viktig for å finne frem til det som er interessant i store datamengder. Forskning på dette området handler både om utvikling av områdespesifikke metoder, f.eks. innen biofag, medisin, språkvitenskap eller geofag, og utvikling av generiske metoder innen matematikk, statistikk og informatikk. Skal man lykkes er det viktig å finne en god balanse mellom investeringer i generiske aktiviteter og områdespesifikke aktiviteter. Metode- og verktøyutviklingen er tverrfaglig og krever kompetanse fra flere fagdisipliner, kompetanse som både må utvikles og koordineres.

Det neste steget er å presentere resultatene på en form som kan forstås og tolkes, og i stadig større omfang gjøres dette gjennom å presentere bilder og animasjoner (visualisering). Det finnes selvfølgelig også andre måter å presentere resultater på. Ut av

dette kommer det ny innsikt. Forskingen på disse områdene gjennomføres som oftest i en stor rundgang. Innsikt gir nye data som kan arkiveres for å skape ny innsikt. Ny innsikt gir informasjon slik at metoder og presentasjonsformer kan forbedres, osv. Pilene i Figur 1. indikerer dette.



Figur 1: Datadrevet forskning – en slags “big data pipeline”!

Anvender man dette, f.eks. innenfor livsvitenskapene, handler det bl.a. om å utføre beregninger for å forstå prosesser og fenomener i kroppen, analyser av bilder (ultralyd, MR, mm), og studier av genmateriale fra mennesker, dyr og planter. I en eller annen form sentreres en stadig større del av forskningen i hele den vitenskapelige bredde rundt data. Når det er sagt vil fortsatt eksperimenter være viktig, ikke minst for å generere data og for å teste hypoteser basert på analyserte data. Med dette bilde som utgangspunkt er det viktig med langsiktig finansiering av eInfrastrukturen, god organisering av forskerstøtte og kunnskaps- og kompetanseutvikling.

Økonomi og finansiering

Arbeidsgruppen antar at de fysiske regne- og lagringsressurser i all hovedsak vil bli finansiert gjennom UiOs deltakelse i den nasjonale ordningen, ordinære investeringer og prosjekter finansiert av eksterne kilder (for eksempel Forskningsrådets infrastrukturprogram). Investeringsbehovet der UiO selv må bidra mye i årene som kommer antas å være utvikling av systemer og mekanismer for håndtering av til dels store datamengder, og utvikling av menneskelige ressurser som på vegne av forskningen ved UiO (og partnere) kan hjelpe forskerne med å utnytte lokal, nasjonal og internasjonal eInfrastruktur.

Situasjonen per i dag er at UiO deltar med ca. 20 mill. kroner per år fra 2015 i den nasjonale satsingen på eInfrastruktur i Norge. Denne satsingen er planlagt for 10 år med en evaluering halvveis. Den nasjonale satsingen kobler også UiO til nordiske og europeiske satsinger innen eInfrastruktur. Videre bruker USIT ca. 10 mill. kroner per år til eInfrastrukturelle tjenester for hele UiO. Et grovt anslag tilsier at fakulteter (og institutter) bruker anslagsvis 15 mill. kroner per år på egne eInfrastrukturelle tjenester. Det totale omfang ved UiO er derfor estimert til 45 mill. kroner per år. Arbeidsgruppen mener at dette ikke er tilstrekkelig for at UiO skal være konkurransedyktige på området, dvs. tilby våre forskere eInfrastrukturelle løsninger som sikrer deltagelse i forskningsfronten innen viktige fagområder. UiO mangler

også tilstrekkelige mekanismer for strategiske satsinger, styring og prioriteringer innenfor eInfrastruktur.

Arbeidsgruppen foreslår derfor at det øremerkes ytterligere 15 mill. kroner per år som deles ut til satsinger koordinert på tvers av fakulteter og at satsingen på USIT ikke reduseres i forhold til dagens nivå. Et typisk eksempel på viktigheten av USIT er at USIT har tilstrekkelig ressurser til å videreutvikle løsninger for sikker håndtering av sensitive data. Arbeidsgruppen foreslår at disse midlene i første omgang avsettes for en periode på 5 år fra 2016. Det er naturlig å se for seg at UiOs andel av satsingen på bioinformatikk tas fra disse midlene. (Oslo Universitets-sykehus bidrar allerede i denne satsingen.)

Organisering

Arbeidsgruppen foreslår videre at de koordinerte satsingene organiseres som nettverk etter en hub-node struktur slik man nå legger opp til innenfor bioinformatikk. Hub'ene skal ha ansvaret for koordineringen av innsatsen, herunder sikre metodeforskning, kunnskapsutvikling og kompetanseoverføring mellom hub og noder. Hub'ene skal typisk bygge opp kunnskap og kompetanse som er generisk for hele strukturen og sikre god tilgang til eInfrastrukturelle ressurser som leveres av USIT (og den nasjonale eInfrastrukturen). Hub'ene må ligge på et veldefinert vertsinstitutt som både kan sikre det vitenskapelige og organisatoriske fundamentet. Vertsinstituttet kan søke om eInfrastrukturmidler på vegne av Hub-node-strukturen, blant annet for å opprette tidsbegrenset vitenskapelig forskningsledelse i informatikk der det er nødvendig. Nodene vil typisk være enkeltforskere og forskningsgrupper som trenger eInfrastrukturelle tjenester. Hub - node strukturen foreslås for å oppnå effektiv utnyttelse av ressurser, kunnskapsutvikling og kompetanseoverføring mellom ulike deler av UiO innenfor et gitt tema (bioinformatikk, språkdata, videodata, osv.). USIT vil typisk kunne defineres som en node i en slik struktur på linje med noder ute i forskningsmiljøene og slik integreres i hub'ene. Vi foreslår denne organiseringen særlig for å ivareta forskningens iboende dynamikk der kontinuerlige nyutvikling av metoder for beregninger og håndtering av store datamengder er viktig for vitenskapelige fremskritt.

eInfrastrukturutvalg ved UiO

Det foreslås videre at det etableres et koordineringsutvalg for IT i forskning, og vi foreslår at dette koordineringsutvalget får navnet *eInfrastrukturutvalget*. eInfrastrukturutvalget skal ha ansvaret for å følge opp UiOs investeringer i den nasjonale eInfrastrukturen, overvåke USITs og enhetenes eksisterende satsing på eInfrastruktur, samt fordele de foreslåtte frie midlene (15 mill. kroner) til tidsbegrensede satsinger, i hub – node-strukturer eller spesifikke prosjekter. Arbeidsgruppen anslår at en årlig investering på 15 mill. kroner vil være tilstrekkelig til å støtte 4-7 hub - node strukturer avhengig av disse strukturenes størrelse og behov. eInfrastrukturutvalget bør ledes fra rektoratet og ha 5-7 medlemmer. Det foreslås at USIT har sekretariatet.

Forskning og utdanning henger tett sammen og det er naturlig at eInfrastrukturutvalget også har ansvar for den delen av utdanningen som benytter eInfrastrukturen.

eInfrastrukturutvalget skal ikke gi støtte til rene vitenskapelige stillinger (eller forskerstillinger) utover tidsbegrensede vitenskapelige forskningsleder- og koordineringsstillinger i informatikk nevnt over.

2. Bakgrunn for arbeidet og mandat

2.1 Forankring

Overordnet mål for IHR og Mål 4 i Strategi 2020 er:

"Universitetet i Oslo skal forvalte sine samlede ressurser offensivt, slik at de bidrar til å understøtte kjerneaktivitetene."

Strategi 2020 har også som mål (Mål 1) at:

"Universitetet i Oslo skal fremme grensesprengende forskning, være et topp europeisk forskningsuniversitet og en etterspurt internasjonal samarbeidspartner."

Strategi 5 for målet er at "forskning i internasjonal front skal understøttes av strategiske investeringer og samarbeid om forskningsinfrastruktur". I universitetets rullerende årsplaner finner vi dette igjen som "videreutvikling av forskningsinfrastruktur" (tiltak #5 i hhv. UiOs årsplan 2013-2015 UiOs årsplan 2014-2016).

23. oktober 2012 vedtok universitetsstyret:

"Universitetsdirektøren skal i samarbeid med ledelsen ved fakulteter og institutter fremme forslag til Universitetsstyret om hvordan støtte til IT i undervisning og forskning skal organiseres ved UiO" (V-SAK 6 Saksnr. 2012/7588, "IHR - forslag fra plangruppene om tiltak for å realisere hovedmålene").

Det er delegert til IT-direktøren å utarbeide forslag. I saksframlegget til universitetsstyret fremheves det:

"Tilbakemeldingene hva angår håndtering av innmeldte behov for nye tjenester viser at det er behov for å skille klarere mellom administrative tjenester og tjenester til primærvirksomheten."

"Det er behov for klargjøring av hvordan støtte til IT i undervisning og forskning best kan utøves og hva USITs rolle skal være på dette området. Dette skal skje i samarbeid med enhetene."

På bakgrunn av dette er arbeidsgruppen for IT i forskning ved UiO etablert.

2.2 Arbeidsgruppens medlemmer

Arbeidsgruppen nedsatt av IT-direktøren ble konstituert 26. februar 2014 med følgende medlemmer:

Morten Dæhlen, dekan MatNat fak. (leder)
Thorbjørn Nordbø, fakultetsdirektør (fungerende), HF
Fridtjof Mehlum, forskningssjef, NHM
Lex Nederbragt, førsteamanuensis, IFI, NSC og CEES
Kirsti Klette, professor, UV fak., ILS
Per Grøttum, professor, Med.Fak.
Tarjei Havnes, førsteamanuensis, Økonomisk institutt, SV

Thomas Evensen, underdirektør, Underavdeling for virksomhetstjenester, USIT
Hans A. Eide, seksjonssjef, Seksjon for IT i forskning, USIT (sekretær)

2.3 Arbeidsgruppens mandat

Det overordnede spørsmålet gruppen skal belyse er på grunnlag av nevnte strategiske føringer og styrevedtak:

Hvordan skal UiO organisere sin IT-virksomhet slik at den fremmer og er aktiv bidragsyter til grensesprengende forskning og blir et attraktivt forskningsuniversitet på internasjonalt nivå?

Av dette følger en rekke problemstillinger rundt universitetets IT-virksomhet til støtte for forskning som forslaget skal gjøre rede for, blant annet:

- Hva mener vi med IT-støtte til forskning? Hva er en hensiktsmessig avgrensning mellom denne og tilgrensende bruksområder for IT? Herunder forhold og avgrensning til forskningsadministrative systemer.
- Hvordan skal virksomheten organiseres: ansvar, roller og kompetanse (og hvilket ansvar, hvilke roller og hvilken kompetanse trengs i de ulike ledd)?
- Sammenheng med ny organisering og finansiering av nasjonal e-infrastruktur
- Forhold til forskningsinfrastrukturutvalget.
- Forhold til IT i utdanningen og forskningsbasert undervisning.
- Organisering i nærhetsmodellen. Hvordan skal IT-organisasjonen samhandle med fagmiljøer, ledelse og sentraladministrasjon?
- Eksternt finansiert virksomhet og deltagelse internasjonale prosjekt
- Strategisk styring og finansiering.

Arbeidsgruppen kan trekke på tidligere gjennomført arbeid:

- [Materiale](#) fra workshop med sentrale interessenter innen forskning 21. desember 2011.
- [Diskusjonsrapport](#) levert av arbeidsgruppen "Lokal-Sentral" til plangruppen for IHR-IT 03. februar 2012.
- "[Organisering og standardisering universitetets IT-virksomhet](#)", vedtatt i universitetsstyret 23. oktober 2012.

Arbeidsgruppe leveranse er et forslag til hvordan organisere støtte til IT i forskning ved UiO. Dette betyr i praksis et organisasjonskart over aktører og interessenter, med distribusjon av roller, ansvar, og kompetanse, og med saksgang ved ulike typer beslutninger, alt dette underbygget av beskrivende og drøftende tekst, og med vurderinger av styrker og svakheter. Leveransen omfatter også eventuelle varianter og alternativer, som på samme vis må veies og vurderes.

2.4 Arbeidsgruppens arbeid

Arbeidsgruppen innså tidlig at oppgaven ville bli krevende, grunnet bredden på forskningen ved universitetet, mangfoldet hva gjelder behov og saksområdets kompleksitet. Saksområdet lagring, arkivering, deling og gjenbruk av forskningsdata er så omfattende at det er nedsatt en egen arbeidsgruppe for dette. Prosessen med etablering av en ny nasjonal ordning for eInfrastruktur til forskning har pågått samtidig med arbeidsgruppens arbeid. Samlet har disse

og flere andre faktorer medført at arbeidsgruppens rapport er forsinket, men samtidig er det arbeidsgruppens oppfatning at rapporten på grunn av dette er bedre fundamentert og anbefalingene tilsvarende styrket.

Målgruppen for dette notatet er beslutningstakere ved institusjonen, og formålet er å gi et grunnlag for beslutning vedrørende organisering og finansiering av IT i forskning ved UiO. Det er viktig å understreke at IT i forskning ikke kun handler om å tilby IT-løsninger til forskningen, men også handler om at forskningen i seg selv endres.

I rapporten brukes begrepet *eInfrastruktur* i samme betydning som "IT i forskning". Selv om vi her primært omtaler eInfrastruktur i forhold til forskning, er også eInfrastrukturen i ferd med å bli svært viktig for store deler av utdanningen ved Universitetet i Oslo.

Arbeidsgruppens rapport er strukturert med sammendrag og anbefalinger først. I det som følger beskriver arbeidsgruppen først en visjon for IT i forskning ved UiO (kapittel 3) som så operasjonaliseres med konkrete behovsrettede forslag til tiltak i kapittel 4. Kapittel 5 er en beskrivelse av dagens eInfrastruktur. Kompetanse og forholdet til utdanningen er så viktig at arbeidsgruppen har valgt å gi dette et eget kapittel (kapittel 6). Ved å sammenholde dagens situasjon med visjonen har arbeidsgruppen kommet fram til forslag til tiltak som er beskrevet i kapittel 7. Organisatoriske og økonomiske forhold drøftes i kapittel 8 og 9.

Det har vært avholdt 7 møter i arbeidsgruppen.

3. Visjon for IT i forskning ved UiO

Arbeidsgruppen har fått den viktige oppgaven å se på hvordan IT ved UiO best kan organiseres for å støtte forskningen frem mot 2020. En optimal organisering må kunne ivareta forskningens dynamiske karakter, og dette er utgangspunktet for vår visjon.

Den første komponenten i arbeidsgruppens visjon, er erkjennelsen om at forskning er iboende dynamisk. Det søkes etter ny viten i ytterkant av den etablerte, på stadig nye områder og med det siste av metoder og verktøy. Dette er metoder og verktøy som i seg selv er resultat av forskning, og i stadig utvikling. De som først og best tar i bruk nye metoder og verktøy i forskningsprosessen, har de beste forutsetninger for å produsere grensesprengende resultater.

De siste 20-30 årene har IT vært en driver i stadig større deler av forskningen. Denne utviklingen har de siste årene akselerert, og vil fortsette i mange år fremover. Det er i dag nærmest utenkelig å ikke benytte til dels svært avanserte IT-verktøy og -metoder i forskning. UiOs suksess som forskningsuniversitet avhenger derfor også av hvor effektivt IT utnyttes, hvilke ressurser, støttefunksjoner og tjenester som finnes, og ikke minst, hvordan IT innrettes og leveres til UiOs fagmiljøer. Skal UiO fremme grensesprengende forskning, være et topp europeisk forskningsuniversitet og en etterspurt internasjonal samarbeidspartner (UiO Strategi 2020, Mål 1), må universitetets IT-ressurser anvendes effektivt og offensivt slik at de bidrar til å understøtte forskningen, i tråd med IHR og Mål 4 i Strategi 2020.

I arbeidsgruppens visjon er hele bredden av forskningen ved UiO understøttet av og integrert med IT på en slik måte at forskerne har tilstrekkelig tilgang på eInfrastruktur, inklusive tjenester og kompetanse, lett tilgjengelig i et dynamisk samspill med lokale og eksterne ressurser. Dynamikken må ivaretas gjennom et tett samarbeid mellom forskerne og de menneskelige ressursene i eInfrastrukturen, dvs. personer med spisskompetanse innen IT, og evne til å sette seg inn i forskernes problemstillinger. På samme tid som forskernes IT-kompetanse må dyrkes, må IT i forskning ved UiO organiseres slik at eInfrastrukturen strekker seg ut og oppsøker forskerne, senker inngangsterskelen og effektiviserer forskningsprosessen.

Like viktig som at forskningen evner å være dynamisk og grensesprengende, er det at IT-ressursene, tjenestene, tilgjengelig kompetanse og eInfrastruktur er dynamiske og aktuelle. IT i forskning består i stor grad av komponenter som er felles og kan deles av hele organisasjonen, og av komponenter som går på tvers av flere fagfelt og disipliner. Samspillet dette medfører tilfører de enkelte fagfelt nye impulser og egenskaper, samt nye metoder som er utslagsgivende for de enkelte fagområder. Når kompetanse og innovasjon blir ivaretatt, er IT-synergier på tvers svært ofte utløsende for ny og grensesprengende forskning.

Den andre komponenten i visjonen er en datasentrisk tilnærming til IT i forskning. Data, i sine mange former, inngår i nær sagt all forskning, både som utgangspunkt for de studiene som gjennomføres og som resultater av forskningen. Data kommer fra mange ulike kilder og samles inn på mange forskjellige måter, fra laboratorieutstyr og sensornettverk, til videoopptak og sosiale medier. Data behandles, analyseres og visualiseres. Data lagres, systematiseres og katalogiseres. Data publiseres og arkiveres. Data skal finnes igjen, og det produseres data om dataene – metadata.

Data vokser organisk, og datamengden vokser eksponentielt. De eInfrastrukturelle utfordringene forbundet med lagring og overføring av store datamengder utgjør i seg selv problemstillinger som trenger spisskompetanse og avanserte løsninger. Forskning drives i stadig større omfang av tilgangen på til dels enorme datamengder fra ulike kilder, og datadrevet forskning har blitt en vesentlig del av hverdagen i mange fagmiljøer. Det skjer over hele verden. Betydningen av å kunne håndtere data øker i takt med datamengdene. Både UiO som institusjon og fagmiljøer må etablere strategier for bærekraftig håndtering av data.

Visjonens tredje komponent er den menneskelige dimensjon: tverrfaglig kompetanse og kunnskapsutvikling. En ideell forskerutdanning gir forskere tilstrekkelig kompetanse til å utnytte IT i forskning. IT-støttet problemløsning må inn fra begynnelsen av utdanningsløpet, noe Det matematisk-naturvitenskapelige fakultet har tatt på alvor gjennom prosjektet "Computing in Science Education" og tidlig introduksjon til bruk av IT-verktøy i hele bredden av fakultetets studieportefølje.

Samtidig må eInfrastrukturen, avanserte IT-ressurser og -tjenester, tilbys på en inkluderende og lett tilgjengelig måte, som ikke krever at mottaker investerer uhensiktsmessig mye tid, med nærhet til individer med spisskompetanse som kan tilby tilpasset rådgivning, og med støtte i alle faser av forskningsprosessen. Forskerne må holdes oppdatert på tilgjengelig eInfrastruktur og ny teknologi, tilbys etterutdanning, kurs og hjelp til å utnytte mulighetene.

Det er også tilfellet at flere forskere eller forskergrupper sitter med kompetanse som kunne vært bedre utnyttet, lokalt på samme enhet, men også på tvers av enheter og fag. Slike ressurser kan utnyttes dersom både enhetene og UiO sentralt var mer bevisst den strategiske betydningen av IT i forskning og organiserte lokal støtte til IT i forskning, kompetansegrupper eller større nettverk. De menneskelige ressursene i eInfrastrukturen, støtteapparatet, må organiseres slik at disse raskt kan settes inn i nye anvendelsesområder og bygge opp forståelse for forskernes behov og problemstillinger. Samlet er det de menneskelige ressursene som vil medføre en effektiv utnyttelse av tid, kompetanseheving og eInfrastruktur, og kvalitetsheving i forskningen.

I forskningsfronten vil ikke de nødvendige metodene innenfor informatikk alltid eksistere. En sentral del av visjonen for IT i forskning er derfor at det legges til rette for forskning om og utvikling av nye informatiske metoder.

4. Operasjonalisering av visjonen

Bredden i forskningen ved UiO, så vel som dens dynamiske natur, gjør det lite hensiktsmessig å tenke seg én sentralisert organisering av støttefunksjoner og ressurser til IT i forskning. IHR og mål i Strategi 2020 om at UiOs IT-ressurser skal anvendes effektivt og offensivt taler for en kombinasjon av sentrale løsninger og løsninger distribuert til fagmiljøene. En balanse må finnes, og organiseringen må ha en innebygget dynamikk. Helhetlig og strategisk forvaltning av denne strukturen er en nøkkelfaktor, og suksess handler om tydelig ledelse og eierskap/forankring på alle nivåer i organisasjonen.

eInfrastruktur, ressurser og tjenester, som i sine egenskaper dekker behov som er felles for store deler av universitetet, er mest effektivt utnyttet gjennom en sentralisert organisering. Eksempler her er felles tungregneanlegg og ressurser for håndtering og sikker forvaltning av forskningsdata, men også kompetanse av generisk karakter som kan benyttes av alle bør ligge sentralt. Flere fellestjenester blir ivaretatt gjennom UiOs deltakelse i den nasjonale ordningen for eInfrastruktur til forskning, i andre tilfeller gjennom UiOs deltakelse i lignende internasjonale prosjekt, og i noen tilfeller av UiO selv.

Nedenfor operasjonaliserer vi visjonen gjennom tre tiltakstyper:

- Oppbygging av IT-infrastruktur for tjenester og teknologi for forskning
- Etablering av kompetanse-hub'er og andre samarbeidsløsninger
- Styrking av USITs kapasitet, bredde og beredskap for prosjektstøtte

4.1 IT-infrastruktur: tjenester og teknologi for forskning

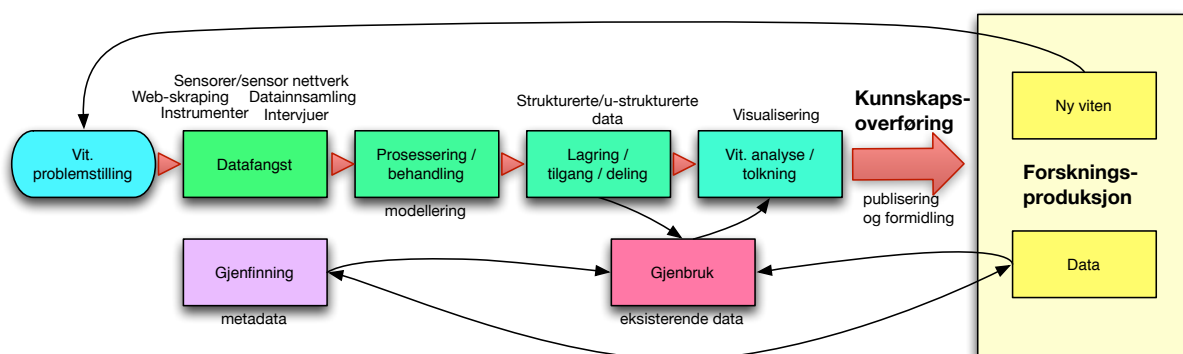
Forskningsdata

Praktisk talt all forskning involverer data, av mange typer og på ulike steg i forskningsprosessen, helt til formidling og undervisning. Det genereres enorme mengder

data, både som resultat av forskning eller som mulig kildemateriale for forskning. Det er allerede etablert at simuleringer, beregningsvitenskap, er et tredje paradigme innen vitenskapelig metode, ved siden av teoretisk utledning og empiri. Bl.a. Jim Gray og flere andre snakker om data-intensiv vitenskap som et *fjerde* paradigme. Gordon Bell m.fl. oppsummerer dette i artikkelen *Beyond the Data Deluge*.² Det handler om ny viten og kunnskap gjennom analyse av store datamengder, og paradigmat gir opphav til en egen profesjon: data scientist.

Forskning drives i stadig større omfang av tilgangen på store datamengder, og datadrevet forskning har blitt en vesentlig del av hverdagen i tilnærmet alle fagmiljøer i hele den vitenskapelige bredde. Begreper som “big data”, “information driven society” og “open data access” går igjen, og de henviser til hva store datamengder i ulike former betyr og kan bety for næringsutvikling spesielt og samfunnsutvikling generelt. Innen forskning (og utdanning) blir data i digital form stadig viktigere, og det skjer i tilnærmet hele den vitenskapelige bredden. Innen naturvitenskap, medisin og teknologi har det vært slik lenge, men de siste årene har dette slått gjennom innenfor humanistiske og samfunnsvitenskapelige fag, og med en kraft det var vanskelig å forutsi for kun få år siden. Bedre og billigere målemetoder (sensorer) som fanger data i stor hastighet fra ulike kilder, og kraftige motorer for beregninger i stor skala genererer en dataflom som gir både utfordringer og muligheter. Det samme gjelder menneskelig aktivitet på internett.

Figur 2 illustrerer den datasentriske forskningsprosessen. Forskningskvalitet og -produksjon vil lide dersom en eller flere av elementene i denne prosessen ikke er tilstrekkelig ivaretatt. Både de enkelte forskningsprosjekt, så vel som rammene disse jobber innenfor, eInfrastrukturen, må innrettes for effektiv bruk av data. Fordi forskningsdata, med tilhørende infrastruktur og tjenester, transenderer de enkelte forskningsprosjekt som genererte dem, har IT i forskning, ved UiO og i større sammenheng, et særskilt ansvar for å understøtte elementene i den datasentriske forskningsprosessen.



Figur 2: Data i forskningsprosessen

² Bell, G., Hey, T., Szalay, A., Computer science. beyond the data deluge. Science 323, 5919 (2009), 1297–8

Hvor kommer data fra? Kildene for data kan grovt sett deles i tre hovedkategorier:

- Instrumenter og sensorer (eller målesystemer), som er en eller annen innretning som fanger data fra naturen eller et menneskeskapt system. I dette bildet kan f.eks. sosiale medier ses på som en redskap for fangst av data.
- Beregninger som utføres av datamaskiner er i seg selv en datagenererende aktivitet og står for en stor del av de data som brukes i forskningen.
- Arkiver eller databaser der data er samlet for et formål brukes i stadig større omfang også til andre formål, oftere og oftere integrert med data fra beregninger og sensorer.

Videre er det et stort og økende behov for integrasjon av data fra ulike kilder, inkludert tilgjengelige datakilder fra store deler av verden. Innhenting av data omfatter behovene for å kunne effektivt samle inn og registrere nye data. For eksempel er det behov for:

- Innsamling av surveydata fra internett og mobile plattformer, som ivaretar behovene for personvern, ideelt sett på en måte som greit kan forklares til respondenten
- Utvikling og implementering av grensesnitt og programmer for eksperimenter som henter inn data
- Automatiserte metoder for tekstanalyse, med grenseflate mot maskinlæring og kunstig intelligens
- Automatiserte metoder for innhenting av data fra internett

Av elementene i den datasentriske forskningsprosessen som i hovedsak må understøttes av prosjektuavhengige mekanismer er:

- Arkivering / preservering av forskningsdata med tilstrekkelige og strukturerte metadata
- Mekanismer for søk og gjenfinning
- Lett tilgjengelige mekanismer for deling av ulike typer forskningsdata med klare, enkle retningslinjer for gjenbruk (bruks- og opphavsrettigheter)

Det er også et overordnet behov for å systematisere, forenkle og sammenfatte prosedyrer og systemer for innhenting av data. Siden innhenting og integrasjon av data normalt er prosjektbasert, er denne kunnskapen ofte lokal og fragmentert. Dette gjelder spesielt prosedyrer for ivaretagelse av personvern for alle typer data (surveydata, Intervjuedata, observasjonsdata, videodata, loggdata og tilsvarende). Videre er det behov for oversikt over og tilgang til "utstyrspool" og tekniske løsninger for innhenting av ulike typer data, for eksempel surveydata via internett, loggdata via internett, tekstdata, eller videodata (opptaksutstyr og tekniske løsninger), samt løsninger og systemer for innhenting av data via sosiale medier.

Store datamengder er krevende å håndtere også med tanke på overføring. Egne løsninger kreves når datastrømmer og -filer blir så omfattende at vanlige overføringsmekanismer, som web eller e-post, ikke strekker til. Følgelig er det behov for tilstrekkelig nettverkskapasitet og tilpassede overføringsmekanismer. Når de aktuelle data er noenlunde statiske, kan problemstillingen løses med replikering av data lokalt. Dette har fordelen at man i tillegg til rask tilgang til data er med på å sikre data fra katastrofale hendelser, men ulempene er dataintegritet (er data de samme som originalen til enhver tid?) og behov for lagringsressurser.

For håndtering av eksisterende dataserier og samlinger er det behov for blant annet tjenester og kompetanse for databaser og statistikk. Det er behov for støtte til digitalisering av flere ulike typer data, gjenstander og kildemateriale. Allerede digitaliserte data er varierte, for eksempel lyd- eller bildemateriale og 3D-scans, og det er behov knyttet til spesiell programvare eller løsninger.

Andre behov grenser nært opp til administrative oppgaver. Ulike former for LIMS, "Laboratory Information Management System" for sporing av prøver og analyser på laboratorier, er det økende behov for. Likeledes verktøy og løsninger for samarbeid, publisering og prosjekt- og økonomistyring. Noen av disse behovene har sin årsak i utfordringer med datadrevet forskning, for eksempel store mengder prøver eller pasienter i medisinske forskningsprosjekt. Andre bunner på økende fokus og krav fra finansierer og administrativ linje om styring og rapportering. Et konkret eksempel på det siste er Norges forskningsråds leiestedsmodell³.

Lagring, arkivering, gjenfinning – Open Access

For å dra nytte av de store datamengdene i forskning, og for å kunne reprodusere og verifisere andres resultater, og bygge videre på disse uten å gjenta dyre eksperimenter eller undersøkelser, er det et fundamentalt behov for strukturert tilgang til utvalgte, eksisterende forskningsdata. Forskningsmyndigheter, så vel som forskerne selv, har i senere år fått bevissthet rundt tematikken, og det er nå mange steder en tydelig holdning og politikk om åpen tilgang til og deling av forskningsdata.

Forskningsrådet har utarbeidet en policy for tilgjengeliggjøring av forskningsdata fra offentlig finansiert forskning⁴, og krever også at det redegjøres for hvordan data skal håndteres i prosjektsøknader ("data management plan").

Dermed er det behov for systemer som sikrer gode og hensiktsmessig løsninger for lagring, indeksering (gjenfinnbarhet) og langtidslagring av data (arkivering) som legger til rette for gjenbruk. Slike systemer må ivareta:

- Krav til et minimum av overordnede metadata (meta-metadata)
- Krav til sikker lagring og personvern hensyn av ulike slag
- Mulighet for felles bruk / gjenbruk av data (især en utfordring for kvalitative data)
- Mulighet for replikasjonsstudier
- At de understøtter tilgang til tidligere analyser allerede gjort på dataene

I stor grad er dagens praksis, i den grad vi kan kalle det praksis, at måten dette håndteres på i beste fall er prosjektspesifikt, i verste fall personbasert og/eller privatisert – noe som bidrar til at nye prosjekt ofte starter på bar bakke og finner egne løsninger slik at problematikken

³ Leiestedsmodellen, eller TDI-modellen, er etablert for å harmonisere og synliggjøre totale kostnader i forskningsprosjekter i UH-sektoren. TDI-modellen er beskrevet på sidene til Universitets- og høyskolerådet: http://www.uhr.no/ressurser/temasider/totalkostnadsmodell_-_tdi

⁴ Forskningsrådets policy for åpen tilgang til forskningsdata er beskrevet på http://www.forskningsradet.no/no/Artikkel/pen_tilgang_til_forskningsdata/1254001013535?lang=no

forsterkes. Følgene er at det skapes hindrer for effektiv gjenbruk og datautnyttelse, samt at det blir vanskelig å etablere effektive søk- og arkiveringsfunksjoner.

Det er nedsatt en egen arbeidsgruppe som skal utrede lagring og deling av forskningsdata ved UiO. Arbeidsgruppen ledes av forskningsdekan Svein Stølen og startet sitt arbeid i oktober 2014 og skal levere sin rapport våren 2015. Arbeidet skal ses i sammenheng med og bygge på herværende rapport.

Håndtering av sensitive data

Forskning med sensitive data vurderes som et spesielt strategisk område for UiO fremover. Forskning som involverer sensitive data er underlagt lover og regler (NSD, Personvernombudet, REK) som det må tas hensyn til, i tillegg til at det krever utvikling og bruk av egne tekniske løsninger. I takt med datautviklingen beskrevet over er det derfor et vesentlig behov for IT-løsninger for å kunne håndtere og forske med slike data. UiO er per i dag i førersetet når det gjelder løsninger for håndtering av sensitive data, men det kreves ressurser for å videreutvikle den løsningen som er etablert samt øke kapasiteten da behovet er stort og økende.

USIT har i samarbeid med flere partnere etablert basis eInfrastruktur for sensitive forskningsdata, Tjenester for Sensitive Data (TSD, se kapittel 5). For flere forskningsprosjekt er det nå mulig å realisere forskning som tidligere var tilnærmet utenkelig grunnet kostnader ved etablering av egne løsninger. Dette er et godt eksempel på den forskjell strategisk satsing på eInfrastruktur utgjør.

Selv om en sentral plattform er etablert med TSD, er det fortsatt behov for utvikling av de tekniske løsningene som plattformen er bygget på, og ikke minst selve tjenestespekteret *innenfor* plattformen. På grunn av den store pågangen er det allerede en ressursutfordring innenfor eksisterende rammer og behov for tilførsel av ressurser til drift, utvikling, infrastruktur og støttetjenester.

Fremover bør derfor UiO, sammen med partnere, satse på å videreutvikle eInfrastrukturen for sensitive data. Ikke minst på grunn av de samfunnsmessige gevinstene den medisinske og helsefaglige forskningen utgjør, men også med tanke på at UiO her har et fortrinn, innen innovasjon og som internasjonal partner.

Regnekraft

Ved alle fakultetene forskes det med større datamengder, med analytiske eller statistiske modeller som øker i kompleksitet med både datamengdene og større muligheter for regnekraft. Behov for regnekraft er spesielt knyttet til simuleringsmetoder, men også til analyse og estimering av komplekse modeller på store datamengder. Det er også et overlapp her til behovet for tilpassing og effektivisering av programmer, digital arbeidsflyt og analyseverktøy nevnt under.

Analyse og behandling av spesielle typer data, for eksempel video, krever spesiell regnekraft (for eksempel grafikkprosessorer), og må ofte være i nærhet av datakildene. Det er derfor behov for kompetanse og støttetjenester for slikt utstyr.

Behovene ovenfor går igjen også for forskning som involverer sensitive data.

Tilgang til tilstrekkelig regnekraft vil i hovedsak komme gjennom den nasjonale eInfrastrukturen, og UiO må være en aktiv og pådrivende partner i dette samarbeidet for å sikre at strategiske hensyn ivaretas. I noen tilfeller kan det være nødvendig med lokale investeringer, for eksempel for å ivareta spesielle lokale behov, men det må være et mål å få også denne type regnekraft ivaretatt gjennom det nasjonale systemet.

4.2 Kompetanse-hub'er og andre samarbeidsformer

Hub – node organisering

Det finnes flere forskningsområder og fagfelt som benytter de samme IT-metodene i sin forskning. Her kan effektiv og offensiv utnyttelse av ressurser oppnås ved at aktørene, enheter, institutter, sentre eller forskergrupper, går sammen om å organisere og finansiere ressurser. En slik modell kaller vi gjerne hub – node (alternativt sentra – drabant) fordi hub'en konsentrerer og forsterker ressursene for nodene den betjener, samtidig som den er et forskningssentrum i seg selv. Kompetanseressurser i en hub kan frikjøpes i kortere eller lengre tidsrom og benyttes i forskningsprosjekt, og dette vil i mange tilfeller være en viktig finansieringskilde for hub'er. Avansert bruk av informatikk i forskningen betyr også at de informatiske metodene i seg selv kan ligge i forskningsfronten og trenge egne vitenskapelige utviklingsprosesser. For å lede dette forskningsarbeidet og den felles kompetanseutviklingen i informatikk vil det oftest være behov for en informatiker med relevant vitenskapelig bakgrunn som faglig leder av hub'en. En viktig grunnlagsinvestering vil derfor være å tilby vertsmiljøet for hub'en en slik stilling, eller andel derav, for en avgrenset periode.

Hub'er medvirker også til bredere kompetansedeling gjennom seminarserier, journal clubs, og brukerstøtte (helpdesk). Hub'er har i tillegg en kompetanseutviklende funksjon ved at den bygger kompetanse innenfor de aktuelle IT- områdene. Hub'er vil selv ha ansvaret for koordineringen av aktiviteter og ressurser, og spesielt for kompetanseoverføring mellom hub og noder.

Hub'er bygger opp kompetanse som er generisk for hele strukturen, og er med på å sikre god tilgang til eInfrastrukturelle ressurser som leveres av USIT (og den nasjonale eInfrastrukturen). I hub – node-modellen er nodene typisk enkeltforskere og forskningsgrupper som trenger eInfrastrukturelle tjenester. Fordi ressurser deles (pooles) sørger hub – node-strukturen for effektiv utnyttelse av ressurser og kompetanseoverføring mellom ulike deler av UiO innenfor et gitt tema (bioinformatikk, språkdata, videodata, osv.). I tillegg til sine noder er det naturlig at en hub knytter spesielt sterke bånd til fellestjenestene. På denne måten oppnås effektiv og god utnyttelse av fellesressursene uten at fagspesifikke føringer i for stor grad belastes disse ressursene. Samvirke med fellesressurser (USIT) vil

kunne defineres som en node i en slik struktur, på linje med noder ute i forskningsmiljøene, og slik integreres tett med hub'ene.

Hub'ene kan bli en del av forskerutdanningen ved at nye stipendiater tilbys poenggivende introduksjonskurs i metoder og verktøy. De nasjonale forskerskolene følger denne tankegangen.

Et eksempel på en hub – node organisering er den planlagte hub'en for bioinformatikk. Organiseringen her kan fungere som en mal for andre satsinger der forskning, utdanning og utvikling av eInfrastrukturelle løsninger må kombineres for å få til optimal utnyttelse av de ressurser som er nødvendige for å utføre forskning på høyt internasjonalt nivå. Gjennom hub'en for bioinformatikk vil det bli etablert et strukturert utdanningstilbud som vil sørge for at det utdannes individer med høy kompetanse på området. Notat om bioinformatikk-satsingen finnes som eget vedlegg til denne rapporten.

Etablering av hub'er krever aktiv kartlegging av forskningen ved UiO for å identifisere tverrfaglige eInfrastrukturelle utfordringer som kan danne utgangspunkt for en hub. Dette kan gjøres gjennom tverrfakultære strukturer eller nettverk, for eksempel ved at hvert fakultet identifiserer personer med ansvar for oversikt over fakultetets forskning med spesielt fokus på eInfrastruktur, metoder og utfordringer. Når en potensiell hub er identifisert trengs det tilrettelegging og koordinering for å definere og etablere en felles plattform for hub'en. Denne vil så være utgangspunkt for søknader og innhenting av finansiering, og endelig etablering av selve hub'en. Satsingsmidler vil virke som insentiver for etablering av hub'er og bør tilføres for en periode på mellom tre til fem år. Bærekraftige hub'er bør være selvfinansierende etter en etableringsfase.

Andre samarbeidsformer

Enkeltstående aktører, fortsatt i betydningen enheter, institutter, sentre eller forskergrupper, med egne, fagspesifikke behov, kan også dra fordel av tettere samarbeid med fellesressursene, selv om ressurser til IT i forskning organiseres lokalt eller innenfor en hub – node-struktur. Det kan dreie seg om utnyttelse av profesjonell IT-kompetanse, innkjøpskompetanse eller stordriftsfordeler, samt mer effektive støtte- og utviklingsfunksjoner rettet mot egne forskere. Fordelen er at nærhet til den enkelte forsker ivaretas samtidig som avstanden til fellestjenestene, som mange forskere opplever, reduseres. IT til forskning ved UiO bør organiseres for å legge til rette for dette.

Det dynamiske i forskningen betyr at vi aldri helt vet hva vi mangler, hvilke behov vi må forsøke å dekke opp rundt neste sving. Ved å gi eierskap til IT i forskning til hub'er og enkeltaktører der det ligger til rette for dette, vil dynamikken bli bedre ivaretatt. Fellesressurser vil i større grad være av generisk natur, og således ikke kunne være like dynamiske – selv om endrings- og tilpassingsevne er viktig også her. Med en tettere integrering med forskningen unngår vi at fellesressursene blir sett på som en passiv leverandør av tjenester til IT i forskning fremfor en integrert del av forskningen.

4.3 Styrking av USITs kapasitet, bredde og beredskap for prosjektstøtte

Prosjekt og utviklingsstøtte

IT, det være seg teknologi, metoder eller implementering av løsninger, er en vanlig utfordring i forskningsprosessen. Allerede i designfasen av et forskningsprosjekt er det kritisk at forskerne har den nødvendige kunnskap om IT-komponenter som inngår. Når forskerne ikke har den nødvendige forståelse, eller mangler informasjon om muligheter, kan et forskningsprosjekt ende opp med å levere resultater av dårlig kvalitet, overskride tids- eller økonomiske rammer, eller rett og slett feile. Det som mangler er ofte bistand i kortere tidsrom, gjerne på ulike stadier i forskningsprosessen. Slik bistand kan komme fra den sentrale eInfrastrukturen eller et kompetansesentra, som for eksempel hub – node-strukturen, skissert foran.

Det er nokså vanlig at forskningsprosjekt involverer utvikling av systemer eller programvare for gjennomføring av prosjektet. Det er ofte ikke tilgang til slik kompetanse innenfor prosjektet og derfor nødvendig å involvere ekstern hjelp. Ved å satse på en sentral eInfrastruktur hvor slike ressurser finnes, inklusive menneskelige ressurser med egen forskerbakgrunn, kan kritisk kompetanse og utviklingsressurser deles mellom universitetets forskningsprosjekter. Prioriteringsorganet som her foreslås kan styre tildeling av slike ressurser, koordinert med den sentrale eInfrastrukturen hvor slike ressurser må finnes, dvs. ved USIT. Slik støtte kalles ofte avansert brukerstøtte, og tildeles også etter søknad fra den nasjonale eInfrastrukturen. Det vil også være naturlig at de enkelte prosjekt bidrar til dekning av merkostnader knyttet til støtten.

Tilpassing og effektivisering av programvare

Beregningorientert forskning blir stadig mer krevende kompetansemessig. Aktuelle utfordringer kan være utnytting av parallelle arkitekturer, programmering for spesielle arkitekturer som akseleratorer, behov for kompleks arbeidsflyt (workflows), og effektiv håndtering av store datamengder. Fordi dette krever spesialkompetanse, er det behov for støttefunksjoner, enten forskerne utvikler programvare selv eller benytter eksisterende løsninger.

For å takle noen av disse utfordringene er det etablert egne kurs og studieretninger, men en forsker i et ikke-relatert felt har ikke tid, og ofte ikke interesse, til å bygge egen spisskompetanse. Behovet er ofte også kortvarig i en avgrenset del av forskningsprosessen.

Behov for støtte innen modellering/programmering omfatter den praktiske implementeringen av teoretiske eller statistiske modeller på eksisterende data. Her er behovene først og fremst på programmeringssiden, og dreier seg i hovedsak om bistand og praktisk hjelp med programmering og effektivisering, vanligvis for analyse av kvantitative data, men enkelte institutter melder også om behov for automatisert analyse av kvalitative data. På dette området er det både behov for direkte støtte, men kanskje like mye støtte til kompetansebygging hos den enkelte forsker og i forskermiljøene.

Ser vi på typen henvendelser til USIT i sammenheng med forskningsrelatert programvare er det i stor grad forespørsler om installasjon/tilgang til programvare og tilpassinger. Det er nesten like mange applikasjoner (programmer) og programbibliotek som det er forskningsproblemstillinger. Programvare endres og oppdateres til stadighet, og i tillegg kommer programvaren i ulike versjoner for ulike arkitekturer og versjoner av operativsystem. Behovet for støtte i denne sammenheng er stort og krevende. Også her vil den skisserte hub-node strukturen kunne spille en sentral rolle.

Tilgang til programvare omfatter også lisensiert og proprietær programvare. Mekanismer for finansiering og administrasjon av tilgang og lisenser kommer da i tillegg til tiltakene beskrevet i avsnittet ovenfor.

For mye brukt programvare er det behov for organisert kursing, sentral dokumentasjon og hjelpetjenester. Ansvar og ressurser for dette må legges ved fellestjenestene, og ved hub'er der det er naturlig.

Brukergrensesnitt, digital arbeidsflyt og analyseverktøy

Brukergrensesnitt, digital arbeidsflyt og analyseverktøy – inkludert opplæring i disse – omfatter behov rundt den praktiske implementeringen av ulike grensesnitt, arbeidsflater, programmerings-, behandlings- og analyseverktøy.

Analyse- og dataorientert forskning er etter hvert så kompetansekrevende at det er behov for brukergrensesnitt som forenkler og effektiviserer oppgavene. Slike behov har blant annet resultert i utvikling av web-baserte arbeidsflater og portaler. Et eksempel er tematiske beregningsportaler som gjør det mulig å gjøre avanserte analyser med beregningsressurser uten å ha kjennskap til terminalbruk/kommandolinjer og annet for bruk av regneark. Slike portaler har flere fordeler:

- Forenklet brukergrensesnitt
- Mulighet for å sette sammen digital arbeidsflyt og lagre disse
- Mulighet for å dele digital arbeidsflyt med andre
- Abstraksjon av kompleksitet forbundet med effektivisering av kode og bruk av ulike typer beregningsressurser og arkitekturer
- Lett tilgang til dokumentasjon og annen tematisk informasjon
- Lett tilgang til delte datakilder og mulighet for å dele data
- Administrative funksjoner for bokføring (resource accounting) og rapportering
- Mulighet for styring av tilgang og autentisering

En viktig utfordring ved denne typen portaler er at løsningene lett kan bli statiske, knyttet til programvaren som eksisterer når portalen bygges opp, og dermed miste dynamikken som er så viktig for at forskningen skal kunne drives framover.

5. Status IT i forskning ved UiO anno 2014

Formålet med dette kapitlet er å gi en tilstandsbeskrivelse av IT i forskning ved UiO. Den er ikke uttømmende eller fullstendig, men tilstrekkelig til å relatere dagens situasjon til visjonen, og dermed mulige forbedringer og alternative organisasjonsformer.

5.1 Fellestjenester

Universitetet i Oslo, med sine ca. 27 000 studenter og 6 000 ansatte, er støttet med en sentral IT avdeling, USIT, og lokal IT-støtte ved fakultetene, museene og enkelte enheter. USIT har ansvar for å levere infrastrukturen og IT-tjenestene og -løsningene UiO trenger for å nå sine mål for utdanning, forskning, formidling og innovasjon. Administrative støttefunksjoner er også en viktig del av dette.

En stor del av USITs ressurser går med til å levere infrastruktur og driftstjenester som danner basis for hele spekteret av aktiviteter ved universitetet. Universitetets to største primærvirksomheter, utdanning og forskning, har egne behov for IT-støtte, med egne tjenester levert av USIT. USIT anslår selv at av driftsbudsjettet går 40% til forskning, 30% til utdanning og 30% til administrative tjenester. Dvs. at ca. 50 millioner årlig av driftsbudsjettet går til forskning.

Fellestjenester til forskning leveres på grunnlag av basis IT infrastruktur og tjenester tilsatt tjenestespesifikke komponenter. En fellestjeneste som beregningstjenesten er bygget på basiskomponenter som maskinromsdrift, nettverk- og lagringsinfrastruktur mv. USIT er i dag organisert i to underavdelinger, en for drift og en for virksomhetsutvikling. I *Underavdeling for virksomhetsutvikling* (UAV) finner vi seksjoner som støtter de ulike delene av primærvirksomheten, herunder *Seksjon for IT i forskning* (ITF) og *Seksjon for IT i utdanning* (ITU). Selv om ITF er ansvarlig for beregningstjenesten er denne leveransen igjen avhengig av tjenester fra resten av USIT, spesielt *Underavdeling for IT-drift*.

Seksjonen for IT i forskning gir råd om digitale løsninger i forskning og sørger for løsninger, tjenester og kompetanse til universitetets forskere og forskningsmiljøer innen metodiske hovedretninger, inkludert museenes forskning og samlingsdatabaser. Seksjonen bidrar til UiOs strategi for området IT i forskning og representerer UiO i nasjonale og internasjonale prosjekter og program som omfatter e-infrastruktur og IT-tjenester til forskning. ITF utgjør omlag 24 årsverk og består av tre grupper:

- Gruppe for forskningsinfrastruktur (FI)
- Gruppe for forskningstjenester (FT)
- Gruppe for datafangst og samlingsforvaltning (DS)

Gruppe for forskningsinfrastruktur (FI) har ansvar for utførelse av oppgaver innen nasjonale e-infrastruktur tjenester for beregninger og lagring av forskningsdata. Aktivitetene omtales som bidragsvirksomhet og er på oppdrag av Norges Forskningsråd via UNINETT Sigma2 (se neste avsnitt). Gruppen har også ansvar for UiOs engasjement i andre nasjonale, nordiske og europeiske satsninger, samarbeidsprosjekt og tiltak på området. Alle årsverk i FI er

bundet i bidragsvirksomheten og regnes derfor ikke som frie, dvs. de kan ikke uten videre (om)disponeres til andre formål.

Gruppe for forskningstjenester (FT) har ansvar for løsninger, IT-tjenester, rådgiving og avansert brukerstøtte til universitetets forskere og forskningsmiljøer innen sentrale metodiske hovedretninger, primært innen tekniske, naturvitenskapelige og medisinske fag. Som for FI er gruppens stillinger i stor grad bundet opp til bidragsvirksomheten.

Gruppe for datafangst og samlingsforvaltning (DS) har ansvar for løsninger, IT-tjenester, rådgiving og avansert brukerstøtte til universitetets forskere og forskningsmiljøer innen sentrale metodiske hovedretninger, primært innen humaniora og samfunnsvitenskapelige fag. Gruppen har ansvar for tjenester for nettbasert datainnsamling, og utvikler databaser og applikasjoner for forvaltning og tilgjengeliggjøring av strukturerte data (samlinger). Et viktig oppdrag innenfor DS er samarbeidstiltaket MUSIT (Universitetsmuseenes IT-organisasjon), en sammenslutning av landets universitetsmuseer. Gjennom samarbeidstiltaket støtter gruppen forskningen ved universitetsmuseene og drifter universitetsmuseenes digitale kultur- og naturhistoriske samlinger og arkiv.

I den grad det er mulig å snakke om en kompetanseprofil til fellestjenestene for IT i forskning vil det være at den er reflektert i de tilgjengelige tjenestene, og i høy grad er basert på forskererfaring og i mindre grad ren IT-teknisk kompetanse. Nærmere 2/3 av de ansatte i Seksjon for IT i forskning har egen forskerbakgrunn (innehar graden PhD.). Fagfeltene disse ansatte kommer fra spenner vidt selv om det er flest med bakgrunn fra naturvitenskapelige fagområder. Erfaringen er at det er en stor fordel med egen forskererfaring i funksjonen IT-støtte til forskning. Det gjør det lettere å sette seg inn i forskernes problemstillinger og kommunisere med forskere. Å benytte ansatte uten primært IT-faglig bakgrunn i et IT-teknisk krevende område er mulig fordi de underliggende grunnfunksjonene, eksempelvis nettverk og maskinromsdrift, leveres av andre deler av USIT med IT-faglig kompetanse på disse områdene. Denne modellen utnytter ressurser bedre og resulterer i tjenester av høyere kvalitet sammenlignet med IT-sentre som leverer IT til forskning alene.

På et hensiktsmessig nivå kan porteføljen av dagens felles IT-tjenester og støtte til forskning grupperes i følgende hovedkategorier:

- Beregningstjenester
- Lagringstjenester
- Tjenester for sensitive forskningsdata
- Prosess-støtte, inklusive datainnsamling og visualisering
- Applikasjonsstøtte
- Metodestøtte
- Generell brukerstøtte
- Avansert brukerstøtte
- Kurs og undervisningsstøtte

Innenfor hver av disse hovedkategoriene finnes flere enkelt-/spisstjenester avhengig av detaljingsnivået. I det følgende gis kun en oppsummerende beskrivelse av hver hovedkategori.

Beregningstjenester betegner kategorien av tjenester som involverer beregningsressurser av ulike slag. Her inngår UiOs to større regneanlegg, Abel og Colossus, henholdsvis for data av ikke-sensitiv og sensitiv natur. Det er ulike tekniske løsninger for beregninger, eksempelvis for minnekrevende applikasjoner, parallelle applikasjoner eller store mengder data. Måten brukerne benytter beregningsressursene på varierer fra tjeneste til tjeneste, fra kommandolinje til tematiske web-portaler.

Behovet for regnekraft av de fleste typer er i dag i stor grad tilfredsstilt gjennom den nasjonale ordningen sammen med muligheten for tilgang til PRACE-anlegg, kombinert med UiOs egne og prosjektspesifikke ressurser. Etersom behovene for regnekraft øker vil ressurs situasjonen forverres og prosjekter vil måtte konkurrere og redusere sin aktivitet. Sigma2 vil, på vegne av NFR sammen med de fire eldste universitetene, investere i nye regneressurser når behovene tilsier det – se nedenfor.

Spesielle, prosjektspesifikke behov for regnekraft krever delvis egne investeringer. Dagens praksis er at prosjektene finansierer investeringer og etableringskostnader, mens USIT foretar og dekker innkjøp, installasjon og drift. Denne praksisen har fungert rimelig godt, men oppgavene for USIT kommer i tillegg til USITs øvrige oppgaver og utfordrer ressurs situasjonen. Ved vedvarende (lokale) prosjektbaserte ressursbehov relatert til infrastruktur praktiseres derfor frikjøp.

Lagringstjenester omfatter på samme måte flere ulike tjenester for datalagring og -håndtering. Tjenester er rettet mot ulike former for forskningsdata, for eksempel strukturerte data, ustrukturerte data, lyd- og videodata. Også her er ressursene skilt av lovpålagte og sikkerhetsmessige grunner når det dreier seg om sensitive data. Lagringstjenestene omfatter også backup og arkivtjenester, samt tjenester for oversendelse eller deling av store eller mange filer.

Innenfor enkelte fagområder (for eksempel klima, astronomi), eller store, langsiktige forskningsprosjekt (for eksempel CERN, Human Brain Project), er lagringsfunksjoner og strukturer etablert, eller i ferd med å etableres. Flere program for forskningsdata av mer overordnet karakter er også etablert. Disse ivaretar i ulik grad flere av de nevnte funksjonene, ofte for ulike typer data (for eksempel EUDAT, NorStore, NSD). Det er likevel en stor mengde forskningsdata, som vi gjerne kan kalle "the long tail", som ikke har noe strukturert eller hensiktsmessig tilbud.

USIT har etablert TSD, **Tjenester for Sensitive Data**, en plattform for sikker håndtering av alle typer data med lov- eller selvpålagte krav til informasjonssikkerhet. Felles for slike data er streng autentisering og autorisasjon, for å styre hvem som aksesserer hvilke data og hva som gjøres med disse. Strengt tilgangsregime omfatter ikke bare brukere, men også dem som er operatører. TSD har vært utfordrende å utvikle fordi sikkerhetsaspektet ofte er

uforenelig med brukerønsker og etablerte driftsrutiner. I TSD-løsningen til USIT er dette implementert med fysisk skjermet infrastruktur uten mulighet for informasjonslekkasje. Det som gjør TSD unik, i nordisk og kanskje større sammenheng, er at systemet håndterer flere samtidige prosjekt, med totalt adskilte brukere innenfor rammen av samme løsning. Det er således mulig å tilby tjenestene fra en delt ressurs for sensitive data. TSD-plattformen kan ta i mot og huse sikre ekvivalenter av tjenester som finnes for "vanlige" data, i tillegg til egne nye tjenester muliggjort av løsningen, og muliggjør også sikker informasjonsutveksling mellom systemer innenfor TSD og utenfor. TSD plattformen fordrer eget driftsmiljø og -rutiner, involverer teknologi som krever egen og dedikert kompetanse, men er tilgjengelig fleksibel og generell nok til at den utgjør en sentralisert, sikker, skalerbar og oversiktlig løsning for sensitive data.

Prosess-støtte, inklusive datainnsamling og visualisering er støttefunksjoner og tjenester til ulike deler av forskningsprosessen. Det finnes tjenester for datainnsamling (spørreskjema, online journaler eller dataregistrering for feltarbeid), integrasjon med instrumenter og sensornettverk, datahåndtering m.v. ITF gir også råd i søknadsprosesser og bidrar ofte direkte i eksternt finansierte prosjekter.

Applikasjonsstøtte er installasjon, tilgjengeliggjøring, tilpassing, vedlikehold, kompetanse og lisenshåndtering for et bredt spekter av programvare i ulike sammenhenger, for eksempel for beregninger eller analyse av data.

Metodestøtte involverer rådgiving, kurs og kompetanse innen metodiske hovedretninger. Støtte innen statistikk og kvalitative metoder er vel etablerte og kjente tjenester, men det gis også støtte i anvendelse av ulike numeriske metoder og algoritmeutvikling (se også avansert brukerstøtte).

Generell brukerstøtte er betegnelsen på generell direkte støtte til brukere i anvendelse av IT i forskningen eller i forbindelse med bruk av infrastruktur og tjenester, for eksempel beregningstjenesten. Henvendelser kommer gjerne inn gjennom en av flere tematiske virtuelle helpdesker. Det er en egen vaktordning for førstelinjesupport og fordeling av henvendelser, og et stort antall henvendelser håndteres daglig. Brukerstøtten er svært mangfoldig, også når det gjelder kompleksitet og ressursbruk. En henvendelse, eller en sammenhengende serie med henvendelser, som er svært ressurskrevende kan bli klassifisert som avansert brukerstøtte.

Avansert brukerstøtte er dypere brukerstøtte som er kompetanse- og ressurskrevende, og som i mange tilfeller medfører utviklings- eller støtteoppgaver som varer i uker eller måneder. Avansert brukerstøtte gis etter avklaring omkring ressursbruk og -tilgang og kompensasjon for denne. Tidligere var parallellisering eller utvikling av vitenskapelig kode en vanlig form for avansert brukerstøtte. Nå er det mer og mer vanlig at avansert brukerstøtte gis i forbindelse med etablering, utvikling og bruk av avanserte it-systemer og tjenester.

Kurs og undervisningsstøtte omfatter kurs i regi av ITF, rettet mot brukere av ulike tjenester, applikasjoner og verktøy, støtte til UiOs undervisning der dette er relatert til

forskningstjenestene eller til kurs i regi av forskermiljø. ITFs egne kurs i bruk av beregningstjenesten og aktuelle verktøy holdes to ganger i året og er svært populære blant nye og uerfarne brukere. I tillegg kjøres kurs rettet mot grupper av brukere ved behov eller etter forespørsel fra spesifikke forskermiljø/-grupper. Støtte til undervisning involverer ofte tilgang til regneressurser eller lignende for studenter i forbindelse med gjennomføring av undervisningen.

5.2 Nasjonal eInfrastruktur

I Norge er det etablert et nasjonalt samarbeid innen eInfrastruktur for forskning. Norges Forskningsråd sammen med de fire eldste universitetene er partnere og finansierer ordningen, men hele UH- og institutt-sektoren kan benytte seg av eInfrastrukturen etter gitte kriterier. Ressurser tildeles etter vitenskapelig vurdering og omfatter regnekraft, datalagring og støttetjenester.

Ordningen administreres og koordineres gjennom et heleid datterselskap av UNINETT, UNINETT Sigma2 AS. Denne konstruksjonen muliggjør avtaler mellom partene. Sigma2 gjør investeringer og gir oppdrag innen drift og brukerstøtte i forbindelse med ressursene. Jevnlig finansieres også avansert brukerstøtte hvor støtten gis av personell ved en av de fire driftssentrene. Videre representerer Sigma Norge og koordinerer deltakelse i internasjonale eInfrastruktur prosjekt (se neste avsnitt).

Den nasjonale eInfrastrukturen er dominert av to prosjekt; Notur, for beregningsressurser og tilhørende tjenester, og NorStore for lagringsressurser og -tjenester.

Notur, eller *Norwegian Metacenter for Computational Science*, forvalter og tilbyr eInfrastruktur for beregningsbasert forskning. Enkeltforskere eller forskergrupper tildeles regnetid etter søknad og faglig vurdering på ett eller flere av anleggene. Hvilket anlegg en tildeling gjelder er uavhengig av hjemmeinstitusjon ettersom det er ulike egenskaper ved anleggene som gjør dem hensiktsmessige for ulike typer oppgaver (applikasjoner). Det er søknadsrunder to ganger i året og tildelingen gis fra et halvt til to år om gangen. Mottaker må rapportere jevnlig til Notur på bruken av ressursene. Tildelingen gjøres av Ressursfordelingskomiteén (RFK), et uavhengig fagfellepanel. UiOs forskere utgjør med god margin den største gruppen av brukere, både i antall og ressursbruk.

Ved UiO er det regneklyngen Abel som inngår i Notur. Anlegget omfattes av en leveranseavtale mellom UiO og Sigma som bla. spesifiserer stillinger til drift og brukerstøtte.

NorStore-prosjektet dekker eInfrastruktur for forskningsdata, forvaltning og arkivering av slike data. Prosjektet startet først i 2007/2008 med relativt små ressurser, men er på vei til å bli på størrelse med Notur budsjettmessig og i omfang grunnet utviklingen på området. For tiden er det meste av NorStores lagringsressurser ved UiO i forbindelse med AstraStore anlegget. Tildelingen gjøres av RFK på en lignende måte som for tildelingen av regnetid innen Notur. NorStore tilbyr lagringsplass og arkivering, men også et økende antall tjenester for datahåndtering (data management). NorStore har vært med å finansiere utvikling av Tjenester for Sensitive Data (TSD).

Fra 2015 er den nasjonale eInfrastruktur-ordningen organisert annerledes. Bakgrunnen er svakheter ved den tidligere organiseringen hvor regneanleggene ble statisk lagt til hvert av de fire eldste universitetene og investeringer ble gjort mer eller mindre uten strategisk styring på nasjonalt nivå. Det resulterte bla. i at investeringene til dels ble gjort innenfor samme tidsrom, noe som har ført til vekselvis høy grad av overkapasitet, og mot slutten av levetiden til anleggene, underkapasitet. Det viktigste i den nye nasjonale ordningen, **Sigma2**, er kanskje likevel at det blir en kraftig økning av finansieringsnivået. Den langsiktige finansieringen, eller basisfinansieringen, går fra litt under 50 MKr årlig til 75 MKr årlig, eller mer, avhengig av prosjektfinsiering. Av dette bidrar universitetene med 50 MKr, og UiO er både den største bruker av og bidragsyter til systemet, og midlene kommer fra universitetet sentralt inklusive USIT.

I tillegg skal det som tidligere søkes midler fra NFRs infrastrukturutlysninger for jevnlig investeringer i nasjonal eInfrastruktur, dvs. beregnings- og lagringsressurser. Ettersom man har funnet at dagens finansieringsnivå for beregningsressurser er omtrent tilstrekkelig gitt teknologiutviklingen, og i hvert fall hvis investeringer hadde vært strategisk fordelt over tid, vil økningen i finansieringsnivået fra 2015 i stor grad gå mot ressurser for datalagring og -håndtering. Det forventes også at finansieringen av brukerstøtte og andre tjenester vil øke.

I den nye ordningen er den koordinerende enheten også et datterselskap av UNINETT, UNINETT Sigma2 AS, men til forskjell fra tidligere er det strategiske ansvaret for investeringer nå lagt tydelig til selskapet og dets daglige leder. Universitetene og NFR bidrar med sin finansiering, og selskapet skal benytte midlene på best mulig måte for nasjonal eInfrastruktur og forskning. Selskapet står i så henseende fritt til å anvende midler slik at nasjonal forskningsproduksjon og -kvalitet maksimeres over total kostnader.

Det nye selskapet investerer i eInfrastruktur og legger ressursene til driftssentre som gir best vilkår økonomisk, teknisk og med hensyn på utbytte. Sigma2 vil også være eier av infrastrukturen, i motsetning til dagens ordning der universitetene eier anleggene. Leveranseavtaler etableres mellom driftssentrene og det nye selskapet og alle driftskostnader fullfinansieres. Det inngår derfor ikke egenandeler fra driftssentrene. Et universitet kan ha *driftsansvar* for et anlegg uten at anlegget er lokalisert ved universitetet. Brukerstøtte vil på tilsvarende måte fullfinansieres og legges som i dag til universitetene for å være nær forskerne, uavhengig av om universitetene også er driftssenter. Finansieringen som tidligere ble tilført nasjonal eInfrastruktur i form av egenandeler fra universitetene kan dermed i den nye ordningen sees på som å være en del av de direkte overføringene til det nye selskapet, Denne innretningen er med på å synliggjøre alle kostnader forbundet med nasjonal eInfrastruktur.

En følge av den nye organiseringen kan bli at det ikke vil være noen nasjonal infrastruktur, for eksempel en regneressurs, ved UiO (begrepet eInfrastruktur inneholder også menneskelige ressurser, men slike vil alltid være tilstede ved UiO i den nye ordningen, og derfor brukes ordet infrastruktur i denne sammenhengen for å avgrense problemstillingen til maskinvare). Organiseringen av IT til forskning ved UiO må derfor være slik at den sørger for å utnytte den nasjonale eInfrastrukturen på best mulig måte uavhengig av hvor de fysiske ressursene befinner seg. Arbeidsgruppen har tatt spesielt hensyn til dette i sine vurderinger.

5.3 Internasjonal eInfrastruktur

Flere internasjonale eInfrastruktur-prosjekt og aktører er av stor betydning for mange av UiOs forskere. Deltakelse i slike prosjekt gir tilgang til ressurser, tjenester og kompetanse man ellers ikke ville hatt tilgang til. Samtidig er UiO som partner i slike prosjekt, enten det er via individuelle forskergrupper eller en større enhet som USIT, med på å tilføre ressurser, levere tjenester og bidra med egen kompetanse i et større hele. Slik aktivitet er del av IT i forskning, og er med på å fremme forskning så vel som utdanning ved UiO, og bidrar til å gjøre UiO til et internasjonalt og attraktivt forskningsuniversitet.

Det er flere eksempler på internasjonale eInfrastruktur-prosjekt, eller prosjekt med stor komponent av IT i forskning, hvor UiO deltar. Et prosjekt som har pågått lenge er aktiviteten mot CERN og deltakelse i Worldwide LHC Computing Grid (WLCG). UiOs deltakelse her har gitt direkte faglig utbytte og eInfrastrukturelle ressurser. Den faglige aktiviteten er selvfølgelig størst ved Fysisk institutt, men Institutt for informatikk og USIT er også involvert. USIT utfører oppgaver i WLCG på vegne av NFR og Norge, og synergieffektene med øvrig eInfrastruktur og kompetansetilførsel som følge av aktiviteten er store. Andre viktige eInfrastruktur prosjekt som bør nevnes i denne sammenhenger er ELIXIR ("The European life-sciences Infrastructure for biological Information") og CLARIN ("Common Language Resources and Technology Infrastructure").

UiO ved USIT deltar i, og utfører også nasjonale oppgaver innen, FP7 prosjektene PRACE – Partnership for Advanced Computing in Europe, og EUDAT – European Data Infrastructure. Førstnevnte gir våre forskere tilgang til kompetanse og regneressurser av en størrelse som ikke finnes innen den nasjonale eInfrastrukturen, mens sistnevnte gir tilgang til lagringsressurser og -tjenester, men kanskje viktigst, er med på å integrere våre forskningsdata og samlinger med tilsvarende infrastruktur internasjonalt, slik at forskningsdata kan preserveres, gjenfinnes, deles og brukes i annen forskning. Begge prosjektene videreføres i Horizon 2020.

I Nordisk sammenheng er det verdt å nevne NeIC, Nordic eInfrastructure Collaboration, som er et prosjekt under Nordforsk med egen finansiering fra Forskningsrådene i Norden. NeIC investerer ikke i infrastruktur, men bruker insentiver til å fremme Nordisk samarbeid innen eInfrastruktur og koordinerer aktiviteter mellom landene. En viktig oppgave for NeIC er koordinering av den Nordiske Tier-1 aktiviteten mot CERN og deltakelse i WLCG. Forskere kan få hjelp og bidrag fra NeIC til etablering av fellesnordiske samarbeidsprosjekt innen eInfrastruktur. Et eksempel på dette er Tryggve prosjektet, hvor målet er å etablere og integrere eInfrastruktur for håndtering av sensitive data innen biomedisin på tvers av landegrensene.

6. Kompetansebehov

Verdien av eInfrastrukturen avgjøres av forskernes evne til å benytte den. Det hjelper eksempelvis ikke med kraftige regneressurser eller avanserte IT-verktøy dersom disse er av en slik karakter at de ikke dekker forskernes behov eller er vanskelige å ta i bruk. Slike sider

av eInfrastrukturen ivaretas med fokus på brukerbehov, men like viktig er det at brukerne besitter den nødvendige kompetansen. Behovet for kompetanseutvikling og kompetansedeling er stort, men det er ulikt fordelt over fagområdene ved UiO.

Noen fagretninger er i dag så IT-tunge at det blir press på fagplanen på grunn av IT-relatert opplæring. Helsefagene er et eksempel. Her er det et stort behov for målrettet opplæring og kurs, både under og etter endt utdanningsløp. I tillegg går utviklingen så raskt at det er et konstant behov for etterutdanning innen de aller fleste disipliner.

6.1 Kompetanseutvikling og kompetansedeling

Det er et stort spenn i utgangspunkt når det gjelder kompetanse på IT i forskning blant UiOs forskere, og selv om noen fagområder er mer IT-tunge enn andre finner vi dette spennet igjen innenfor alle fagfelt. For enkelte kan det være en utfordring å lese inn data i et regneark og gjøre enkel analyse, mens andre har behov for kompetanse innen avansert parallellprogrammering for grafikkakseleratorer. Dette tilsier at det ikke er mulig å tilby kurs og opplæring for alle mulige behov. Å tilegne seg ny kompetanse er i tillegg ofte tidkrevende, og i mange tilfeller trengs kunnskapen kun i korte deler av et forskningsprosjekt.

En tilnærming til problemstillingen er å utnytte både kompetanseutvikling og kompetansedeling, og å gjøre dette på ulike nivåer. I denne tankegangen ligger ansvaret for kompetanseutvikling av mer generell eller gjennomgående karakter ved fellestjenestene, mens den mer fagspesifikke kompetanseutviklingen tillegges kompetanse-hub'er og enhetene. Kompetansedeling handler om bistand, rådgivning og medvirkning, og spenner fra alminnelig brukerstøtte, via avansert brukerstøtte og prosjektarbeid, til direkte deltakelse i forskningsprosjekter.

Kompetanse-hub'er har en særskilt viktig funksjon i et slikt scenario. En hub gjør det mulig å konsentrere tverrfaglig kompetanse innenfor faglige eller metodiske hovedretninger, og å dele og videreutvikle denne kompetansen. Hub'er kan dedikere tid og stillinger til målrettet kursaktivitet, av personer som bruker sin øvrige tid med de samme metodene og problemstillingene i sitt daglige virke. Sammenlignet med en ukoordinert struktur (ukoordinerte noder) gir dette en langt mer effektiv ressursutnyttelse enn hvis noder hver for seg avsetter mindre ressurser. Hub'ene kan bli en del av forskerutdanningen ved at nye stipendiater tilbys poenggivende introduksjonskurs i metoder og verktøy. De nasjonale forskerskolene følger denne tankegangen. Kompetanseressurser i en hub kan frikjøpes i kortere eller lengre tidsrom og benyttes i forskningsprosjekt, og dette vil i mange tilfeller være en viktig finansieringskilde for hub'er. Hub'er medvirker også til kompetansedeling gjennom seminarserier, journal clubs, og brukerstøtte (helpdesk).

Tverrgående kompetanseutvikling kan også foregå utenom enheter, hub'er og fellestjenestene, men gjerne i samarbeid med disse. Her kan modellen følges som har blitt utviklet av Software Carpentry⁵, en frivillig organisasjon som hjelper forskere (stipendiater og postdocs) til å bli mer effektive i beregningsbasert forskning, eksempelvis i bruk av

⁵ software-carpentry.org

kommandolinje, programmering og utvikling (samarbeidsverktøy, versjons kontroll, m.m.). Gjennom to-dagers workshops gis forskere en innføring i 'best practices' innen beregningsorientert forskning og vitenskapelig programmering. Tilbudet passer spesielt godt for forskere og teknisk personale med kompetansebehov på dette området og som etterutdanning. Det har hittil vært avholdt to slike workshops på UiO. MatNat fakultetet ønsker å organisere Software Carpentry workshops med jevne mellomrom, og øke antallet instruktører for å kunne utvide tilbudet til alle forskere som har behov. Computing in Science Education ser på Software Carpentry som en naturlig forlengelse av sitt arbeid. Data Carpentry⁶ er et lignende initiativ som arrangerer workshops for å bygge kompetanse innen data-dreven forskning og store datasett, men det har ikke vært avholdt Data Carpentry workshops ved UiO foreløpig. Initiativene Software og Data Carpentry er en god modell for kompetansedeling og –utvikling for forskere ved UiO.

6.2 Forholdet til "IT i utdanning"

De to arbeidsgruppene IT i forskning og IT i utdanning leverer uavhengige rapporter, men som beskrevet over er IT i utdanningen vesentlig for senere effektiv utnyttelse av eInfrastruktur, og den samme eInfrastrukturen er en del av IT i utdanningen. Selv om vi her primært omtaler eInfrastruktur i forhold til forskning er også eInfrastrukturen i ferd med å bli svært viktig for utdanningen ved Universitetet i Oslo (UiO). Dette kapitlet beskriver noen potensielle sammenhenger mellom disse to områdene.

UiO er langt fremme på å bruke beregninger i undervisningen på laveregradsstudier. Undervisningsmetoden kalles Computing Science in Education (CSE). Metoden går ut på at studentene allerede fra første semester lærer å løse problemer fra fagområdene med beregninger. Et eksempel er å bruke beregninger til å lære om gravitasjonskraften. I tradisjonell undervisning bruker studenter papir og blyant og kan bare studere gravitasjonsloven med enkle oppgaver som kuler som renner nedover et skråplan. I CSE-modellen kan nybegynnerstudentene laste ned informasjon fra NASA og lære hvordan romfartøyet Rosetta, i en ti år lang ferd gjennom en rekke gravitasjonsfelt i solsystemet vårt, kunne sendes fra Jorda til kometen 67P/Tsjurjumov-Gerasimenko.

Dette kalles også forskningsnær utdanning. Når forskningen bringes inn i eller nærmere utdanningen, oppnås opplæring i bruk av forskningsverktøy, og studenter kan delta/bidra direkte til forskningen. Forskningsnær IT i utdanningen kan således føre til mer forskning på et gitt område, eksempelvis ved nettdugnad ("crowdsourcing"), produksjon av rådata som kan benyttes i forskningen.

Som herværende arbeidsgruppe beskriver arbeidsgruppen for IT i utdanningen et behov for utdanning innen bruk av IT-faglige verktøy, hvordan ulike verktøy effektivt kan brukes i forskningen, hva mulighetene er for å få en bedre forskning med bedre bruk av IT-verktøy. Et eksempel er å undervise om nye verktøy gjennom portaler, for eksempel Lifeportal, UiOs analyseportal for livsvitenskapene.

⁶ datacarpentry.org

Å bringe eInfrastruktur inn i utdanningen er et viktig tiltak for å styrke IT i forskning, og det bringer undervisningen nærmere forskningen slik at den gjøres mer relevant, både for studentene og for fremtidige arbeidsgivere. Prosjektet "Computing in Science Education" på Det matematisk-naturvitenskapelige fakultet er et godt eksempel på bruk av beregningsverktøy og data i utdanningen, og dette prosjektet utvides i disse dager til også å omfatte studieprogrammer med mindre matematikk, for eksempel innen biovitenskap og geofag..

7. Forslag til tiltak

Arbeidsgruppen har kommet fram til de følgende tiltakene ved å holde visjonen for IT i forskning og den behovsorienterte operasjonaliseringen av denne (kapittel 3 og 4) opp mot beskrivelsen av dagens situasjon (kapittel 5). En vurdering av de kompetansemessige behovene og forhold til utdanningen (kapittel 6) har resultert i egne tiltak.

7.1 Utnyttelse av eksisterende eInfrastruktur

For at IT i forskning ved UiO skal ha tilstrekkelig regnekraft er det viktig at den nasjonale ordningen for eInfrastruktur ivaretar UiO-forskernes behov og, like viktig, at UiOs forskere er i stand til å benytte ordningen. Det første krever at UiOs forskere melder inn sine behov nasjonalt, og at UiOs representanter i ordningens styrende organer er informert om og understøtter disse. Det siste er dels et ansvar som ligger i implementeringen og leveransen av nasjonal eInfrastruktur, og dels et ansvar som ligger i implementeringen av IT i forskning ved UiO.

For å hjelpe UiOs forskere i å utnytte den nasjonale ordningen for regnekraft er det behov for mindre lokale regneressurser, for utvikling av programvare, opplæring, og for å kvalifisere og kvalitetssikre prosjekt som ønsker å utnytte nasjonale ressurser. Det er også lokale behov for regnekraft i forbindelse med undervisning, men det er sagt at nasjonale anlegg så langt det lar seg gjøre også skal dekke slike behov.

Arbeidsgruppen anbefaler følgende tiltak for å sikre effektiv utnyttelse av eksisterende nasjonal og internasjonal eInfrastruktur:

- Det etableres hub – node-konstellasjoner ved UiO der dette er hensiktsmessig. Disse vil variere i størrelse og omfang i takt med dynamikken i forskningen. Noen vil forsvinne etter en tid, mens andre vil utvikle seg til egne enheter og/eller bli en del av fellestjenestene. Hub'ene integreres med fellestjenestene ved hjelp av delte stillinger eller ved gjennomføring av konkrete fellesprosjekter. Organisering og finansiering av hub – node konstellasjoner, og hvordan slike kan etableres, diskuteres nærmere i kapittel 8.
- USIT må aktivt medvirke til at UiOs forskere utnytter den nasjonale eInfrastrukturen og mulighetene i internasjonale eInfrastrukturprosjekt. Dette må involvere brukerrettete tiltak, informasjon, dokumentasjon, brukerstøtte og kurs, og samtidig aktiv deltakelse og involvering i de samme eInfrastrukturprosjektene.

- Hub'er og noder (forskermiljøer) må kobles tettere med fellestjenestene (dvs. USIT), hvor nasjonal og internasjonal elinfrastruktur ligger, ved at personal fra hub'er og noder integreres med USITs aktiviteter på området. Slike individer skaper en forbindelse mellom sine miljøer og fellestjenestene og bidrar til mer effektiv utnyttelse av elinfrastrukturen uten at fellestjenestene trenger oppbemanning og spisskompetanse innen stadig skiftende forskningsfelt. Dette kan gjøres ved å gi de nødvendige tilknytningene til USIT og oppgaver innen oppdragene samtidig som slike individer er primært ansatt ved og finansiert fra sine enheter.

7.2 Forskningsdata

En egen arbeidsgruppe er nedsatt for å se på tiltak rettet mot forskningsdata og datahåndteringsproblematikken. Denne arbeidsgruppen vil gjøre en kartlegging av situasjonen med tanke på lagring og deling av forskningsdata ved UiO, gå grundig inn i problemstillingene og komme med egne anbefalinger. Herværende arbeidsgruppe vil likevel anbefale følgende overordnede tiltak:

- UiO må etablere en policy for tilgjengeliggjøring og åpen tilgang til forskningsdata (Open Access) som tilsvarer den etablert av Norges forskningsråd. Det må utvikles en lagrings- og tilgangsstrategi som tar hensyn til aktuelle behov og kostnader ved ulike lagrings- og dataformer.
- Forskere må bevisstgjøres problemstillingene rundt datalagring, deling (Open Access) og gjenfinnbarhet, for eksempel ved etablering av forskerkurs om metadata, rettigheter (lisenser og sitering), arkivering og søkemetoder. Det bør satses på undervisning der studenter finner, bruker og deler data.
- Fellestjenester som gir enkle og sikre løsninger for lagring, deling og gjenfinning av forskningsdata må utredes og etableres.
- UiOs nettverkinfrastruktur må dimensjoneres internt og utad med tilstrekkelig ytelse for å håndtere rutinemessig overføring av store datamengder.

7.3 Kompetanse og kompetanseutvikling

Kompetanse og kompetanseutvikling innen IT i forskning er, som diskutert i kapittel 6 og flere andre steder i denne rapporten, av fundamental betydning for våre forskeres evne til å utnytte elinfrastrukturen og derigjennom utslagsgivende for kvaliteten i forskningen. Ved å utnytte elinfrastruktur godt, ved å utvikle og ta i bruk nye metoder og verktøy, vil forskningen ved UiO kunne nå lenger på den internasjonale arena. Konkrete tiltak på området er:

- Det må gis større plass til bruk av IT-verktøy og IT-støttet vitenskapelig problemløsning i utdanningen, og spesielt i forskerutdanningen. Kompetanse hub'ene vil ha sentrale roller for å sørge for at dette skjer, se for øvrig omtale av kompetansehub'enes rolle i kapittel 4 og 6.
- Det må satses på etterutdanning av forskere, så vel som støttepersonell, i bruk av elinfrastruktur og IT-verktøy. Fellestjenestene må etablere et velutviklet kurstilbud som retter seg mot forskere med ulik bakgrunn og forutsetning, og det må gjøres i samarbeid med hub'er og enheter.

- Det må satset på tverrgående initiativ for etterutdanning av forskere, for eksempel etter modell av Software Carpentry.
- Hub'ene selv må integreres i poenggivende studieløp i forskerutdanningen ved å tilby poenggivende kurs i metoder og verktøy.
- Det må skapes grunnlag for nødvendig metodeforskning og kompetanseutvikling i informatikk, og koordinering av disse aktivitetene, gjennom tidsavgrenset finansiering og engasjement av vitenskapelige ledere knyttet til hub'ene.

7.4 Støttebehov for IT i forskning og nærhetsmodellen

Et tilbakevendende behov fanget opp av arbeidsgruppen er forskergruppers ønske om nærhet til personressurser med fagtilpasset kompetanse og oversikt over ressurser og tjenester som faktisk finnes. Der det er personer knyttet til en enhet eller forskergruppe med særskilt ansvar for slike oppgaver er utnyttelsen av IT i forskning generelt på et høyere nivå enn der hvor det ikke er slik støtte tilgjengelig. Det er avdekket manglende kunnskap om hvilke tjenester og ressurser for IT i forskning som er tilgjengelige, sentralt ved USIT eller andre steder, selv om denne informasjonen finnes tilgjengelig på nettsider eller i andre former.

Selv når forskerne kjenner til slike ressurser og tjenester er det noen ganger også slik at man ikke kan eller vil benytte seg av dem, fordi de ikke er enkle å bruke, ikke er tilpasset det aktuelle problemet eller fordi det er knyttet byråkratisk merarbeid eller kostnader til det. I noen tilfeller bunner disse hindringene i misoppfatninger, som igjen bunner på manglende kunnskap og/eller informasjon. Dette fenomenet, en opplevelse av avstand mellom forskeren og eInfrastrukturen, har arbeidsgruppen etter hvert begynt å omtale som et gap.

Gapet representerer et bredt spekter av behov som på billedlig vis kunne sies å være dekket av at hver enkelt forsker eller forskergruppe hadde en egen omnikompetent og ressurssterk assistent til IT i forskning med direktelinje til USIT. Oftest er behovet for tett oppfølging et midlertidig behov i en kortere eller begrenset fase i forskningsprosessen. Dette gjør at slike ressurser kan deles.

I større forskningsprosjekt kan det være behov for utvikling av større IT-løsninger, noe som igjen krever egne utviklingsprosjekt med ressurser fra USIT. De samme IT-løsningene må ofte driftes i lengre tid. Presset på USITs utviklings- og driftsressurser kan bli for stort, særlig fordi universitetets øvrige funksjoner, administrasjon og studier, legger beslag på de samme ressursene. De tilgjengelige ressursene må koordineres og utnyttes best mulig, og derfor trengs overordnede mekanismer for strategisk koordinering og prioritering. For å unngå at drift opptar unødvendige ressurser, og for å effektivisere utvikling, er det nødvendig med en overordnet IT-arkitektur og forvaltning av systemer.

Tiltak:

- Fellestjenestene må gis tilstrekkelig handlingsrom (kapasitet) til å være dynamiske og tilpassingsdyktige, slik at vi unngår at fellestjenestene blir en passiv leverandør av statiske tjenester og heller blir en aktiv og integrert del av IT i forskning. Dette oppnås

ved å gjøre fellestjenestene prosjektorientert med bred kunnskapsmessig basis (metode, utvikling, web, nett, beregninger, database, billedbehandling osv.), noe som igjen vil gjøre det mulig å etablere gode prosjektgrupper. En mulighet er å etablere en "taskforce", i sammenheng med et prioriterende organ, som kan gå inn, være med forskerne å analysere og definere behov og eventuelt etablere skreddersydde prosjektgrupper.

- Kompetansemessig må fellestjenestene (USIT men også hub'er) gjøres i stand til å bidra med grunnleggende IT-kompetanse kombinert med spisskompetanse (innen metodiske hovedretninger og teknologi, eksempelvis bruk av akseleratorer). Forskergruppene kan da kombinere dette med sin faglige kompetanse (eksempelvis knyttet til nye algoritmer) i utvikling og gjennomføring av forskningsprosjekt. En slik kompetanse- og oppgavedeling er et godt eksempel på integrert samarbeid
- Enkeltaktører (hub'er, noder og enhetene) oppfordres til å etablere egne ressurser (dedikerte personer) for IT i forskning. Disse kobles i størst mulig grad med fellestjenestene, for eksempel ved å integrere disse med USIT, slik at det oppnås nærhet til egne forskere og deres problemstillinger og behov, og samtidig en forbindelse til fellestjenestene og eInfrastrukturelle ressurser.
- Det bør etableres et nettverk, *eInfrastrukturnettverket*, for IT i forskning hvor alle enheter er representert. Dette nettverket vil være et rådgivende organ og kommunikasjonskanal for både fellestjenestene (USIT) og eInfrastrukturutvalget (se neste punkt og kapittel 8). eInfrastrukturnettverket må bestå av personer med kompetanse om IT i forskning og oversikt over egen enhets forskningsaktiviteter og – strategi.
- Iverksetting av større IT utviklingsprosjekt eller satsinger må koordineres, også i forhold til annen slik aktivitet ved UiO. Ressursmessige konsekvenser og kostnader ved drift av nye systemer og tjenester må utredes før utviklingsprosjekt eller satsinger iverksettes. Det må derfor etableres et strategisk koordineringsutvalg, *eInfrastrukturutvalget*, som prioriterer og forvalter ressurser for IT i forskning. Dette er nærmere beskrevet i det neste kapitlet.

8. Organisering av IT til forskning ved UiO

Som denne rapporten viser er IT i forskning ved UiO stort, omfattende og i stadig endring. Måten vi fremover organiserer dette på må ta høyde for både det dynamiske aspektet ved forskning og det at området IT i forskning i seg selv er i konstant endring. Nye behov initieres av forskning (nye idéer), nytt vitenskapelig/teknisk utstyr og økt tilgang på data fra ulike kilder. Tilgjengelige ressurser må benyttes på mest mulig effektiv måte. Dette krever strategisk styring og prioritering.

8.1 Roller og ansvar

Når vi skal diskutere organisering av IT til forskning ved UiO må vi ha et klart bilde av hva som dette omfatter. Vi kan definere eInfrastruktur på fire forskjellige nivåer; internasjonalt, nasjonalt, institusjonelt og lokalt ute i fagmiljøene. Bredden av behov, for eksempel basistilgang til regnekraft, lagring og programvare, dekkes mest effektivt gjennom sentrale fellestjenester. NFR er med på å finansiere den nasjonale eInfrastrukturen, og UiO oppnår mest effektivt en basis eInfrastruktur ved å delta i den nasjonale ordningen. Når behovene blir spissede, rettet mot spesifikke brukergrupper eller fagområder, er det ikke lenger snakk om felles eInfrastruktur, men likevel behov som må adresseres. Helt ytterst (eller øverst) er den enkelte forsker med sine spesielle behov.

USIT har ansvar for felles IT til forskning ved UiO. Aktivitetene innbefatter alt som er beskrevet i kapittel 5. Utover basistjenester er det ved USIT i dag rundt 22 årsverk dedikert IT i forskning (museenes samarbeidstiltak MUSIT, er her ikke medregnet). Av disse er i overkant av 12 årsverk knyttet opp i bidragsoppdrag i forbindelse med nasjonal eInfrastruktur og deltakelse i internasjonale prosjekter (per 2014). Disse noe over 20 årsverkene utgjør også i all hovedsak de felles menneskelige ressursene for IT i forskning. Med disse fellesressursene skal flest mulig av UiOs forskere få sine basisbehov dekket, primært ved å utnytte ressurser og tjenester innen samarbeidsprosjektene, spesielt den nasjonale eInfrastrukturen.

Totalt utgjør kostnader til felles eInfrastruktur i størrelsesorden 30 MKr årlig (fom. 2015 i ny nasjonal ordning). Av dette går omlag 20 MKr årlig til fellesressurser integrert med nasjonal og internasjonal eInfrastruktur (samarbeidsprosjektene). Det øvrige går til fellesressurser som ikke dekkes av, eller er del av, dette. Til det siste hører eInfrastruktur, dvs. tjenester, kompetanse, maskin- og programvare, for sensitive forskningsdata (TSD), for statistikk og kvalitative metoder, og videre tilsvarende for lyd- og bildebehandling, karttjenester, støtte for samlinger (databaser) og ulike former for datainnsamling. Hvilke fellesbehov som kan tilfredsstilles innenfor eksterne samarbeidsprosjekter, og hvilke fellesbehov UiO alene vil adressere, vil være konstant endring. For eksempel forventer vi at flere tjenester for sensitive data etter hvert vil tilbys gjennom de nasjonale prosjektene.

Når behov knyttet til IT i forskning er spesifikke og spisset mot en avgrenset brukergruppe, vil ansvaret, inklusive ansvar for finansiering, ligge hos brukergruppen selv, men *i forlengelsen* av fellestjenestene. På samme måte som fellestjenester for IT i forskning står på skuldrene til USITs øvrige tjenester, skal således hub - node og miljøspesifikk eInfrastruktur kunne nå enda lenger ved å utnytte felles eInfrastruktur.

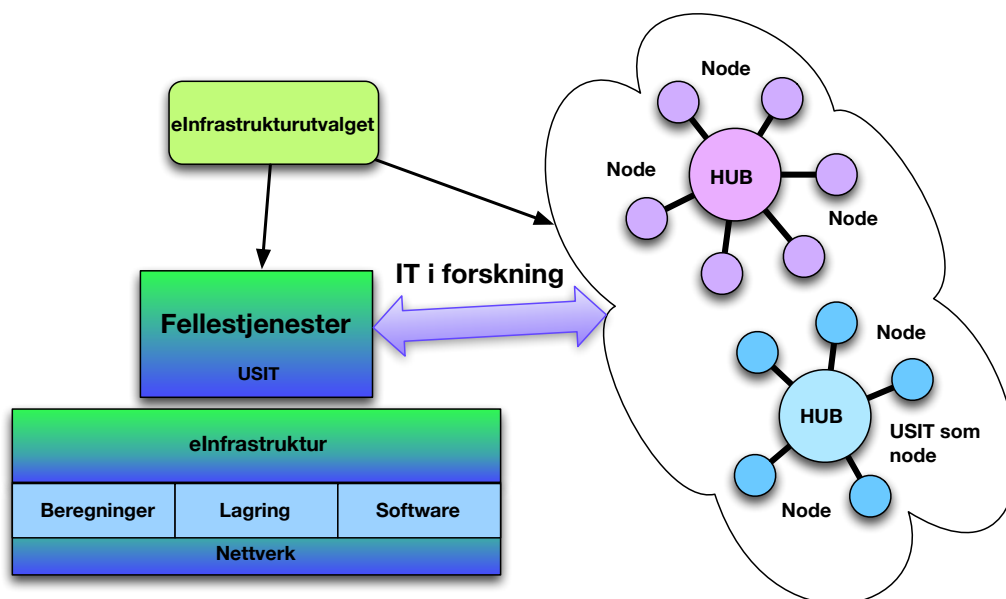
Kapittel 7.4 beskriver gapet som oppleves av forskere ved mange enheter. Det kan være fristende å si at ansvaret for å ta i bruk IT i forskning ligger hos forskeren selv, at forskeren aktivt må finne fram til de nødvendige støttefunksjonene og tilegne seg tilstrekkelig kompetanse. Men eInfrastrukturen, i ordets videste betydning, må også komme den enkelte forsker i møte. Når kompleksiteten øker i takt med teknologiutviklingen blir dette stadig viktigere, og det blir i økende grad et behov for at eInfrastrukturen, dvs. IT-folkene, går inn i et tettere samarbeid med forskerne om å utarbeide løsninger på deres utfordringer og forskerproblemstillinger. Dette krever nærhet til forskerne, IT-kyndige med egen

forskererfaring, teknologiovervåking, kurs og formidling. Her er det et delt ansvar mellom fellesressursene, hub'er og enhetene.

8.2 Hub'er og noder

Dedikerte, langsiktige stillinger og et faglig miljø av en bærekraftig størrelse er nødvendig for å bygge opp ledende spisskompetanse innen IT i forskning for et spesifikt fagområde. Dette er utenfor rekkevidde for enkeltstående prosjekt, og på grunn av kostnader og behov for tverrfaglighet, for ett enkelt institutt eller SFF. Det er også utenfor oppgaveområdet for USIT som skal levere fellestjenester til hele bredden av forskning ved UiO.

Ut fra erfaringene med Computational Life Science (CLS) initiativ ved MN og lignende sentre ved andre institusjoner, ser arbeidsgruppen at en hub - node-organisering, der forholdene ligger til rette, vil adressere en rekke av de behov som har framkommet. Slik organisering startes med en satsing innenfor fag- eller metodiske områder som har de rette forutsetninger for å etablere et bærekraftig, selvfinansierende miljø for bygging av spisskompetanse og ekstra utnyttelse av IT i forskning for de deltagende prosjekter, forskergrupper og enheter.



Figur 3: Illustrasjon av hub - node modellen og forholdet til fellesressurser og -tjenester (USIT)

Biostatistikk-senteret og CLS er inspirasjonskilder for en konkret anbefaling fra denne arbeidsgruppen om å opprette en hub for bioinformatikk. Dette vil være et senter som vil gi store avkastninger, både ved UiO og regionalt, innen livsvitenskap og spesielt medisinsk forskning. Et slikt senter er tenkt som et informatisk tungt fagmiljø, hvor forskningstygde og spisskompetanse sikrer kvalitet. UiO, OUS og Helse Sør-Øst vil være sentrale partnere. Det

må bygges opp en kritisk masse (profesjonalisering, flere faste og spissede vitenskapelige stillinger) for at hub'en skal bli attraktiv og effektiv. Et senter for bioinformatikk kan etableres der CLS er, dvs. 10. etasje i Ole Johan Dahls hus (ifi2), og etter hvert i det nye Life Science bygget. I tillegg til forskningsfunksjoner vil en viktig funksjon til et senter for bioinformatikk være å tilby tverrfaglig utdanning og forskningsbaserte, avanserte tjenester.

For å sørge for at sentrene har en brobyggende funksjon mot sine noder og fellestjenestene, er planen å benytte delte stillinger hvor stipendiater og postdoktorer har primær ansettelse og en stillingsandel ved sin node eller i hub'en, kombinert med en stillingsandel ved hhv. hub'en eller en node, eller ved USIT. Nodene, kan være både storbrukere (prosjekt) som ønsker å spisse sine egne bioinformatikk og dataressurser, eller mindre brukere, som kommer til å være mer avhengige av direkte service og ressurser fra hub'en. En modell er at bioinformatikere i hub'en blir rekruttert til prosjekter og avansert brukerstøtte i disse mindre nodene.

8.3 Koordinering og prioritering

Arbeidsgruppen foreslår at det etableres et strategisk koordineringsutvalg for IT i forskning ved UiO, og at dette utvalget får navnet *eInfrastrukturutvalget*. Utvalget bør ledes av prorektor for forskning, og forvalte en årlig strategisk bevilgning for IT i forskning. USIT er utvalgets sekretariat. UiO har et egen utvalg for forskningsinfrastruktur, og arbeidsgruppen har vurdert om dette utvalget også kunne ta det overordnede koordinerende ansvaret for eInfrastruktur. Arbeidsgruppen er imidlertid av den oppfatning at eInfrastrukturens brede nedslagsfelt på UiO og behovene for nasjonal og internasjonal koordinering krever et dedikert utvalgt. Arbeidsgruppen foreslår derfor et eget eInfrastrukturutvalg. Arbeidet i dette eInfrastrukturutvalget må selvfølgelig koordineres godt med investeringer i annen infrastruktur som produserer data som krever eInfrastrukturelle løsninger.

eInfrastrukturutvalget må ha ansvar for kunderollen i UiOs deltakelse i den nasjonale eInfrastrukturen, og sørge for at behov og strategiske prioriteringer for forskningen ved UiO i størst mulig grad ivaretas innenfor ordningen.

Videre må eInfrastrukturutvalget ha ansvar for prioritering av felles og fagspesifikke ressurser rettet mot slike behov som ikke er dekket innenfor nasjonal eInfrastruktur. En viktig strategisk dimensjon er at slike ressurser kan sette UiO i forkant innenfor fremvoksende disipliner, eksempelvis digital humaniora og forskning som involverer sensitive data, eller innen nye metodiske retninger, for eksempel "big data". eInfrastrukturutvalget skal iverksette strategiske utviklingsprosjekter og investeringer, basert på prosjektforslag, og ivareta koordinering mellom disse og UiOs administrasjon, inklusive USIT, og enhetene. Modellen som foreslås har således likheter med Strategisk koordineringsgruppe for administrative IT-systemer (SKAIT).

eInfrastrukturutvalget skal også medvirke til oppretting av hub'er som beskrevet over. Dette vil være strategiske satsinger med tre til fem års varighet for etablering av hub'er basert på prosjektforslag. Slike satsinger kan innebære investeringer, men vil i større grad finansiere stillinger.

UiO må tydeliggjøre sitt institusjonelle ansvar og etablere retningslinjer og mekanismer for håndtering av UiOs åpne forskningsdata. Hvordan dette skal implementeres antas utredet av arbeidsgruppen for lagring og deling av forskningsdata ved UiO.

Forskning og utdanning henger tett sammen, og det er naturlig at eInfrastrukturutvalget også har ansvar for den delen av IT i utdanning som er eInfrastruktur. eInfrastrukturutvalget skal ikke gi støtte til rene vitenskapelige stillinger (eller forskerstillinger), med mindre det er nødvendig med brofinansiering på inntil fem år for å bygge opp helt ny kompetanse på et gitt område med et bredt nedslagsfelt på UiO. Det finnes i dag noen få slike områder, for eksempel innen "data science".

9. Ressursbehov og økonomiske forhold

Som i dag vil anslagvis 30 millioner fortsette å gå med til felles eInfrastruktur ved UiO. Dette er deltagelsen i den nasjonale (og internasjonale) eInfrastrukturen (ca. 20 Mkr) og lokale fellesressurser (10 Mkr). I tillegg vil enhetene bidra med egen, fagspesifikk eInfrastruktur som tidligere. Innenfor denne kategorien kan medregnes eksempelvis eInfrastruktur som NHMs DNA-laboratorium, HFs fellestjenester og UVs videoanalysefasiliteter. Spesielt MN fakultetet og medisinsk fakultet har mye fakultet- og instituttspesifikk eInfrastruktur. Avhengig av hva som regnes med anslår arbeidsgruppen at dette representerer 15–30 millioner kroner årlig.

eInfrastrukturutvalget bør disponere 15 millioner kroner per år fra 2016 på vegne av helheten. Disse summene er helt avgjørende for at UiO i fremtiden skal være et internasjonalt ledende forskningsuniversitet.

Arbeidsgruppen foreslår videre at midlene fra eInfrastrukturutvalget fordeles etter søknader av ulike kategorier. En kategori er satsingene på etablering av hub-node-strukturene, med støtte for inntil 5 år, for eksempel for 3 år med mulighet for 2 års forlengelse. En annen kategori kan være satsinger (investeringer og/eller utviklingsprosjekt) for spesifikke formål. Et eksempel vil være at USIT tilføres prosjektersurser for utvikling av nye løsninger, på linje med det allerede gjennomførte prosjektet for etablering av løsninger for sikker håndtering av sensitive data, gjerne med medfinansiering fra partnere.

Arbeidsgruppen antar at de fysiske regne- og lagringsressursene i all hovedsak vil bli finansiert gjennom den nasjonale ordningen, ordinære investeringer og prosjekter finansiert av eksterne kilder (for eksempel Forskningsrådets infrastrukturprogram). Investeringsbehovet der UiO selv må bidra mye i årene som kommer antas å være utvikling av systemer og mekanismer for håndtering av til dels store datamengder og utvikling av menneskelige ressurser som på vegne av forskningen ved UiO (og partnere) kan utnytte lokal, nasjonal og internasjonal eInfrastruktur,

Fagspesifikke hub'er skal på sikt ideelt sett være selvfinansierende, dvs. finansieres av prosjekt (ekstern finansiering), brukere og/eller enheter. USIT har et spesielt ansvar for å sørge for at fellesressursene og nasjonal eInfrastruktur benyttes til å støtte opp om sentra

slik at disse kan tiltrekke seg ekstern finansiering, og rette sin finansiering inn mot spisskompetanse og -ressurser, ikke basistjenester.