

# Behov og finansieringsstrategi for nasjonal e-infrastruktur for forskning for perioden 2020 – 2030

2019.12.17 (endelig)

*«På IKT-området er det særlig to områder som krever oppmerksomhet fremover når det gjelder forskningsinfrastruktur, både i nasjonal og internasjonal sammenheng. Det første er kapasitet og kompetanse innenfor e-infrastruktur, og det andre er behovet for å få etablert bærekraftige forretningsmodeller for datainfrastrukturer.*

*E-infrastruktur er IKT-baserte infrastrukturer som dataarkiver for lagring av store datamengder, høykapasitets datanettverk og tilhørende tjenester som autentisering og autorisering, verktøy for effektiv arbeidsflyt og programvare for simulering og analyse av data. Sist, men ikke minst omfatter det regneressurser for store beregninger, såkalt tungregning (High Performance Computing (HPC)). Tungregning er et viktig verktøy for å møte store vitenskapelige og samfunnsmessige utfordringer, blant annet innenfor marin forskning, klimaforskning og helseforskning. EU-kommisjonen ønsker å samarbeide med sentrale europeiske land, deriblant Norge, om partnerskapet EuroHPC.»*

Stortingsmelding 4 (2018-2019)  
Langtidsplan for forskning og høyere utdanning 2019–2028  
S. 65

## Oppsummering og hovedkonklusjoner

For at Norge skal kunne nå sine forskningspolitiske mål må Norge investere i e-infrastruktur minst på linje med de beste av land det er naturlig å sammenligne seg med. For å nå dette målet, vil dagens norske investeringer måtte økes til rundt 250 MNOK årlig. Samtidig vil utviklingen innenfor bruk av e-infrastruktur kunne endre seg hurtig, og Norge må være i stand til å tilpasse seg slike internasjonale endringer. Denne rapporten foreslår en ny finansieringsmodell som har en slik nødvendig fleksibilitet.

Utvalgets forslag til en bærekraftig og langsiktig finansieringsmodell for nasjonal, generisk infrastruktur har følgende hovedelementer:

- Norge må investere minst like mye i e-infrastruktur som de ledende blant forskningsnasjonene vi ønsker å sammenligne oss med: Danmark, Finland, Nederland, Sverige og Østerrike. Utvalget mener det er gode grunner til at Norge bør ha en enda sterkere satsing.
- Grunnfinansieringen av nasjonal, generisk e-infrastruktur bør dekke minst 80% av de totale kostnadene for disse tjenestene.
- Spissytelse og investeringer i nye e-infrastrukturarkitekturer og tjenester bør finansieres gjennom den konkurranseutsatte arenaen i INFRASTRUKTUR-programmet, og bør ikke overstige 20% av de totale kostnadene for e-infrastruktur.
- Brukerfinansiering kan inngå i grunnfinansieringen, men utvalget anbefaler at nivået på brukerfinansiering ikke overstiger 10%. Brukerfinansiering vil dels komme fra salg av kommersielle tjenester, men også gjennom Forskningsrådet i form av prosjektfinansiering til større prosjekter og SFFer med stort forbruk av e-infrastruktur tjenester (bidragsmodell).
- Dagens finansierende institusjoner (Forskningsrådet og de fire eldste universitetene) må øke sine bidrag til grunnfinansieringen av de nasjonale e-infrastruktur-tjenestene for å dekke egne økte behov for slike tjenester.
- Det bør være en bredere basis av institusjoner som bidrar til grunnfinansieringen, både fra UH og instituttsektoren. Selv om utvalget mener at hensynet til effektiv ressursutnyttelse tilsier at alle institusjoner bør bidra til grunnfinansieringen, har utvalget forståelse for utfordringene knyttet til lav basisfinansiering av de teknisk-naturvitenskapelige instituttene. Utvalget mener det er viktig at instituttsektoren også bidrar til grunnfinansieringen, og ber Norges Forskningsråd innlede en dialog med relevante institutt og departement om hvordan dette best kan gjøres.
- Grunnfinansieringen av e-infrastruktur-tjenester avtales for en periode på tre år med finansierende institusjoner, og skal inkludere den forventede veksten i grunnfinansieringen i avtaleperioden.
- Institusjoner som over en tre-årsperiode har et forbruk som svarer til mer enn 1% av de totale kostnadene for nasjonale e-infrastruktur-tjenester (tungregning og/eller datalagring) bør bidra til grunnfinansieringen av infrastrukturen. Ved etablering

av den nye finansieringsmodellen, anbefaler utvalget at fordeling mellom institusjoner baseres på forbruket de foregående tre år, og at fordelingen revideres hvert tredje år deretter etter samme prinsipp.

- Større forsknings- og forvaltningsprogram bør inkludere e-infrastrukturkostnader i programmets totalbudsjett som bidrag til grunnfinansieringen av den nasjonale e-infrastrukturen. Utvalget foreslår at styret i NFR ivaretar dette ansvaret for midler som kanaliseres gjennom Norges Forskningsråd, og at det ivaretas gjennom tildelingsbrevene for andre forskningsprogram og forvaltningsoppgaver.
- Modellen med et aksjeselskap for organisering av nasjonale e-infrastrukturtenester anbefales videreført. Selskapet evalueres hvert sjette år. En positiv evaluering viderefører selskapet for seks nye år. Ved en negativ evaluering må selskapet komme opp med en ny organisasjonsmodell i samarbeid med finansierende institusjoner. I ytterste konsekvens kan denne prosessen lede til en anbefaling om avvikling av selskapet. Selskapet bør gis inntil tre år på å avvikle virksomheten etter at finansierende institusjoner endelig har avvist nye forslag til organisasjonsmodell.

## Innhold

<b>Oppsummering og hovedkonklusjoner .....</b>	<b>3</b>
<b>Innledning.....</b>	<b>6</b>
<b>Rapportens avgrensinger .....</b>	<b>9</b>
<b>Behov for e-infrastruktur for forskning og forvaltning.....</b>	<b>11</b>
<i>Nasjonal organisering .....</i>	<i>14</i>
<b>Ambisjonsnivå .....</b>	<b>16</b>
<b>Finansieringsmodeller.....</b>	<b>19</b>
<i>Dagens finansieringsmodell.....</i>	<i>20</i>
<i>Alternativer for en ny finansieringsmodell.....</i>	<i>24</i>
<i>Utvalgets anbefalinger .....</i>	<i>26</i>
Mulighet: Økt grunnfinansiering fra deltakende institusjoner.....	27
Mulighet: Økt tverrsektorielt bidrag til grunnfinansieringen gjennom NFR/departement.....	28
Mulighet: Økt brukerfinansiering gjennom betaling for særskilte behov .....	30
<i>Utvalgets anbefaling av ny finansieringsmodell.....</i>	<i>32</i>
<i>Utvalgets forslag til fordelingsmodell.....</i>	<i>35</i>
<b>Styring og organisering .....</b>	<b>37</b>
<b>Oppfølging av arbeidsgruppens anbefalinger .....</b>	<b>37</b>
<b>APPENDIKS 1: Om arbeidsgruppens arbeid .....</b>	<b>37</b>
<i>Arbeidsgruppens medlemmer .....</i>	<i>38</i>
<i>Arbeidsgruppens mandat.....</i>	<i>38</i>
<i>Arbeidsprosess.....</i>	<i>39</i>
<b>APPENDIKS 2: Om e-infrastruktur.....</b>	<b>41</b>
<i>Formål.....</i>	<i>41</i>
<i>Eksisterende forskningsinfrastruktur.....</i>	<i>42</i>
<i>Behov for nyetablering, oppgradering og/eller samordning .....</i>	<i>44</i>
<i>Relasjon til andre områder.....</i>	<i>44</i>
<b>APPENDIKS 3: Forskningsprogrammene/Departementene sin bruk av Sigma2 sin e-infrastruktur .....</b>	<b>45</b>
<b>APPENDIKS 4: Begrepsdefinisjoner .....</b>	<b>48</b>

APPENDIKS 5: Akronymer .....	49
APPENDIKS 6: Vurdering av høringsinnspill og endringer etter høring .....	50
APPENDIKS 7: Oversikt over forbruk fordelt på vitenskapelige disipliner .....	55

## Innledning

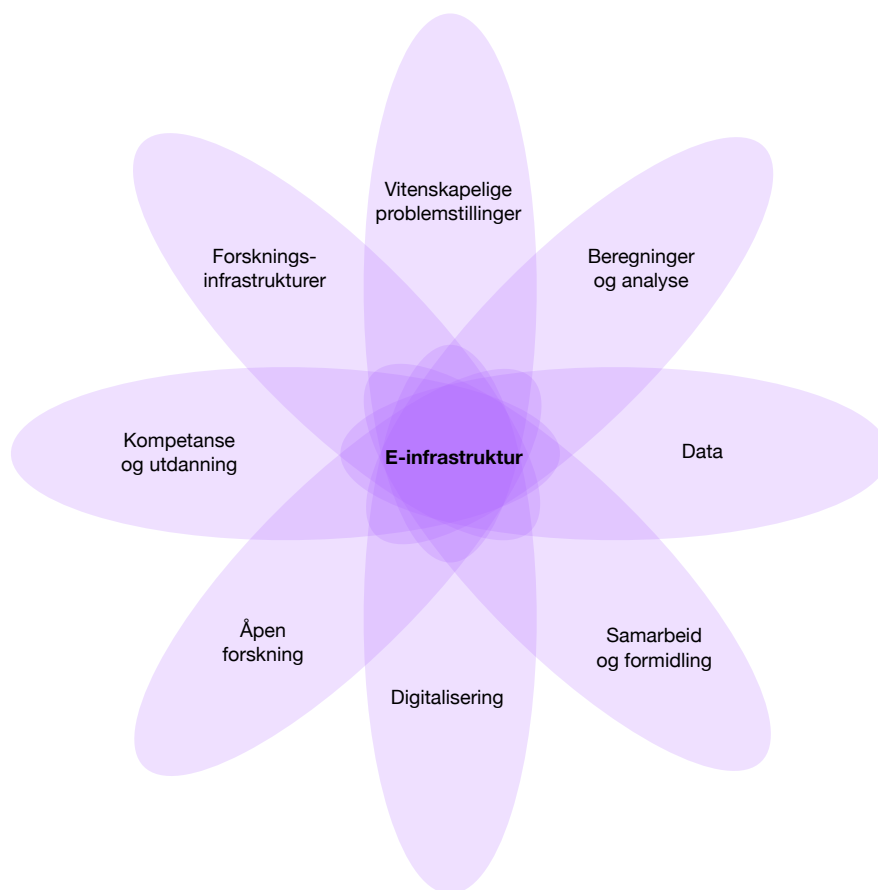
Når værmeldingen indikerer at man bør planlegge for en søndagstur, eller orkanvarsler fører til masse-evakueringer, så er det avanserte beregninger på bakgrunn av store mengder observasjonsdata som ligger til grunn. Gjennom avanserte analysealgoritmer og beregninger kan selskaper som Facebook og Google bygge på all den elektroniske informasjon vi lagrer til å skreddersy reklamer og nyheter som svarer til våre personlige interesser. Med avanserte bildebehandlingsverktøy og numeriske metoder kan vi avgjøre om føflekker er ondartede eller godartede. Gjennom marine observasjonsdata og kompliserte beregningsmodeller kan vi bedre vår forvaltning av marine ressurser. Felles for alle disse anvendelsene er at de forutsetter at dataene kan lagres og overføres, og at beregningene utføres, på en velfungerende e-infrastruktur.

I dag er e-infrastruktur helt nødvendig for vitenskapelig fremskritt innenfor nærmest alle fagfelt. Grensesprengende forskning baseres i stadig større utstrekning på beregninger og avanserte analyser av store datamengder, ofte gjennom anvendelse av nye informatikk-metoder som kunstig intelligens. E-infrastruktur er også essensielt i digitalisering og i bevaring og anvendelse av de økende mengdene data som fremskaffes med offentlig finansiering. Observasjonsdata lagret gjennom lange tidsserier og med økende oppløsning i rom og tid, sammen med modeller og analyse, blir stadig viktigere for bærekraftig forvaltning av naturressurser og for å ivareta samfunnssikkerhet, eksempelvis i planleggingen av fremtidig bosetting og infrastruktur i lys av endringer i ras- og flomfarer som en følge av klimaendringer. E-infrastruktur gjør det også mulig for forskere å samarbeide over store avstander, bygge på hverandres forskning og å forske på tvers av fagfelt (åpen forskning).

### Hva er e-infrastruktur?

E-infrastruktur for forskning er IKT-baserte infrastrukturer som muliggjør avansert og samarbeidsorientert forskning. E-infrastruktur for forskning omfatter utstyr, drift og relaterte tjenester for tungregning, datalagring, programvaresystemer og høyhastighetsnettverk samt verktøy for effektiv arbeidsflyt og programvare for simulering og analyse av data. Begrepet e-infrastruktur benyttes også om digitale registre og databaser, samt verktøy og tjenester for å sikre og gjøre disse tilgjengelige.

Se også Appendix 2 *Om e-infrastruktur*.



*Figur 1: E-infrastruktur er en integrert del av flere dimensjoner ved forskning, utdanning og forvaltning, og er en forutsetning for å ligge i front innenfor de ulike prosesser som nå bidrar til å endre det digitaliserte samfunn.*

Norge har siden midten av 80-tallet med ujevne mellomrom investert i ny e-infrastruktur. Begrunnelsen for denne nasjonale satsingen har vært at dette gjør det mulig å bygge og ta i bruk hele landets kompetanse på området. Store, nasjonale innkjøp og et helhetlig perspektiv på behovene for felles generiske løsninger i kombinasjon med denne spisskompetansen har bidratt til god ressursutnyttelse. Det har bidratt til å heve kompetansen over hele landet i relevante disiplinorienterte fagmiljøer som bruker e-infrastrukturtenester, og det har sørget for at norske forskningsmiljøer har kunnet ta i bruk den fremste internasjonale e-infrastrukturen. Resultatet av denne målrettede satsingen er at norske forskningsmiljøer som bruker e-infrastruktur produserer forskning som jevnt over er av høyere kvalitet enn gjennomsnittet for norsk forskning ifølge analyser gjort av Ressursfordelingskomiteen i Sigma2. Flere av disse miljøene er også i den internasjonale forskningsfronten. Ser vi på den historiske utviklingen, så viser den at det som er dagens høyteknologiske løsninger blir morgendagens normalsituasjon, og disse investeringene har dermed sørget for at Norge har hatt den nødvendige kompetansen til å ligge langt fremme på digitalisering og automatisering, ikke bare innen forskning, men også gjennom ringvirkninger i hele samfunnet.

I dag er generisk e-infrastruktur for beregningsorientert forskning og lagring av store datamengder organisert nasjonalt gjennom Uninett Sigma2 AS (heretter Sigma2) i

partnerskap med UiB, UiO, NTNU og UiT (de fire breddeuniversitetene) og knyttet sammen med forskningsnettlet forvaltet av Uninett. Opprettelsen av Sigma2 ved årsskiftet 2014-2015 var resultat av et sterkt økende behov for investeringer i e-infrastruktur og forståelsen av at dette bare kunne løses effektivt i fellesskap. Stadig større datamengder og økende krav til lagring og sporbarhet av beregningsresultater bidro også sterkt til behovet for en ny organisering av e-infrastrukturtenester. Lagring av vitenskapelig data krever en annen form for langsiktighet knyttet til investeringer og finansieringer enn det som var mulig før opprettelsen av Sigma2.

Med Sigma2 ble organisasjon og finansiering tilpasset behovet for langsiktighet og nivået på finansieringen økt. I den relativt korte perioden som har gått etter at selskapet ble etablert har imidlertid behovene fortsatt å øke langt ut over det som ble forutsatt ved opprettelsen av Sigma2. Dagens modell viser seg ikke å være i stand til å ta høyde for disse hurtige endringene, og som konsekvens er den nåværende finansieringsmodellen i ferd med å bryte sammen. En betydelig del av norsk forskning, deriblant noen av våre fremste forskningsmiljø, står i fare for ikke lenger å kunne konkurrere i den internasjonale forskningsfronten. Dette må forventes å ha konsekvenser for fremtidig kompetanse innenfor analyse av store datamengder og utnyttelse av komplekse beregningsmodeller, noe som vil påvirke alle fagområder hvor dette står sentralt i den faglige utviklingen.

Mange offentlige forvaltningsoppgaver har behov både for lagring av store datamengder og for avanserte beregninger basert på slike datasett. Med bedre beregningsmodeller for prediksjon av ulike hendelser med store samfunnsmessige konsekvenser som for eksempel vurdering av ras- og flomfare og prediksjon av ekstrem- og romvær, vil vi forvente økt bruk av e-infrastruktur også for samfunnskritiske funksjoner hvor også tid-til-svar (*time-to-solution*) vil kunne være en kritisk suksessfaktor. Det kan være gode grunner til at slike behov bygges på toppen av en generisk e-infrastruktur hvor det gis anledning til å løse samfunnskritiske problemstillinger på kort varsel. Det vil gjøre det mulig å investere i betydelig større e-infrastrukturanlegg, og på den måten sikre kortest mulig beregningstid, samtidig som det sikrer en effektiv utnyttelse av kostbar e-infrastruktur til andre formål i perioder hvor det ikke er slike beregningsbehov.

Hovedmålgruppen for dette notatet er derfor beslutningstakere ved norske institusjoner som driver med forskning og/eller forvaltning, Forskningsrådet, samt departementer som finansierer forskning eller har forvaltningsoppgaver med store behov for e-infrastruktur.



## Rapportens avgrensinger

Begrepet e-infrastruktur omfatter mer enn infrastruktur for tungregning (High Performance Computing - HPC) og datalagring. I tillegg til all IKT-infrastruktur som benyttes spesielt til forskning, medregnes også programvare, verktøy, tjenester og ikke minst brukerstøtte i ulike former for å sikre at e-infrastrukturen brukes mest mulig effektivt, og at forsknings- og forvaltningsoppgaver kan utnytte de mulighetene som den nyeste e-infrastrukturen åpner for.

Høsten 2018 og våren 2019 pågår det fire parallelle prosesser som har betydning for det fremtidige landskapet for e-infrastruktur for forskning, forvaltning og (høyere) utdanning i Norge:

1. Arbeidsgruppen etablert av de fire breddeuniversitetene i samarbeid med Forskningsrådet og Sigma2 for organisering og finansiering av nasjonal e-infrastruktur for forskning og forvaltning frem mot 2030 (herværende rapport).
2. Unit<sup>1</sup>-prosjektet *Plan 21* for utvikling av handlingsplaner, med bl.a. delprosjekt for forskning (og samtidig med etablering av finansieringsmekanismer og en styringsstruktur med Digitaliseringsstyre<sup>2</sup> og fagutvalg<sup>3</sup>).
3. Et datalagringsprosjekt i regi av Uninett, som ser på brukerbehov og muligheter for effektivisering og samordning i hele sektoren (ikke bare forskning)<sup>4</sup>.
4. På oppdrag fra KD har Forskningsrådet, i samarbeid med Unit og representanter for forskningsinstitusjonene, pågående arbeid med å utrede og gi råd om bærekraftige finansieringsmodeller for drift av forskningsdatainfrastrukturer og forskningsdataarkiver<sup>5</sup>.

Samtidig er også Direktoratet for e-helse godt i gang med Helsedataprogrammet<sup>6</sup> (HDP), hvor målet bl.a. er bedre utnyttelse av helsedata i forbindelse med forskning og etablering av Helseanalyseplattformen (HAP).

De fire ovenfor nevnte prosessene og HDP har varierende grad av overlapp med hverandre når det gjelder formål/virkefelt, tidshorisont, arbeidsplaner og sammensetning av arbeidsgruppene. Lagring og tilgjengeliggjøring av forskningsdata er et gjennomgående tema i alle prosessene. Den nasjonale digitaliseringsstrategien er felles utgangspunkt.

---

<sup>1</sup> <https://www.unit.no/about/om-oss/>

<sup>2</sup> <https://www.unit.no/digitalisering/digitaliseringsstyret/mandat-for-digitaliseringsstyret/>

<sup>3</sup> <https://www.unit.no/digitalisering/fagutvalgene/>

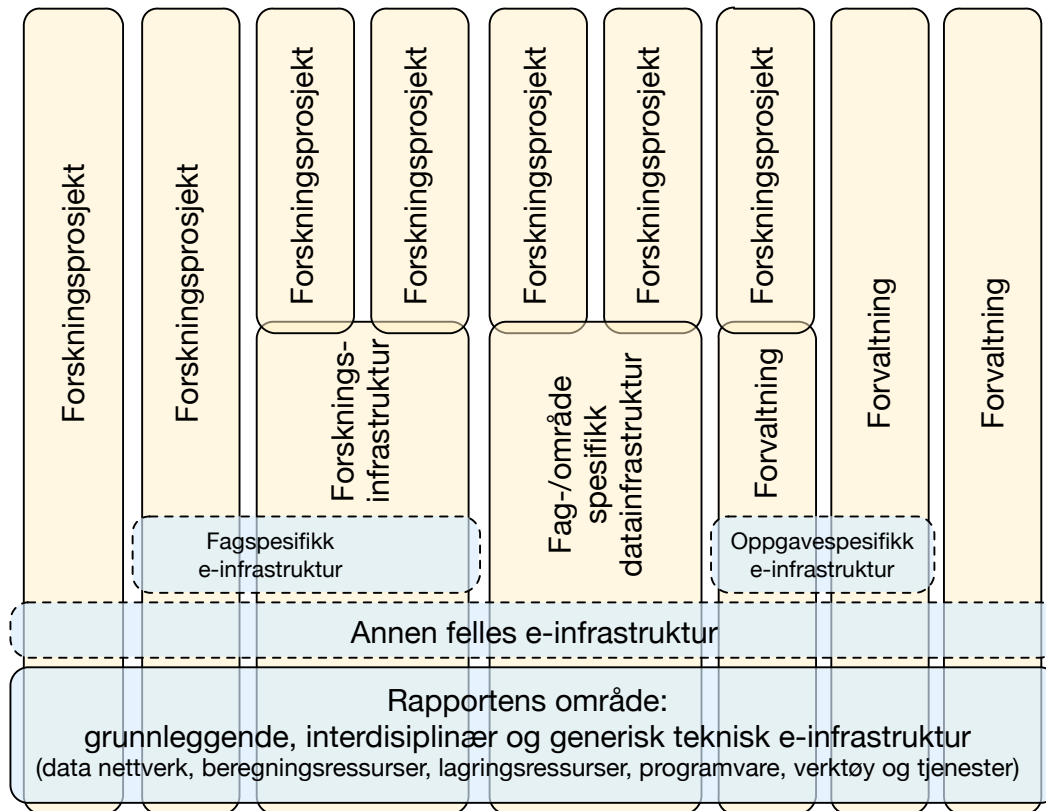
<sup>4</sup> <https://www.uninett.no/lagring>

<sup>5</sup> <https://www.regjeringen.no/no/dokumenter/nasjonale-strategi-for-tilgjengeliggjoring-og-delning-av-forskningsdata/id2582412/>

<sup>6</sup> <https://ehelse.no/helsedataprogrammet>

Samtidig som det er nødvendig med samordning for å unngå dobbeltarbeid eller sprikende anbefalinger fra prosessene, gir situasjonen også mulighet for utnyttelse av synergier og komplementering/arbeidsdeling.

Denne rapporten diskuterer finansieringsordninger for investering i nasjonal e-infrastruktur for forskning og forvaltning i Norge frem mot 2030. Dette betyr drift og bruk av e-infrastrukturen, utvikling og etablering av verktøy for HPC og lagring av store datamengder, samt investering i avansert brukerstøtte. Denne utredningen avgrenses derfor til underliggende generisk e-infrastruktur som tillater å støtte både enkeltprosjekt og andre e-infrastrukturer som i seg selv bruker de generiske e-infrastrukturtenestene som beskrives her. Den vitenskapelige programvaren som prosjektene har utviklet eller kjøpt for sin forskning er ikke en del av den generiske infrastrukturen. Et velfungerende forskningsnett, som er essensielt for en velfungerende infrastruktur for tungregning og lagring, er ikke en del av denne utredningen. Utgangspunktet er faktiske og forventede behov for e-infrastruktur av generisk karakter som legger til rette for å bli benyttet av fagspesifikke og administrative tjenester og prosesser. Den generiske e-infrastrukturen kan ha støttefunksjoner for datakuratering som går på tvers av disipliner, mens direkte støtte for datakuratering krever fagnærhet og organiseres derfor som oftest av fagspesifikke infrastrukturer eller nettverk av slike. Dette betyr at kuratering av data, som kan forventes å bli av økende betydning etter hvert som mengden data øker og antallet fagfelt som lagrer og bruker data øker, ikke vil omfattes av denne rapporten. I mange tilfeller vil e-infrastruktur også være i behov av fagnærhet, f.eks. e-infrastruktur som ikke kan deles utenfor et bestemt forskningsprosjekt eller som er knyttet til en bestemt forskningsinfrastruktur, og vil derfor kreve en annen organisering og løsning enn for de generiske tjenestene som behandles i denne rapporten.



Figur 2: Generisk/felles E-infrastruktur er tverrgående og "horisontal". Rapporten diskuterer finansiering av grunnleggende generisk og teknisk e-infrastruktur som beskrevet i teksten over.

## Behov for e-infrastruktur for forskning og forvaltning

Vitenskapelige beregninger og datadrevet vitenskap har for lengst blitt akseptert som en tredje vitenskapelig arbeidsmåte, som skiller seg fra og som komplementerer de tradisjonelle vitenskapelige metoder med eksperimenter og teoriutvikling. Beregninger gir mulighet for å skape mer nøyaktige og realistiske teoretiske modeller og gjør det mulig å analysere og identifisere nye teoretiske sammenhenger fra de store datamengdene som genereres fra for eksempel gensekvensering, satellittobservasjoner, og fra store mengder av små sensorer som en del av Internet of Things (IoT). Analyse av store datamengder påvirker også hvordan menneske-maskin og menneske-menneske-interaksjoner skjer, noe sosiale medier og smarttelefoner er gode eksempler på.

I dag er vitenskapelige beregninger og analyse av store datamengder en essensiell og integrert del av hvordan forskning utføres innenfor de fleste disipliner, som for eksempel klimaforskning, havforskning, biovitenskap/medisin, helse og omsorg, digitalisering og maskinlæring, språkanalyse, økonomi, transport og materialvitenskap, for å nevne noen

forskningsområder. Forskningen gir på sin side nye verktøy og grunnlag for innovasjon som kan brukes til å bedre forvaltningen av våre nasjonale ressurser.

Ta for eksempel innsamling av data fra skip, satellitter og ubemannede farkoster eller observasjonsplattformer. Disse dataene gir rike muligheter for å forstå det komplekse samspillet som eksisterer mellom ulike fysiske parametere som temperatur, saltinnhold, havstrømmer og lignende i havområdene våre, og hvordan dette i sin tur kan påvirke biologien og artsmangfoldet. I kombinasjon med avanserte beregningsmodeller vil dette gi bedre verktøy for en bærekraftig forvaltning av naturressursene i norske havområder, inkludert kommersielle fiskebestander.

Når data fra havområdene kombineres med atmosfæremålinger vil dette gi bedre datagrunnlag for klimamodellering. I tillegg til økte behov for observasjonsdata med stadig høyere oppløsning i rom og tid, vil klimamodellene i fremtiden kreve stadig høyere numerisk oppløsning, og dermed stadig mer kostbare beregninger, for å kunne gi den nødvendige *lokale* innsikten i hvordan klimaendringene påvirker farene for flom, ras og ekstremvær<sup>7</sup>. Slik informasjon er kritisk for at vi skal kunne gjøre gode samfunnsmessige investeringer i infrastruktur som er rustet for de forventede effektene av klimaendringer. På samme måte vil modellering av romvær være essensielt for å unngå lammelse av samfunnskritisk infrastruktur ved kraftige solstormer.

Bruk av store mengder data fra trafikken i kombinasjon med stor regnekraft vil kunne gjøre det mulig å optimalisere trafikkflyten og på den måten gi store samfunnsgevinster ved redusert trafikkø, og som en følge av dette, bedre luftkvalitet og bedre helse. Tsjeckia har for eksempel gjennom flere år investert i slike systemer for å optimalisere trafikkflyten både på motorveier og i bysentra<sup>8</sup>.

Som en del av det grønne skiftet vil livssyklusanalyser og materialflytsanalyser være sentrale elementer for å få oversikt over den totale miljøpåvirkningen av ulike produkter fra vugge til grav. Slik kunnskap er viktig for en bærekraftig utvikling, men vil kreve både store datamengder og stor beregningskraft.

Innenfor helseforskning har det i de senere år skjedd en revolusjon gjennom mulighetene for gensekvensering og avansert teknologi for å danne bilder av det indre av menneskekroppen. Dette danner utgangspunkt for persontilpasset medisin og nye avanserte behandlingsformer. Metodene genererer store mengder data, som må analyseres med krevende beregninger og oppbevares etter strenge lover for personvern.

De viktige og nødvendige investeringer Norge gjør innenfor nasjonal infrastruktur for forskning<sup>9</sup>, skaper også indirekte et økende behov for e-infrastruktur. Dette følger av at

---

<sup>7</sup> Globale klimamodeller har for eksempel i dag en horisontal oppløsning på rundt 100 km, som er for grovt til å fange opp fenomener som tornadoer. Til dette kreves det en langt finere oppløsning, med tilsvarende økte behov for beregningskraft.

<sup>8</sup> <https://www.it4i.cz/3885-2/research-teams/it-for-disaster-and-traffic-management/?lang=en&lang=en>.

<sup>9</sup> <https://www.forskningsradet.no/prognett-infrastruktur/Forside/1224697900450>.

de eksperimentelle forskningsinfrastrukturene vil generere store datamengder som må forvaltes i henhold til gjeldende strategier<sup>10</sup> og som ofte krever avanserte beregninger for å kunne analyseres og utnyttes fullt ut.

Utvalget har i arbeidet med denne rapporten valgt å ikke selv sammenstille et vitenskapsgrunnlag med begrunnelser for behov for fremtidige investeringer i nasjonal e-infrastruktur. Imidlertid er mange momenter som ble trukket frem i eSOP (eScience Opportunities Panel) sin rapport fra 2012 fortsatt gyldige<sup>11</sup>. Videre har utvalget gjennomført en prosess med bred involvering av ulike norske aktører innenfor e-infrastruktururområdet (se Appendiks 1), og har på den måten skaffet seg god oversikt over hvordan ulike aktører ser på utviklingen og behovene for e-infrastruktur i fremtiden. Utvalget har også basert seg på ferske analyser gjort på europeisk nivå, som The Scientific Case for Computing in Europe utviklet av Partnership for Advanced Computing in Europe (PRACE)<sup>12</sup>, samt grunnlaget for EU-kommisjonens beslutning om betydelige investeringer i kunstig intelligens, med et forslag om 7 milliarder Euro investert i perioden 2021-2027<sup>13</sup>. Utvalget viser også til en impact-analyse av Engineering and Physical Sciences Research Council (EPSRC) sine investeringer i e-infrastruktur<sup>14</sup>. Rapporten viser at *Return on Investments (ROI)* på disse e-infrastrukturinvesteringene ligger et sted mellom 6,5:1 og 19,5:1. Utvalget mener impact-analysen har stor overføringsverdi til norske forhold hvis vi ser bort ifra ROI for industri/næringsliv. Selv om vi antar en lavere ROI for industrien i Norge, så vil ROI fortsatt være på 4,5:1 og 13,4:1. Det er med andre ord en betydelig merverdi for forskning og utdanning knyttet til investeringer i e-infrastruktur.

I forbindelse med etableringen av EuroHPC JU ble det gjort et omfattende arbeid for EU-kommisjonen med å utrede status på dette området i Europa og verden forøvrig<sup>15</sup>. I fremlegget til kommisjonen fremgår det at HPC anses som grunnleggende for å adressere de store vitenskapelige, økonomiske og samfunnsmessige utfordringene Europa står overfor. Av utfordringer Europa står overfor pekes det bl.a. på:

- a) Europa har uforholdsmessig få HPC-anlegg blant de kraftigste i verden, og ingen HPC-anlegg blant de 10 kraftigste hvis man ser bort fra Sveits. I tillegg er man helt avhengig av ikke-europeisk teknologi og leverandører for slike anlegg.
- b) Noen europeiske forskere har behov for regnekraft som går langt over det som er tilgjengelig innenfor EU. I tillegg er det å bruke tjenester utenfor EU-området i

---

<sup>10</sup> <https://www.regjeringen.no/no/dokumenter/nasjonal-strategi-for-tilgjengeliggjoring-og-delning-av-forskningsdata/id2582412/>

<sup>11</sup> <https://www.forskningsradet.no/globalassets/publikasjoner/1253978594086.pdf>

<sup>12</sup> <http://www.prace-ri.eu/third-scientific-case/>.

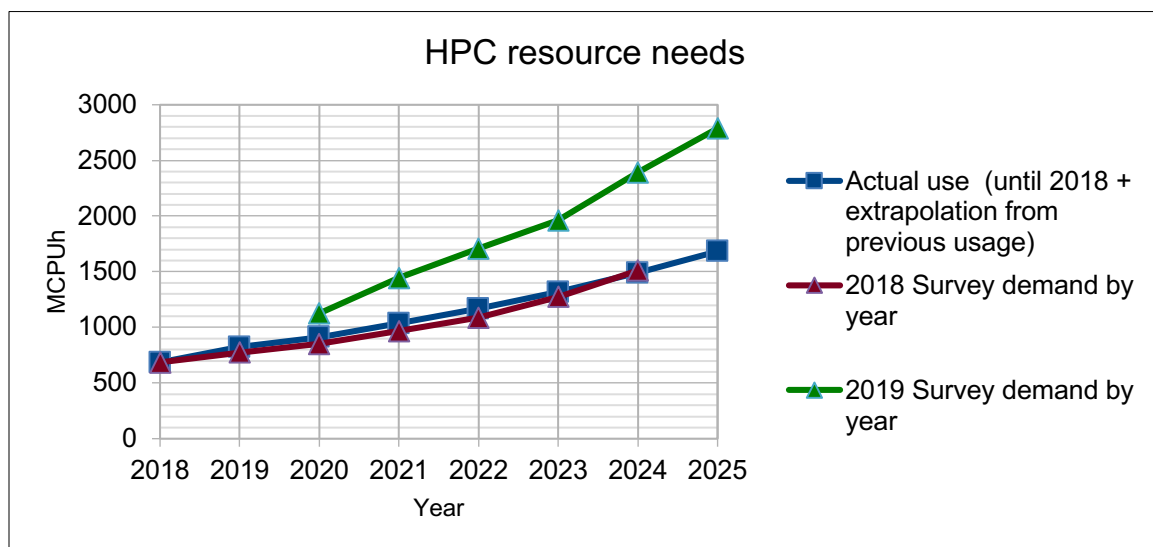
<sup>13</sup> [https://ec.europa.eu/commission/news/artificial-intelligence-2018-dec-07\\_en](https://ec.europa.eu/commission/news/artificial-intelligence-2018-dec-07_en).

<sup>14</sup> *The impact of EPSRC's investments in High Performance Computing Infrastructure. Final report.* London Economics (2019).

<sup>15</sup> Council Regulation on Establishing the European High Performance Computing Joint Undertaking, COM(2018) 8 final, 2018/0003 (NLE)

- mange tilfeller et problem i forskningsprosjekter som involverer persondata eller andre former for sensitive data, f.eks. forretningshemmeligheter.
- c) Investeringene som gjøres er i liten grad koordinert. Det er mangelfull strategi og implementering på området. I tillegg er anvendelse av HPC i industrien langt bak det man f.eks. ser i USA, Japan og Kina. Regionalt sett og sammenlignet med disse landene er det en underinvestering i HPC i Europa i størrelsesorden 500-700 millioner Euro per år.
  - d) Det investeres for lite i utvikling av teknologi og produkter for HPC, sett relativt til andre land og økonomier. Det er også i for liten utstrekning samarbeid mellom offentlig finansiert forskning og innovasjon og kommersielle aktører.

Behovskartlegging for HPC hos norske forskere som ble gjennomført sommeren 2019 gir følgende bilde:



Figur 3: Utvikling av HPC behov i Norge

Figuren viser dels at forventet økning i behov for regneressurser blant norske forskere har økt betraktelig fra 2018 til 2019, og at dette behovet langt overgår det behovet som lå til grunn for Sigma2 sin søknad til INFRASTRUKTUR-programmet høsten 2018.

### Nasjonal organisering

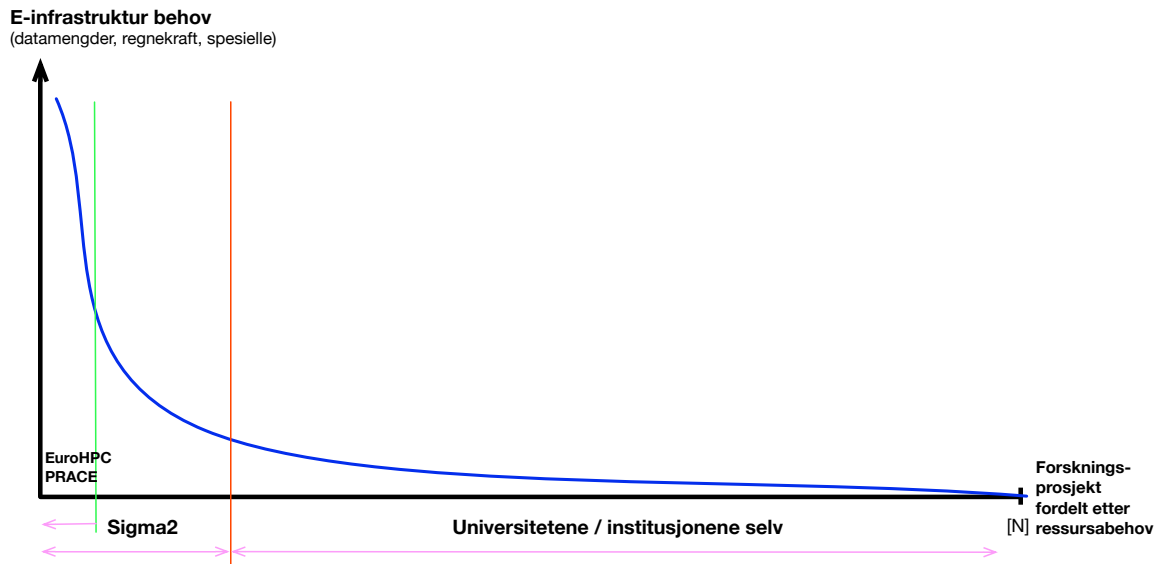
Historisk har tungregnesystemer alltid vært delt på tvers av disipliner og tilgjengelige for flere forskningsprosjekt. Det store flertall av forskningsprosjekt har hver for seg ikke behov som sikrer 100% utnyttelse av et regneanlegg gjennom dets levealder. Dette kan bare oppnås når flere utnytter den samme ressursen. Utnyttelsesgraden av dagens delte e-infrastruktur er således svært høy, noe som er naturlig gitt ressursbegrensningene som denne rapporten bl.a. adresserer. Delte ressurser kan gjøres større og derigjennom gjøre det mulig å håndtere større problem, og man får stordriftsfordeler med innkjøp, drift og opprettholdelse av kompetanse.

En sentral problemstilling er hva som skal organiseres og finansieres nasjonalt versus hva som er institusjonenes eget ansvar. Selv om det ikke finnes et skarpt skille, er strategien

at det er mest effektivt å organisere nasjonalt den type e-infrastrukturressurser som krever store investeringer og som er komplekse og kompetansekrevende å drifte. Dette kan illustreres som i Figur 4 hvor de største og mest komplekse behovene søkes håndtert i fellesskap nasjonalt, eventuelt internasjonalt, mens mindre og enklere behov håndteres lokalt ved institusjonene. Dette medfører videre at den største andel av prosjekter vil få sine e-infrastruktur-behov håndtert lokalt gjennom institusjonene.

Dæhlen-utvalget<sup>16</sup> trakk frem metasenter konseptet - en delt, nasjonal e-infrastruktur – som et vesentlig element i en kosteffektiv utbygging av norsk tungregning, hvor den fysiske e-infrastrukturen lokaliseres til et fåtall driftssentra, men hvor kompetansen bygges ved alle de sentrale forskningsutførende institusjonene. Flere institusjoner fremhevet også i høringsuttalelsene til Dæhlen-rapporten viktigheten av å bruke de kompetansemiljøene som finnes ved de fleste institusjoner som har en betydelig bruk av e-infrastruktur tjenester til å bygge et nasjonalt, distribuert brukerstøtte og kompetansemiljø (metasenter).

Utvalget slutter seg til de vurderinger som lå til grunn for Dæhlen-utvalget sin analyse og oppfatter at høringsinnspillene til Dæhlen-utvalgets rapport fra institusjoner som er store brukere av e-infrastruktur tjenester slutter opp om en nasjonalt koordinert e-infrastruktur med en distribuert modell for bruk og oppbygging av kompetanse. Utvalget har tatt dette som utgangspunkt for sitt arbeid med en ny nasjonal finansieringsmodell og viser ellers til Dæhlen-utvalget sin rapport om merverdiene av en nasjonal organisering.



Figur 4: Skjematisk fremstilling av ansvar og rollefordeling mellom lokal (ved institusjonene), nasjonal og internasjonal e-infrastruktur. Denne rapporten handler om organisering og finansiering av generisk og delt e-infrastruktur til venstre for den røde vertikale streken og involverer både nasjonale og internasjonale løsninger.

<sup>16</sup> <https://titan.uio.no/sites/default/files/unmanaged/samlet-versjon-med-alle-vedlegg.pdf>

## Ambisjonsnivå

Hva som er riktig nivå for investering i nasjonal e-infrastruktur er et nøkkelspørsmål. I og med at e-infrastruktur er et sentralt verktøy for forskningen innenfor de aller fleste fagområder<sup>17</sup>, så kan spørsmålet i stor grad omformuleres til hvilket ambisjonsnivå man ønsker for norsk forskning. Dette vil omfatte både grunnforskning og anvendt forskning og forskning som skjer i norske bedrifter. Utvalget finner ingen grunn til å anta at norske forskningsmiljøer har mindre bruk og behov for e-infrastrukturtenester enn sine utenlandske kollegaer, snarere tvert imot.

I regjeringens forskningsbarometer for 2018 måles kvaliteten på norsk forskning opp mot land av sammenlignbar størrelse og forskningsinnsats<sup>18</sup>, spesifikt våre nordiske naboer Sverige, Danmark og Finland, samt Nederland og Østerrike. Ambisjonen må være at norsk forskning bør være sammenlignbar med de beste av disse. For at norske forskere skal ha de nødvendige forutsetninger for å kunne nå dette målet, må etter utvalgets oppfatning norske investeringer i e-infrastruktur være sammenlignbare med investeringsnivået i de beste av disse landene. Utvalget mener også at de store norske havområdene med rike naturressurser og Norges internasjonale ansvar for å bidra til å forstå og dempe effektene av klimaendringer gjennom vårt store nærvær i Arktis tilsier at Norge bør ha et enda høyere ambisjonsnivå.

Som beskrevet innledningsvis inngår det mange ulike elementer i begrepet e-infrastruktur. Når vi skal måle oss mot andre land må dette gjøres på en slik måte at det er sammenlignbart og innenfor avgrensingen av denne rapporten. Når det gjelder tungregning, så har det gjennom mange år blitt publisert en liste over de til enhver tid kraftigste regneanlegg i verden<sup>19</sup>. Investeringer i denne typen regneanlegg utgjør i dag en stor del av det totale finansieringsbehovet, samtidig som kostnadene ved datalagring forventes å stå for den største veksten i årene som kommer. Investeringer i tungregning i forhold til landenes BNP eller befolkningstall vil per i dag allikevel gi en god målestokk på hvordan ulike land ser på viktigheten av investeringer i e-infrastruktur for forskning. Vi har her kun tatt utgangspunkt i åpne akademiske regneanlegg fra denne listen, og ettersom Norge har uforholdsmessig stor BNP, valgt å sammenligne relativt til landenes befolkningstall.

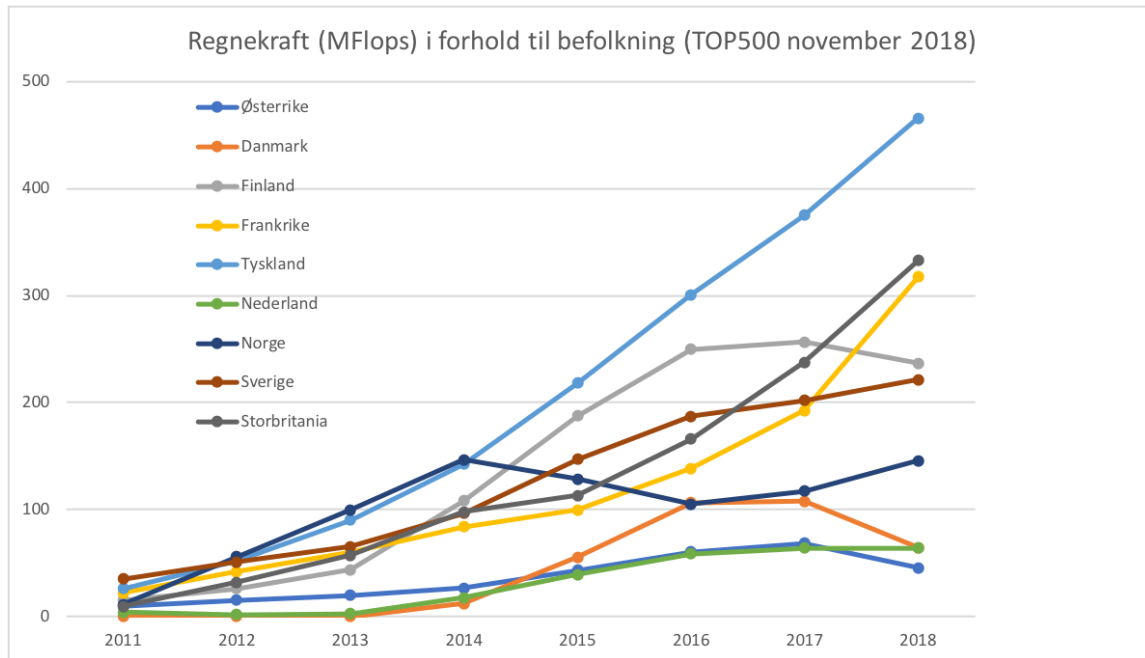
---

<sup>17</sup> Se appendiks 7 for oversikt over fordeling av ressursforbruk mellom fagområder/ vitenskapelige disipliner når det gjelder HPC og storskala lagring.

<sup>18</sup> Forskningsbarometeret 2018. Kunnskapsdepartementet.

<sup>19</sup> <https://www.top500.org/>





Figur 5: Regnekraft per innbygger i utvalgte land basert på Top500 listen over åpne akademiske regneanlegg. For å unngå for store svingninger som følge av investeringer, er regnekraften i et gitt år snittet av regnekraften de tre foregående år, dvs. at regnekraft i 2015 er snittet av åren 2013, 2014 og 2015.

Figur 5 viser regnekraft per innbygger midlet over en treårsperiode for våre fem referanseland<sup>20</sup>. Fra år til år gir dette et direkte mål på investeringer i tungregneanlegg, og antagelig et godt bilde av nivået på finansiering av e-infrastruktur i de samme landene. I figuren har vi også inkludert data for noen europeiske land hvor investeringer i e-infrastruktur har blitt gitt særlig prioritet for å plassere disse landene i den internasjonale forskningsfronten på dette området (Storbritannia, Frankrike og Tyskland). Det er verdt å merke seg at Norge frem 2014 var på linje med Tyskland i investeringer per innbygger, men har siden sakkraftig akterut. Sveits er ikke inkludert i denne figuren, da Sveits sine investeringer omsatt til regnekraft per innbygger er nesten fem ganger så store som Tyskland. Sveits betrakter investeringer i utdanning, forskning og utvikling som essensielle for velstanden i landet. Dette inkluderer store investeringer i forskningsinfrastruktur som HPC er en viktig del av<sup>21</sup>.

<sup>20</sup> Danmark og Østerrike skiller seg ut med tilsynelatende veldig lavt nivå på investeringer. Det har sammenheng med både måten investeringer i e-infrastruktur gjøres på og at investeringene som gjøres hver for seg ikke kommer med på listen over de 500 kraftigste regneanleggene i verden. Vi vil derfor i det følgende se bort fra disse i listen over referanseland.

<sup>21</sup> Utsagnet er hentet fra lederen av Sveits sitt nasjonale HPC senter, Thomas Schulthess hos CSCS.

Det er flere særnorske forhold som kan trekkes frem som argumenter for at Norge bør betrakte e-infrastruktur som så strategisk viktig at en forsterket satsing på linje med Tyskland, Storbritannia eller Sveits, bør vurderes:

- De enorme naturressursene i norske havområder – marine, olje/gass og mineraler – krever god forvaltning, noe som igjen krever store datamengder og komplekse matematiske modeller.
- Betydningen for samfunnsplanlegging og samfunnssikkerhet at vi kan beregne forventede effekter av klimaendringer, og vår evne til hurtig å vurdere konsekvensene av ekstremvær.
- Vår geografiske beliggenhet som gir et unikt fortrinn i innsamlingen av data fra satellitter i polare baner og de muligheter som ligger i innhenting og analyse av data fra ulike fjernmålingskilder.
- Mulighetene som åpner seg ved å kombinere biobankene og ny teknologi for blant annet DNA-sekvensering med de gode norske helseregistrene.
- Norge er et teknologisk modent samfunn og har derfor et meget godt utgangspunkt for utvikling av morgendagens næringsliv, et næringsliv som i stor grad vil bygges på ulike former for bruk av e-infrastruktur.

Disse forholdene taler for at investeringer på linje med toppnivået i Sverige, Finland, Nederland og Østerrike må betraktes som et minimumsnivå for Norge, og at det er gode grunner for å ha et enda høyere ambisjonsnivå.

Om vi, basert på BNP per innbygger for EU-regionen, overfører det EU-kommisjonen mener er nivået for underinvesteringer til norske forhold, vil det tilsi at graden av underfinansiering i Norge tilsvarer 130 – 190 MNOK per år. Dette svarer til mer enn en dobling av investeringer i e-infrastruktur relativt til nivået for 2017-2019.<sup>22</sup>

Behovet for økte investeringer i e-infrastruktur gjør det maktpåliggende at investeringene gjøres kosteffektive. Selv om ulike forsknings- og forvaltningsoppgaver har unike behov for e-infrastrukturtenester, vil alle disse bygges på toppen av en generisk e-infrastruktur bestående av tungregnerressurser, datalagringsressurser og spisskompetanse innenfor drift og anvendelse av slike generiske e-infrastrukturer.

Utvalget mener at kosteffektivitet, og det nasjonale ambisjonsnivået, best kan nås gjennom å ha en felles nasjonal plattform for generisk e-infrastruktur. Forsknings- og forvaltningsinstitusjoner med offentlig finansiering bør derfor bidra til denne nasjonale e-infrastrukturplattformen, og på den måten fokusere på sine faglige behov, fremfor å bygge opp egne generiske e-infrastrukturer.

Selv om plattformen er nasjonalt koordinert, kan den fortsatt være distribuert hvis dette gir best kosteffektivitet. Den nasjonale plattformen vil også kunne ha en helhetlig tilnærming til hvor stor grad av investeringene som skal skje nasjonalt versus internasjonalt, og om investeringene bør skje i UH-sektoren eller gjennom private tilbydere. De forslag som fremmes i denne rapporten må derfor ikke oppfattes som å

---

<sup>22</sup> Sigma2 bruker i snitt ca. 132 MNOK per år til e-infrastruktur.

antydde behov for investeringer i norske regnearbeid ved norske universitets- og høyskolemiljø, men snarere peke på et investeringsbehov som kan løses gjennom kommersielle og internasjonale tilbydere, herunder bruk av skytjenester. Et førende prinsipp vil derimot være at investeringene må kunne svare på brukerbehovene og gi kosteffektive løsninger. Utvalget vil allikevel peke på merverdien i at noe kompetanse opprettholdes nasjonalt i forhold til å møte nye teknologiske utviklingstrekk innen datalagring og beregningsvitenskap.

## Finansieringsmodeller

En robust finansieringsmodell bør ha flere og om mulig uavhengige kilder. E-infrastruktur må jevnlig skiftes ut og oppgraderes, i takt med den teknologiske utviklingen og for å kunne møte nye og økende behov. Dette krever jevnlig investeringer. Samtidig må e-infrastrukturen driftes, kompetanse bygges, data vedlikeholdes, og tjenester og støtte tilbys brukere. Modellen må dekke både investerings- og driftskostnader. Teknologiske muligheter kommer og går i akselererende tempo, mens data og kompetanse representerer langsiktige verdier. En finansieringsmodell må gi forutsigbarhet, men samtidig ha fleksibilitet overfor utviklingen innenfor teknologi og brukerbehov.

Ulike finansieringsmodeller kan rettes inn mot ulike deler av denne dynamikken, f.eks. en grunnfinansiering for forutsigbarhet og langsiktighet, og en mer mulighetsorientert komponent for rask pilotering av ny teknologi.

Finansieringsmodellene kan deles i tre kategorier:

1. Grunnfinansiering<sup>23</sup> fra forskningsinstitusjoner som universiteter, forsknings- og forvaltningsinstitutter, Forskningsrådet og offentlige instanser med signifikant forbruk eller ut fra disse institusjonenes strategiske prioriteringer.
2. Brukerfinansiering.
3. Finansiering fra åpne konkurransearenaer.

Grunnfinansiering danner grunnlaget for langsiktighet og en bærekraftig organisering, først og fremst ved å etablere grunnleggende infrastruktur og organisatoriske rammer som igjen kan implementere basistjenester, hente inn brukerfinansiering og søke konkurranseutsatt finansiering. For forskningsinstitusjoner vil bidrag til denne typen finansiering reflektere institusjonens strategiske valg gjennom å gi sine forskere og studenter tilgang til ressurser som det ikke er effektivt å stille til rådighet lokalt, samtidig som den gir institusjonen medbestemmelse. Den kan også understøtte institusjonenes ansvar for ivaretagelse og tilgjengeliggjøring av forskningsdata ut over prosjekters levetid. Tilsvarende kan andre offentlige instanser, f.eks. med forvaltningsoppgaver som genererer store mengder data, eller som av andre grunner bruker eller drifter datainfrastrukturer, bidra med grunnfinansiering for å lagre og gjøre sine data tilgjengelige for forskning i den nasjonale e-infrastrukturen. Det er kostnadseffektivt og

---

<sup>23</sup> Med grunnfinansiering forstås her et langsiktig finansielt bidrag som kommer uten bindinger.

kvalitetsfremmende at forskningsinfrastrukturer og forskningsprosjekt i størst mulig grad deler underliggende generisk e-infrastruktur og tilhørende spisskompetanse. Den nasjonale e-infrastrukturen er i tillegg utgangspunkt for deltakelse i en rekke internasjonale aktiviteter som er med på å strekke lagmodellen for ressurs- og kompetansedeling ut over nasjonalt nivå.

Brukerfinansiering er en mekanisme for å få avstemt finansiering med faktiske behov. Med brukere kan vi her både forstå enkeltforskere eller større forskningsgrupper, men også offentlige forvaltningenheter som igangsetter aktivitet som krever e-infrastrukturressurser. Brukerfinansiering vil vanligvis dekke driftskostnader, mens investeringskostnader må dekkes av grunn- og konkurranseutsatt finansiering. I tillegg kan brukerfinansiering medvirke til bevisstgjøring av kostnader hos brukere og derigjennom større grad av nøkternhet i forhold til planlegging og gjennomføring av prosjekter. Uten brukerfinansiering er det i stor grad volumet av tilgjengelige ressurser og innbyrdes konkurranse om disse ressursene, basert på vitenskapelig kvalitet, som setter grenser.

Det er allikevel utfordringer med brukerfinansiering. Først og fremst gjelder dette forutsigbarhet og langsiktighet. For det andre er det nødvendig med en rettferdig modell og effektive mekanismer for kanalisering av finansiering denne veien, for å gi like brukerkostnader og effektiv administrasjon. For det tredje er brukerfinansiering reaktiv i forhold til investeringer som trenger avsetninger i forkant. Sist så forutsetter brukerfinansiering et visst volum i bruk av de tjenester som etableres, jf. kostnader ved å etablere og drifte ulike leiesteder og brukerfinansierte tilbud. Utvalget vil også peke på at brukerfinansiering ofte vil være knyttet til prosjekter av begrenset varighet (3-5 år), og at slike prosjekt ikke kan ivareta myndigheters og finansierende institusjoners krav om lagring av data for lengre tid, for eksempel i forhold til behov for kunne reprodusere resultater. Slik langtidslagring vil måtte dekkes av grunnfinansieringsmekanismer.

Den konkurranseutsatte finansieringen fra Forskningsrådets INFRASTRUKTUR-program medfører faglig og eksternt evaluering - i tillegg til forskningspolitisk prioritering - av nivået på investeringer til e-infrastruktur. Dette gir fleksibilitet og mulighet for mer dynamiske justeringer ut over nivået på grunnfinansieringen. Den relativt hyppige syklusen av utlysninger til dette programmet gir mulighet for å være i rimelig takt med behovsutviklingen. Konkurranseutsatt finansiering er også bedre egnet til pilotering av e-infrastruktur basert på ny teknologi, i kontrast til grunnfinansiering som bør anvendes mot et omforent nivå av etablerte og grunnleggende e-infrastrukturteknologier. Konkurranseutsatt finansiering er imidlertid ofte organisert som prosjekter med tidsavgrensede finansiering (3 – 4 år), med de dilemmaene dette skaper (organisatorisk, kompetanse- og ressursmessig) for langsiktig implementering og drift av infrastruktur som krever hyppig fornyelse.

## Dagens finansieringsmodell

Det at grunnfinansieringen har ligget fast siden oppstart har vist seg å være en utfordring med dagens finansieringsmodell. Ettersom behovet for e-infrastruktur har økt, har den nasjonale e-infrastrukturen blitt stadig mer kritisk avhengig av suksess i kampen om konkurranseutsatt finansiering. Brukerbetaling er innført, men har ikke kunnet ta høyde

for de økte behovene. Figur 5 viser utviklingen av de ulike typene finansiering, og viser at mens grunnfinansieringen har vært konstant siden etableringen av Sigma2, så har den konkurranseutsatte finansieringen økt dramatisk i samme periode. Resultatet er at det opprinnelige forholdet mellom grunnfinansiering og konkurranseutsatt finansiering er i ferd med å bli snudd på hodet i forhold til utgangspunktet 2015-2017.

Grunnfinansieringen fra universitetene har heller ikke vært indeksregulert selv om kostnaden for de tjenester Sigma2 kjøper fra de samme universitetene har vært det.

Finansieringen av Sigma2 har i perioden 2015-2018 bestått av:

MNOK	2015	2016	2017	2018	2019
Forskningsrådet årlig (grunnfinansiering)	29.0	29.0	25.0	25.0	25.0
Universitetene årlig (grunnfinansiering)	50.0	50.0	50.0	50.0	50.0
Konkurranseutsatt		37.5	37.5	57.9	57.9
Konkurranseutsatt NeIC				5.2	5.3
Brukerbidrag			0.3	4.8	7.4
Internasjonal	1.5	2.3	6.9	3.3	5.6
SUM	80.5	118.8	119.7	146.2	151.2

*Tabell 1: Utvikling i ulike finansieringsbidrag for Sigma2 (tall for 2019 er estimat)*

Den gjeldende modellen for Sigma2 er lagt opp med 10 års horisont med evaluering hvert femte år (såkalt 5+5 evaluering). Ved hver evaluering skal fordelingsnøkkelen for grunnfinansiering fra partnerinstitusjonene revideres. Prosjektets langsiktige finansiering kan justeres hvert femte år gjennom at nye avtaler for en 5+5 års-periode vil inngås. Det gjelder også grunnfinansieringen fra Forskningsrådet som i dag er på 25 MNOK.

Vi ser i dag at (den eksisterende) grunnfinansieringen alene ikke er tilstrekkelig til å finansiere et minimumsnivå av tjenester. I en situasjon hvor man ikke lykkes med å sikre konkurranseutsatt finansiering, står man i fare for ikke å kunne opprettholde helt nødvendige basistjenester. Institusjoner vil da i verste fall tvinges til å løse disse behovene hver for seg, noe som fører til en totalt sett mindre kosteffektiv løsning.

Et annet problem i dagens finansieringsmodell er selve størrelsen på den konkurranseutsatte finansieringen som baserer seg på søknader til Forskningsrådets INFRASTRUKTUR program som har utlysninger annethvert år. Finansiering av

#### **Om nasjonal satsing på forskningsinfrastruktur (INFRASTRUKTUR programmet)**

Mye av forskning krever kostbar infrastruktur, eksempelvis elektronmikroskop, maskiner for sekvensering av gener og måling av partikler i luft eller vann. Dette er av så stor betydning for norsk forskning at det lenge har vært et eget program administrert av Forskningsrådet, INFRASTRUKTUR, med utlysninger annen hvert år.

Hovedmålet med Forskningsrådets finansiering av nasjonal forskningsinfrastruktur er at norske forskningsmiljøer og næringsliv skal ha tilgang til relevant og oppdatert infrastruktur som understøtter forskning av høy kvalitet, som i sin tur vil bidra til å møte samfunnets kunnskapsutfordringer. Ansvarsfordelingen mellom forskningsinstitusjonene og Forskningsrådet innebærer at Forskningsrådet skal bidra til å samordne investeringene i de nasjonale forskningsinfrastrukturene. Gjennom INFRASTRUKTUR skal Forskningsrådet sikre en god kobling mellom finansiering av forskningsinfrastruktur og øvrig forskningsfinansiering, samt en helhetlig vurdering av balansen mellom nasjonal investering og deltakelse i internasjonale forskningsinfrastrukturer.

Bevilgninger til nye infrastrukturprosjekter som har en kostnadsramme på 2-200 millioner kroner vedtas av Forskningsrådets styre. Beslutninger om internasjonalt forskningssamarbeid, herunder norsk deltakelse i etableringen av prosjekter i ESFRI Roadmap, som innebærer betydelige og varige forpliktelser knyttet til investeringer og medlemskontingenter fattes på departementsnivå.

Nasjonale forskningsfasiliteter som innebærer investeringer over 200 millioner kroner, vil håndteres på departements- eller regjeringsnivå, gjerne etter råd fra Forskningsrådet. Fortrinnsvis er dette midler som må komme i tillegg til den faste posten til forskningsinfrastruktur på statsbudsjettet.

Se også Appendiks 2 *Om e-infrastruktur*.

INFRASTRUKTUR kommer fra en fast sektorovergripende post i statsbudsjettet via

Kunnskapsdepartementet.

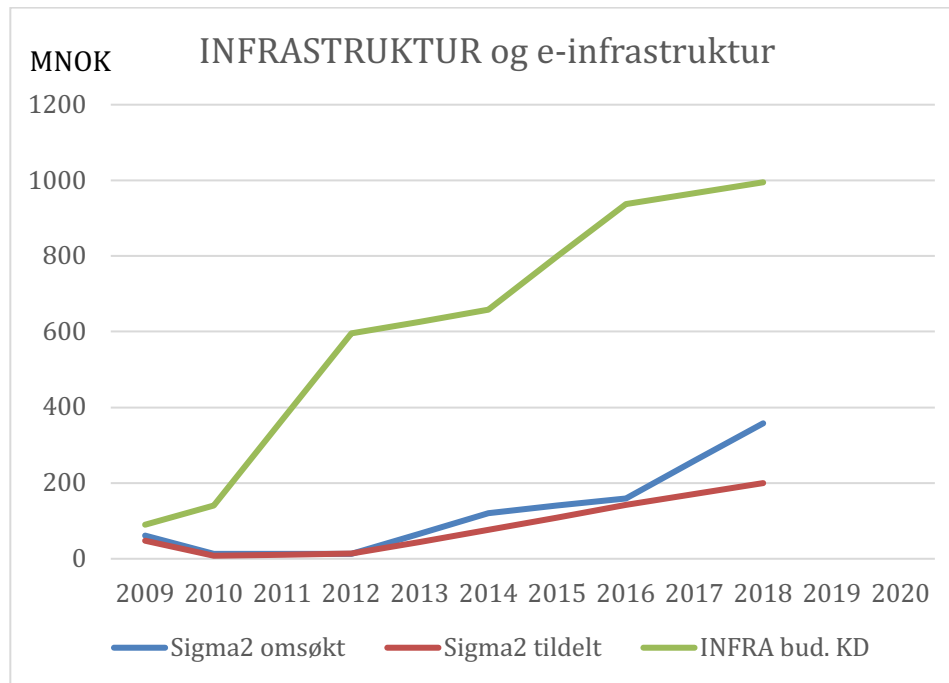
INFRASTRUKTUR finansiering retter seg mot etablering eller oppgradering av alle typer forskningsinfrastruktur av nasjonal viktighet slik disse begrepene er definert i strategien *Verktøy for forskning*.

INFRASTRUKTUR kan bidra til finansiering av drift av infrastrukturene i en oppstartsfasen, men vil være restriktiv med å gi støtte til langsiktig grunnfinansiering av infrastrukturenes drift.

Forskningsrådets INFRASTRUKTUR program har hatt varierende rammer fra utlysning til utlysning, men er i størrelsesorden 800-1000 MNOK. Dette skal finansiere etablering av ny forskningsinfrastruktur i Norge, og en typisk tildeling er på 50 MNOK.

For inneværende utlysning har Sigma2 søkt om 358 MNOK fra INFRASTRUKTUR programmet. Infrastrukturbehov over 200 MNOK krever i tillegg til god faglig evaluering en godkjenning på departementsnivå. Det omsøkte beløp fra Sigma2 er basert på nøkterne analyser av rapporterte behov, som også viser at

utviklingen kommer til å fortsette. Figur 6 viser utviklingen i omsøkt og innvilget beløp i forhold til rammer for utlysninger i INFRASTRUKTUR programmet



Figur 6: Utvikling av økonomisk ramme for INFRASTRUKTUR, omsøkt og tildelt andel til Sigma2, to-årlig. Tildeling fra 2018 utlysning ikke kjent.

Selv om forhandlingene ikke er endelig avsluttet, så vil ikke søknaden få innvilget et beløp over grensen på 200 MNOK, altså så vidt mer enn halvparten av det omsøkte beløp. Uten annen finansiering vil dermed Norge potensielt se en videre svekkelse av sin internasjonale posisjon innen beregnings- og databasert forskning og forvaltning.

Modellen som ble innført i 2014 og som skulle sikre langsiktighet og forutsigbarhet i investeringene i e-infrastruktur, viser seg dermed allerede nå å ha utfordringer, noe som reflekteres i Norges posisjon innen tilgjengelig tungregning i Figur 5. Problemene med dagens finansieringsmodell danner bakgrunnen for behovet for å diskutere alternative modeller for finansiering av e-infrastruktur.

Utfordringene med dagens modell kan kort oppsummeres i følgende punkter:

- Utilstrekkelig: Dagens modell er ikke finansielt bærekraftig sett i forhold til behovsutvikling og nasjonalt ambisjonsnivå.
- For stor andel konkurranseutsatt finansiering: Med dagens ramme for INFRASTRUKTUR-programmet utgjør behovene til e-infrastruktur en for stor andel av dette budsjettet, og nærmer seg 50% av rammen for typiske utlysninger. Dette skaper uforutsigbarhet og innbyrdes ekskluderende konkurranse med øvrig forskningsinfrastruktur som søker mot det samme programmet. Samtidig er generisk e-infrastruktur ofte en forutsetning for øvrig fagspesifikk forsknings- og forvaltningsinfrastruktur.

- Ingen systematisk tilnærming for å hente inn grunnfinansiering fra andre enn de fire breddeuniversitetene og NFR. I dag er det institusjonene selv som må ta initiativ for å bidra til grunnfinansieringen. Styringsmodellen er i tillegg knyttet til grunnfinansieringen.
- Ingen helhetlig tverrsektoriell/tverrprogrammatisk implementering av brukerbidrag.
- Ingen dynamikk i forholdet mellom konkurranseutsatt finansiering og grunnfinansiering. Grunnfinansieringen har ligget fast mens den konkurranseutsatte delen har økt uforholdsmessig mye for å kunne dekke et sterkt økende behov.
- Alle investeringer i den nasjonale generiske e-infrastrukturen dekkes av Kunnskapsdepartementets budsjett. Det er uklart i hvilken grad det er dialog mellom departementene når større satsinger som krever behov for økte investeringer i e-infrastruktur igangsettes. Det mangler en helhetlig tilnærming til felles behov på tvers av ulike sektorer/departement.
- I liten grad bidrag fra offentlig forvaltning. Dette er i tillegg med på å hemme utnyttelse av offentlig finansierte data til forskning, og kan også bidra til at forskningen og forvaltningens felles behov for e-infrastruktur tjenester løses på en totalt sett lite kosteffektiv måte, for eksempel ved at det bygges opp parallelle strukturer med de samme behovene for spisskompetanse.

### Alternativer for en ny finansieringsmodell

I lys av utfordringene som har blitt identifisert med dagens finansieringsmodell, kan vi trekke frem følgende ønskede elementer i en ny finansieringsmodell:

- Modellen må være robust i forhold til behovsutvikling og konjunkturer.
- Forholdet mellom grunnfinansiering, brukerfinansiering og konkurranseutsatt finansiering må være i rimelig innbyrdes balanse over tid.
- Modellen må kunne legge til rette for langsiktig planlegging og forpliktelser.
- En grunnfinansiering for generisk e-infrastruktur som sikrer at Norge kan nå sine forskningspolitiske ambisjoner, og med bidrag fra alle institusjoner med forbruk over et visst nivå.
- En styringsstruktur som inkluderer bidragsytere på en organisatorisk hensiktsmessig måte gjennom å ta hensyn til størrelsen på bidrag.
- Mekanismer for harmonisering av finansiering til generisk e-infrastruktur relativt til øvrig forskningsinfrastruktur og forskningsprosjekter med behov for e-infrastruktur.
- Tverrsektorielt bidrag fra alle relevante departement, inklusive bidrag fra offentlig forvaltning der det er naturlig og/eller formålstjenlig (f.eks. tilgjengeliggjøring av offentlige data til forskning og tjenester som krever tilgang til eller beregninger på slike data).



- Helhetlig tverrsektoriell/tverrprogrammatisk politikk og implementering av mekanismer for bidrag fra alle brukere av e-infrastrukturen.

Behovene for generisk e-infrastruktur er i dag tverrsektorielle. Denne trenden vil forsterkes fremover med økende mengder data i alle sammenhenger, datadrevet forskning og nye metodiske retninger som f.eks. maskinlæring og kunstig intelligens. Gode finansieringsmodeller må ta hensyn til dette og utnytte muligheter for flere offentlige kilder til finansiering.

Forskningsrådet vedlikeholder et veikart for forskningsinfrastrukturer som har vært gjennom en evaluering som bekrefter disse infrastrukturenes viktighet for norsk forskning. Nasjonal, generisk e-infrastruktur har ligget fast på dette veikartet, og vil kunne konkurrere om støtte fra INFRASTRUKTUR-programmet ved hver utlysning. Vi vet at behovet for e-infrastruktur er vedvarende og grunnleggende for forskning og at e-infrastruktur trenger jevnlig utskifting. Et vesentlig spørsmål er derfor hvor stor andel av den generiske e-infrastrukturen som bør dekkes via denne konkurranse-arenaen – som skal dekke mange behov for nasjonal forskningsinfrastruktur. Dersom dagens finansieringsmodell ikke endres, vil den generiske e-infrastrukturen ha behov for en stor andel av midlene tilgjengelig i dette programmet.

I en bærekraftig modell må økte investeringsbehov i stor grad kunne håndteres gjennom økning i grunn- og brukerfinansiering, og ikke som i dag gjennom stadig større søknader til åpne konkurransearenaer. Her ligger også en mulighet for tverrsektorielt samarbeid, herunder samfinansiering fra flere departementer.

En finansieringsmodell hvor det er en grad av kobling mellom nivået på grunnfinansiering og konkurranseutsatt finansiering er naturlig, f.eks. ved at deltakende institusjoner jevnlig justerer sitt bidrag til grunnfinansiering i forhold til et omforent nivå for det som regnes som grunnleggende generisk e-infrastruktur. Tildelte konkurranseutsatte midler kan da holdes i rimelig innbyrdes balanse med grunnfinansieringen. Alle institusjoner som bidrar til grunnfinansieringen, må etablere en prosess for justering av grunnfinansieringen. Dette vil samtidig kunne styrke den strategiske forankringen ved institusjonene. Utfordringer kan være de deltakende institusjonenes interne budsjettprosesser og redusert mulighet for langtidsbudsjettering, samt det at institusjonene kan ha ulike ambisjonsnivå som i tillegg kan endre seg over tid. Av denne grunn vil det med en slik modell være naturlig å revidere innbyrdes fordeling mellom deltakende institusjoner etter prinsipper som institusjonene slutter seg til.

I dagens modell fordeles de fire breddeuniversitetenes årlige grunnfinansiering på 50 MNOK etter historisk forbruk. Inntil nylig har andre institusjoners bruk av ressursene vært liten. Sigma2 kjøper i tillegg driftstjenester tilbake fra de fire breddeuniversitetene i størrelsesorden opp mot 50% av bidraget. Rollen som både kunde og leverandør gjør at de fire breddeuniversitetene har et spesielt nært forhold til nasjonal e-infrastruktur, noe som også gjenspeiles i dagens organisasjonsform.

I en ny finansieringsmodell kan bidraget til grunnfinansieringen økes ved at flere institusjoner bidrar i h.h.t. forbruk. Ut fra de seneste trender vil da SINTEF, NORCE, Nansensenteret og Akvaplan-Niva samt Universitetet i Stavanger og NMBU kunne komme inn med mindre andeler relativt til de fire breddeuniversitetene. Bidragsytende institusjoner vil forvente innflytelse, f.eks. gjennom representasjon i Sigma2 sitt styre.

Disse legitime ønskene må samtidig veies opp mot at et for stort styre kan bli ineffektivt. En løsning kan være rullerende representasjon. Vi viser her til kapittelet om utvalgets forslag til organisering. Samtidig er det viktig å peke på at kundemøtene mellom Sigma2 og de finansierende institusjonene vil være en viktig arena for å gi strategiske styringssignaler til selskapet.

Selv om brukerfinansiering er veletablert for bruk av mange forskningsinfrastrukturer, så er dette i liten utstrekning en vanlig finansieringsmodell for e-infrastruktur internasjonalt. Ikke minst gjelder dette for nasjoner som er i den internasjonale forskningsfronten, som USA og Sveits. Disse landene tilbyr gratis regneressurser ikke bare til sine egne forskere, men også til internasjonalt ledende forskningsmiljøer, alt basert på faglige kvalitetsvurderinger. Begrunnelsen for dette er dels at dette skjerper kravene til egne forskere, men også betydningen av å bygge opp kompetansemiljø innen avanserte e-infrastrukturtenester som kan støtte de fremste internasjonale forskningsmiljøene. Det bør samtidig påpekes at USA og Sveits har henholdsvis 1,7 og nesten 5 ganger så mye regnekraft per innbygger som det fremste landet i Figur 5, Tyskland.

En utfordring med brukerfinansiering av e-infrastrukturtenester er at noen av de fremste prosjektene som opererer på disse nasjonale eller internasjonale e-infrastrukturene bruker en betydelig andel av den totale kapasiteten<sup>24</sup>. For disse miljøene finnes det ikke prosjektfinansieringsmekanismer som kan dekke de reelle kostnadene ved denne bruken. Samtidig vinner mange andre brukermiljø på mulighetene for å ta i bruk de store nasjonale e-infrastrukturene, selv om disse brukermiljøene hver for seg står for en såpass liten andel av bruken at det ikke er kosteffektivt å administrere brukerbetaling. Ved at tilgang til e-infrastrukturressurser i liten grad baserer seg på betalingsevne/-villighet, gjør også at tilgang i all hovedsak baseres på faglige kvalitetskrav. Analyser gjort av Ressursfordelingskomiteen i Sigma2 viser av norsk beregningsbasert forskning i større grad publiserer på nivå to enn snittet av norsk forskning.

## Utvalgets anbefalinger

Denne rapporten har synliggjort at det er et akutt behov for økte investeringer i e-infrastruktur for forskning og forvaltningsoppgaver. Utvalget har tatt utgangspunkt i Dæhlen-utvalgets analyse og konklusjon at en felles nasjonal modell for finansiering og drift av generiske e-infrastrukturtenester, inkludert (avansert) brukerstøtte, er en mer kost-effektiv løsning enn at hver enkelt institusjon bygger opp slike tjenester på egen hånd. Dette underbygges av impact-analysen av EPSRC sine investeringer i beregningsressurser<sup>14</sup>.

Denne konklusjonen forsterkes av at bruken av dagens generiske e-infrastruktur er langt mer mangfoldig enn de fem institusjoner som i dag finansierer disse e-infrastrukturtenestene (for Forskningsrådets del gjennom Kunnskapsdepartementets budsjett). Vi har i Appendiks 3 oppsummert prosentvis bruk av dagens e-

---

<sup>24</sup> Verdien av HPC-ressursene som de 3 største prosjektene forbrukte i 2018 var på henholdsvis 9.1, 5.4 og 3.96 MNOK. (2018.1 og 2018.2 for NN2834K, NN2345K og NN4654K)

infrastrukturtenester (tungregning og datalagring) fordelt på ulike institusjoner og forskningsprogram. Vi merker oss allikevel at denne listen ikke gir en uttømmende oversikt over bruk av e-infrastrukturtenester: Universitetet i Agder har for eksempel investert i en ny lokal superdatamaskin, mens Havforskningsinstituttet inntil nylig i all hovedsak har driftet egne datalagringsløsninger.

Utvalgets klare anbefaling er at grunnfinansieringen bør økes slik at den bringer Norge opp på linje med toppnivået blant forskningsnasjoner vi ønsker å sammenligne oss med, Sverige, Finland og Nederland. Det bør vurderes om Norges spesielle forsknings- og forvaltningsbehov gjør at ambisjonsnivået bør økes ytterligere. Det er verdt å merke seg at Norge inntil 2014 gjorde investeringer på linje med det fremste landet i Figur 5, Tyskland.

Vi vil først diskutere ulike modeller for å øke grunnbevilgningen, før vi presenterer vårt forslag til en helhetlig finansieringsmodell.

### **Mulighet: Økt grunnfinansiering fra deltakende institusjoner**

Grunnfinansieringen kan økes gjennom å øke bidraget fra de fire breddeuniversitetene, fra Norges Forskningsråd og/eller ved å hente inn bidrag fra flere institusjoner. Selv om de fire breddeuniversitetene fortsatt dominerer bruken av nasjonal e-infrastruktur, står institusjoner som i dag ikke bidrar til finansieringen, for 15% av det totale tungregneforbruket på de nasjonale ressursene, og nær 30% av den totale datalagringskapasiteten. Hvis en for stor andel av de nasjonale ressursene brukes av andre enn finansierende institusjoner, vil dette kunne svekke oppslutningen om de nasjonale løsningene blant dagens finansierende institusjoner. Det kan derfor være naturlig at det defineres en nedre grense for størrelse på bruk av nasjonale e-infrastrukturtenester og at det forventes at institusjoner bidrar hvis deres bruksandel av ressursene overstiger denne grensen, eksempelvis 1% av totalkostnadene for e-infrastrukturen. Alternativt må man vurdere en topris-løsning for bruk av de nasjonale e-infrastrukturtenestene, men dette vil gi en langt mer komplisert og uforutsigbar finansieringsmodell.

Hvis institusjoner som har et visst minimumsforbruk vil bidra til grunnfinansiering av de nasjonale e-infrastrukturtenestene, vil dette samtidig gi disse institusjonene bedre strategisk innflytelse på utviklingen av e-infrastrukturtenestene. Sigma2 gjennomfører i dag styringsdialogsmøter med alle finansierende institusjoner, og disse samtalen, sammen med brukerundersøkelser, danner grunnlaget for det strategiske arbeidet i selskapet.

Samtidig er det viktig at institusjonene selv faktisk har en strategi for e-infrastruktur, og spesielt for hvordan institusjonenes lokale e-infrastruktur skal samspille med den nasjonale e-infrastrukturen, slik at det blir en hensiktsmessig ansvars- og arbeidsdeling mellom disse og som gir mest mulig effektiv utnyttelse og best mulig brukeropplevelse.

En bredere basis av finansierende institusjoner vil bidra til å fordele de økte kostnadene på flere aktører. En videre avlastning vil kunne komme ved at Forskningsrådet også øker sitt bidrag til grunnfinansieringen. Forskningsrådet sin grunnfinansiering er ment å delfinansiere kostnadene for både små og store aktører i UH- og institutt-sektoren. Det er dermed naturlig at Forskningsrådet også øker sin grunnbevilgning som en følge av veksten i e-infrastrukturbehov. Erfaringene med brukerbetalning at disse er krevende å

implementere i Forskningsrådets egen portefølje, og Forskningsrådet bør derfor vurdere en annen tilnærming for å dekke økt bruk av nasjonale e-infrastrukturtenester i forskningsprosjekter de finansierer.

Denne rapporten har vist at dagens modell ikke har hatt tilstrekkelig fleksibilitet til å endre grunnfinansieringen, da fem år er en for lang tidshorison. Det er derfor viktig at en fremtidig finansieringsmodell er i stand til å justere grunnfinansieringen i tråd med økte behov og i takt med den nasjonale ambisjonen for finansiering av e-infrastruktur. Samtidig må modellen være forutsigbar og ikke komme med for store administrative omkostninger.

En måte å ivareta disse ulike hensynene er at alle finansierende institusjoner inngår tre-årige budsjettavtaler som inkluderer en forventet budsjettvekst i avtaleperioden. Ved overgangen mellom budsjettperiodene justeres budsjettet i tråd med den faktiske veksten i det internasjonale ambisjonsnivået. Fordelingen av grunnfinansieringen mellom institusjonene justeres på basis av forbruket i foregående tre-årsperiode (se avsnitt om fordeling av kostnader mellom institusjoner). Det historiske forbruket de siste tre år bør også danne utgangspunkt for fordelingen mellom institusjonene når en slik finansieringsmodell innføres.

En justering av det norske ambisjonsnivået i forhold til den definerte internasjonale referanse kan gjøres i forbindelse med en internasjonal evaluering av Sigma2 som utvalget foreslår finner sted hvert sjette år. Den internasjonale evalueringen vil gi institusjonene et godt faktagrunnlag for å vurdere om ambisjonsnivået man har lagt seg på gir de nødvendige faglige gevinster og svarer til Norges ambisjoner på ulike fagområder.

En alternativ modell vil være at institusjonene selv definerer sitt ambisjonsnivå og forventede bruk av e-infrastrukturtenester. Utvalget er av den formening at dette gir en vesentlig mindre forutsigbar modell både for institusjonene (som vanskelig kan vite hva faktisk forbruk vil være i årene som kommer) og for Sigma2 som i mindre grad kan fokusere på en helhetlig strategi for norsk e-infrastrukturtenester, spesielt for lengre tidsperioder. Utvalget har allikevel forståelse for at institusjoner som i all hovedsak baserer sin aktivitet på eksterne prosjekter i mindre grad kan forventes å bidra til infrastrukturen basert på et historisk forbruk som har sin basis i en annen stadig skiftende ekstern prosjektportefølje.

### **Mulighet: Økt tverrsektorielt bidrag til grunnfinansieringen gjennom NFR/departement**

I dagens finansieringsmodell bidrar Forskningsrådet med 25 MNOK årlig til grunnfinansieringen av Sigma2, i tillegg til de midler som tildeles gjennom den konkurranseutsatte arenaen. Begge tildelinger kommer fra Kunnskapsdepartementet sin bevilgning til Norges Forskningsråd. Samtidig er det en lang rekke forskningsprogram (blant andre KLIMA, ENERGIX, PETROMAKS, NANO21, BIOTEK21) og offentlige forvaltningsinstitutt (f.eks. Met og Havforskningsinstituttet) med finansiering fra andre departement som er betydelige brukere av nasjonale e-infrastrukturtenester. I dagens finansieringsmodell finansieres disse aktivitetene utelukkende over

Kunnskapsdepartementets budsjett, dels gjennom forskningsrådsbevilgningen, men også dels gjennom bidraget til grunnfinansieringen fra de fire breddeuniversitetene.

Rapporten har allerede vist at dagens behov for e-infrastruktur utfordrer programmet for finansiering av infrastruktur, hvor søknaden om e-infrastruktur nå nærmer seg halve rammen for INFRASTRUKTUR utlysningen. Veksten i behovene for e-infrastruktur for å løse sektorutfordringer både innen forskning og forvaltning har dermed en direkte effekt på Norges investering i forskning, siden alle slike behov må dekkes over Kunnskapsdepartementets budsjett. Det er viktig at sektordepartementenes behov for økt kunnskap eller deres forvaltningsoppgaver koordineres godt med Kunnskapsdepartementet når dette involverer betydelig bruk av e-infrastruktur tjenester.

Utvalget mener at i tråd med prinsippene som ligger til grunn for brukerbetaling, bør store forsknings- og forvaltningsoppgaver finansiert av ulike departement inkludere omkostningene for investeringer i nasjonal e-infrastruktur. I hvilken form et slikt bidrag skal komme bør diskuteres nærmere med aktuelle departementer/ Forskningsrådet/forvaltningsinstitutter. I en slik dialog må det tas høyde for at departementene finansierer forsknings- og forvaltningsoppgaver både gjennom Forskningsrådet og gjennom egne tildelingsbrev og forskningsprogrammer.

Utvalget presiserer at selv om disse forslagene kan gi inntrykk av økte omkostninger for de ulike departementene, vil disse omkostningene uansett måtte dekkes hvis departementene i stedet skal bygge opp egne e-infrastruktur tjenester. Utvalget er av den oppfatning at størst grad av kosteffektivitet oppnås ved at et selskap (Sigma2) har det overordnede ansvar for all nasjonal generisk e-infrastruktur. Domene-spesifikke behov, som for eksempel behovet for datakuratorer innenfor områder som klima, helse etc. vil fortsatt måtte dekkes av fagspesifikke infrastrukturer, de enkelte institusjoner og departementer, men bør bygges på toppen av den felles nasjonale generiske e-infrastrukturen.

Vi diskuterer her kort ulike modeller for hvordan tverrsektorielle bidrag til finansieringen av nasjonal e-infrastruktur kan organiseres.

- 1. Direkte bevilgning:** Departementene kan gi direkte bevilgninger til Sigma2 eller gjennom bevilgningsbrevet til Norges Forskningsråd gi spesifikke tilskudd for å dekke omkostningene for e-infrastrukturbehov som departementenes forskningsprogram eller forvaltningsinstitusjoner har. Erfaringene fra den gangen tungregning var bevilget direkte over statsbudsjettet gjør at utvalget er usikker på om en slik ordning gir tilstrekkelig handlingsrom for å kunne håndtere hurtig voksende e-infrastrukturbehov innenfor forsknings- eller forvaltningsoppgaver som hører til departementenes porteføljer.
- 2. E-infrastrukturbevilgninger besluttet av styret i Forskningsrådet:** For bevilgninger som kanaliseres gjennom Forskningsrådet, kan styret i Forskningsrådet sette av nødvendige midler til basisinvesteringer i infrastruktur for de aktiviteter departementene igangsetter og som krever e-infrastruktur. Omfang av slike bevilgninger kan baseres på enkelte programområders historiske

bruk og i dialogstyringen mellom forskningsrådet og de enkelte departement, samt gjennom Forskningsrådets budsjettforslag og de ulike departementers bevilgninger til Forskningsrådet.

En slik modell vil sikre en helhetlig tilnærming til finansieringen av e-infrastruktur som tar høyde for de behov ulike departement har for forskning og forvaltningsoppgaver som krever e-infrastrukturtenester. Ulempen med modellen er at man viderefører en modell hvor e-infrastruktur oppfattes å komme “uten kostnad” for det enkelte porteføljestyre, og som dermed ikke vil kunne se helheten i sin portefølje. Utvalget er likevel av den oppfatning at dette er den beste løsningen, men at det vil kreve en tett dialog med de enkelte porteføljestyrene, da strategiske valg som fører til økt bruk av e-infrastruktur vil påvirke deres portefølje og handlingsrom.

- 3. Porteføljestyre/forvaltningsinstitusjoner avsetter midler til forventede e-infrastrukturbehov.** De enkelte porteføljestyrene i Forskningsrådet og institutt med forvaltningsoppgaver som bruker nasjonale e-infrastrukturtenester avsetter midler til grunnfinansiering av e-infrastruktur fra sine respektive budsjett ut ifra et historisk forbruk. Disse midlene overføres direkte fra det enkelte porteføljestyre/institutt til Sigma2. Alternativt kan porteføljestyrene avsette midler i sine programutlysninger til å dekke omkostninger for bruk av e-infrastrukturtenester, eller infrastrukturtenester mer generelt, slik det ble gjort for eksempel i siste SFF utlysning. Dette siste alternativet diskuteres i mer detalj i neste seksjon om brukerbetaling. Fordelen med at porteføljestyrene avsetter midler basert på historisk forbruk eller gjennom brukerbetaling er at kostnadene for bruk av nasjonale e-infrastrukturtenester synliggjøres som en del av de totale kostnadene ved forsknings- og forvaltningsoppgavene til porteføljestyret/instituttet.

Utvalget vil anbefale en kombinasjon av grunnbevilgning vedtatt i styret i Norges Forskningsråd og basert på dialog mellom Forskningsrådet og departement som har store e-infrastrukturbehov både til forskning og til forvaltning, og brukerbetaling for prosjekter over en viss minimumsstørrelse på totalbudsjettet (jmf. neste seksjon). Utvalget vil anbefale at grunnfinansieringsbidraget fra de ulike departementene fastsettes på samme måte som fordeling mellom institusjoner, dvs. basert på historisk forbruk de foregående tre år og ut ifra et veldefinert ambisjonsnivå for norske investeringer i e-infrastruktur, og NFR sin andel av denne totalrammen.

#### **Mulighet: Økt brukerfinansiering gjennom betaling for særskilte behov**

Dagens finansieringsmodell åpner for brukerbetaling. Vi må her skille mellom to ulike typer brukerbetaling: 1) Brukerbetaling for dedikert aksess, enten knyttet til forvaltningsoppgaver med tidskrittisk behov for resultater eller for oppdragsforskning utført for eksempel av forskningsinstitutter; 2) Brukerbetaling for prosjekt med spesielt store behov for e-infrastrukturtenester. For den første gruppen vil pris dekke «Total Cost

of Ownership» (TCO), dvs. både investerings- og driftskostnader, og prosjektene vil primært bli vurdert på grunnlag av om de tilfredsstillende minstekrav til vitenskapelig kvalitet. For den andre gruppen vil prosjektene konkurrere om tilgang på linje med andre prosjekter og bli vurdert ene og alene ut fra vitenskapelig kvalitet, og de vil ha tilgang til ressursene på lik linje med andre prosjekt, men betaler i henhold til Forskningsrådets leiestedsmodell kun for driftskostnader. Betaling vil kun bli aktuelt hvis prosjektene blir funnet å ha høy nok vitenskapelig kvalitet og dermed blir gitt adgang til e-infrastrukturtenestene.

Sigma2 har i dag veldefinerte regler for brukerbetaling for store datamengder, da dette er å betrakte som dedikert tilgang. For tungregning har Sigma2 satt som prinsipp at det for prosjekter med en totalramme som overstiger 15 MNOK forventes egenbidrag for bruk av nasjonal e-infrastruktur tilsvarende driftskostnadene (kategori 2). Dette beløpet er valgt for å ha en god balanse mellom administrative kostnader og forventet størrelse på behov for e-infrastrukturtenester. Per i dag er det kun to sentre for fremragende forskning som bidrar med denne type brukerfinansiering for e-infrastruktur.

De nasjonale e-infrastrukturtenestene brukes i liten grad til oppdragsforskning og slik bruk vil uansett måtte begrenses til 20% av totalbruk for ikke å utfordre prinsippet om egenregi som må ligge til grunn for at UH-sektoren skal kunne løse sitt behov for e-infrastrukturtenester gjennom et aksjeselskap.

Det kreves i dag ikke brukerbetaling for EU prosjekter for å unngå en konkurransevridning vs. andre europeiske prosjekter, da brukerbetaling for e-infrastrukturtenester ikke er vanlig i andre europeiske land.

I sum betyr dette at brukerbetaling alene neppe vil være i stand til å dekke forventede økninger i behov for finansiering av e-infrastrukturtenester.

Fordelen med brukerbetaling på prosjektnivå er at kostnadene ved disse tjenestene synliggjøres både for porteføljestyret, forskerne og de finansierende departementene. Det vil også bidra til at forskerne er mer nøkterne i sin bruk av e-infrastrukturtenester, kanskje spesielt i forhold til hvilke typer og kvaliteten på data som lagres.

En potensiell ulempe med brukerbetaling er at disse kostnadene kan være betydelige for spesielt e-infrastrukturintensive forsknings- eller forvaltningsprosjekter. Innenfor en fastlagt prosjektramme kan det bli få ledige midler igjen til andre deler av forskningsprosjektet.<sup>25</sup> Disse ulempene kan reduseres ved at egne infrastrukturmidler settes av som kun kan brukes til e-infrastruktur, eller infrastruktur mer generelt, og kun av prosjekter med slike behov. Det kan også settes et øvre tak på den årlige brukerbetalingen, selv om prosjektet gis anledning til å bruke ytterligere ressurser. Dette var modellen som ble brukt i den siste SFF utlysningen.

---

<sup>25</sup> Fotnote 24 viste for eksempel at det de tre prosjektene med størst forbruk av e-infrastrukturtenester har et årlig forbruk som svarer til 40-90% av den totale rammen for tre-årige FRIPRO prosjekter.

Brukerbetaling skaper en finansiell usikkerhet for Sigma2, samtidig som det er en mekanisme som sørger for økte budsjetter for e-infrastruktur når nye forskningsprosjekter igangsettes etter å ha nådd opp i konkurranse. Fordi brukerbetalingen vil avhenge av hvorvidt disse prosjektene er av høy nok faglig kvalitet, vil det bli vanskeligere for Sigma2 å drive langsiktig strategisk planlegging av investeringer i kompetanse og utstyr. Den finansielle risikoen for Sigma2 reduseres delvis ved at de fleste forskningsprosjekter har minst en tre- til fem-årig tidshorisont.

Andre ulemper ved stor grad av brukerbetalning er at grunnlaget for den nasjonale samfinansieringsmodellen utvannes, og at utstrakt bruk av brukerbetalning kommer med betydelige administrative kostnader som dermed delvis vil kunne redusere kosteffektiviteten i den nasjonale modellen.

### Utvalgets anbefaling av ny finansieringsmodell

Utvalget mener at dagens situasjon hvor rundt 60% av finansieringen er konkurranseutsatt skaper stor grad av usikkerhet og stor risiko for tap av viktig kompetanse og en reell fare for at viktige forsknings- og forvaltningsoppgaver ikke kan gjennomføres. Denne risikoen må også sees i lys av de nasjonale føringene om langsiktig lagring av forsknings- og forvaltningsdata, som krever en finansieringsmodell som sikrer langsiktighet.

Samtidig vil konkurranse, også opp mot annen infrastruktur, bidra til å sikre god kvalitet i spissytelser og nye arkitekturer, og at dette gir en merverdi utover den periodiske evalueringen av Sigma2 sine leveranser.

Utvalget mener at for å sikre langsiktighet samtidig som man ivaretar at norske forskningsmiljø skal være konkurransedyktige med de nasjoner vi ønsker å sammenligne oss med, så bør grunnfinansieringen utgjøre minst 80% av Sigma2 sitt totale budsjett. De resterende 20% bør komme fra det konkurranseutsatte INFRASTRUKTUR-programmet. Dette vil sikre at investeringer i nye maskinarkitekturer eller spesielle spissytelsesbehov er faglig kvalitetssikret og konkurransedyktig mot andre nasjonale infrastrukturbehov.

Utvalget har definert et internasjonalt referansenivå som bør danne grunnlaget for det nasjonale ambisjonsnivået for finansiering av e-infrastruktur, og som vil kreve en økning per i dag på 130-190 MNOK i forhold til dagens finansiering. Hvorvidt dette ambisjonsnivået bør dekkes av grunnfinansieringen eller om også den konkurranseutsatte finansieringen bør inngå for å nå dette ambisjonsnivået kan diskuteres. Utvalget mener at de særnorske forvaltningsbehov og Norges sterke posisjon innenfor beregningsbasert og dataintensive forskningsfelt, samt Norges ambisjon innenfor hurtig voksende felt som kunstig intelligens, tilsier at grunnfinansieringen alene bør svare til det nasjonale ambisjonsnivået for investeringer i e-infrastruktur. Selv om noen av høringsuttalelsene ga uttrykk for et ønske om at dette behovet ble godt dokumentert ut ifra norske behov, så tolker utvalget høringsuttalelsene som en villighet til å strekke seg mot et slik ambisjonsnivå under forutsetning at dette er et reelt behov. Høringsuttalelsene fra deler



av instituttsektoren antyder at selv dette er et for lavt ambisjonsnivå. Utvalget står fast på at et ambisjonsnivå på rundt 250 MNOK årlig bør være grunnlaget for grunnfinansieringen de neste tre år, og at konkurranseutsatt finansiering kommer på toppen av dette.

I en modell hvor institusjonene, Forskningsrådet og de ulike departementene bidrar til grunnfinansieringen basert på historisk bruk, så forventer utvalget at omfanget av brukerfinansiering vil bli meget begrenset, og primært være rettet mot oppdragsforskning og spesielt store (e)-infrastrukturbehov. Utvalget ser allikevel at brukerfinansiering kan være en måte for Forskningsrådet å bidra til grunnfinansieringen av norsk e-infrastruktur. Høringsuttalelsene var delt på dette punkt, med gode argumenter både for og imot økt grad av brukerbetaling. Utvalget mener allikevel at Sigma2 sine muligheter for langsiktig planlegging vil svekkes hvis brukerfinansieringen utgjør mer enn 10% av driftskostnadene, dette gjelder spesielt hvis store deler av brukerbetalingen baseres på konkurranseutsatt finansiering, dvs. gjennom prosjekter som når opp på ulike konkurransearenaer.

For å nå det ambisjonsnivået som utvalget mener Norge bør ha for investeringer på e-infrastruktur og i lys av utvalgets vurdering av mulighetene som ligger i ulike finansieringsmekanismer, så mener utvalget at grunnfinansieringen av nasjonal e-infrastruktur bør finansieres tverrsektorielt og tverrinstitusjonelt. En slik finansieringsmodell bør ha tre hovedkomponenter, som kort oppsummeres nedenfor, samt i Figur 7. For å konkretisere disse forslagene til et kronebeløp, så tar utvalget utgangspunkt i at det nasjonale ambisjonsnivået utvalget Norge bør ha i dag, vil kreve en årlig grunnfinansiering på rundt 250 MNOK.

- Dagens finansierende institusjoner (Forskningsrådet og de fire breddeuniversitetene) vil måtte forvente å øke sine bidrag til grunnfinansieringen av de nasjonale e-infrastrukturtenestene for å dekke sine egne økte behov. Basert på tallmaterialet i Appendiks 3, vil de fire breddeuniversitetene måtte øke sitt årlige bidrag til 120 MNOK årlig siden de fire breddeuniversitetene står for 80% av forbruket fordelt på institusjoner.
- Det bør etableres en bredere basis av institusjoner som bidrar til grunnfinansieringen, både fra UH og instituttsektoren. Utvalget foreslår at institusjoner som over en tre-årsperiode har et forbruk som er større enn 1% av totalkostnadene for den nasjonale e-infrastrukturtenesten (tungregning og/eller datalagring) bør bidra til grunnfinansieringen av denne.<sup>26</sup> Grensen på 1% er satt som en avveining mellom byråkrati for å kreve inn penger og et rimelig bidrag fra disse institusjonene. Basert på at disse nye institusjonene står for omlag 20% av bruken av nasjonale e-infrastrukturressurser til interne prosjekt i UH-sektoren og

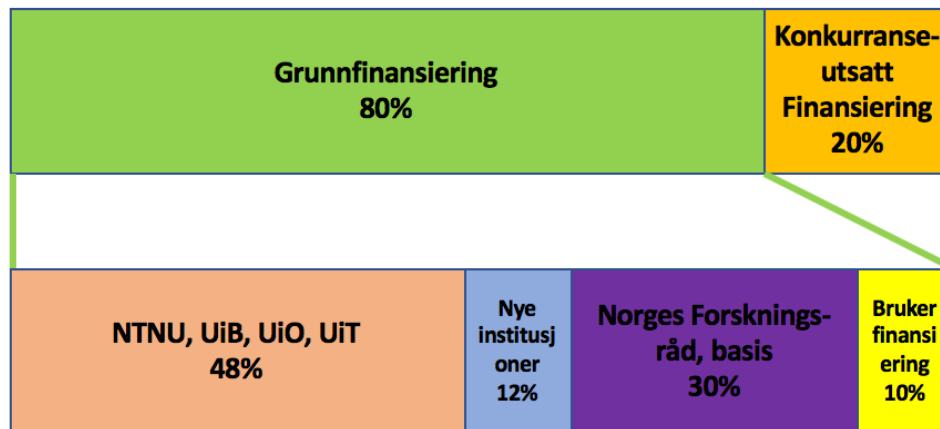
---

<sup>26</sup> Basert på forbruk de siste tre år, så er det 6 nye institusjoner som bør bidra til grunnfinansieringen (se tabell 2): SINTEF AS, Nansen Environmental and Remote Sensing Center, Akvaplan-NIVA AS, Uni Research AS, Universitetet i Stavanger, og CICERO.

instituttsektoren, så vil deres årlige bidrag utgjøre 30 MNOK, som da fordeles på disse 6 nye partnerinstitusjonene. Det bør åpnes for at Universitetet i Agder også blir en aktiv partner i den nasjonale e-infrastrukturen på grunn av sin satsing innen maskinlæring og regneressurser knyttet til dette.

- Større forsknings- og forvaltningsprogram bør vurdere sine behov for e-infrastruktur og inkludere disse kostnadene i programmets totalbudsjett. Utvalget mener dette bidraget, gjennom Forskningsrådet sin grunnfinansiering, i sum bør utgjøre 100 MNOK, hvorav en begrenset andel inn til 15-20 MNOK kan finansieres gjennom prosjektbasert grunnfinansiering.
- Utvalget ser at deler av instituttsektoren har en inntektsstrøm gjennom prosjektfinansiering samtidig med lav basisfinansiering, som i kombinasjon med kostnadene for bruk av e-infrastruktur gjør at et bidrag til grunnfinansieringen basert på historisk forbruk er vanskelig å ivareta. I stor grad vil slik finansiering komme fra Norges Forskningsråd gjennom prosjekter (og basis). Utvalget har vurdert at instituttsektorens bidrag til grunnfinansieringen fortrinnsvis kan løses gjennom en økt basisfinansiering øremerket e-infrastrukturfinansiering, eller at instituttsektoren gis anledning til å søke om dekning for e-infrastrukturkostnader i alle prosjekter etter den modellen som brukes for dekning av infrastruktur i dagens SFF-ordning, dvs. i stor grad gjennom en brukerfinansieringsmodell. Utvalget anbefaler at Norges Forskningsråd innleder en dialog med relevante institutt og departement om hvordan dette best kan løses i praksis.

I sum gir dette en årlig grunnfinansiering på rundt 250 MNOK. Ytterligere 50-60 MNOK årlig kan så skaffes tilveie gjennom den konkurranseutsatte INFRASTRUKTUR-arenaen.



*Figur 7: Utvalgets forslag til finansieringsmodell. Forholdet mellom de fire breddeuniversitetene og nye institusjoner er basert på nå-situasjonen. Norges Forskningsråd kan forventes å stå for størstedelen av bidraget til brukerfinansiering, selv om det kan forventes økte kommersielle bidrag over tid.*

Et bredere basis for grunnfinansieringen vil også kreve noen justeringer av dagens praksis for styring og styresammensetning. Dette diskuteres nærmere nedenfor. Utvalget vil imidlertid påpeke at de årlige møtene mellom Sigma2 og de enkelte finansierende

institusjoner vil være en viktig arena for å sikre at institusjonenes behov og strategiske prioriteringer legges til grunn for Sigma2 sin investerings- og utviklingsstrategi.

Et konkret eksempel på kostnader ved en fordeling som nevnt over, med grunnfinansiering på 250 MNOK (jf. Figur 7):

<b>Fordeling av finansiering på 250 MNOK årlig:</b>	<b>SUM</b>	<b>250</b>
BOTT	48%	120
Andre (Forvaltning, Institutt og UH)	12%	30
Forskningsrådet	40%	100

Tabell 2: Fordeling av finansiering på 250 MNOK årlig

En grunnfinansiering på 250 MNOK vil i en slik modell gi at institusjoner utenom NFR bidrar med 150 MNOK. Fordelingen av finansieringsansvar for den nasjonale e-infrastrukturen basert med utgangspunkt i forbruk de siste tre år, vises i Tabell 3:

Eksempel på kostnadsfordeling som tar hensyn til institusjoners ulike prosentvise bruk av ressurser		Institusjonelle bidrag MNOK	150			
		Fordeling mellom HPC og lagring	70%	30%		
Organisasjoner sin prosentvise forbruk for 2015-2018	% bruk av HPC-ressurser	% bruk av Lagrings-ressurser	105	45	SUM MNOK	% av bidrag fordelt på alle
			Fordeling i MNOK for HPC	Fordeling i MNOK for lagring		
Universitetet i Oslo	36.6%	18.8%	38.38	8.46	46.84	31%
Norges Teknisk-Naturvitenskapelige Universitet	22.9%	3.0%	24.01	1.33	25.34	17%
Universitetet i Bergen	21.7%	49.3%	22.74	22.19	44.93	30%
SINTEF AS	5.8%	3.1%	6.06	1.40	7.46	5%
UiT - Norges arktiske universitet	3.4%	0.0%	3.54	0.00	3.54	2%
Nansen Environmental and Remote Sensing Center	2.1%	5.6%	2.22	2.53	4.75	3%
Akvaplan-NIVA AS	1.5%	3.0%	1.52	1.37	2.89	2%
Uni Research AS	1.2%	3.0%	1.28	1.37	2.64	2%
Universitetet i Stavanger	0.9%	0.0%	0.91	0.00	0.91	1%
CICERO - Senter for klimaforskning	0.6%	1.2%	0.60	0.54	1.13	1%
Norges miljø- og biovitenskapelige universitet	0.0%	1.0%	0.00	0.45	0.45	0%
Andre (Forvaltning, Institutt og UH)	3.6%	11.9%	3.74	5.36	9.10	6%
<b>SUM</b>	<b>100%</b>	<b>100%</b>	<b>105.00</b>	<b>45.00</b>	<b>150.00</b>	

Tabell 3: Fordelingen av finansieringsansvar for den nasjonale e-infrastrukturen i eksempelet med utgangspunkt i forbruk de siste tre år.

### Utvalgets forslag til fordelingsmodell

Vi har i beskrivelsen av utvalgets anbefalte finansieringsmodell ikke problematisert hvordan kostnadene skal fordeles mellom de ulike finansierende institusjoner,

forskningsområder og departement. Utvalget vil her kort diskutere tre mulige fordelingsmodeller, selv om disse ikke kan sees helt uavhengig av finansieringsmodellen: 1) forventet bruk; 2) faktisk bruk; 3) en fordelingsnøkkel basert på historisk bruk.

- Den første av disse fordelingsprinsippene er krevende da det er vanskelig for institusjoner og forskningsområder å forutsi faktiske behov samt å vurdere hvorvidt deres søknader vil nå opp i konkurransen med andre søknader om bruk av e-infrastruktur. Modellen vil også skape stor finansiell usikkerhet for Sigma2 og gjøre langsiktig planlegging av investeringer i kompetanse og utstyr vanskelig.
- Det andre fordelingsprinsippet vil også kunne skape usikkerhet i finansieringen for Sigma2. Dette er også en krevende modell for de finansierende institusjonene, da de ikke *a priori* vil kunne forutsi faktisk forbruk av e-infrastruktur. For både Sigma2 og de finansierende institusjonene er det også en utfordring at betaling vil måtte skje etterskuddsvis.
- Det tredje fordelingsprinsippet gir forutsigbarhet både for Sigma2 og for de forskningsfinansierende institusjonene, men faktisk forbruk vil ikke nødvendigvis svare direkte opp mot andel av finansieringen. Samtidig kan man forvente at evt. avvik vil kompenseres i en ny finansieringsperiode hvor andelen justeres basert på det historiske forbruket.

Utvalget vil anbefale den siste av disse modellene og foreslår at grunnfinansieringen av Sigma2 avtales for en periode på tre år, og inkluderer en forventet vekst av grunnfinansieringen i denne perioden for å kunne opprettholde det nasjonale ambisjonsnivået, jf. Figur 5. Den enkelte institusjons andel av grunnfinansieringen fastsettes på bakgrunn av gjennomsnittlig forbruk de tre foregående år og fordelingen mellom institusjonene ligger fast for en tre-årsperiode. Ved innføring av en slik modell, så bør fordeling mellom institusjoner også baseres på historisk forbruk.

I forbindelse med revidering av fordelingen av kostnader mellom institusjonen vurderes det også om nye institusjoner bør bidra til grunnfinansieringen fordi deres totale forbruk av tungregnings- og/eller datalagringstjenester har oversteget 1% av totalkostnadene, og tilsvarende om enkelte institusjoner ikke lenger forventes å bidra til grunnfinansieringen (men de kan fortsatt gis anledning til å bidra).

Totalrammen for grunnfinansieringen justeres hvert tredje år i tråd med endringer i finansieringsbehovet som en følge av det nasjonale mål som er satt for (grunn)finansieringen av de nasjonale e-infrastrukturtenestene.

Utvalget foreslår at Sigma2 evalueres som i dagens modell, men hvert sjetten år i stedet for hvert femte år. Et nytt element i evalueringen bør være om det nasjonale ambisjonsnivået er et godt mål, eller om dette bør justeres med bakgrunn i internasjonale utviklingstrekk og eventuelle særnorske forhold.

En positiv evaluering vil medføre at modellen videreføres for nye seks år. Ved en negativ evaluering må selskapet komme opp med en ny organisasjonsmodell i samarbeid med finansierende institusjoner. I ytterste konsekvens kan denne prosessen lede til en

anbefaling om avvikling av selskapet. Selskapet bør gis inntil tre år på å avvikle selskapet etter at finansierende institusjoner endelig har avvist nye forslag til organisasjonsmodell.

## **Styring og organisering**

Utvalget har diskutert alternative organisasjonsformer til dagens modell med et aksjeselskap eid 100% av Uninett AS. Utvalget kan ikke se at det finnes andre modeller som vil kunne bidra til en bedre organisering av de nasjonale e-infrastrukturtenestene for forskning, selv om det for eksempel blir en større bredde i antallet finansierende institusjoner. I det følgende vil vi derfor ha som utgangspunkt dagens Sigma2 og kun kommentere de justeringer som må gjøres som en følge av endringer i finansieringsmodellen.

I dagens organisering oppnevnes Sigma2's styre av generalforsamlingen representert ved Uninett AS. Oppnevning av styremedlemmer tar utgangspunkt i forslag på styremedlemmer fra de finansierende institusjoner (de fire breddeuniversitetene pluss Forskningsrådet) samtidig som man ivaretar kravene til styrets totale kompetanseprofil og kjønns sammensetning. Frem til nå har de fire breddeuniversitetene hatt et styremedlem hver.

Den generelle prosessen for oppnevning av styre har fungert godt og bør videreføres. Generalforsamlingen bør ta sikte på at de største brukermiljøene/finansiørene er representert i styret (institusjoner som står for mer enn 15-20% av bruken) og at enkelte av de mindre institusjonene også er representert i styret, samtidig som helheten i styrets kompetanse gis betydelig vekt. Uavhengig av styrerepresentasjon, så vil mye av den direkte strategiske utvikling styres gjennom styringsdialogen mellom Sigma2 og alle institusjoner som bidrar med grunnfinansieringen av selskapet.

## **Oppfølging av arbeidsgruppens anbefalinger**

Utvalget ber om at Forskningsrådet følger opp utvalgets anbefalinger i dialog med relevante aktører, slik at en felles forståelse og tilslutning til den nye finansieringsmodellen kan oppnås. Ambisjonen må være at en ny, bred finansieringsmodell basert på utvalgets modell er på plass innen sommeren 2020, slik at budsjettkonsekvensene kan inkluderes i alle finansierende institusjoners budsjett for 2021.

## **APPENDIKS 1: Om arbeidsgruppens arbeid**

Arbeidsgruppen ble etablert av de fire breddeuniversitetene i samarbeid med Forskningsrådet og UNINETT Sigma2 sommeren 2018. Hensikten var å se nærmere på utfordringene og foreslå en helhetlig bærekraftig investerings- og driftsmodell for

nasjonale e-infrastrukturtenester som er robust nok til å kunne håndtere fremtidige behov for slike tenester.

### Arbeidsgruppens medlemmer

Arbeidsgruppen har bestått av

Institusjon	Representant	
NTNU	Torbjørn Svendsen	Terese Løvås
UiB	Knut Børve	Inge Jonassen
UiO	Kirsti Klette	Morten Dæhlen
UiT	Kenneth Ruud (leder)	Camilla Brekke
Sigma2	Gunnar Bøe	Hans A. Eide
<i>Forskningsrådet</i>	<i>Ulrike Jaekel (sekretariat)</i>	

### Arbeidsgruppens mandat

Mandatet pr. 16. Oktober 2018

Arbeidsgruppens oppgave er å foreslå en oppdatert og helhetlig nasjonal strategi for videreutvikling og finansiering av nasjonal e-infrastruktur for forskning og høyere utdanning i Norge i perioden 2020-2030.

Arbeidsgruppen skal utarbeide og begrunne et realistisk estimat for behovet for slik e-infrastruktur fra alle relevante brukermiljøer innenfor ikke-kommersiell forskning og høyere utdanning, inklusive relevant forvaltning. Dette estimatet skal baseres på eksisterende nasjonale og internasjonale rapporter i tillegg til konsultasjon med relevante FoU-miljøer.

Strategiarbeidet må inkludere problemstillingens faglige bredde med utgangspunkt i nåværende og fremtidige brukermiljøer.

Strategiarbeidet skal utarbeide forslag til operasjonelle modeller som bidrar til kosteffektivitet gjennom prosesser for å vurdere kommersielle og internasjonale tilbydere opp mot kostnadene ved oppbygging og drift av nasjonale ressurser, samtidig som norske forskere gis mulighet til å ligge i forskningsfronten innen sine fagområder.

Arbeidsgruppen skal forankre sitt arbeid hos alle relevante institusjoner gjennom møter med både institusjonenes ledelse og relevante brukermiljøer.

Strategien for finansiering skal kunne legge til rette for forutsigbar og robust investering, fornyelse/oppgradering, drift og videreutvikling av den nasjonale e-infrastrukturen og relaterte tenester.

Arbeidsgruppen skal i sitt arbeid vurdere bærekraftige finansieringsmodeller inkludert

justerbare og robuste mekanismer for kostnadsdeling mellom nasjonale finansierer (departementet og Forskningsrådet), institusjonene, og de prosjekter og forskergrupper som er brukere av e-infrastruktur og de relaterte tjenestene.

Arbeidsgruppen skal hente inn høringsuttalelser på en første versjon av rapporten som utarbeides og eventuelt innarbeide relevante innspill fra høringsrunden.

Arbeidsgruppen kan selv revidere mandatet i løpet av utvalgets arbeid.

## Arbeidsprosess

Utvalget har totalt avholdt syv møter, supplert med noen Skype-møter i slutfasen av arbeidet. I tillegg har utvalget gjennomført fire møter rettet mot ulike aktører som har interesse for eller behov av e-infrastruktur:

1. Et møte som involverte andre institusjoner og organisasjoner som leverer ulike digitale tjenester til forsknings- og høyere utdanningssektoren i Norge. Fokus i dette møtet var på å avklare avgrensninger av ansvarsområder for de ulike tjenestetilbyderne.
2. Et møte rettet mot eksisterende brukere av dagens e-infrastruktur og brukermiljøer som vi forventer i fremtiden i økende grad vil være i behov av infrastruktur tjenester. I tillegg til innledninger av ledende internasjonale eksperter på tungregning og dataintensive anvendelser og deres fremtidsperspektiver, så var fokus på å avklare norske forskningsmiljøers fremtidige behov for e-infrastruktur tjenester, herunder behovet for avansert bruker støtte.
3. Institusjonsledere fra både universitets- og høyskolesektoren og instituttsektoren. I tillegg til å få avklart hvorledes de ulike institusjonene ser på betydningen av e-infrastruktur tjenester for egen strategisk satsing, så var fokus rette mot å avklare hvilke tjenester de ulike institusjonene ønsker løst nasjonalt og hvilke tjenester de ulike institusjonene ønsker å tilby lokalt.
4. Det siste møtet involverte relevante departement, direktorat og forvaltningsinstitusjoner, dels for å informere om gruppens arbeid og for å få en forståelse for deres behov og tanker rundt organisering og finansiering av egne e-infrastruktur behov.

Vi gir her bare en kort versjon av utvalgets konklusjoner fra disse møtene, for å synliggjøre utvalgets forståelse av fremtidige behov for e-infrastruktur og avgrensning av utvalgets mandat.

Utvalget finner at det ikke er noen andre nasjonale aktører som tilbyr e-infrastruktur tjenester for tungregning og for lagring og behandling av store datamengder. Det finnes ulike aktører som tilbyr digitale tjenester for forskning og høyere utdanning, men disse fokuserer på lagring av mindre datamengder. For disse aktørene er det relevant å vurdere å bruke den generiske e-infrastrukturen som dette utvalget har fokus på som en underliggende e-infrastruktur, hvor andre mer dedikerte tjenester kan implementeres, jf. Figur 4.

Brukerbehovene for e-infrastruktur tjenester vil fortsette å øke, og dagens tilbud er ikke tilstrekkelig for å møte brukerbehovene. Brukermiljøene er i økende grad villig og

interessert i å tilpasse eksisterende programmer på nye maskinarkitekturer gitt at nødvendig brukerstøtte er tilgjengelig. Mange av de programmene som brukes og utvikles av norske forskningsmiljøer er av betydelig størrelse og med en lang utviklingshistorie. Tilpasning eller videreutvikling av disse programmene for nye arkitekturer vil kreve betydelige investeringer i form av personale med høyt spesialisert kompetanse både innenfor fagområdet og innenfor tungregning/datalagring.

E-infrastruktur har ulik oppmerksomhet hos ledelsen ved de ulike institusjonene. Der dette har oppmerksomhet i ledelsen, så betraktes e-infrastruktur som kritisk for institusjonens mulighet for å hevde seg i den internasjonale forskningsfronten. Institusjonene uttrykker i sin alminnelighet tilfredshet med dagens nasjonale organisering, og utvalget fikk ingen klare tilbakemeldinger på oppgaver som institusjonene sterkt ønsker skal løses lokalt.

En første versjon av rapporten ble sendt ut på høring før sommeren 2019. Utvalget har så revidert dette utkastet basert på høringsinnspillene (se Appendiks 6). Dette danner så grunnlaget for denne rapporten.



## APPENDIKS 2: Om e-infrastruktur

Denne teksten er hentet fra Forskningsrådets web side<sup>27</sup> om e-infrastruktur.

Elektronisk infrastruktur (e-infrastruktur) omfatter verktøy og tjenester for de fleste fagområder, men er spesielt viktig for forskning som krever omfattende beregninger eller genererer store mengder data. E-infrastruktur omfatter også digitale registre og databaser. Klimaforskning, biologi, bioinformatikk, medisin, kjemi, fysikk, materialvitenskap, energiforskning og lingvistikk er eksempler på ulike fagområder som alle benytter e-infrastruktur.

### Formål

E-infrastruktur omfatter grunnleggende tjenester som brukes innenfor de fleste fagområder, og utgjør et fundament for mange andre forskningsinfrastrukturer. E-infrastruktur er spesielt viktig for forskning som krever omfattende beregninger eller genererer store mengder data gjennom simulering og analyse, som klimaforskning, biologi, bioinformatikk, medisin, kjemi, fysikk, materialvitenskap, energiforskning og lingvistikk. Sensitive data må også kunne håndteres på en god måte.

Målene for e-infrastruktur er tredelt:

- Levere tjenester til forskningsprosjekter og andre forskningsinfrastrukturer
- Sikker lagring og tilgjengeliggjøring av data i tråd med de internasjonale FAIR-prinsippene
- Levere områdespesifikk e-infrastruktur

De store verdiene som investeres i å fremskaffe og analysere data krever at data blir sikret, får økt verdi gjennom katalogisering og generering av metadata og at dataene blir gjort tilgjengelig for andre brukere i tråd med FAIR-prinsippene. Tilgang til og effektiv bruk av e-infrastruktur for alle fagområder er derfor en av hjørnesteinene i dataintensiv forskning.

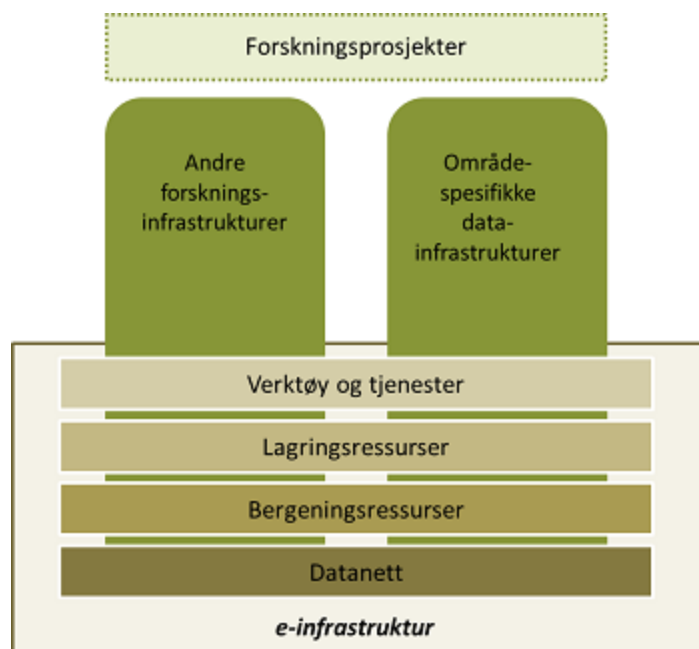
Bedre tilgang til forskningsdata styrker kvaliteten på forskningen, både fordi resultater enklere kan valideres og etterprøves, og fordi data kan brukes på nye måter og i kombinasjon med andre data. Åpen tilgang til forskningsdata bidrar til færre dupliseringer og unødvendig dobbeltarbeid og kan legge til rette for mer tverrfaglig forskning. Åpen tilgang til forskningsdata er et prioritert område både nasjonalt og internasjonalt. I 2017 lanserte Kunnskapsdepartementet en [Nasjonal strategi for tilgjengeliggjøring og deling av forskningsdata](#).

I 2017 reviderte [Forskningsrådet sin policy for åpen tilgang til forskningsdata](#) som skal bidra til at forskningsdata blir gjort tilgjengelige for relevante brukere, på like betingelser,

---

<sup>27</sup> <https://www.forskningsradet.no/prognett-infrastruktur/Artikkel/Einfrastruktur/1253998286437?lang=no>

til lavest mulig kostnad og i tråd med de internasjonale FAIR-prinsippene for økt dataverdi. FoU institusjoner skal vurdere om prosjekter som mottar finansiering fra Forskningsrådet skal utarbeide en datahåndteringsplan. Hensikten med en datahåndteringsplan er å planlegge for hvordan forskningsdataene skal sikres, ikke bare underveis i prosjektet, men også for fremtidig gjenbruk. Retningslinjene i policyen gjelder alle data i prosjekter som er finansiert av Forskningsrådet – med noen unntak. Ulike e-infrastrukturer har utviklet digitale verktøy som forskningsprosjekter kan benytte for å generere datahåndteringsplaner.



Figur 8: E-infrastruktur leverer grunnleggende tjenester til forskning og forvaltning

### Eksisterende forskningsinfrastruktur

Norske forskningsinstitusjoner har i dag en god og kostnadseffektiv samordning av e-infrastruktur for forskning og høyere utdanning innenfor mange fagområder. UNINETT AS utvikler og driver det norske høyhastighetsnettverk for forskning og utdanning, som forbinder mer enn 200 norske institusjoner og over 300 000 brukere, og knytter dem opp mot internasjonale forskningsnett. Selskapet eies av Kunnskapsdepartementet og driver ikke-kommersielt. Tilknytningen til forskningsnettet er basisen for de fleste andre tjenester levert av UNINETT og er essensiell for en velfungerende infrastruktur for tungregning og lagring.

Datterselskapet UNINETT Sigma2 AS (Sigma2) har ansvar for å anskaffe, drifte og videreutvikle den generiske nasjonale e-infrastrukturen for tungregning og datalagring i Norge. I perioden fra 2016 – 2019 skiftes de fire nasjonale tungregningsanleggene som ble anskaffet i 2012 ut og erstattes med to nye regneanlegg (E-INFRA ved UNINETT Sigma 2). Frem til 2015 har regnekraft og datalagringsløsninger vært tilbudt som to separate tjenester for forskning i Norge, og regneanleggene og datalagringsanleggene har vært fysisk adskilt. Med større datamengder og stadig flere forskningsprosjekter som baserer seg på analyse av data er det nå behov for en tettere kopling av beregnings- og

datalagringsanleggene, slik at analyse av data kan skje der dataene befinner seg. Derfor omfatter de pågående investeringene i nytt utstyr også en omstrukturering til en datasentrisk e-infrastruktur: Den nye datalagringsinfrastrukturen, NIRD (nasjonal infrastruktur for forskningsdata), som erstatter forgjengeren Norstore er nå direkte knyttet til regnearbeidene, noe som legger til rette for en mer effektiv levering av tjenester for dataanalyse og visualisering. NIRD gir lagringsressurser med årlige kapasitetsoppgraderinger, datasikkerhet gjennom lagring på to fysiske lokasjoner, støtte for flere lagringsprotokoller og migrasjon til tredjeparts skytilbydere.

Gjennom et tett samarbeid med de fire eldste universitetene tilbyr Sigma2 flere relaterte tjenester innenfor tungregning og datalagring til norsk universitets- og høyskolesektor og til andre forskningsorganisasjoner med finansiering fra det offentlige. Videre leder og koordinerer Sigma2 norsk deltakelse i internasjonalt samarbeid innenfor e-infrastruktur, som Nordic e-Infrastructure Collaboration (NeIC), Partnership for Advanced Computing in Europe (PRACE) og European Data Infrastructure (EUDAT).

Innenfor enkelte områder som behandler personsensitive data er det behov for løsninger som ivaretar krav til sikring av data samtidig som forskerne skal ha tilgang til å analysere dataene. Slike løsninger tilbys blant annet gjennom Tjenester for Sensitive Data (TSD) som drives og utvikles i samarbeid mellom Universitet i Oslo og Sigma2. Det er bevilget midler fra Nasjonal satsning på forskningsinfrastruktur til investering i nytt utstyr for både beregnings- og datalagringsanleggene for personsensitive data (TSD).

Norsk Senter for Forskningsdata (NSD) arkiverer, tilrettelegger og formidler data til forskningsmiljøer, både nasjonalt og internasjonalt, og utarbeider teknologiske løsninger slik at forskningssektoren får åpen tilgang til forskningsdata. NSD er personvernombud for alle landets universiteter, de fleste høyskoler og flere helseforetak og forskningsinstitutter. Siden 2003 har NSD vært etablert som et aksjeselskap eid av kunnskapsdepartementet. NSD har fått midler fra Nasjonal satsning på forskningsinfrastruktur til Norwegian Open Research Data Infrastructure (NORD-i) som er en løsning for lagring og tilgjengeliggjøring av forskningsdata.

Andre generiske datainfrastrukturer som kan nevnes er UiT Open Research Data som er et åpent forskningsdataarkiv etablert av Universitetet i Tromsø. Infrastrukturen er tilgjengelig både for forskere ved UiT og for andre institusjoner og enkeltforskere. Videre er BIBSYS BIRD et generisk verktøy for lagring, dokumentasjon, deling og publisering av forskningsdata, utviklet av BIBSYS (som nå inngår i det nye organet for høyere utdannings- og forskningstjenester) i samarbeid med Handelshøgskolen BI.

Det finnes også mange fagspesifikke datainfrastrukturer som tilbyr tjenester rettet mot særskilte behov hos ulike brukermiljøer. Disse fagspesifikke infrastrukturene er tilpasset data som skal gjøres tilgjengelig innenfor de ulike fagområdene. For å oppnå mest mulig gjenbruk av allerede innsamlede data, er det avgjørende at det finnes gode infrastrukturer som gjør det enkelt å finne fram til relevante data og koble ulike datasett sammen. Mer informasjon om fagspesifikke datainfrastrukturer er beskrevet under de ulike områdestrategiene.

## Behov for nyetablering, oppgradering og/eller samordning

Stadig bedre måle- og sensorteknologi, mer omfattende målinger, økt fokus på datadrevet forskning og mer avanserte verktøy for dataanalyser medfører en økning i behovet for tungregning og lagring og tilgjengeliggjøring av store mengder forskningsdata. Dette gjelder ikke bare fagområder som tradisjonelt har forholdt seg til store datamengder. Stadig flere forskningsfelt generer eller bruker store mengder data.

Regneanlegg må skiftes ut om lag hvert fjerde år fordi de slites og dermed blir for dyre i drift. Sigma2 har derfor et pågående arbeid med å skifte ut og oppgradere både beregnings- og datalagringsanleggene for forskning i Norge. Gjennom en framskriving av historisk etterspørsel og forespørsel fra nye brukergrupper beregner Sigma2 hvilke regnekapasiteter de nye anleggene bør ha for å kunne ta vare på behovene fra norske forskere.

I tråd med økt etterspørsel for tungregning og datalagring er det forventet at etterspørselen for håndtering av sensitive data vil øke sterkt i årene fremover.

## Relasjon til andre områder

Forskningsrådet oppfordrer til samarbeid mellom aktørene ved etablering av tjenester for datahåndtering, slik at vi i størst mulig grad kan utnytte investeringer som allerede er gjort. Dette kan skje gjennom prosjektsamarbeid eller direkte bruk av eksisterende tjenester. Slikt samarbeid er ikke begrenset til nasjonale løsninger. Innenfor noen områder vil det være mest naturlig å samarbeide om internasjonale datainfrastrukturer, slik mange av ESFRI prosjektene innenfor datahåndtering er eksempler på.

Forskningsrådet vil normalt ikke gi bidrag til investeringer i, og drift av, regneressurser for store beregninger dersom ikke investeringene er koordinert med eller foretas av Sigma2. Forskningsmiljøer som har behov for regneressurser anbefales først å kontakte Sigma2, for å avklare om behovene kan dekket gjennom eksisterende eller planlagte investeringer i regi av Sigma2. Ved søknad om nye nasjonale forskningsinfrastrukturer som har behov for lagringsressurser eller beregningsressurser, forutsetter Forskningsrådet at prosjektansvarlig har en dialog med Sigma2 om hvordan disse behovene kan ivaretas og at kostnadene legges inn i budsjettet for den omsøkte infrastrukturen.

De internasjonale FAIR-prinsippene er utarbeidet som et sett av retningslinjer for å tilrettelegge for økt dataverdi. FAIR er et akronym for ordene findable, accessible, interoperable og reusable. Data og metadata bør være gjenfinnbare, tilgjengelige, gjenbrukbare og kunne håndteres maskinelt.

## APPENDIKS 3: Forskningsprogrammene/Departementene sin bruk av Sigma2 sin e-infrastruktur

All bruk av Sigma2 sine e-infrastruktur ressurser skjer gjennom prosjekter og disse må oppgi hvilken finansiering de har. For de prosjektene som har sin finansiering fra Forskningsrådet kan man dermed finne hvilket program/utlysning finansieringen kommer fra og dermed også hvilke departementer som indirekte har finansiert prosjektet. Denne informasjonen hentes automatisk ut fra Forskningsrådet sin prosjektdatabase.

En sammenstilling av dette gir følgende oversikt:

Departementene sin %-vise forbruk av ressurser gjennom forskningsprogrammer	%-vis bruk av HPC-ressurser	%-vis bruk av Lagringsressurser
KoM - Klima- og miljødepartementet	6.9%	19.3%
KD - Kunnskapsdepartementet	39.5%	17.1%
KD-S - Kunnskapsdepartementet - sektorovergripende	18.2%	29.5%
OoE - Olje- og energidepartementet	9.3%	2.0%
SD - Samferdselsdepartementet	0.6%	0.2%
LoM - Landbruks- og matdepartementet	0.5%	0.6%
NFD - Nærings- og fiskeridepartementet	7.9%	2.9%
AoS - Arbeids- og Sosialdepartementet	0.3%	0.1%
UD - Utenriksdepartementet	0.4%	0.5%
MOD - Kommunal og moderniseringsdepartementet	0.0%	0.1%
FD - Finansdepartementet	0.0%	0.0%
HOD - Helse- og Omsorgsdepartementet	0.0%	0.4%
Annen eller ukjent	16.3%	27.3%
	100%	100%

Departementer som har 0.0% har finansiert prosjekter som har brukt infrastrukturen, men da i mindre grad. I kategorien Annen eller ukjent er prosjekter med EU-finansiering, annen ekstern finansiering eller internt finansiert av universitetene.

Ut ifra tabellen over kan det se ut som om Forskningsrådet sine forskningsprogrammer dominerer norsk bruk av e-infrastruktur, da det kun er 16-27% av bruken som ikke er knyttet til et av Forskningsrådet sine program. Prosjekter som oppgir at de har støtte fra et eller flere NFR-program vil fullt ut tilskrives de ulike departement i tabellen over, selv om de fleste av disse også kan ha en betydelig bruk knyttet til internt finansierte forskningsprosjekt. Vi kan grovt sett anta at prosjektene som oppgir finansiering fra Forskningsrådet har en lik fordeling av bruk på eksternt og internt finansierte forskningsprosjekter. Basert på et forslag på 250 MNOK i ny grunnfinansiering, der fordelingen mellom HPC og lagring er henholdsvis 70% og 30%, så vil andelen finansiering knyttet til Annen eller ukjent være på 49 MNOK. Det betyr at det gjenstår

201 MNOK til fordeling mellom NFR og UH- og instituttsektoren. Som nevnt over, så er finansieringen av forskningsprosjekter nokså likt fordelt mellom intern og ekstern finansiering, som betyr at 201 MNOK deles likt mellom NFR og UH- og instituttsektoren. «Annen eller ukjent» vil i praksis ofte være intern finansiering og vil ansvaret for denne andelen tilfalle UH- og instituttsektoren. Dette betyr at en fordeling mellom NFR og UH- og instituttsektoren for et nivå på rundt 250 MNOK vil være henholdsvis 100 MNOK og 150 MNOK.

Det betyr det at UH- og instituttsektoren står for rundt 60% av finansieringen, og Forskningsrådsprosjekter for rundt 40%.

Når det gjelder ulike organisasjoner sin bruk av ressursene i Sigma2 sin e-infrastruktur er den her gjengitt i tabellen nedenfor. Her er det bare tatt med organisasjoner som bruker mer enn 1% av HPC eller lagringsressurser.

<b>Organisasjoner sin %-vise forbruk for 2015-2018</b>	<b>%-vis bruk av HPC-ressurser</b>	<b>%-vis bruk av Lagringsressurser</b>
Universitetet i Oslo	36.6%	18.8%
Norges Teknisk-Naturvitenskapelige Universitet	22.9%	3.0%
Universitetet i Bergen	21.7%	49.3%
SINTEF AS	5.8%	3.1%
UiT - Norges arktiske universitet	3.4%	0.0%
Nansen Environmental and Remote Sensing Center	2.1%	5.6%
Akvaplan-NIVA AS	1.5%	3.0%
Uni Research AS	1.2%	3.0%
Universitetet i Stavanger	0.9%	0.0%
CICERO - Senter for klimaforskning	0.6%	1.2%
Norges miljø- og biovitenskapelige universitet	0.0%	1.0%
Andre	3.6%	11.9%
	100%	100%

Følgende liste viser programmer/utlysninger i Forskningsrådet med prosjekter som bruker eller har brukt Sigma2 sin e-infrastruktur. Et prosjekt kan ha finansiering fra flere programmer.

BIA  
BIOBANK  
BIONÆR  
BIOTEK2021  
CERN  
CLIMIT  
DEMO2000  
ENERGIX  
EUROCORES  
EUROSTARS  
EVITA  
FME  
FMETEK  
FORINFRA  
FORNY20  
FRIBIO  
FRIBIO2  
FRIHUMSAM  
FRIMED2  
FRIMEDBIO  
FRINAT  
FRINATEK  
GASSMAKS  
HAVBRUK2  
HAVBRUKS  
HAVKYST  
HELSE-EU  
IKTPLUSS  
INDNOR  
INTPART  
ISPATTEK  
IS-TOPP

KLIMAFORSK  
KOSK-II  
MARINFORSK  
MAROFF-2  
MAT-SLF  
MILJØ2015  
MILJØFORSK  
NANO2021  
NANSEN  
NATURNAER  
NEVRONOR  
NORDSATS  
NORKLIMA  
PETROMAKS  
PETROMAKS2  
PETROSENTR  
POLARPROG  
PSYKISK  
RBGRUNTEKN  
RENERGI  
ROMFORSK  
SFF  
SFI  
SIPHINIFES  
SKATT  
STAMCELLER  
STIPINST  
STRAT-MAT  
SYNKNØYT  
UNI-MUSEER  
VERDIKT  
YFF

## APPENDIKS 4: Begrepsdefinisjoner

### E-infrastruktur

[Forskningsrådets definisjon] E-infrastruktur for forskning er IKT-baserte infrastrukturer som muliggjør avansert og samarbeidsorientert forskning. E-infrastruktur for forskning omfatter utstyr, drift og relaterte tjenester for tungregning, datalagring, programvaresystemer og høyhastighetsnettverk samt verktøy for effektiv arbeidsflyt og programvare for simulering og analyse av data. Begrepet e-infrastruktur benyttes også om digitale registre og databaser, samt verktøy og tjenester for å sikre og gjøre disse tilgjengelige. Se også appendiks 2 *Om e-infrastruktur*.

### Lagring (datalagring)

Lagring er en mekanisme for å ivareta og tilgjengeliggjøre data. En mer konkret definisjon er vanskelig da begrepet benyttes i ulike sammenhenger. Lagring betyr å ta vare på informasjon (data), men kan ikke sees isolert fra tjenester som tilgjengeliggjør informasjonen. Lagring har således ulike egenskaper ut fra hvordan de lagrede data anvendes. For eksempel er både arkivering og databaser en form for lagring. Lagringen har også ulike egenskaper i forhold til typen data som lagres, eksempelvis åpne versus sensitive data, strukturerte versus ustrukturerte data. Metadata er helt kritisk for å få verdi ut av data som er lagret, og metadata må selv lagres i systemer som er egnet for det.

### Forskningsinfrastruktur

Avansert vitenskapelig utstyr og store utstyrsfasiliteter omfatter utstyr innenfor alle fagområder og inkluderer store nasjonale laboratorier, utstyrsenheter og forskningsinstallasjoner.

### Datainfrastruktur (forskningsdatainfrastruktur)

Forskningsinfrastrukturer med formål å innsamle, bearbeide, lagre, kuratere, tilgjengeliggjøre og/eller forbedre utnyttelse av data.

### FAIR

De internasjonale FAIR-prinsippene er utarbeidet som et sett av retningslinjer for å tilrettelegge for økt dataverdi og videre bruk av forskningsdata. FAIR er et akronym for ordene *findable*, *accessible*, *interoperable* og *reusable*. Forskningsdata skal med andre ord være av en kvalitet som gjør dem tilgjengelige, gjenfinnbare og gjenbrukbare. Videre skal data og metadata kunne håndteres maskinelt og det skal brukes konsistente vokabular.

### HPC – High Performance Computing

På norsk, tungregning eller høyhastighetsberegninger. Brukes som betegnelse for regnekraft som er langt over det som er tilgjengelig på en enkelt kontormaskin/PC.



## APPENDIKS 5: Akronymer

AI	Artificial Intelligence (kunstig intelligens)
BOTT	UiB, UiO, NTNU og UiT (de fire breddeuniversitetene)
EDI	European Data Infrastructure
EOSC	European Open Science Cloud
FAIR	Findable, Accessible, Interoperable, Reusable (om data)
HAP	Helseanalyseplattformen
HDP	Helsedataprogrammet
HOD	Helse og Omsorgsdepartementet
HPC	High Performance Computing
PRACE	Partnership for Advanced Computing in Europe
TCO	Total Cost of Ownership – total kostnad ved anskaffelse og drift

## APPENDIKS 6: Vurdering av høringsinnspill og endringer etter høring

Vi oppsummerer her kort de innspill som har kommet som en del av høringen og som har blitt vurdert og som enten har ført til endringer i dokumentet eller som utvalget har valgt ikke å ta til følge. Det vises for øvrig til innkomne høringsinnspill for en fullstendig oversikt over alle innspill.

Generelt får rapporten støtte for sin analyse og forslag til finansieringsmodell. Overordnede innspill til første utgave av rapporten fokuserer på utvalgets definisjon av ambisjonsnivå hvor en del aktører ønsker at dette i større grad knyttes til dokumenterte norske behov, den tverrsektorielle finansieringen, graden av brukerbetaling samt grunnlaget for finansieringsmodellen. Spesielt er det stor bekymring fra instituttsektoren for at de ikke vil kunne bidra inn i den foreslåtte grunnfinansieringen, og av den grunn vil måtte endre sine fokusområder.

Kort oppsummering av de ulike høringsinnspillene sammen med utvalgets vurdering:

**FFA** mener at finansiering av forskningsinfrastruktur primært bør være et statlig ansvar, og at det bør foretas en avklaring av ansvaret for e-infrastruktur mellom NFR og sektorene. Ulikhet i grunnfinansiering gjør at FFA foreslår at institutter med under 25% basisfinansiering må kunne få godkjent grunnfinansieringskostnadene som prosjektkostnader hos NFR og EU. FFA er generelt positive til brukerbetaling, samtidig som prismodellen ikke må stenge ute små aktører.

*Utvalget mener at selv om e-infrastruktur i prinsippet bør være et statlig ansvar, så mener utvalget at de ulike store aktørene, som i all hovedsak er statlig finansiert, må bidra, i tråd med utvalgets forslag. Utvalget har endret forslaget til finansiering fra aktører i instituttsektoren som bruker mer enn 1% av ressursene til at dette ansvaret legges til NFR. Utvalget ber NFR vurdere om dette ansvaret best ivaretas gjennom brukerbetaling i prosjekter eller som en øremerket økning i basis til institutter med under 25% basisfinansiering, men da som en direkte overføring til Sigma2 fra NFR. Det siste vil gjøre bidraget mer forutsigbarhet i forhold til investeringer.*

**NERSC** mener at norsk e-infrastruktursatsing ikke i tilstrekkelig grad har tatt innover seg de store internasjonale investeringer og satsinger på dette området, og således at utvalgets ambisjonsnivå er for lavt. Samtidig påpeker NERSC at med-finansieringsmodellen foreslått av utvalget ikke er mulig for NERSC, da den årlige utgiften for NERSC med dagens bruk vil overstige selv den største forskningsfinansieringen NERSC har hatt.

*Utvalget mener at det er rom for å diskutere ambisjonsnivået, og viser til vurderingen av innspillet fra FFA når det gjelder med-finansieringsmodell.*

**NTNU** ønsker en bedre begrunnelse for ambisjonsnivået basert på dokumenterte norske behov, og mener videre at utvalget i liten grad har vurdert andre løsninger som for eksempel skyløsninger og andre kommersielle løsninger. NTNU gir tilslutning til fordelingsmodell, men fremhever viktigheten av brukerbetaling for å gjøre brukere

bevisst på kostandene ved infrastrukturen. Samtidig ser NTNU behovet for å unngå en for stor grad av fragmentering av finansieringen. NTNU fremhever også viktigheten av en bred basis og støtte for e-infrastruktur og at lokale, tematiske og nasjonale/generiske infrastruktureer samordnes.

*Utvalget har ikke tatt standpunkt til hvordan e-infrastrukturbehovene skal løses operasjonelt. Dette kan skje både gjennom oppbygging og drift av nasjonale tungregne- og datalagringsanlegg, men kan også involvere internasjonale og kommersielle løsninger. Utvalget har forsøkt presisere at det strategiske ansvar for hvordan behovene løses på en kost-effektiv måte vil ligge hos styret i Sigma2. Utvalget har sett på behovsestimater for fremtidig bruk av e-infrastruktur og dette er nå tatt inn i rapporten.*

**Havforskningsinstituttet** er positiv til utvalgets modell, men ønsker en presisering av om utvalgets modell for fordeling av kostnader for grunnfinansiering videreføres med dagens nøkkel eller om denne vil baseres på historisk bruk.

*Utvalget har presisert i rapporten at nye nøkkel bør baseres på historisk bruk de siste 3 år. Samtidig er det slik at Sigma2 ikke vil ha mandat til å kreve en innbetaling, og dermed må basere seg på at sektoren slutter seg til prinsippet for fordeling og at forhandling må skje med alle interessenter om ambisjonsnivået for en ny finansieringsperiode. Disse forhandlingene vil ta utgangspunkt i Sigma2-styrets anbefaling for ambisjonsnivå, men vil måtte reflektere alle interessenters syn, uavhengig av styrerepresentasjon.*

**NMBU** gir gjennomgående støtte til utvalgets forslag, og fremhever som NTNU behovet for å koordinere fagspesifikke og generiske e-infrastruktureer.

**MET** deler NERSC sin bekymring om at norske forskningsmiljøer sakker akterut i den internasjonale konkurransen innen fag som krever avansert e-infrastruktur. MET mener rapporten er uklar på om fokus er forskning eller forskning og forvaltning, og mener samarbeid er viktig, og at datalagring synes som det mest nærliggende område for samarbeid. MET mener videre at finansieringsmodellen bør reflektere ulike e-infrastrukturbehov, og videre at modellen bør skille bedre mellom investering og drift. MET mener også at modellen bør ha en høyere grad av brukerbetaling (men at langtidslagring skal dekkes fullt ut av grunnfinansieringen). De påpeker også at en modell basert på historisk forbruk er krevende for aktører som baserer sin finansiering på konkurranseutsatte midler. Som NTNU, ønsker MET at rapporten er tydeligere på andre løsninger enn investeringer i egne anlegg.

*Utvalget mener at den beste løsningen er en modell som omfatter både forskning og forvaltning, men det forutsetter at relevante departement deler denne oppfatningen og investerer i nasjonale løsninger fremfor å bygge opp egne løsninger. Som for NTNU, så har utvalget tydeliggjort at det er Sigma2 som har ansvaret for å finne løsninger for å skaffe tilveie nødvendige e-infrastrukturressurser, og at dette innebærer mer enn investering i egne anlegg. Utvalget ser behovet for å videreutvikle finansieringsmodellen med større grad av differensiering, men har ikke fulgt dette videre for samtidig å sikre at*

*modellen er så enkel som mulig. Utvalget deler ikke oppfatningen om at en høyere andel brukerfinansiering er rett vei å gå.*

**UiA** har dekket sine behov for e-infrastruktur tjenester ved lokale løsninger, men er interessert i en fremtidig dialog om økt samarbeid.

**UiO** støtter utvalgets hovedkonklusjoner og trekker frem behovet for lokal bruker støtte i den nye organiseringen. UiO ønsker også en hyppigere evalueringsrate enn hvert 6. år og at en negativ evaluering ikke nødvendigvis skal ha som konsekvens at det nasjonale samarbeidet avvikles, men at dette brukes som grunnlag for å videreutvikle modellen. UiO peker også på behovet for en nasjonalt koordinert tjeneste for kuratering og langtidslagring av forskningsdata.

*Utvalget har justert omtalen av evalueringshyppighet og konsekvensene av en negativ evaluering. En evaluering av redusert omfang (kanskje mer kvantitativ enn kvalitativ) hvert 3. år kan supplere den mer omfattende evalueringen hvert 6. år, og utvalget deler synet at Sigma2 ved en negativ evaluering først gis anledning til å rette opp problematiske områder fremfor umiddelbart å styre mot avvikling. Utvalget ser behovet for kuratering og langtidslagring av forskningsdata, og at en nasjonal løsning også kan være ønskelig for dette. Utvalget mener derimot at dette ligger utenfor utvalgets mandat.*

**SSB** deler FFA sitt syn om at finansieringen av nasjonal e-infrastruktur bør være statlig og videre at den bør være uavhengig både av institusjoner og forskningsråd. De deler videre et syn av at bruk av infrastrukturen bør involvere et element av brukerbetaling.

*Utvalget noterer seg innspillet fra SSB, men har ikke gjort nevneverdig justeringer på bakgrunn av dette innspillet alene.*

**UiB** støtter rapportens hovedkonklusjon om at det er behov for en opptrapping på området. UiB peker videre på at implementering av en nasjonal løsning må ta et helhetsperspektiv da en stor andel av beregnings- og lagringsbehov krever volum, men ikke nødvendigvis spissytelse. UiB ønsker en høyere andel brukerfinansiering også for å styrke kostnadsbevisstheten, og foreslår et mål om 20% brukerfinansiering, samtidig som enkel tilgang for småbrukere må ivaretas.

*Utvalget noterer seg innspillet fra UiB og har, som for andre høringsinnspill, forsøkt å tydeliggjøre at helhetsperspektivet skal ivaretas av styret i Sigma2. Utvalget vil på dette tidspunkt ikke heve ambisjonsnivået for andelen brukerfinansieringen, men noterer at dette kan bli høyere enn rapporten antyder hvis NFR velger å øke sitt bidrag primært gjennom prosjektfinansieringsmodeller.*

**Simula** støtter hovedelementene i rapporten, og ønsker ikke et høyere nivå på brukerbetaling. Som UiO mener Simula at en negativ evaluering ikke uten videre må føre

til en fase som tar sikte på en avvikling. Simula påpeker videre at utviklingen innen maskinlæring/kunstig intelligens vil føre til økte e-infrastrukturbehov fra industrien gjennom instituttsektoren, og at finansieringsmodellene må gi rom for slik aktivitet. Simula tar også opp utfordringer knyttet til mangel på programmer for som støtter utvikling av moderne algoritmer for moderne infrastruktur.

*Utvalget har ikke valgt å endre finansieringsmodellen ut fra forventningene om økt aktivitet knyttet til industri-finansierte prosjekter. Dagens modell åpner for slik aktivitet og det legges klare rammer for omfanget av slike beregninger gjennom kravene til egenregi. Utvalget presiserer at Sigma2 må følge utviklingen nøye og vurdere helheten slik at alle brukerbehovene kan bli dekket på en best mulig måte. Norges deltagelse i EuroHPC kan gi rom for økt industriell deltagelse. Utvalget anerkjenner utfordringer knyttet til mangel på forskningsprogrammer for utvikling av programvare for moderne infrastruktur, men anser at dette ligger utenfor utvalgets mandat.*

**Clarino** peker på behovet for støtte til prosjekter som har små, men kritiske e-infrastrukturbehov.

*Utvalget noterer seg innspillet, men har ikke gjort endringer i rapporten. Utvalget er av den oppfatning at institusjonene må ta ansvar for slike behov.*

**NSD** slutter seg til hovedprinsippene som er foreslått i rapporten, og gir noen innspill på deres rolle i e-infrastrukturlandskapet.

*Utvalget har justert tekst og figurer slik at disse best mulig reflekterer NSD sin rolle i e-infrastrukturlandskapet.*

**UNIT** mener at sett i et bredt samfunnsperspektiv er summene som det legges opp til i finansieringsmodellen ikke spesielt høye for en nasjonal e-infrastruktur. Unit etterspør en klargjøring av de ulike finansieringstyper og viser til begreper som brukes av UNIT og NFR samt til Helsedataprogrammets definisjoner. UNIT peker også på, som FFA og NERSC, de særlige finansieringsutfordringer som gjelder for instituttsektoren. Unit etterspør også helseforetakene/sykehusene sin plass i e-infrastrukturlandskapet. Unit etterspør også enkelte klargjøringer for deler av rapporten, og at det presiseres at kuratering av data finansieres og at bruk av skytjenester vil kunne redusere behovet for langsiktighet i grunnfinansieringen, da en slik løsning i større grad vil kunne håndtere variasjon i behov.

*Utvalget har presisert at det er Sigma2 som har ansvaret for å finne en optimal implementering av e-infrastrukturtenestene, slik at ulike brukerbehov dekkes. Vi har klargjort noen av de momentene som UNIT har trukket frem som uklare. Helseforetakene/sykehusene er potensielt viktige brukere av nasjonale e-infrastrukturtenester, og dette må følges nøye. Samtidig er kravene til sikre løsninger langt større for slike sensitive data enn mer generelt, og det ikke gitt at felles løsninger kan utarbeides. Utvalget er enig i at kuratering av data er viktig, men mener dette krever nærhet til fag, og at dette derfor ikke vil kunne betraktes som generiske e-*

*infrastrukturtenester, men snarere fagspesifikke tenester som må finansieres på andre måter.*

**UIT** slutter seg til hovedprinsippene i rapporten. UiT tror ikke en økt andel av brukerfinansiering er realistisk samtidig som en skal forvalte regelmessige fornyinger av infrastrukturen, da dette krever forutsigbarhet. Ber om at investeringer i e-infrastrukturtenester sees i en helhet mot digitaliseringen som skjer i sektoren som helhet, og spesielt i regi av UNIT.

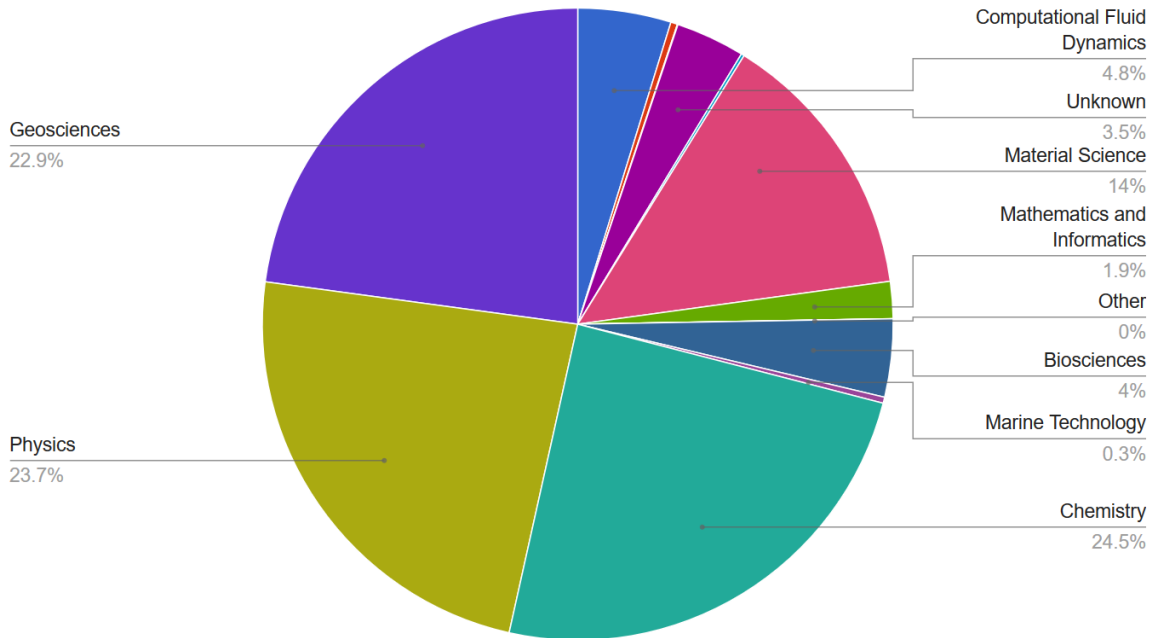
***Utvalget noterer seg** innspillet fra UiT og har ikke gjort spesielle endringer på basis av denne. Utvalget forutsetter av behovet for å koordinere generisk infrastruktur med tematiske eller fagspesifikke infrastrukturer løses i dialog mellom Sigma2 og UNIT.*

**NFR** slutter seg til utvalgets problembeskrivelse og at en ren konkurransemekanisme ikke er velegnet for generisk e-infrastruktur. NFR peker på viktigheten av en god samhandling mellom den nasjonale e-infrastrukturen og lokale og/eller tematisk løsninger, samtidig som at forskjellene mellom disse, også i finansiering, er tydelig. For å tydeliggjøre nedslagsfeltet for bruken av e-infrastruktur, ber NFR om oppdaterte data brutt ned noe mer finmasket.

***Utvalget noterer seg** NFR sin generelle tilslutning og ønske om presiseringer, og har tydeliggjort dette i den reviderte rapporten.*

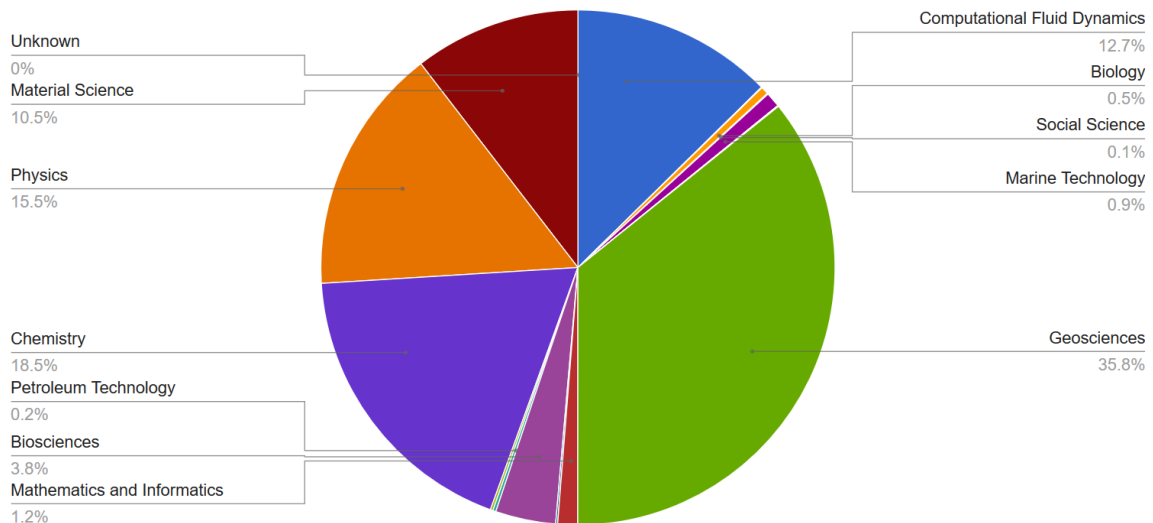
## APPENDIKS 7: Oversikt over forbruk fordelt på vitenskapelige disipliner

HPC Usage - Primary Field of Science



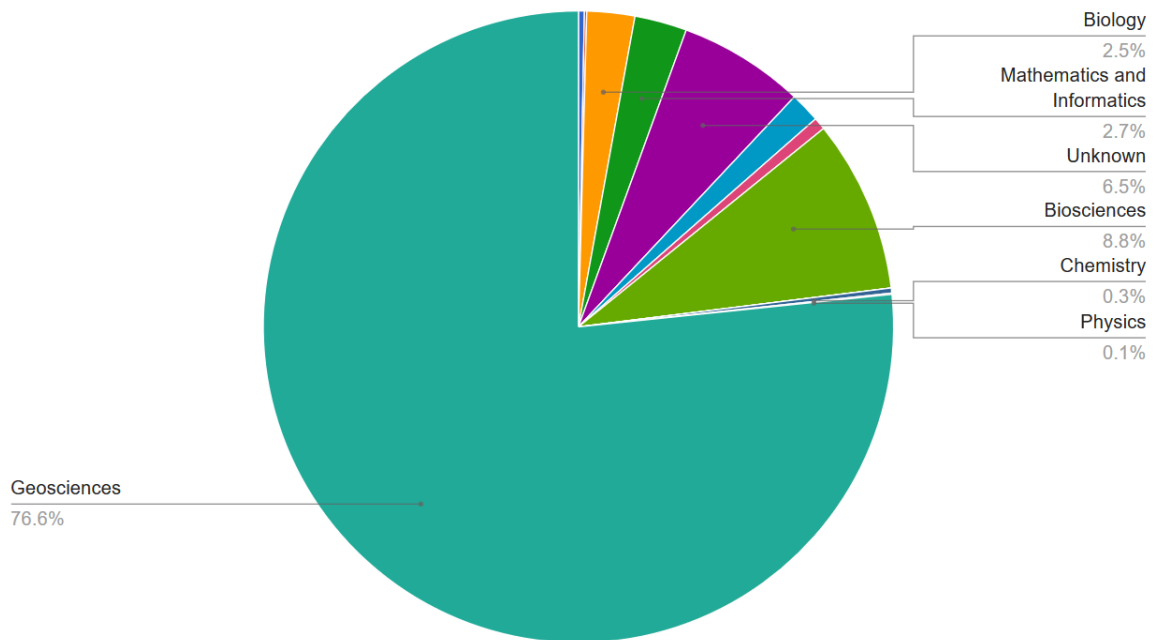
Figur 9: HPC forbruk i 2015.1 fordelt på vitenskapelig disiplin

HPC Usage - Primary Field of Science



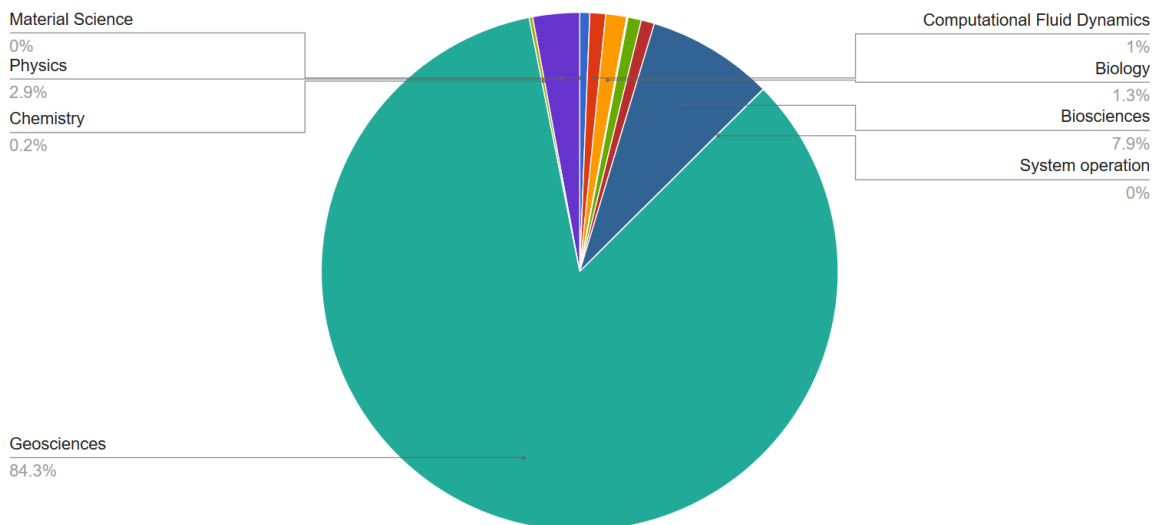
Figur 10: HPC forbruk i 2019.1 fordelt på vitenskapelig disiplin

**Storage Usage - Primary Field of Science**



*Figur 11: Lagring forbruk i 2015.1 fordelt på vitenskapelig disiplin*

**Storage Usage - Primary Field of Science**



*Figur 12: Lagring forbruk i 2019.1 fordelt på vitenskapelig disiplin*

Tabell 4 under viser en oppsummering i den relative endringen for ulike vitenskapelige domeners bruk av infrastrukturen. Det er viktig å merke seg at gjennom disse årene har



Primary field of Science	% HPC usage		% change	% Storage usage		% change
	2015	2019		2015	2019	
Biology		0.5	New	2.5	1.3	-48%
Biosciences	4.0	3.8	-5%	8.8	7.9	-10%
Chemistry	24.5	18.5	-24%	0.3	0.2	-33%
Computational Fluid Dynamics	4.8	12.7	165%		1	New
Geosciences	22.9	35.8	56%	76.6	84.3	10%
Marine Technology	0.3	0.9	200%			
Material Science	14.0	10.5	-25%			
Mathematics and Informatics	1.9	1.2	-37%	2.7		
Petroleum Technology		0.2	New			
Physics	23.7	15.5	-35%	0.1	2.9	2800%
Social Science		0.1	New			
Unknown	3.9			6.5		

Tabell 4: Relativ (innbyrdes) endring i bruk for ulike vitenskapelige domener.

den totale mengden ressurser økt. Derfor kan et domene som tilsynelatende har en nedgang i andel forbruk, faktisk likevel ha en økning i absolutt forbruk.

Oversikt over ulike type brukere	2015	2019	%-vis utvikling
Totalt antall brukere	1381	1909	138%
Totalt antall interne brukere (ved vertsinstitusjoner)	890	1154	130%
Totalt antall eksterne brukere	491	755	154%
Antall studenter (Master, Bachelor)	204	57	28%
Antall PhD-student	265	244	92%
Antall forskere (fast ansatte, post doc., mm.)	571	570	100%
Antall brukere fra industri/næringsliv	93	141	152%
Totalt antall prosjekter	299	460	154%

Tabell 5: Endring i ulike brukergrupper fra 2015 - 2019 (2019 ikke avsluttet da tabellen ble laget og det kan medføre at tallene blir marginalt høyere).