

Anbefalinger til en sammenhængende e- infrastruktur til fremme af dansk forskning

for perioden 2018-2028, set i lyset af forskningens digitalisering og globalisering

Indholdsfortegnelse

1	Forord – digitalisering kræver forsknings e-infrastruktur	4
2	Overordnede anbefalinger	5
3	Potentiale i eScience og Open Science.....	6
4	Overblik – elementerne i forskningens e-infrastruktur	9
4.1	<i>Den geografiske sammenhæng mellem elementerne.....</i>	9
5	Internationale tendenser og e-infrastrukturer	11
5.1	<i>Europæisk Science Cloud</i>	11
5.2	<i>Den generelle Databeskyttelsesforordning</i>	12
5.3	<i>Computing.....</i>	12
5.4	<i>Datamanagement og lagring.....</i>	12
5.5	<i>Forskningsnet, sikkerhed og autentifikation</i>	13
5.6	<i>Kompetencer</i>	13
6	Situation, udfordringer og anbefalinger for e-infrastrukturelementerne	15
6.1	<i>Datamanagement og – lagring.....</i>	15
6.2	<i>Computing.....</i>	17
6.3	<i>Software og applikationer.....</i>	19
6.4	<i>Forskningsnet</i>	20
6.5	<i>Sikkerhed og autentifikation</i>	20
6.6	<i>Kompetencer</i>	21
7	Konklusion og samlede anbefalinger	23
7.1	<i>Økonomioversigt</i>	26
7.2	<i>Afsluttende bemærkninger.....</i>	29
8	Referenceliste	30
	Bilag 1: Videnskabelige temaer.....	33
	Bilag 2: Sammenligning af investeringer og økonomiske forhold omkring e-infrastruktur	37
	Bilag 3: e-IRG Roadmap 2016 – recommendations for national governments/funding agencies	41

1 Forord – digitalisering kræver forsknings e-infrastruktur

Digitaliseringen af data breder sig. Det forandrer vores hverdag og verden og giver nye muligheder for økonomisk vækst og udvikling indenfor industri og samfund.

Forskningen er godt i gang med at udnytte digitaliseringens muligheder. Fra forskning i de meget store datamængder fra Large Hadron Collider hos CERN til data indsamlet i store mængder fra Sociale medier, data genereret fra gen-sekventering indenfor sundhedsforskning, samt data opnået ved digitalisering af skriftlige kilder og arkiver. Mængden af digitale data vokser eksponentielt indenfor alle forskningsområder. For at sikre den bedst mulige anvendelse og samfundsmæssige udnyttelse af de enestående muligheder for forskning og udvikling, er der behov for en national strategi.

Forskning ved anvendelse af de digitale data, kaldes eScience og kræver en understøttende e-infrastruktur, specielt stor computekraft til beregning og simulering, Internationalt går udviklingen mod større åbenhed og deling af data, og anvendelse af disse i forskningen under betegnelsen "Open Science" – Åben Videnskab.

I april 2016 lancerede EU Kommissionen "European Cloud Initiative" (EU, 2016) der skal være med til at sikre at forskning, erhvervsliv og samfundet som helhed kan udnytte datarevolutionen. Ministerrådet vedtog på deres møde den 17. maj 2016 (Ministerråd, 2016), aktivt at arbejde for at gøre åben videnskab til virkelighed. European Cloud Initiative er en del af EU's store initiativ om Digital Single Market.

I Forsknings- og Innovationspolitisk redegørelse 2016, understregede Uddannelses- og forskningsministeren, at regeringen ønsker at tage aktiv del i udviklingen mod en mere åben videnskab – Open Science (UFM, Forsknings- og innovationspolitisk redegørelse, 2016)

Udnyttelse af potentialet i Open Science kræver en understøttende infrastruktur, der håndterer transport, opbevaring og behandling af data. Den digitale infrastruktur, der sikrer dette; e-infrastruktur, omfatter blandt andet lagringsfaciliteter, supercomputere til behandling af data og dedikerede netværksforbindelser til transport af data.

De danske forskernes adgang til at udnytte mulighederne i Open Science afhænger blandt andet af, at der er en sammenhængende e-infrastruktur lokalt, nationalt og internationalt, som er let at gå til, og som understøtter de forskellige behov indenfor de videnskabelige områder.

DeiC - Danish e-infrastructure Cooperation, er den danske organisering, der varetager det nationale og internationale samarbejde på e-infrastruktur området, og har blandt andet til opgave at formulere strategier, der er med til at sikre at eScience i Danmark udvikler sig bæredygtigt i Danmark.

Dette dokument indeholder DeiCs anbefalinger til at sikre en sammenhængende dansk e-infrastruktur for perioden 2018-2028 – til fremme af dansk forskning set i lyset af den digitale og internationale udvikling.

25. april 2017

Børge Obel

Steen Pedersen

Bestyrelsesformand

Direktør

2 Overordnede anbefalinger

Med basis i den forsknings- og innovationspolitiske redegørelse december 2016 og EU-ministerrådsbeslutning den 27. maj 2016, som begge indikerer en dansk opbakning bag fremme af Open Science, tager dette dokument som udgangspunkt, at Danmark

- vil fremme eScience indenfor alle videnskabelige områder, og dermed give de danske forskere samme muligheder, som deres internationale kolleger
- vil udnytte de muligheder, der åbnes for hele samfundet ved at satse på Open Science

Den nødvendige e-infrastruktur til understøttelse af eScience og Open Science opnås ved at alle danske forskere sikres velkendt og ubesværet adgang til sammenhængende og sikre e-infrastrukturløsninger.

Kun ved en koordineret indsats kan den førerposition Danmark i dag har i "det digitale samfund" fastholdes og udbygges inden for samfundsudviklingen i det, der kaldes den 4. industrielle bølge. Det kræver investering i såvel kompetenceudvikling som infrastruktur.

DeiC har følgende konkrete anbefalinger omkring e-infrastrukturen

1. Fortsat understøttelse af aktiviteter, der sikrer Open Science og eScience kompetencer
2. Investering i nationale supercomputing anlæg
3. Øget investering i datamanagement af forskningsdata
4. Sikring af forskningsnettet som en fortsat "state-of-the-art" forskningsinfrastruktur
5. Øget satsning på sikkerhed i e-infrastrukturen
6. Aktiv deltagelse i internationalt samarbejde

Gennemførelse af anbefalingerne kræver:

- a. En 10 årig national vision for e-infrastrukturen
- b. En 10 årig aftale om finansiering
- c. En organisering, der kan understøtte en sammenhængende e-infrastruktur

3 Potentiale i eScience og Open Science

Forskning baseret på digitale data går under betegnelsen eScience; det omfatter indsamling, behandling og anvendelse af videnskabelig information i dataform.

eScience har længe været en integreret del af forskningen indenfor de tekniske og naturvidenskabelige områder. Digitaliseringen betyder, at der også på de øvrige forskningsområder er et stort, - og i vid udstrækning stadig uudnyttet potentiale for banebrydende forskningsresultater.

Indførelse af Open Science vil gøre forskningspotentiallet endnu større, med mulighed for forskning i data på tværs af videnskaber.

eScience er ikke længere kun forbundet med anvendelse af meget stor regnekraft, og det meste forskning i fremtiden vil have elementer af eScience, med forskellige behov for e-infrastrukturunderstøttelse afhængigt af projektets størrelse og fase.

Den optimale understøttelse af det enkelte forskningsprojekts eScience behov vil dermed ske bedst i et samspil af lokale, nationale og internationalt infrastrukturer, der kan tilbyde den funktionalitet, der er ønsket.

Dette afsnit vil forsøge at belyse nogle af de eScience anvendelser, der er i dag, og de muligheder, der kan være i fremtiden, hvis forskningen understøttes af den nødvendige e-infrastruktur for området.

Der har på nuværende tidspunkt ikke været en dansk undersøgelse af behovet for e-infrastruktur for de forskellige videnskaber. Dette afsnit bygger derfor på resultater af vurderinger foretaget i Sverige (Vetenskapsrådet, 2014), Finland (CSC, 2015), Norge (UNINETT, 2017) og England (UK, 2012), og på citater fra forskere fra de forskellige områder, primært fra præsentationer på "National Supercomputing Day, - Perspectives of High Performance Computing" (DeIC, National Supercomputing Day - Perspectives on High Performance Computing, 2016).

Temaerne nedenfor er udvalgt efter Dansk Roadmap for Forskningsinfrastruktur 2015 (Uddannelses- og Forskningsministeriet, 2015).

Energi, klima og miljø

Klimaforandringer, begrænsede ressourcer og tilsvarende udfordringer gør forskning indenfor energi, klima og miljø vigtigt for den fremtidige planlægning og politik på området.

Beregning af vejret, klimasituationen og simuleringer af den fremtidige udvikling kræver supercomputing-ressourcer. Den svenske science case for e-infrastruktur (Vetenskapsrådet, 2014) forudsiger en femdobling af behovet indenfor en femårig periode.

Mængden af data og muligheden for at samarbejde og kombinere med data fra andre relaterede forskningsområder gør også storage- og

"In climate and Earth system science, modelling is the key tool for research. Earth system models help us to study feedbacks in the Earth system on a range of timescales. These feedbacks contribute to the uncertainty in projections of future climate"
Associate professor Guy Schurgers, Science, Københavns Universitet (Schurgers, 2016)

datamanagementudfordringen væsentlig i det fremtidige forskningsarbejde.

DMI placerede i 2015 deres supercomputer på Island af hensyn til strømbesparelsen. Placeringen der kræver en sikker og stabil internetforbindelse til overførsel af data til Danmark, som leveres af det danske forskningsnet i samarbejde med NORDUnet.

Biotek, Sundhed og LifeScience

Sundhedsområdets anvendelse af eScience er i kraftig fremgang, og vigtigheden af forskningen er umiddelbar. Udvikling af lægemidler og medicinske teknologier, programmer til personlig/præcisions medicin, sundere og bedre fødevarer osv., foregår både på universiteterne, hospitalerne, i sektorforskningsinstitutionerne og i den private industri.

Det er et område med mange interessenter og meget samarbejde på tværs af sektorerne, og giver dermed udfordringer til såvel datamanagement, sikkerhed og autentifikation som computing-ressourcer og sikker transport af data.

Mange af de data, der anvendes i forskningen er personhenførbare, og skal håndteres specielt sikkert og fortroligt. Udfordringerne er mange, og en sikker håndtering bør understøttes af nationale tiltag, der sikrer, at såvel forskning som personhensyn tilgodeses.

Det koordinerende organ for registerforskning (KOR)s opgave er blandt andet at bidrage til at forbedre forskernes adgang til og anvendelse af registerdata, herunder bidrage til at koordinere udnyttelse af ressourcer på området og bidrage med koordinerede og langsigtede strategier for udviklingen af registerforskningsområdet i samspil med de forskellige aktører på området. (KOR, 2014).

e-infrastrukturunderstøttelse af forskningsområdet kræver en national indsats med sikre netforbindelser og avancerede computing muligheder.

Det vil være rigtig tosset, hvis Danmark faldt bagud, når vi har så mange data. Det vil andre lande også have om 5, 10, 20 år, så det gælder om at udnytte den fordel vi har i dag. Det er vigtigt at få tegnebogen frem og få lavet nogle investeringer. Professor Søren Brunak, Københavns Universitet (Supercomputing i Danmark Podcast, 2017)

Fysik og Univers

Forskning indenfor fysik og univers har en længerevarende tradition for anvendelse af eScience. Forståelsen af jordens opståen og universets opbygning og udviklingen af kvantecomputere er alle afhængig af regneanlæg.

Også på dette område tager datamængden til, og behovet for såvel større lagringsfaciliteter, regnekrafter og netforbindelser er stadigt stigende.

For eksempel har Niels Bohr Institutet i København i mange år modtaget data fra Large Hadron Collider (LHC) på CERN i Schweiz. LHC gennemgår løbende forbedringer, og producerer stadig flere data, der skal overføres, behandles og gemmes.

Etableringen af det gigantiske mikroskop ESS (European Spallation Source) i Lund og det sammenhørende Datamanagement og Software Center (ESS-DMSC) i København vil også påvirke behovet for understøttende national e-infrastruktur, for at sikre at forskningspotentialer og mulighederne for de danske forskere udnyttes.

To harvest the full potential of the [ESS] Data Centre here in Denmark, it is important that we anchor it in the existing research- and business environments. Last year we made a long-term ESS strategy. The strategy emphasises the importance of close ties between the Data Centre and Danish universities. (Ulla Tørnes, 2016)

Materiale og Nanoteknologi

Forskning indenfor materiale og nanoteknologi har tradition for anvendelse af supercomputing både hos universiteterne og i industrien.

Vestas havde i mange år den kraftigste computer i Danmark til beregning af for eksempel optimal placering af vindmøller, vindsimuleringer mv.

Professor Ole Sigmund, DTU Mekanik er en af de danske forskere, der har fået tildelt regnetid i det europæiske supercomputer-samarbejde PRACE. Et af forskningsprojekterne gik ud på at optimere fly-vinger på et fx Boeing 777 gennem simuleringer inspireret af bl.a. udformningen af fisk og fugle.

Adgang til at regne på de store internationale anlæg kræver som regel, at forskeren har mulighed for at afprøve sine beregningsprogrammer på lokale eller nationale anlæg inden beregningen igangsættes på det internationale anlæg.

Humaniora og samfundsvidenskab

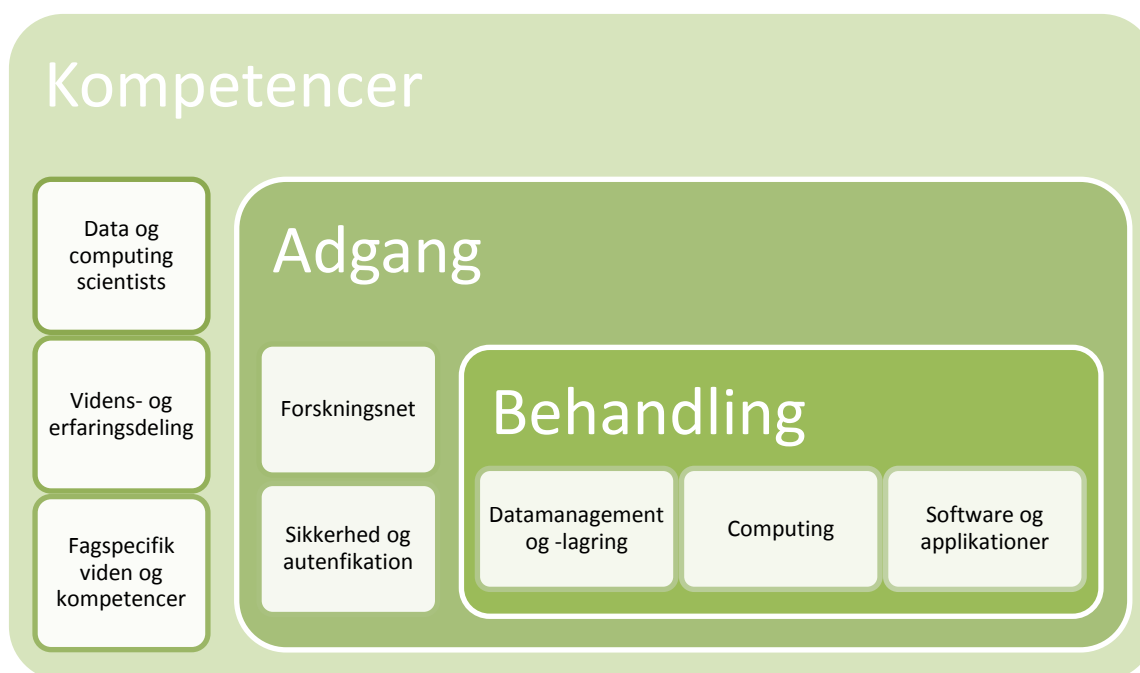
Humaniora og samfundsvidenskab er typisk set som novicerne indenfor eScience. Der eksisterer dog store eScience forskningsprojekter, og det fremtidige behov og potentiale er anerkendt, men kompetencerne og adgangen til den nødvendige infrastruktur er ikke på plads.

Mulighederne for anvendelse af eScience indenfor humaniora og samfundsvidenskab udspringer af den store mængde af data, der er opstået blandt andet efter digitalisering af skriftlige kilder og billedmateriale og indsamling af data fra internettet. For at kunne udnytte de muligheder digitale data giver, er der primært behov for support og kompetenceudvikling indenfor eScienceområdet, i alle faser af forskningen.

*“We get much more data than before, and it is much more complex, and we also want to use it on a much larger geographical scale, and that makes the challenge of handling these large data sets and making them comparable over large distances, one of the major challenges for Archaeology in the future. In my mind, there is no doubt, that **if we are capable of handling these data, it will also open for a complete new picture of the past**”. (Holst, 2014)*

4 Overblik – elementerne i forskningens e-infrastruktur

Dette afsnit giver et overblik over de enkelte elementer i forskningens e-infrastruktur. De enkelte elementer hænger sammen og understøtter hinanden i, hvad der kan betegnes e-infrastrukturens økosystem. Det er et komplekst system med mange dimensioner og interesser med forskellige krav og behov. Figuren nedenfor samler de forskellige elementer i en overskuelig sammenhængende oversigt.



e-infrastrukturens temaer og elementer, inspireret fra (Research Councils UK, 2014)

Der er generel enighed om, at e-infrastrukturen samlet omfatter ovennævnte sammenhængende elementer. Der skal findes og vedligeholdes udstyr til at *behandle* og opbevare data, *der skal være adgang til* data og beregningsressourcer og der skal være *kompetencer* til at anvende udstyr og data. e-infrastrukturen er ikke selve forskningen, men en væsentlig komponent i understøttelsen af megen forskning, specielt eScience og Open Science.

Internationalt bruges betegnelsen "e-infrastructure commons" ofte, (fx (e-IRG, 2017) (SNIC, 2017)) om e-infrastrukturens økosystem. I den betegnelse indgår også de intentioner, der ligger bag European Cloud Initiative/European Data Infrastructure (EU, 2016) om at give adgang til data (Open Science) og muliggøre dette gennem koordinering af de europæiske e-infrastruktur-elementer.

4.1 Den geografiske sammenhæng mellem elementerne

e-infrastruktur har et lokalt, nationalt og internationalt element. En national koordinering og samarbejde sikrer, at elementerne spiller sammen på den mest hensigtsmæssige måde under hensyntagen til såvel ressourceudnyttelse og forskningsresultater.

- Det lokale niveau, som det enkelte universitet (eller et samarbejde mellem flere universiteter) eller forskningsprojekt udvikler og driver, designet til specifikke behov

- Det nationale niveau, hvor e-infrastrukturen stilles til rådighed for de forskere i Danmark, der har behov, som kan løftes her, og hvor behovet er mere generelt og samme e-infrastruktur kan opfylde behovet på tværs af universiteter eller fagområder
- Det internationale niveau, hvor forskerne får adgang til e-infrastruktur, der ikke understøttes lokalt eller nationalt, eller hvor det giver mening at anvende de internationale e-infrastrukturer fremfor at etablere dem lokalt eller nationalt. Forskernes muligheder for at opnå adgang til internationale infrastrukturer faciliteres gennem nationens deltagelse i de relevante internationale organisationer og fora.

5 Internationale tendenser og e-infrastrukturer

Hvor der tidligere har været forskellige leverandører af forskningsnet og computing i de forskellige lande, ses en tendens til, at der dannes samlede organisationer, der leverer og koordinerer det nationale niveau af e-infrastrukturer i fx Norge, Holland, England og Finland.

Også på europæisk plan ses en konsolidering af organisationerne bag de forskningsunderstøttende e-infrastrukturer, som følge af udviklingen på området, og som en konsekvens af kravet i EU's Science Cloud initiativ om koordinering og effektivisering på e-infrastruktur området.

e-IRG, der rådgiver om udviklingen af den europæiske e-infrastruktur set i et globalt perspektiv, har netop udgivet en Roadmap (e-IRG, 2017) med anbefalinger til en udvikling, der kan støtter op om de nye muligheder indenfor forskningen. e-IRG anbefaler blandt andet at de enkelte lande evaluerer finansiering og rammer for deres nationale e-infrastruktur og deltager i udviklingen af det europæiske e-infrastrukturlandskab til fordel for såvel den nationale forskning som det internationale forskningssamarbejde.

Dette afsnit beskriver to EU initiativer, der har indflydelse på udviklingen i Danmark på e-infrastrukturuområdet. Herefter beskrives de internationale fora og organisationer, der i dag er et leverer det internationale niveau af e-infrastruktur og samarbejde og som Danmark samarbejder med i dag.

5.1 Europæisk Science Cloud

EU initiativet **europæiske science cloud (ECI)** (EU, 2016), består af to dele; **European Open Science Cloud (EOSC)** og **European Data Infrastructure (EDI)**.

Den åbne europæiske forskningssky (EOSC) tager sigte på at give Europa en førende rolle i verden inden for videnskabelig datainfrastruktur og at sikre, at europæiske forskere får det fulde udbytte af datadreven forskning. Rent praktisk vil den give 1,7 millioner europæiske forskere og 70 millioner fagfolk inden for videnskab og teknologi et virtuelt miljø, som tilbyder åbne og problemfri tjenester inden for lagring, forvaltning, analyse og genbrug af forskningsdata på tværs af landegrænser og videnskabelige faggrænser. (EUR LEX, 2016)

Formålet med den europæiske datainfrastruktur (EDI) er, at understøtte ambitionerne omkring initiativet om åbne data (EOSC). Det er vurderingen at, Europa har behov for integreret højtydende databehandlingskapacitet i verdensklasse, højhastighedsforbindelser og førende data- og software-tjenester for forskere og andre af sektorens førende brugere (herunder SMV'er) og den offentlige sektor. E-infrastruktur vil give mulighed for at udnytte værdien af big data og digitalisering som standard. Den europæiske datainfrastruktur vil omfatte en forvaltningsstruktur for ledelse og udvikling af datainfrastrukturer og -tjenester samt beslutningstagning inden for finansiering, bæredygtighed og sikkerhed på langt sigt. (EUR LEX, 2016)

Selv om der i begyndelsen fokuseres på forskermiljøet, vil brugergrundlaget blive udvidet til den offentlige sektor og erhvervslivet, hvilket skal muliggøre løsninger og teknologier, der kommer alle områder af økonomien og samfundet til gode.

5.2 Den generelle Databeskyttelsesforordning

EU's databeskyttelsesforordning (The General Data Protection Regulation, GDPR) er vedtaget i april 2016 og skal være fuldt implementeret 25. maj 2018, hvor den afløser persondataloven fra 2000 (Blume, 2016).

Forordningen, der indfører fælles persondataret på tværs af EU's medlemslande, og blandt andet har til formål at sikre fri udveksling af data, må ses som en af forudsætningerne for Open Science og initiativerne omkring European Open Science Cloud.

Den indebærer skærpede krav til og øget myndighedstilsyn med forskningens behandling af personhenførbare data, ligesom universitetsadministrationernes behandling af persondata vedrørende nuværende og tidligere studerende og ansatte er omfattet af reglerne.

Forordningens bestemmelser betyder bl.a. at universiteterne skal opbygge en datarådgiverfunktion (Data Protection Officer, DPO). Institutionerne skal have et beredskab, der gør det muligt at indsamle det nødvendige grundlag for at anmelde eventuelle datalæk og lignende sikkerhedshændelser til Datatilsynet og til de berørte, registrerede personer inden 72 timer efter at en hændelse opdages. Endvidere skal man være parat til med kort varsel at besvare henvendelser fra registrerede ("datasubjekter") om alt, hvad man har registreret om de pågældende.

Indførelsen af GDPR vil sætte et forøget fokus på og forventninger til datasikkerheden i e-infrastrukturen.

5.3 Computing

Danmark er gennem DeiC medlem af **PRACE** (Partnership for Advanced Computing in Europe), som giver danske forskere adgang til computing-ressourcer der er væsentligt større, end dem vi kan tilbyde i Danmark.

Adgangen tildeles ved en ansøgningsproces baseret på peer-review. PRACE samarbejdet har givet flere store danske forskningsprojekter de nødvendige regneressourcer i de seneste år. Det bør gennem dansk medlemskab også i fremtiden sikres, at danske forskere har denne mulighed

I regi af **NeIC** (Nordic e-Infrastructure Cooperation) arbejdes i forskellige projekter på udvikling af fælles nordisk ressourcedeling og erfaringsudveksling, fx deling af regneressourcer (Dellingr) og håndtering af følsomme data (Tryggve). (NeIC Activities, 2017). NeIC samarbejdet giver synergi mellem de nordiske initiativer og arbejder på ressourceoptimering på tværs af Norden.

5.4 Datamanagement og lagring

På datasiden eksisterer en mængde forskellige initiativer og projekter på internationalt plan.

Danmark/DeiC har engageret sig i to af initiativerne; EUDAT (European Data Infrastructure) og RDA (Research Data Alliance).

EUDAT er et EU finansieret initiativ, der arbejder på at sikre e-infrastrukturløsninger i hele datas livscyklus til understøttelse af forskernes behov på området.

Initiativet er primært erfaringsudveksling omkring løsninger på området. Det vil sige at det enkelte land/e-infrastrukturleverandør selv skal installere og drive e-infrastrukturen, men kan samarbejde med de øvrige EUDAT medlemmer om udvikling og initiativer

RDA er et globalt initiativ, der understøtter deling af forskningsdata gennem implementering af teknologier og praksis, der gør det muligt at dele data på tværs af grænserne og øvrige forhindringer. RDAs aktiviteter er primært udgivelse af rapporter og anbefalinger, der udarbejdes af forskellige arbejdsgrupper.

Danmark deltager desuden i **Knowledge Exchange**-samarbejdet gennem DEFF. Knowledge Exchange samarbejder om at støtte udviklingen af digital infrastruktur, der muliggør forskning og uddannelse med åben adgang til publikationer, data og software på tværs af grænserne.

5.5 Forskningsnet, sikkerhed og autentifikation

Forbindelsen fra det danske forskningsnet til de internationale forskningsnet og en række kommercielle parter varetages på Danmarks vegne af det fællesnordiske samarbejde **NORDUnet A/S**, hvor den danske stat på linje med de øvrige nordiske lande er hovedaktionær.

Forbindelsen fra **NORDUnet** går blandt andet til det pan-europæiske forskningsnet **GÉANT**, der forbinder forskningsnet i 43 europæiske lande og derigennem servicere omkring 50 millioner brugere indenfor alle forskningsområder. Derudover har **GÉANT** forbindelser til omkring 60 øvrige forskningsnet i resten af verden, som Asien, Afrika, USA og Sydamerika.

Forbindelsen til de internationale forskningsnet er nødvendig for at give de danske forskere mulighed for at deltage i internationale forsknings-samarbejder. Forskningsnettene giver sikkerhed for stabile højkapacitetsforbindelser på tværs af grænserne,

Det europæiske samarbejde mellem forskningsnet under **GÉANT** står bag eduroam, infrastrukturen, der giver forskere (studerende og øvrige universitetsansatte) nem adgang til trådløst net både på eget campus og på rejser til samarbejdspartnere over hele verden. Eduroam findes i dag i mere end 70 lande.

GÉANT faciliterer også samarbejde mellem de nationale e-infrastrukturleverandører indenfor sikkerhed og autentifikation.

I **TF-CSIRT** samarbejdet sikres koordinering mellem sikkerhedsorganisationerne hos e-infrastrukturleverandørerne og øvrige relevante organisationer.

SIG-ISM (Information Security Management) er et forum for erfaringsudveksling for forskningsnettenes sikkerhedsansvarlige.

Deltagelse i det internationale arbejde og håndteringen af sikkerheden i den nationale e-infrastruktur, varetages i Danmark af **DKCERT**, der er en service fra DeiC.

DKCERT er medlem af den internationale organisation **FIRST** (Forum of Incident Response and Security Teams), der har ca. 370 CERT'er og CSIRT'er fra hele verden som medlemmer.

Internationalt samarbejde omkring autentifikation sker i samarbejdet **REFEDS**, hvor der arbejdes på udvikling af fælles standarder og politikker på området. Deltagerne kommer fra såvel forskningen som e-infrastrukturerne.

GÉANT samarbejdet har udviklet den internationalt anvendte service **eduGAIN**, der sikrer sammenhæng med nationale tjenester som fx danske **WAYF**.

5.6 Kompetencer

Erfaringsudveksling om udvikling af eScience og kompetencerne på området sker på europæisk niveau i **PLAN-E** (Platform for National eScience Centers in Europe).

Formålet med PLAN-E er at samle eScience centre fra Europa om at koordinere fortsat innovation i videnskabelige metoder og anvendelsen af e-infrastruktur. PLAN-E udgav i 2016 en hvidbog om eScience, potentiale og understøttelse (PLAN-E, 2016)

I det nordiske **NeIC** samarbejde arbejdes med kompetenceudvikling i aktiviteterne (NeIC Activities, 2017)

- Pool Competencies (PoCo), der har til formål at opnå bedre udnyttelse af de nordiske eScience kompetencer.
- CodeRefinery, der har til formål at gøre forskere bedre til at udvikle software og udnytte de muligheder, ny hardware og software tilbyder. En del af projektet går ud på at undervise forskere i moderne systemudviklingsmetoder og værktøjer.

6 Situation, udfordringer og anbefalinger for e-infrastrukturelementerne

For at udnytte potentialet i eScience og Open Science bør forskere i Danmark have let adgang til en sammenhængende e-infrastruktur, som er kendetegnet ved at

- den drives og udvikles koordineret, uden silotænkning og begrænsninger
- forskerne får adgang via en simpel og velkendt proces
- den matcher forskernes behov indenfor forskningsfelt og niveau

Forskernes behov for e-infrastruktur og dens sammensætning varierer ikke alene mellem de 5 videnskabelige hovedområder, men også mellem de enkelte forskergrupper/faglige felter.

De enkelte sammenhængende elementer af e-infrastrukturen, som skitseret i øko-system-figuren er nærmere beskrevet i de efterfølgende afsnit. Den

nuværende danske situation og udfordringerne også i et internationalt perspektiv beskrives. De enkelte afsnit afsluttes med anbefalingerne for det specifikke e-infrastrukturelement.

”Danske forskere har brug for adgang til de mest moderne faciliteter, så de også i fremtiden kan være blandt de førende i verden”. (Uddannelses- og Forskningsministeriet, 2015)

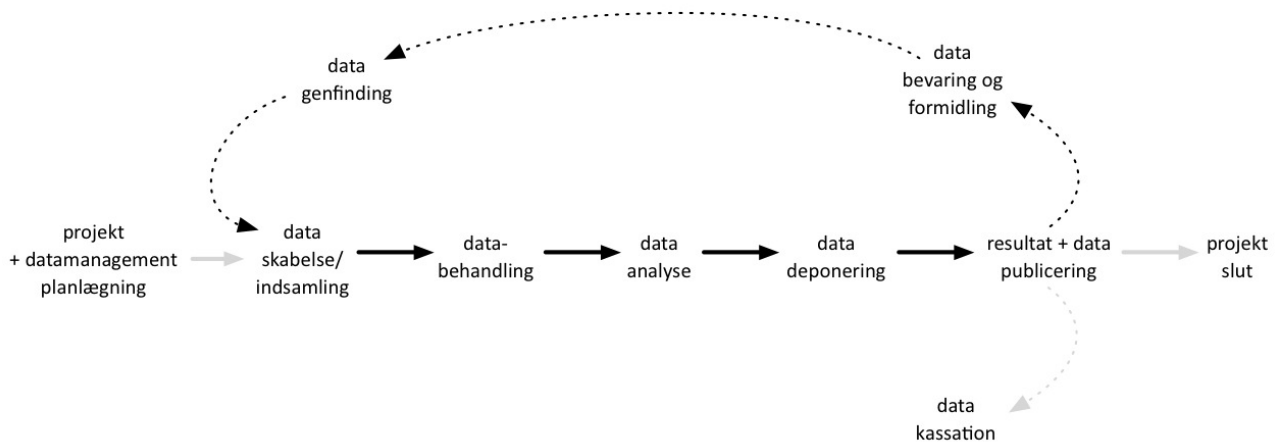
The e-IRG recommends strongly, that e-Infrastructure coordination and consolidation on the national level is embraced in full force in every European country. A strong European e-Infrastructure is dependent on strong national building blocks. (e-IRG, 2017)

6.1 Datamanagement og -lagring

Open Science, der fokuserer på åbenhed om videnskabelige resultater og deling af data på tværs af fagligheder og i nye sammenhænge, kræver fokus og ekspertise på håndteringen af data i hele datas livscyklus (se nedenstående figur) på tværs af de videnskabelige områder.

Data er forskelligartede af natur, og det vil ikke være alle data, der kan være ”åbne” i forståelsen gratis eller frit tilgængelige. Inden for Open Science taler man om FAIR-data, som dækker over at data kan findes (Findable), kan tilgås (Accessable), er kompatible og kan udveksles på tværs af systemer, evt. efter viderebearbejdning (Interoperable), og at de kan genbruges (reusable).

Forskningsdatas livscyklus



6.1.1 Status:

Universiteterne, DEFF og DeiC vedtog i 2014 en National Strategi for Forskningsdata Management (DEFF og DeiC, 2014), hvor der i fællesskab blev sat 40 mio.kr af til at igangsætte aktiviteter i perioden 2015-2018. Arbejdet er bredt forankret og det er lykkedes at tilgodese det nationale perspektiv, ligesom strategien er accepteret og har opnået legitimitet blandt samtlige involverede organisationer, fremgår det af status fra DM LedelsesCAB (DM LedelsesCAB, 2016)

Statusrapporten konkluderer blandt andet:

- De etablerede pilotprojekter er alle væsentlige for at udbrede og kvalificere datamanagement i Danmark, og de medvirker derfor alle til opfyldelsen af den nationale strategi
- Der måtte fravælges flere pilotprojektkandidater, der også ville have kunnet bidrage i meget væsentlig omfang, ligesom alle de valgte pilotprojekter ikke har fået fuld medfinansiering, hvilket peger på, at der med fordel kunne anvendes flere midler på området.
- Der er en fornuftig fremdrift i pilotprojekterne, men projekterne vil kunne lide under, at de involverede ressourcer typisk kun kan allokere tid i begrænset omfang. En øget medfinansiering fra centralt hold vil kunne modvirke dette forhold.
- Det er opnået at universiteterne og bevaringsinstitutionerne medfinansierer aktiviteterne under strategien i det planlagte forhold 1:1.
- Koordinationen af og udvekslingen mellem de forskellige aktiviteter i nationalt regi og lokale aktiviteter vil med fordel kunne styrkes for både at undgå overlappende aktiviteter og for at sikre yderligere videndeling.
- Den samlede indsats på området åbne data er ikke tilstrækkelig, og det bør drøftes, hvordan den kan øges og især målrettes understøttelsen af de generelle principper og politikker, der udspringer af det internationale samarbejde.
- Registerforskningen i Danmark er af en sådan strategisk vigtighed, at der bør være et særligt fokus på datamanagement-discipliner, der har relevans for registerforskningen. Det bør gennem

strategiperioden løbende specifikt vurderes, om der iværksættes og gennemføres tilstrækkelige initiativer på området.

- Det første års arbejde med den nationale strategi har afsløret et stort uopfyldt behov for initiativer på datamanagement-området, og at vi i Danmark har endnu et stykke vej at gå, før vi kan sammenligne os med de bedste, hvad datamanagement angår.

Det nationale initiativ har således kun i nogen grad forbedret situationen, når det kommer til lagringsinfrastruktur. Der er uddelt 5 mio. kr. til understøttelse af fire infrastrukturpilotprojekter, men ressourcemæssigt vurderes det ikke at være pilotprojekter, som vil kunne komme i mål til at blive en bæredygtig infrastruktur.

Når det handler om etablering af politik og support til management af forskningsdata, så er Danmark til gengæld bedre hjulpet af National Forum for Forskningsdata Managements tre aktiviteter: Samarbejdsaktiviteter, train-the-trainer-aktiviteter og understøttelse og deling af viden om lokale datamanagement-fora.

Men den nødvendige e-infrastruktur og de nødvendige hjælpemidler til bearbejdning af aktive data mangler, hvis de danske forskere skal deltage i udviklingen mod Open Science og FAIR data, i såvel et nationalt som internationalt samspil.

Det vil imidlertid kræve en national ledelse og finansiering og et medejerskab, som kan sikre enighed om teknologi og politik, der muliggør dansk deltagelse i den europæiske dagsorden.

Det er væsentligt at se på området i et bredt perspektiv og inddrage interessenter fra fx bevaringsinstitutionerne og registerforskningen i arbejdet.

6.2 Computing

Computing området dækker den beregningshardware, der bruges til modellering, simulering, dataanalyse og visualisering. Det spænder fra desktop-computere over mindre og/eller specialiserede lokale HPC-anlæg til nationale supercomputere og store internationale regneark som fx er til rådighed gennem ansøgninger til PRACE samarbejdet. Området inkluderer også Cloud løsninger, som kan være bygget ovenpå de nævnte regneark, primært med det formål at lette anvendelsen for nye brugergrupper eller for at øge fortrolighed og sikkerhed i databehandlingen.

Lidt større computerinstallationer tiltrækker sig ofte en del opmærksomhed, idet omkostningerne til såvel investering som drift er ganske betydelige, og levetiden sjældent overstiger 4 – 5 år. Der er derfor behov for en løbende dialog mellem forskere og driftsorganisationer, for hele tiden at sikre, at de samlede tilgængelige national computerressourcer har den rette tekniske sammensætning, så behovene kan opfyldes samtidig med at der opnås en maksimal udnyttelse af investeringen over tid.

6.2.1 Status

Supercomputing landskabet i Danmark i dag består af en række lokale anlæg, der servicere specifikke forskningsprojekter eller – områder, og i begrænset omfang er åbent for andre forskere/ forskningsområder. Finansieringen af disse anlæg sker typisk gennem forskningsrådsbevillinger, mens driften (strøm, varme, gulvplads) finansieres af universitetet.

Der er siden 2012 etableret 3 nationale computerinstallationer, finansieret af universiteterne og Det Kgl. Bibliotek, Århus med tilskud fra DeiC. Driften på de nationale anlæg finansieres af brugerne.

Det har ikke været muligt at tilfredsstille de forskellige brugergruppers behov og forventninger med det tilgængelige investeringsbudget, så der har været prioriteret på tre typer af brugere:

1. DeiC National HPC Centre, SDU, Abacus 2.0 (<https://vidensportal.DeiC.dk/Abacus>) er etableret som et generelt og traditionelt High Performance Computer center. Der er forskellige beregningsmuligheder og plads til en række forskellige typer af brugere – dog med det nuværende set-up - ikke de helt ”store brugere” - som beslaglægger en større del af maskinen i længere tid ad gangen. SDU, AU og AAU har bidraget til etablering og udgør de primære brugere af anlægget. Her er tale om en samlet investering på 45 MDKK.
2. DeiC Nationale Life Science Supercomputer, DTU, Computerome, (<https://vidensportal.DeiC.dk/Computerome>) er specielt indrettet til de behov der er ved beregninger og især datahåndtering indenfor Life Science området. Her har DTU og KU bidraget til etableringen og forskere herfra udgør også pt. de fleste brugere. Den samlede investering her er 40 MDKK.
3. DeiC Nationale Kulturarvscluster, Det Kongelige Bibliotek, Århus, (<https://vidensportal.DeiC.dk/Kulturarvscluster>), er særligt indrettet til at give adgang til de nationale samlinger af digitale kulturarvsdata på Det Kongelige Bibliotek i Århus, samtidig med at anlægget er udstyret med den type programmer, som især benyttes af forskere indenfor humaniora og samfundsvidenskab. Den samlede investering er her på 3 MDKK.

Alle tre anlæg repræsenterer en start på nationale beregningsanlæg, og der har været behov for indkøring af forskellig art, men det er vurderingen at især de to større anlæg er kommet rigtig godt i gang, og at de brugere, som anlæggene er designet til, og som har fået adgang til at benytte anlæggene, er meget tilfredse. Der kan konstateres en god anvendelsesgrad og en positiv udvikling i brugerskaren.

Kulturarvsclusteret har haft en del startvanskeligheder p.g.a. leverandørvanskeligheder, men forventes at komme rigtig i gang i løbet af foråret 2017 med flere projekter, der står klar.

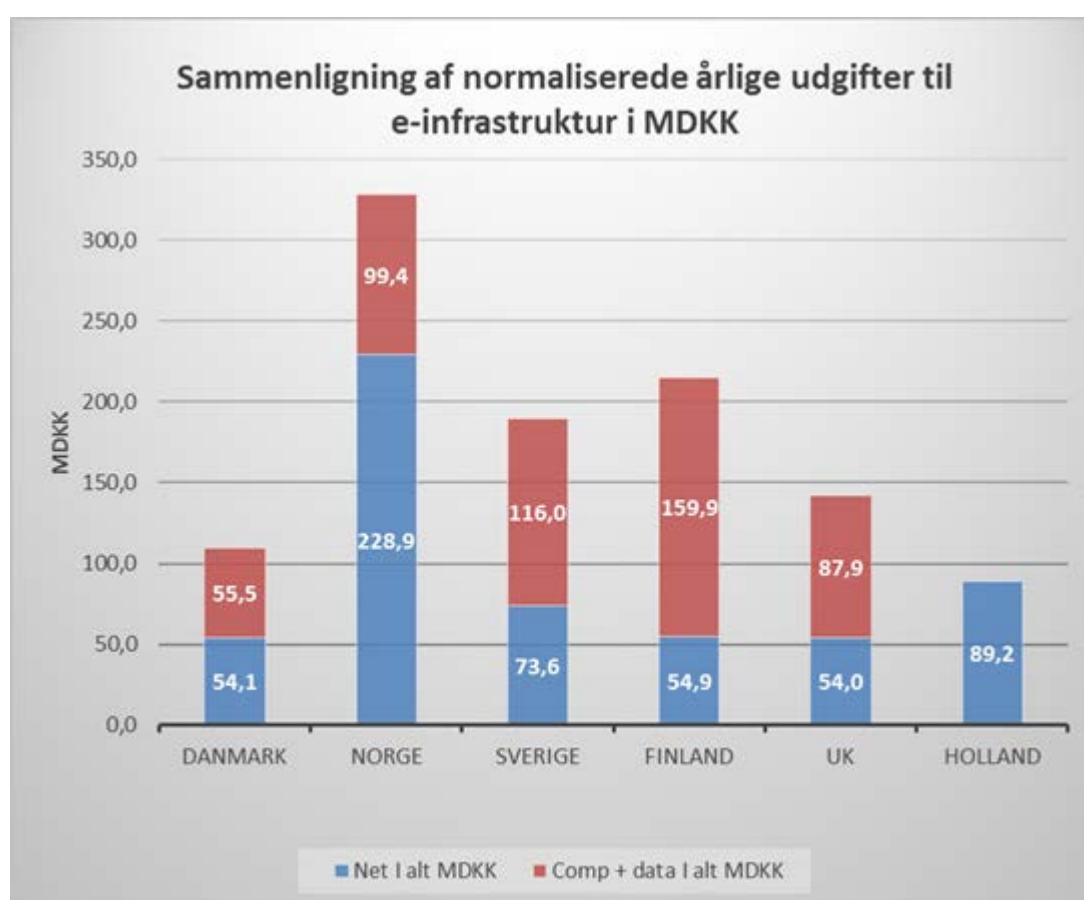
Det har ikke været muligt på tilstrækkelige vis at tilfredsstille en del af de traditionelle storbrugere af større regnearbejde i Danmark med de tre ovenstående nationale anlæg. (Evalueringsrapport, 2015). Disse storbrugere kommer primært fra naturvidenskab og teknisk videnskab. Adgangen til det danske nationale anlæg er styret af økonomi, hvor disse forskere tidligere har fået tildelt ressourcer gennem forskningsansøgninger og forskningsmerit gennem forskningsrådene til brug af på de lokale anlæg og på internationale anlæg.

Tildeling af ressourcerne på de nationale anlæg på baggrund af økonomi har blandt andet baggrund i ønsket om at udbrede eScience til nye brugergrupper. Der har været meget forskellig praksis på de enkelte universiteter mht. at give økonomiske ressourcer til brug på de nationale anlæg, som har haft betydning for udnyttelsen af anlæggene.

Adgang til meget stor regnekraft formidles gennem medlemskab af PRACE, og tildeles gennem ansøgninger. Enkelte danske forskere har også fået tildelt regnetid på fx store amerikanske regnearbejde.

I landene omkring os har der gennem flere årtier været en udbredt tradition for opbygning og vedligeholdelse af nationale computing og storage anlæg. Flere steder kan man tale om at lokale computing centre har vokset sig til selvstændige enheder.

De nationale forhold hos vores naboer er tydeligt illustreret ved de årlige omkostninger, der her benyttes på beregnings- og data infrastrukturen, som er gjort op i bilag 2. Der er individuelle forhold, der gør sig gældende for det enkelte land, men fælles for alle er, at niveauet af nationale aktiviteter er mindst en faktor 2 større end det danske, som det fremgår af nedenstående figur fra bilag 2, hvor de samlede udgifter i MDKK, normaliseret med hensyn til befolkningstal er vist henholdsvis netværk og computing og storage. Hvis Danmark skal leve op til ambitionerne omkring eScience og open Science kræver det en betydelig indsats over en længere periode.



6.3 Software og applikationer

Rækken af forskningsunderstøttende software og applikationer er lang og spænder over såvel samarbejdsværktøjer som avancerede beregningsprogrammer til databehandling.

Nogle kræver specialiseret viden om og indsigt i det videnskabelige område det retter sig imod, mens andre er mere mainstream, og kan anvendes af flere eller alle forskningsgrupper.

En naturlig del af den samlede forskningsinfrastruktur er tjenester til datadeling og samarbejde. Det kan være platforme til distribution og udveksling af medieindhold, platforme til understøttelse af arbejdsprocesser og infrastruktur til videokonferencer og generelle værktøjer til transmission og deling af data.

For software og applikationer, der er rettede mod specielle behov og videnskaber, kan det i mange tilfælde stadig give mening nationalt set, at sikre en platform og fora til udveksling af erfaringer og best practise på området.

NeIC (Nordic e-Infrastructure Collaboration) arbejder i initiativet CodeRefinery (NeIC) med at etablere en software udviklings-infrastruktur med teknisk ekspertise og træningsaktiviteter for at imødekomme et voksende behov for personer med kendskab til programmering til supercomputing og datahåndtering.

6.4 Forskningsnet

Forskningsnet er et højkapacitetsnetværk, specielt udviklet og drevet til fordel for forskning. I langt de fleste lande i verden sikrer forskningsnet, at forskningsprojekter og –samarbejde kan udveksle ideer, data og resultater. Forskningsnet forbinder forskningsfaciliteter og sikrer datastrømmen fra dem, også når de står i ørkenen i Mexico, på Svalbard, - eller for den sags skyld i det danske vindmølletestcenter i Høvsøre i Vestjylland.

Stort set alle civiliserede lande i verden har et forskningsnet, som godt nok er forbundet til internettet, men er en separat infrastruktur og en separat organisation, som har det formål at lave netinfrastruktur, der er optimeret efter andre parametre end de kommercielle net. På forskningsnettene opererer man ikke med fyldte linjer og overbooking, men har høj kvalitet og en ledig bufferkapacitet, således at ingen forskere længere behøver at sende kasser med harddiske, og således at realtidsanvendelser kan ske umiddelbart uden forudgående booking og tilkøb. I Europa er det f.eks. kun Letland og Bosnien-Hercegovina, der ikke har et separat forskningsnet.

Forskningsnettene leverer direkte punkt-til-punkt forbindelser på tværs af nationale grænser til geografisk distribuerede forskningsprojekter.

Forskningsnettene leverer typisk også testbeds til forskning i fremtidens internetteknologier.

6.4.1 Status

Det danske forskningsnet vil være "State-of-the-art" på området efter en igangværende opgradering, der fuldføres medio 2017. Finansieringen af denne og andre opgraderinger sker gennem den løbende brugerbetaling.

6.5 Sikkerhed og autentifikation

Sikkerhed og brugerautentifikation dækker over de nødvendige politikker og procedurer, der sikrer fortrolighed, integritet og tilgængelighed for data under transport og ved behandling og lagring i den øvrige e-infrastruktur.

6.5.1 Status

Truslerne, dels fra den stadigt stigende, grænseoverskridende cyberkriminalitet (EUROPOL, 2016), (Cybersikkerhed, 2017), dels fra privat og statsstøttet spionage mod dansk forskning, (Cybersikkerhed, 2016) gør det nødvendigt at skærpe indsatsen for at sikre e-infrastrukturen og de data, der transporteres, behandles og lagres mod tab af fortrolighed, dataintegritet eller tilgængelighed.

Universiteterne er generelt meget bevidste om sikkerhedsudfordringerne, og udveksler viden og erfaringer på området i CISO-forum, der er en arbejdsgruppe under Danske Universiteters CIO-forum.

DKCERT, der er en service fra DeiC deltager i dette samarbejde, og varetager den løbende overvågning af sikkerheden i e-infrastrukturen.

Det forventes at krav i blandt andet persondataforordningen vil betyde at flere af elementerne i e-infrastrukturen skal kunne sikkerhedscertificeres/-mærkes for at kunne indgå i forskningssamarbejdet, og dermed også vil kræve flere ressourcer til sikkerheds(sam-)arbejdet.

Let og uhindret adgang på den ene side og sikker identifikation og autorisation på den anden side er en forudsætning for at bruge forskningsinfrastrukturene effektivt.

For at få adgang til tjenester internt på sit universitet, bruger en dansk forsker væsentligst den online-identitet (typisk brugerid og password) universitetet forsyner ham med. Med login-tjenesten WAYF (gen)bruges den samme identitet til en lang række andre danske tjenester og med det internationale samarbejde eduGAIN (mellem alle landes modstykker til WAYF), udstrækkes dette yderligere til en stor mængde internationale tjenester.

Det kan handle om alt fra søgning i tidsskriftsdatabaser, biblioteksarkiver og laboratoriesystemer, – herunder tilgangen til CERNs ressourcer - til konferencetilmeldinger.

Med WAYF kan man - ud over brugeridentiteten - overføre yderligere oplysninger (attributter) fra universitet til tjeneste. Dette findes ikke i automatiseret form på tværs af platforme med nogen af de andre udbredte metoder til identifikation og adgangsautorisation.

“The expanding network of Higher Education and Research facilities through inter-federation, whilst extremely valuable for collaboration and online security at large, exposes an inviting new vector of attack for malicious actors. A single compromised account may provide an entry point to this global network of resources linking thousands of organisations. How can we coordinate a response spanning countries and continents? How can we build trust between organisations in our communities? We propose a Security Incident Response Trust Framework to do just this.” (Hannah Short (CERN), 2016)

6.6 Kompetencer

En absolut nøgleaktivitet for at sikre gevinsterne af Open Science og eScience, i de kommende år er, at sikre udbygning af kompetencer på disse forholdsvis nye områder.

Der er et stort behov for at universiteterne sætter udviklingen af de nye fagligheder i system og skaber nye uddannelser, som på alle niveauer kan tilfredsstille den voksende efterspørgsel overalt i samfundet.

6.6.1 Status

eScience-kompetencerne i Danmark befinder sig primært i faglige nicheområder, både i lokale forskningsprojekter på universiteterne, men også nationalt og internationalt indenfor større forskningskonsortier.

Flere af universiteterne har på det seneste etableret eSciencecentre og –komiteer, der arbejder på tværs af de videnskabelige områder, der skal sikre synergi og vidensoverførsel mellem projekter og forskningsområder.

For at understøtte datamanagement og forudsætningerne for Open Science har flere af universiteterne også etableret datamanagement-rådgivningsfunktioner, der skal hjælpe forskerne på vej.

DeiC har, blandt andet i samarbejde med DigHumLab, arbejdet på at sikre vidensdeling på tværs af universiteter og videnskabelige områder, gennem afholdelse af møder, workshops og konferencer.

Kompetencer til at anvende e-infrastrukturen i forskningen er en grundlæggende forudsætning for at opnå forskningsresultater. eScience og Open Science har behov for data-forskere og eScience-ingeniører for at bringe videnskaben videre. Universiteterne udbyder uddannelse på området, hvor behovet vil vokse væsentligt over de kommende år.

Integration af eScience og Open Science i alle uddannelser og kurser på PhD-niveau og til efteruddannelse er væsentligt for at kunne udnytte mulighederne for forskningen.

Også industrien har hårdt brug for kompetencer til at omsætte deres data-ressourcer til vækst – og disse kompetencer kan kun universiteterne levere.

7 Konklusion og samlede anbefalinger

Danmark ønsker at udnytte de potentialer for nye forskningsresultater og samfundsudvikling, som ovenstående eksempler fra de videnskabelige områder viser ligger i at anvende eScience og understøtte Open Science.

En fortsat satsning på områderne kræver en understøttende e-infrastruktur, som er på højde og i samspil med den internationale e-infrastruktur. Det kræver en klar og sammenhængende vision og ambition for elementerne i e-infrastrukturens økosystem, en langsigtet finansieringsplan og en samlet organisering, som kan understøtte dette i et samarbejde med de øvrige aktører på området.

Den europæiske e-Infrastructure Reflection Group (e-IRG) anbefaler, at de enkelte lande evaluerer finansiering og rammer for deres nationale e-infrastruktur og deltager i udviklingen af det europæiske e-infrastrukturlandskab til fordel for såvel den nationale forskning som det internationale forskningssamarbejde

E-infrastrukturen har et lokalt, nationalt og internationalt niveau, der skal koordineres for at sikre forskeren let adgang til en sammenhængende e-infrastruktur, der understøtter den enkelte forskningsopgave. Det kræver en aktiv indsats og involvering i de internationale e-infrastruktursamarbejder, at sikre den vidensudveksling, der giver den danske e-infrastruktur et internationalt niveau, og sikrer de danske forskere adgang til de internationale infrastrukturer.

Danmark er rigtig godt med på flere af e-infrastrukturområderne, mens det vil kræve en ekstra indsats, at nå op på et niveau, der kan understøtte forskerne i den internationale konkurrence.

Nedenfor angives DeiCs samlede anbefalinger for en e-infrastruktur, der kan understøtte en ønsket udvikling af eScience og Open Science

Anbefaling 1: Fortsat understøttelse af aktiviteter, der sikrer Open Science og eScience kompetencer

En absolut nøgleaktivitet, for at sikre udnyttelsen af de muligheder som Open Science og eScience giver, er at sikre udbygning af kompetencer på disse forholdsvis nye områder. Der er et stort behov for at sætte udviklingen af de nye fagligheder i system og skabe nye uddannelser, som på alle niveauer kan tilfredsstille den voksende efterspørgsel overalt i samfundet.

Flere af universiteterne er godt på vej med uddannelser og har etableret eSciencecentre og -komiteer, og datamanagement-rådgivningsfunktioner, der skal fremme udviklingen på områderne. Der bør også fortsat sættes på uddannelse af f.eks. eScientists, data scientists/data stewards og videreuddannelse på området af forskere, primært på PhD niveau.

Det er vurderet, at der i Europa er behov for 500.000 data-stewards – en per 20 forsker, hvis potentialet fra Open Science skal udnyttes fuldt ud, (Mons, 2016).

Anbefaling 1.1: Kompetenceudviklingen skal understøttes af en national koordinering af initiativer og erfaringsudveksling i et samarbejde med de lokale initiativer.

Anbefaling 1.2: Den nationale kompetenceudvikling bør koordineres med tilsvarende internationale aktiviteter gennem aktiv deltagelse i relevante organisationer og fora.

Anbefaling 1.3.: Der skal afsættes betydelige midler både nationalt og ikke mindst lokalt.

Anbefaling 2. Investering i nationale supercomputing anlæg

En satsning på udviklingen af eScience og Open Science vil give et øget behov for beregningsressourcer og dermed kræve en fortsat udvikling af nationale computinganlæg, der er tilgængelige for alle forskningsgrupper. Set i et internationalt perspektiv ligger investeringerne i computinganlæg i Danmark lavt.

I dag er der 2 nationale supercomputere, som tilgodeser både et bredt generelt behov for alle videnskaber, og et specifikt behov indenfor LifeScience. Derudover er der et nationalt cluster specielt rettet mod forskning i den digitale kulturarv.

Det har med den nuværende kapacitet vist sig vanskeligt at tilfredsstille alle brugertyper.

(Evalueringsrapport, 2015) En fortsat satsning på udbredelse af eScience vil samtidig betyde, at der i fremtiden vil komme flere brugere med både små og store beregningsbehov, hvilket vil kræve let adgang til betydelige nationale ressourcer.

Anbefaling 2.1: Der bør afsættes beløb i en størrelsesorden, som matcher landene omkring os, som skal dække udvikling, investeringer og drift af de nuværende nationale anlæg.

Anbefaling 2.2: Der bør nedsættes et hurtigt-arbejdende udvalg, der med udgangspunkt i en vurdering af forskernes fremtidige behov fastsætter et hensigtsmæssigt niveau og udformning af nationale supercomputing anlæg, og sikrer koordinering i forhold til det lokale og internationale niveau.

Anbefaling 2.3: Der bør nedsættes et udvalg med repræsentanter fra forskningen og forskningsrådene, der i fællesskab definerer en adgangsmodel til nationale computinganlæg, der sikrer at såvel den erfarne forsker med merit på området, som den (eScience-)uerfarne forsker få let og ubesværet adgang til de nødvendige beregningsressourcer.

Anbefaling 2.4: Fortsat finansiering af aktiv deltagelse i internationalt samarbejde på computing-infrastrukturområdet skal sikre, at danske forskere har adgang til international regnekraft, der ikke kan dækkes af lokale og nationale anlæg, som f.eks. i det nuværende PRACE samarbejde.

Anbefaling 3: Øget investering i datamanagement af forskningsdata

Universiteterne, DEFF og DeiC vedtog i 2014 en National Strategi for Forskningsdata Management, hvor der i fællesskab blev afsat 40 mio. kr. til at igangsætte aktiviteter på området i perioden 2015 – 2018.

Initiativet har på nuværende tidspunkt opnået gode resultater, men der skal fortsat investeres i området, for at sikre den nødvendige lagringsinfrastruktur og hjælpemidler til bearbejdning af aktive data i hele data livscyklus. Den hastige udvikling på Open/FAIR data og Open Science var ikke forudset i strategien.

Anbefaling 3.1: Hvis Danmark vil være med i den internationale udvikling på FAIR data og Open Science området kræver det en national ledelse, koordinering og finansiering samt et medejerskab fra interessenterne, som kan sikre enighed om teknologi og politik.

Anbefaling 3.2: Der skal investeres i lagringsinfrastruktur og programmel, der lever op til de internationale standarder, som sikrer FAIR data i nationalt og internationalt samspil og som understøtter forskernes forskelligartede behov afhængig af datatypologi.

Anbefaling 4: Sikre forskningsnettet som en fortsat "state-of-the-art" forskningsinfrastruktur

En større udbredelse af eScience og understøttelse af Open Science aktiviteterne vil kræve en sikker, stabil og kraftig netforbindelse, som i dag ikke kan garanteres af de kommercielle net. Forskningsnettet er som en

grundlæggende forskningsinfrastruktur udviklet specielt til at opfylde forskningens krav. Forskningsnettet giver forskerne en sikker, stabil og kraftig forbindelse til overførsel af data, adgang til supercomputere og øvrig forskningsinfrastruktur såvel nationalt som internationalt.

Den nuværende teknologi på forskningsnettet er "state-of-the-art" og på niveau med forskningsnet internationalt. Opgradering af forskningsnettet finansieres i dag via en løbende opsparing fra den årlige betaling fra brugerne.

Anbefaling 4.1: Det bør også i fremtiden sikres at forskningsnettet som en grundlæggende forskningsinfrastruktur kan følge den internationale udvikling og være state-of-the-art.

Anbefaling 4.2: Forskningsnettet bør også i fremtiden finansieres direkte af kunderne, og det bør løbende sikres, at der er en hensigtsmæssig balance mellem driftsomkostninger og aktiviteter

Anbefaling 4.3: Forskernes mulighed for internationalt samarbejde og datatransport, bør også fremover sikres gennem aktivt samarbejde med internationale forskningsnet, som NORDUnet og GÉANT.

Anbefaling 5: Øget satsning på sikkerhed i forskningsinfrastrukturen

Sikkerhed og autentifikation dækker over de nødvendige politikker og procedurer, der sikrer fortrolighed, integritet og tilgængelighed for data under transport og ved behandling og lagring i den øvrige e-infrastruktur.

I erkendelse af den øgede trussel fra cyberkriminalitet og cyberspionage, sammen med et større ønske om datadeling og anvendelse af e-infrastrukturen i forskningen i sammenhæng med udbredelse af eScience og Open Science, øger vigtigheden af til stadighed at have fokus på sikkerheden. Det anbefales derfor, at

Anbefaling 5.1 En national koordinering af sikkerheden i e-infrastrukturen bør understøttes.

it-kriminalitet rammer alle, men samarbejde, erfaringsudveksling og advarsler om mulige trusler kan afbøde eller helt forhindre skaderne efter angreb. Der er behov for et stærkt lokalt sikkerhedsarbejde, en national koordinering og et internationalt samarbejde på området.

Anbefaling 5.2 Der bør sættes på at have en stærk og koordineret autentifikationsinfrastruktur, der understøtter sikker adgang til data, også når disse er personhenførbare, og samtidig faciliterer forskernes adgang til samarbejde, også internationalt.

Anbefaling 6: Aktiv deltagelse i internationalt samarbejde

Mange danske forskningsprojekter sker i et internationalt samarbejde. Med udbredelsen af Open Science vil danske forskere i endnu højere grad have behov for at kunne tilgå internationale ressourcer og samarbejde over grænserne.

Anbefaling 6.1 Det bør sikres at e-infrastrukturen også i fremtiden understøtter de danske forskeres mulighed for at deltage i og medvirke til at præge udviklingen af internationale forskningssamarbejde og at de ubesværet kan tilgå internationale ressourcer.

Anbefaling 6.2 Derfor bør aktivt samarbejde med og medlemskab af internationale e-infrastrukturer og interesseorganisationer også fremover sikres finansiering.

--

e-IRG (European Infrastructure Reflection Group) har netop udgivet deres Roadmap 2016 (e-IRG, 2017), der blandt andet indeholder anbefalinger til de nationale regeringer og forskningsråd. Disse anbefalinger er gengivet i bilag 3.

7.1 Økonomioversigt

Anbefalingerne ovenfor indeholder forslag om enten at fastholde eller øge indsatsen for de enkelte aktiviteter. Der er for hvert område skitseret rammer for en samlet national indsats og det internationale engagement. Nedenfor præsenteres de aktuelle økonomiske rammer for de enkelte aktiviteter. Afsnittet sluttet af med en sammenligning med de tilsvarende aktiviteter i nogle af de lande vi normalt sammenligner os med og en præsentation af mulige scenarier for den fremadrettede indsats.

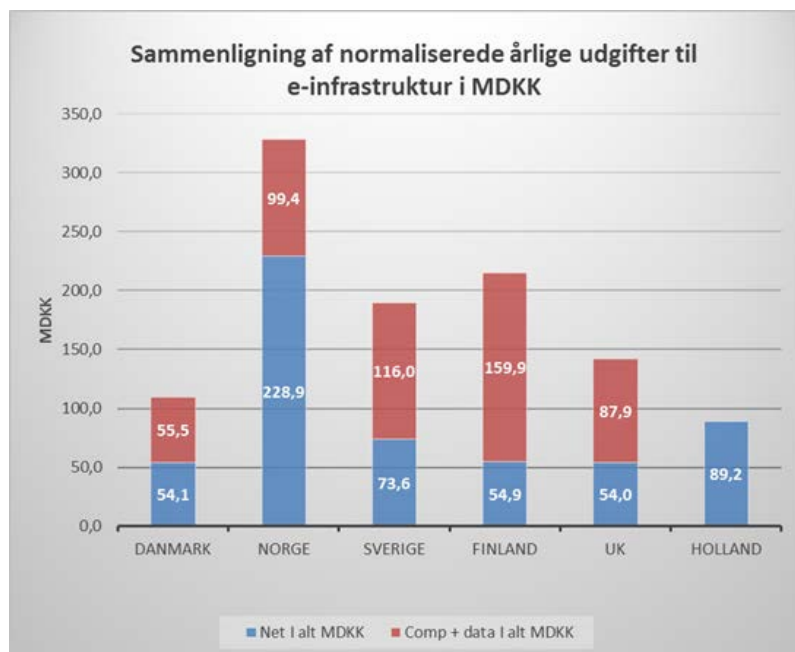
7.1.1 Økonomisk ramme 2013-2016

Aktivitet		Gennemsnitlig økonomisk ramme per år.	
		2013 - 2016	MDKK
Anbefaling 1	Sikre eScience kompetencer		3
Anbefaling 2	National supercomputing		37
Anbefaling 3	Forsknings data management		9
Anbefaling 4	Forskningsnettet		48
Anbefaling 5	Sikker e-infrastruktur		6
Anbefaling 6	Aktiv i internationale fora		7
	Samlet		110

Den nuværende ramme, der har været til rådighed i perioden 2013 til 2016 består af både nationale og lokale midler til såvel drift som investeringer. Heri indgår 50 MDKK til eScience fra Infrastrukturpuljen, som DeiC har forvaltet, fordelt over 4 år, men som ikke længere fra 2017 er til rådighed, DeiCs årlige bevilling på Finansloven, DEFF midler til forskningsdata management sammen med de universitetsmidler, der er benyttet til medfinansiering og til driftsudgifter ifbm. de nationale computeranlæg og til medfinansiering af DEFF aktiviteterne.

Af bilag 2 fremgår de økonomiske aktiviteter sammenlignet med tal fra lande vi normalt sammenligner os med, fordelt på 1) Netværk og 2) computing og storage som overordnede kategorier. Tallene er normaliserede med hensyn til befolkningstallet.

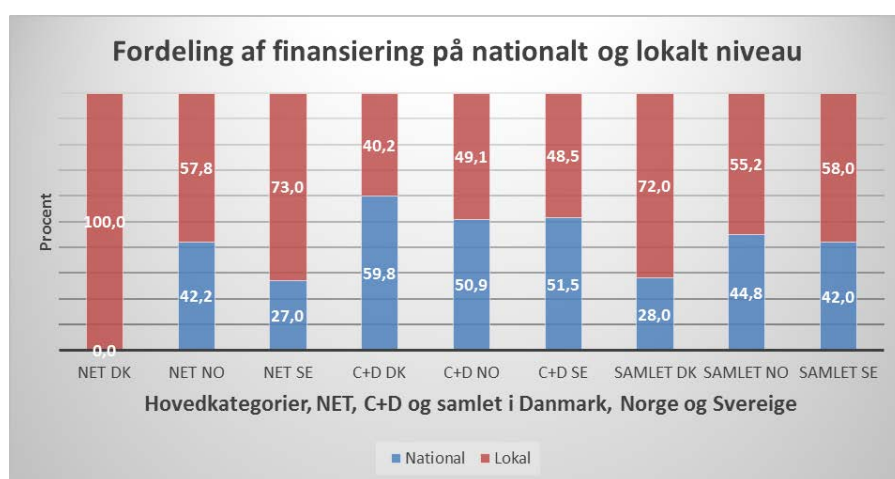
Dette gengives grafisk nedenfor (blå søjle for netværk og rød søjle for computing og storage).



Det fremgår af sammenligningen, at udgiftsniveauet til netværk i Danmark følger niveauet i flere lande, mens udgifter til computing og storage er betydelig lavere i Danmark.

7.1.2 Økonomisk fordeling lokalt og nationalt niveau

Af bilag 2 fremgår også fordelingen af finansiering mellem nationalt og lokalt niveau. Nedenfor illustreret ved den procentvise fordeling af de to kategorier for Danmark, Norge og Sverige. Blå søjle er udgifter på et nationalt budget og rød søjle er udgifter afholdt lokalt. Her viser sammenligningen at det nationale niveau i Norge og Sverige afholder en noget større andel af de samlede udgifter.



Open Science og den europæiske fokus på forskningsdata management giver behov for en øget satsning for alle landene. I Danmark er der herudover også behov for en øget indsats på computingområdet, hvis danske forskere skal have adgang til en e-infrastruktur på niveau med landene omkring os.

De aktiviteter der sammenlignes mellem landene, omfatter udelukkende fælles nationale aktiviteter, og tallene udgør summen af økonomiske bidrag fra universiteter/brugere og fra nationale kasser. Inden for alle seks kategorier foregår der lokale, og dermed i høj grad egenfinansierede, aktiviteter på de enkelte institutioner. Fordelingen og opgaveplacering, og dermed finansiering, varierer fra land til land.

7.1.3 Scenarier

Det fremtidige økonomiske behov afhænger af hvilken vision, der opstilles for udviklingen af eScience og Open Science i Danmark, og dermed hvilken ambition, der fastlægges for e-infrastrukturen.

I det efterfølgende skitseres to mulige scenarier.

- a) Danmark på et niveau som nummer 2 i Norden.
- b) Danmark satser intenst på at blive frontløber, og arbejde frem mod et meget højt internationalt niveau.

Scenarie a): Danmark er på niveau med de andre nordiske lande, i et gennemsnit af de seks kategorier. På dele af datamanagement området har vi mulighed for at være med i front og på netværksområdet er vi allerede på et godt niveau, men der skal tilføres ressourcer til de øvrige områder, som indikeret nedenfor.

Scenarie b): Danmark bestræber sig på at være med i front på flere områder, og grundlæggende har mulighed for at deltage i og dermed præge den internationale udvikling af områderne. Dette vil kræve en koordineret og langsigtet indsats med fokus på udvikling og kompetenceopbygning.

Aktivitet		Økonomisk ramme MDKK/år		
		DeIC 2013-2016	Scenarie a) <i>2'er i norden</i>	Scenarie b) <i>Frontløber</i>
Anbefaling 1	Sikre eScience kompetencer	3	10	20
Anbefaling 2	National supercomputing	37	65	100
Anbefaling 3	Forsknings data management	9	20	30
Anbefaling 4	Forskningsnettet	48	48	60
Anbefaling 5	Sikker e-infrastruktur	6	10	15
Anbefaling 6	Aktiv i internationale fora	7	20	25
	Samlet	110	173	250

Fordelingen af de økonomiske rammer for scenarie a) er en skitse, der tager udgangspunkt i anbefalingerne, hvoraf det fremgår at:

- Økonomien omkring forskningsnettet er allerede på et passende niveau i nordisk regi.
- Der er forslag om at øge indsatsen på alle andre områder.

For scenarie b) skal indsatserne øges yderligere indenfor alle områder i forhold til scenarie a), og her er der især behov for at øge de internationalt orienterede aktiviteter. Her skal der findes en måde at organisere

dette på, så alle interessenter kan indgå i et samarbejde og bidrage til og nyde af denne satsning. Det vil på sigt give mulighed dansk deltagelse i internationale udviklingsprojekter såvel med faglige som organisatoriske ressourcer.

De fremtidige økonomiske rammer for områderne bør fremgå af den vision der opstilles for e-infrastrukturen.

7.2 Afsluttende bemærkninger

Ovenfor er givet status for den danske eScience og e-infrastruktur. På basis af dette og ud fra forventninger om en fremtidig satsning på Open Science er beskrevet en række konkrete anbefalinger.

Udviklingen af eScience og Open Science området går meget hurtigt. I øjeblikket arbejdes i Uddannelses- og Forskningsministeriet på at definere hvad Danmark skal i forhold til Open Science og FAIR data og hvad omkostningerne til FAIR Open Science vil være. Konklusionen af dette arbejde vil kunne vise, at den økonomiske ramme skitseret ovenfor vil være utilstrækkelig. Dette vil bl.a. afhænge af hvad der tænkes lavet på lokalt niveau og hvad der skal dækkes nationalt.

8 Referenceliste

- Andreasen, C. (7. November 2016). *National Supercomputing Day - Perspectives of High Performance Computing*. Hentet fra vidensportal.deic.dk:
<https://vidensportal.deic.dk/sites/default/files/uploads/Casper%20Andreasen.pdf>
- Blume, P. (2016). *Den nye persondataret. Persondataforordningen*. DJØF Forlag.
- CSC, 2. (2015). *The Scientific Case and User Requirements for High-Performance and Data-Intensive Computing in Finland 2017-2021*. Hentet fra CSC.fi: <https://www.csc.fi/-/sciencecase2017-2021>
- Cybersikkerhed, C. f. (November 2016). *Trusselsvurdering: Hackere fra udlandet truer dansk offentlig forskning*. Hentet fra fe-ddis.dk/cfcs: <https://fe-ddis.dk/cfcs/CFCSDocuments/Trusselsvurdering%20-%20%20dansk%20forskning%20-%20december%202016.pdf>
- Cybersikkerhed, C. f. (februar 2017). *Trusselsvurdering: Cybertruslen mod Danmark*. Hentet fra fe-ddis.dk/cfcs: <https://fe-ddis.dk/cfcs/CFCSDocuments/Cybertruslen%20mod%20Danmark.pdf>
- DEFF og DeIC. (December 2014). *National strategi for forskningsdatamanagement*. Hentet fra DeIC: https://www.deic.dk/datamanagement/national_strategi_med_bilag
- DeIC. (2016). *National Supercomputing Day - Perspectives on High Performance Computing*. Hentet fra vidensportal.deic.dk: <https://vidensportal.deic.dk/national-supercomputing-day-nov-7-2016>
- DeIC. (2017). *Dedikeret linje mellem supercomputere letter arbejdet for forskere*. Hentet fra deic.dk: https://www.deic.dk/news-2017-01-20-dedikeret_linje
- DM LedelsesCAB. (November 2016). https://www.deic.dk/datamanagement/DM_LCAB. Hentet fra DeIC.dk: <https://www.deic.dk/sites/default/files/uploads/PDF/Statusrapport%20til%20DeiC%20og%20DEFF%20version%201.1.pdf>
- e-IRG. (11. Januar 2017). *e-IRG to publish Roadmap 2016 as a next step towards the e-Infrastructure Commons*. Hentet fra e-Infrastructure Reflection Group: <http://e-irg.eu/documents/10920/12353/Roadmap+2016.pdf>
- EU. (19. April 2016). *European Cloud Initiative to give Europe a global lead in the data-driven economy*. Hentet fra EU: http://europa.eu/rapid/press-release_IP-16-1408_en.htm
- EUR LEX. (April 2016). *European Cloud Initiative - Building a competitive data and knowledge economy in Europe*. Hentet fra <http://eur-lex.europa.eu>: <http://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX:52016DC0178>
- EUROPOL. (2016). *The Internet Organised Crime Threat Assessment (IOCTA)*. Hentet fra www.europol.europa.eu: <https://www.europol.europa.eu/activities-services/main-reports/internet-organised-crime-threat-assessment-iocta-2016>

- Evalueringsrapport. (December 2015). *DeiC evalueringsrapport 2015*. Hentet fra deic.dk:
<https://www.deic.dk/node/69>
- Evalueringsudvalg. (September 2015). *Evalueringsrapport DeiC*. Hentet fra DeiC.dk:
<https://www.deic.dk/node/69>
- Hannah Short (CERN). (Juni 2016). *Promoting Security and Trust within our Inter-Federated Network*. Hentet fra <https://tnc16.geant.org>: <https://tnc16.geant.org/core/presentation/837>
- HLEG EOSC. (11. Oktober 2016). *Realising the European Open Science Cloud*. Hentet fra <http://ec.europa.eu>:
http://ec.europa.eu/research/openscience/pdf/realising_the_european_open_science_cloud_2016.pdf#view=fit&pagemode=none
- Holst, M. K. (2014). *eScience åbner nye døre for arkæologien*. Hentet fra Vidensportal.deic.dk:
<https://vidensportal.deic.dk/node/12>
- Jacob Davidsen, A. U. (7. November 2016). *BIGVIDEO in the humanities; præsentation på National Supercomputing Day - Perspectives of High Performance Computing*. Hentet fra vidensportal.deic.dk:
<https://vidensportal.deic.dk/sites/default/files/uploads/Jacob%20Davidsen.pdf>
- KOR. (Maj 2014). *Det Koordinerende Organ for Registerforskning*. Hentet fra [Registerforskning.dk](http://www.registerforskning.dk):
<http://www.registerforskning.dk/kor/>
- Ministerråd, E. (27. Maj 2016). *Rådets konklusioner om overgangen til et system for åben videnskab*. Hentet fra <http://data.consilium.europa.eu/>: <http://data.consilium.europa.eu/doc/document/ST-9526-2016-INIT/da/pdf>
- Mons, B. (13.. April 2016). We need 500.000 respected data stewards to operate the European Open Science Cloud. (P. Magazine, Interviewer) YouTube. Amsterdam.
- NeIC Activities. (2017). *Activities*. Hentet fra wiki.neic.no: <https://wiki.neic.no/wiki/Activities>
- NeIC, C. (u.d.). *CodeRefinery*. Hentet fra wiki.neic.no: <https://wiki.neic.no/wiki/CodeRefinery>
- PLAN-E. (2016). *eScience: Insight Out (PDF)*. Hentet fra <https://plan-europe.eu>:
<https://planeurope.files.wordpress.com/2015/03/plan-e-white-paper-version-1-00.pdf>
- Research Councils UK. (2014). *E-Infrastructure Roadmap*. Hentet fra Research Councils UK:
<http://www.rcuk.ac.uk/research/xrcprogrammes/otherprogs/einfrastructure/>
- rfii. (2016). *Rfii Recommendations 2016: Performance through Diversity*. Hentet fra Rat für Informations Infrastrukturen: <http://www.rfii.de/en/category/documents/>
- Schurgers, G. (7.. November 2016). *National Supercomputing Day - Perspectives of High Performance Computing*. Hentet fra vidensportal.deic.dk:
<https://vidensportal.deic.dk/sites/default/files/uploads/Guy%20Schurgers.pdf>

- Simon Rasmussen, A. P. (7. November 2016). *National Supercomputing Day - Perspectives of High Performance Computing*. Hentet fra Vidensportal.deic.dk:
<https://vidensportal.deic.dk/sites/default/files/uploads/Simon%20Rasmussen.pdf>
- SNIC. (2017). *SNIC 2.0 - SNIC after 2017*. SNIC.
- SUNET. (2015). *Verksamhetsberättelse Sunet 2015*. Hentet fra sunet.se: https://www.sunet.se/wp-content/uploads/2017/01/Verksamhetsberättelse-Sunet-2015_9.pdf
- Supercomputing i Danmark Podcast, S. f. (Januar 2017). *Podcast "Hvad er supercomputing i Danmark"*. Hentet fra <https://vidensportal.deic.dk/node/316>
- Tryggve, N. (u.d.). *Tryggve*. Hentet fra wiki.neic.no: <https://wiki.neic.no/wiki/Tryggve>
- Uddannelses- og Forskningsministeriet, R. 2. (December 2015). *Dansk Roadmap for Forskningsinfrastruktur 2015*. Hentet fra Uddannelses- og Forskningsministeriet:
<http://ufm.dk/publikationer/2015/filer/dansk-roadmap-for-forskningsinfrastruktur-2015.pdf>
- UFM. (7. December 2016). *Forsknings- og innovationspolitisk redegørelse*. Hentet fra Uddannelses- og Forskningsministeriet UFM.dk: <http://ufm.dk/publikationer/2016/filer/forsknings-og-innovationspolitisk-redegorelse-2016.pdf>
- UFM. (1. Februar 2017). *Forsk2025 - første samlede udkast*. Hentet fra Uddannelses- og Forskningsministeriet: <http://ufm.dk/forskning-og-innovation/indsatsomrader/forsk2025/forste-udkast-til-forsk2025-katalog/forste-udkast-til-forsk2025-katalog-01-02-2017.pdf>
- UK . (4. januar 2012). *e-infrastructure strategy: roadmap for developing advanced computing, data and networks*. Hentet fra GOV.UK - Department for Business, Innovation and Skills:
<https://www.gov.uk/government/publications/e-infrastructure-strategy-roadmap-for-development-of-advanced-computing-data-and-networks>
- Ulla Tørnes, F. 2. (29. August 2016). *Lund and Copenhagen - stronger together*. Hentet fra Uddannelses- og Forskningsministeriets hjemmeside UFM.dk: <http://ufm.dk/en/the-minister-and-the-ministry/the-minister/speeches/2016/lund-and-copenhagen-stronger-together>
- UNINETT. (2. Februar 2017). *Arbeidsgruppe for IKT-strategi og helhetlige løsninger - IKT-strategi for forskning*. Hentet fra uninett.no: <https://www.uninett.no/sites/default/files/ikt-strategi-uh-forskning-31012017.pdf>
- Vetenskapsrådet, S. 2. (2014). *Swedish science cases for e-infrastructure*. Hentet fra Vetenskapsrådet, vr.se: <https://publikationer.vr.se/en/product/swedish-science-cases-for-e-infrastructure-2/>

Bilag 1: Videnskabelige temaer

Dette bilag indeholder citater og henvisninger til yderligere information om de videnskabelige temaer i afsnit 3.

Energi, klima og miljø

Dansk Roadmap for Forskningsinfrastrukturer 2015: Fagområdet "Energi, Klima og Miljø" dækker et bredt fagligt spektrum, som kobler til flere centrale samfundsudfordringer med stort dansk fokus. Herunder hvordan vi bedst minimerer og håndterer effekterne af klimaforandringer, forvalter vores naturressourcer på ansvarlig vis og understøtter en grøn omstilling af vores samfund og erhverv. Området grænser op til og understøtter udvikling og innovation i det private erhvervsliv inden for særligt vindmølle-, bioenergi og miljøteknologi. (Uddannelses- og Forskningsministeriet, 2015)

A strategic vision for UK e-infrastructure: With increased supercomputing capacity the Met Office would be able to improve short-range weather forecasts, long-range predictions, and climate change projections - all of which would deliver considerable further socio-economic benefit to the UK (UK , 2012, s. 10)

Biotek, Sundhed og LifeScience

Dansk Roadmap for Forskningsinfrastrukturer 2015: Fagområdet "Biotek, Sundhed og Life Science" dækker over et bredt fagområdefelt af stor relevans for udviklingen i den danske befolknings sundhedstilstand, herunder udvikling af nye lægemidler og medicinske teknologier samt bedre og sundere fødevarer. Forskningen foregår på blandt andet universiteter, hospitaler og klinikker samt ved sektorforskningsinstitutioner. Fagområderne grænser op til og støtter udvikling og innovation i det private erhvervsliv omfattende blandt andet lægemiddelindustrien, medico- og bioteknologiske virksomheder samt fødevarerindustrien, som alle er vigtige kilder til dansk beskæftigelse og eksport. (Uddannelses- og Forskningsministeriet, 2015)

Swedish Science Case for e-Infrastructures:

- To enable Swedish scientist to keep up with the forefront of research based on Earth System Models, HPC resource increases of the order of factor five are needed for the next five years.
- Efficient storage with access and analysis capabilities to the huge output data that the community are producing are as important as CPU hours.
- The need for organized and distributed databases for mainly observational data involves basically all established and new user communities. (Vetenskapsrådet, 2014)

Swedish Science Case for e-Infrastructures

Due to technology development and general usefulness, the needs in all areas will increase substantially, with more and larger projects and many new users.

- The current needs have, across the entire field, reached the ceiling of the current national e-Science infrastructures in Sweden. There is thus a need for more storage, computing and support-capacity.
- Improvements in computer technology are needed to cope with the massive increase in data *production*.
- There will be an increasing demand for service functions that enable efficient use of generated big data sets.
- A national e-Science infrastructure must be developed which supports secure analysis, storage and access

to data with clear rules and routines which guarantee that all activities follow current legislation.

- There is a need to develop tools and practices for handling sensitive medical information.

(Vetenskapsrådet, 2014)

Forsk2025 - første samlede udkast: En vigtig forudsætning for forskningen [indenfor personlig medicin] er i den forbindelse samarbejde omkring fælles standarder for den nationale forskningsinfrastruktur for eksempel fælles informations og it-standarder som kan sikre en sammenhængende it-arkitektur på tværs af sekventeringsdata, nationale registre og databaser og øvrige kliniske systemer. (UFM, Forsk2025 - første samlede udkast, 2017)


A strategic vision for UK e-infrastructure:

The next generation of pharmaceutical e-infrastructure therefore needs not only to facilitate large scale clinical and molecular analytics and modelling but also allow secure collaborative working over both public and proprietary data sets. (UK , 2012, s. 8)


DTU

Why supercomputers?


Our dataset: 100 billion pieces of DNA!



100 billion pages from
100s of different books



Tear out all pages
& mix them



Reassemble?
Computerome

Præsentation på National Supercomputing Day - Perspectives of High Performance Computing (Simon Rasmussen, 2016) *Sygdomsbekæmpelse. Gennem forståelse for udviklingen af tidligere dødelige sygdomme som pest og kolera, får man redskaber til bekæmpelse af sygdomme i dag som SARS, Ebola, HIV*

Fysik og Univers

Dansk Roadmap for Forskningsinfrastrukturer 2015: Inden for fagområdet "Fysik og Univers" beskæftiger forskerne sig med at finde svar på de helt store spørgsmål såsom naturen af universets mørke stof og mørke energi, eftersøgningen af nye planetsystemer og muligheden for liv på andre planeter, udforskningen af Jordens magnetfelt, dannelsen af stjerner og galakser ved Big Bang og opklaringen af mysteriet om, hvorfor elementarpartikler og dermed alt i naturen vejer noget. (Uddannelses- og Forskningsministeriet, 2015)

Swedish Science Case for e-Infrastructures

- The current HPC in Sweden corresponds to 45 M-CPU-H/y and the demand will, in five years, rise by a factor of five.
- Increased user support for optimizing software for multi-core systems is needed.
- The network speed needed for observable astronomy will reach 40 Gbit/s in 2020.
- Breakthroughs thanks to e-Science could be the identification of the "Higgs" particle, a working fusion reactor, and the identification of the Universe's dark matter (Vetenskapsrådet, 2014)

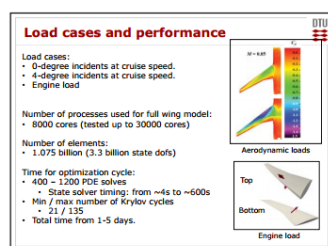
Materiale og Nanoteknologi

Dansk Roadmap for Forskningsinfrastrukturer 2015: Materiale- og Nanoteknologi omhandler begge undersøgelse, karakterisering og manipulering af materialer, inklusive biologiske materialer. Feltet har inden for de senere år udviklet sig til et meget dynamisk og tværfagligt felt, som strækker sig lige fra kemi,

fysik og materialefremstillingsindustri til blandt andet biologi, molekylærbiologi og medicin. Materialeegenskaber såsom holdbarhed, hårdhed, ledeevne og korrosionsmodstand har således eksempelvis betydning for så forskellige opgaver som udvikling og fremstilling af kvalitetskonstruktionsmaterialer til byggerier, fly og biler til funktionelle materialer til lægemidler, brændselsceller og mikroelektronik. Forskningen inden for materiale- og nanoteknologi har derfor også et stort industrielt potentiale i forhold til udvikling af nye produkter og produktionsmetoder inden for blandt andet lægemiddel-, vindmølle- og plastindustri, eksempelvis ved produktion af mere robuste, holdbare og miljøvenlige materialer. (Uddannelses- og Forskningsministeriet, 2015)

Swedish Science Case for e-Infrastructures

- Materials, Chemistry and Nano-science constitute the largest community of HPC users in Sweden and worldwide.
- There is a need for an increase in available computer power of orders of magnitude.
- It will become possible to move into a paradigm of predictive simulation of materials and molecular design.
 - New understanding will emerge from large-scale simulations of materials, nano-systems and chemical phenomena.
- Accurate and realistic calculations of complex and dynamic systems will become possible.
- There is a need for development of methods, models and software.
- There is a need for application experts at different levels. (Vetenskapsrådet, 2014)



Præsentation på National Supercomputing Day -
Large scale structural optimization - insight through HPC
(Andreasen, 2016)

Optimering af flyvinge ved computerberegninger på PRACE

Humaniora og Samfundsvidenskab

Dansk Roadmap for Forskningsinfrastrukturer 2015: Fagområdet "Humaniora og Samfundsvidenskab" beskæftiger sig blandt andet med at finde nye måder hvorpå man kan organisere sig, kommunikere, lære, motivere og optimere processer. Fagområdet bidrager til og er essentielt for vores kulturelle, sociale, politiske og økonomiske liv. Samtidig understøtter og fremmer det innovation i et bredt privat erhvervsliv beskæftiget med eksempelvis udvikling af nye læringsteknologier, konsulentytelser, virksomhedsdrift, kulturformidling, sprog og kommunikation og er en forudsætning for, at danske virksomheder kan agere på et globalt marked med stor kulturel og sproglig forskellighed. (Uddannelses- og Forskningsministeriet, 2015)

Swedish Science Case for e-Infrastructures

Dramatic expansions of digitized data demand improved systems of curation (from collection to dissemination) and greater cooperation across disciplines and national boundaries to ensure common standards and data sharing.

- New methods for data security – including anonymization and copyright protections – are required to enable increased access to personal information and proprietary data for research.
- Investments in human capital will be extensive in order to bring a diverse community of users to levels of expertise and creativity required to fully exploit e-Science possibilities.
- Possible breakthroughs include new understandings of:
 - Cultural heritage & contemporary production
 - Pre-historic social & economic organization
 - Genetic susceptibility to illness & disease
 - Contextual influences on cognitive development and academic performance – Segregation across time and space
 - Political discourses and their transmission across time and place – Intergenerational transmission of inequality (Vetenskapsrådet, 2014)

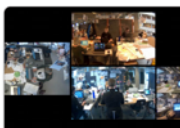
The Scientific Case and User Requirements for High-Performance and Data-Intensive Computing in Finland 2017-2021

Machine learning

The modern society is being radically transformed by the global megatrend of digitalization of all fields of science, industry and the rest of the society. In the core of this revolution is learning from data, alternatively called machine learning, studied e.g. in Acad. Prof. Samuel Kaski's group in Aalto University. Finland is particularly strong in this field which combines computerscience and advanced statistics. The field is extremely computationally intensive, and in particular requires coupling massive data sets to large-scale computational resources. Utilization of this significant asset Finland has, and staying ahead in the development of the extremely rapidly advancing field, requires sufficient computational resources of which a supercomputer is a key part (CSC, 2015)

RENDERING: ME AND MY DESKTOP

- Stitching and rendering video
 - 4-8 cameras
 - 360 camera
- Rendering process of 8 hours of video from 8 camera angles = 90 hours



Præsentation på National Supercomputing Day - BIGVIDEO in the humanities (Jacob Davidsen, 2016)

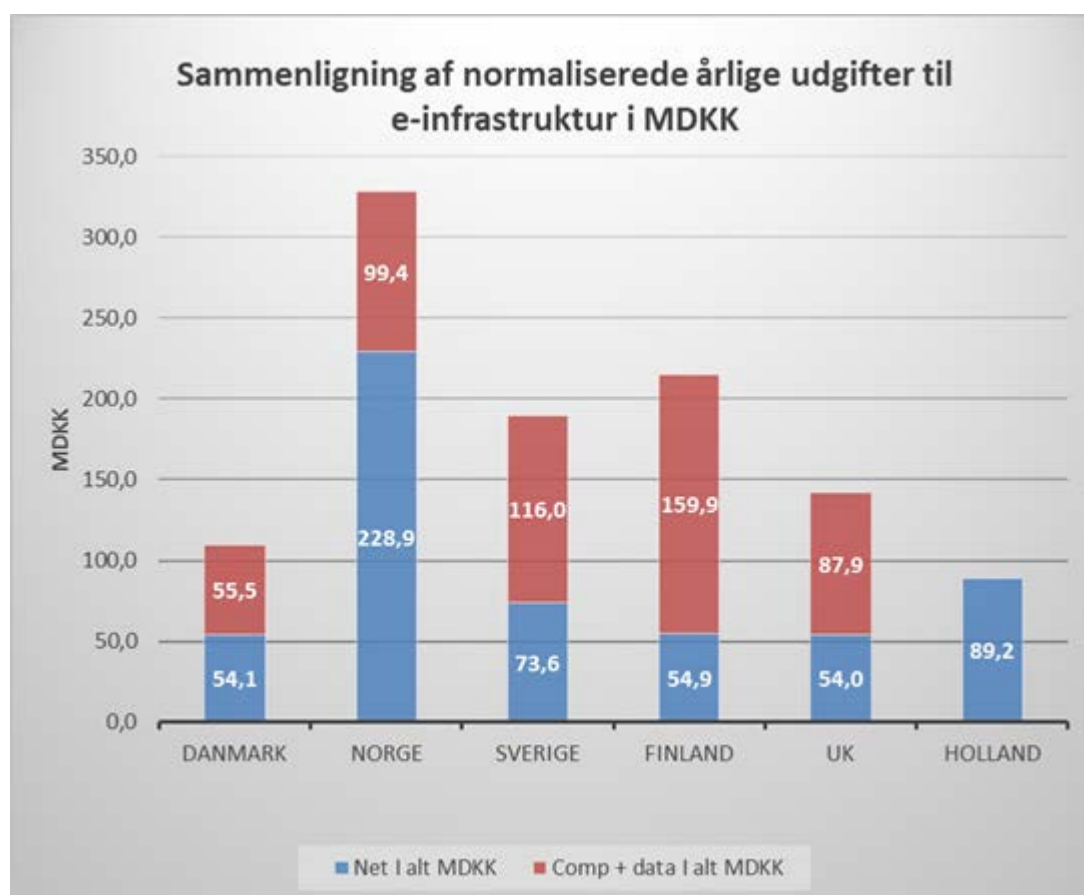
Der forskes i menneskelig interaktion gennem analyse af videotagelser

Bilag 2: Sammenligning af investeringer og økonomiske forhold omkring e-infrastruktur

DeiC 06-03-2017 Steen Pedersen

Sammenligning af investeringer og økonomiske forhold omkring e-infrastruktur i lande omkring os

Der er indhentet aktuelle økonomital fra de lande vi traditionelt sammenligner os med på forskningsområdet. Der er lavet en opdeling i udgifter til netværk (Net) og til computing/datahåndtering (C+D) og i drift og investeringer, der hvor det tilgængelige talmateriale giver mulighed for det. Resultaterne udtrykker den gennemsnitlige årlige udgift pr. kategori i DKK og tallene er herefter normaliseret mht. befolkningstal og sammenlignet for henholdsvis netværk og computing/datahåndtering. En samlet oversigt er vist nedenfor, hvor de samlede udgifter til de to hovedkategorier er sammenlignet. En mere detaljeret opstilling og beskrivelser af de enkelte tal og finansieringsform er angivet herefter. Dette omfatter også en opgørelse over finansieringsfordelingen mellem nationalt og lokalt niveau, hvor disse oplysninger har været tilgængelige.



De aktuelle udgifter omregnet til MDKK er vist her.

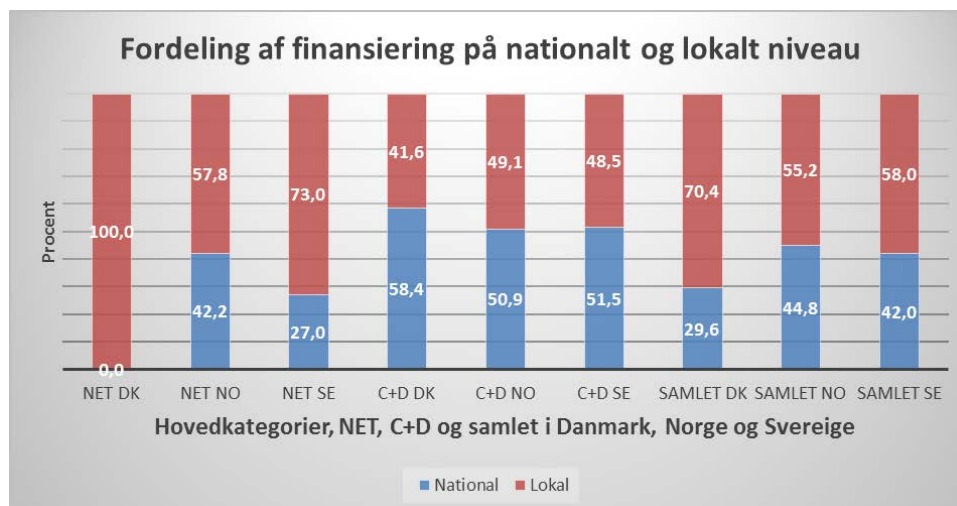
	Udgifter i MDKK/år							
	Net			Computing og data				I alt
	Drift+inv	Int - NORDUnet	I alt	Drift	Investering	Int- PR+NeIC	I alt	e-infra
Danmark	37,0	17,1	54,1	29,0	21,5	5,0	55,5	109,6
Norge	178,8	28,5	207,3	60,0	24,0	6,0	90,0	297,3
Sverige	97,6	28,0	125,6	96,0	96,0	6,0	198,0	323,6
Finland	39,1	13,8	52,9	74,5	74,5	5,0	154,0	206,9
UK	600,0	15,0	615,0	217,5	783,0		1000,5	1.615,5
Holland	253,3	9,7	263,0	NA	NA	NA		

Når udgifterne normaliseres mht. landenes befolkningstal giver det disse resultater, som er vist i figuren.

Befolkning			Årlige udg. normeret mht. befolkning		
Mio	Land/DK		Net I alt MDKK	Comp + data I alt MDKK	e-infra samlet i alt MDKK
5,7	1,0	Danmark	54,1	55,5	109,6
5,2	0,9	Norge	228,9	99,4	328,3
9,8	1,7	Sverige	73,6	116,0	189,5
5,5	1,0	Finland	54,9	159,9	214,8
65,0	11,4	UK	54,0	87,9	141,9
16,8	2,9	Holland	89,2		

Finansieringen foregår i alle lande i et samarbejde mellem det nationale og det lokale niveau. Fordelingerne mellem disse niveauer i det nordiske område fremgår af nedenstående tabel og figur.

	Fordeling af finansiering på nationalt og lokalt niveau							
	Net			Computing og data				I alt
	Drift+inv	Int - NORDUnet	I alt	Drift	Investering	Int- PR+NeIC	I alt	e-infra
DK-national	0,0	0,0	0,0	14,5	12,9	5,0	32,4	32,4
DK-lokal	37,0	17,1	54,1	14,5	8,6	0,0	23,1	77,2
NO-national	59,0	28,5	87,5	19,8	24,0	2,0	45,8	133,2
NO-lokal	119,8	0,0	119,8	40,2	0,0	4,0	44,2	164,0
SE-national	26,4	7,6	33,9	48,0	48,0	6,0	102,0	135,9
SE-lokal	71,2	20,4	91,7	48,0	48,0	0,0	96,0	187,7



Der er hentet tal fra en række forskellige kilder, som beskrives nærmere herunder.

Land	Udgifter til Net drift og investering	Internationalt net samarbejde	Computing og datahåndtering driftsudgifter	Computing og datahåndtering investering	Internationale aftaler
Danmark	Fra DeiC's regnskab for 2016. 100% finansieret af universiteterne	Dansk medlemsbidrag til NORDUnet i 2016. 100% finansieret af universiteterne	Tallet er fremkommet som summen af driftsudgifterne til de tre nationale anlæg + drift af kompetencecenter og datamanagement aktiviteter. Ca. 50% centralt finansieret via DeiC og DEFF, ca. 50% universitets finansieret	DeiC engangsbevilling + medfinansiering fordelt over 4 år. Heraf ca. 60% centralt finansieret og 40% finansieret af universiteterne	Dansk bidrag til NeiC (inkluderer drift af Nordic WLCG Tier-1) og PRACE. Finansieret 100% af centrale midler
Norge	Fra UNINETT regnskab 2015. Bemærk: Norge er et "udstrakt" land og der er relativt mange institutioner koblet på nettet, og der er flere services.	Norsk medlemsbidrag til NORDUnet i 2016. 100% finansieret af centrale midler.	Fra Sigma2 budget, gælder for en 10 årig periode. 67% finansieret af universiteterne og 33% finansieret af forskningsråd	Fra Sigma2 budget, gælder for en 10 årig periode. 100% finansieret af forskningsråd	Deltagelse i NeiC, PRACE, EGI og EUDAT. Finansieret som driftsmidlerne

	Ca. 67% finansieret af brugerne og ca 33% centralt tilskud.				
<i>Sverige</i>	Fra SUnets regnskab. 73% kommer fra brugerne og 27% fra centralt tilskud (SUNET, 2015)	Svensk medlemsbidrag til NORDUnet i 2016. Samme finansiering som net drift	Fra Future SNIC dokument. 50% finansieret af Vetenskapsrådet og 50% af universiteterne	Fra Future SNIC dokument. 50% finansieret af Vetenskapsrådet og 50% af universiteterne	Pt. 100% finansieret af Vetenskapsrådet.
<i>Finland</i>	Fra Géant NREN kompendium	Finsk medlemsbidrag til NORDUnet i 2016	Fra CSC Science Case, som indeholder budget for en fire årig periode. Budgetfordeling uafklaret	Fra CSC Science Case, som indeholder budget for en fire årig periode. Budgetfordeling uafklaret	Deltagelse i NeIC, PRACE, EGI og EUDAT. Omfatter ikke økonomien for CSC deltagelse i et betydeligt antal udviklingsprojekter
<i>UK</i>	Fra Géant NREN kompendium	UK medlemsbidrag til Geant i 2016	Fra UK "e-infrastructure roadmap". Mere end 80% nationalt bidrag	Fra UK "e-infrastructure roadmap". Mere end 80% nationalt bidrag	N.A.
<i>Holland</i>	Fra Géant NREN kompendium	Hollands medlemsbidrag til Geant i 2016	N.A.	N.A.	N.A.

Bilag 3: e-IRG Roadmap 2016 – recommendations for national governments/funding agencies

Kilde: (e-IRG, 2017)

The 2013 White Paper contains the following recommendations for this actor group:

1. Provide a basic funding level for the national e-Infrastructure, in particular devoted to its continuous innovation;
2. Empower and fund national user communities for the use of e-Infrastructure services, enabling them to influence the development of the national e-Infrastructure;
3. Remove existing national regulatory or political constraints for accessing publicly funded eInfrastructures for private research and public-private research ventures;
4. Provide input for the strategy setting and coordination bodies for their national eInfrastructures;
5. Encourage the actors in the national e-Infrastructures to collaborate and join forces with their counterparts in other countries and at EU level.

With regard to these recommendations e-IRG sees many indications of progress. In an increasing number of European countries, e-Infrastructures are appearing as a separate item on national Roadmap for (large) research infrastructures, recognizing the importance of these horizontal infrastructures. The national research funders also increasingly require data management paragraphs in research proposals.

Furthermore e-IRG sees an increase in countries, where the various cornerstones of e-Infrastructure development and provisioning (computing, data, networking), that used to act quite independently and uncoordinated, start to team up. This is driven by the sheer necessity to address the institutions and their users in a more coherent fashion, which requires a coordinated planning and funding of eInfrastructure service development and innovation. In some cases this coordination results in consolidation into a single organisation, with governance models varying between countries (such as the SURF cooperative in NL or the CSC - IT Center for Science (limited company) in Finland).

National e-Infrastructure organisations have a pivotal role. Usually they have strong formal links (such as through memberships) with research institutions and universities, where the 'long tail of science' resides. On the other hand they form the link to European e-Infrastructure projects and organisations. It is nearly impossible to reach this long tail on the European level: you need the intermediate national and institutional connections!

Currently national e-Infrastructure services are quite often not open by default to user groups from other countries and this makes coordinated provisioning at European level difficult.

Specific attention should be devoted to the challenge to serve the long tail of science. It is clear that infrastructure provision is indeed very different depending on to whom one wants to provide infrastructure for. The variation can be described as two continuums:

1. From very high level of technical e-Infrastructure competence, via lower levels to close to zero understanding.
2. From internationally very well organised research communities, via moderately organised to unorganised individual researchers in need of infrastructure.

These various preconditions related to the dichotomy may call for different infrastructure provisions strategies, organisational set-ups, policies and funding schemes. Again, the role of national and local level is very important in this domain, as well as in some cases the one of disciplinary level, and the necessity to integrate infrastructures provided at different levels should be recognised.

The e-IRG recommends strongly, that e-Infrastructure coordination and consolidation on the national level is embraced in full force in every European country. A strong European e-Infrastructure is dependent on strong national building blocks.

However, given always tight budgets for research, governments and funding agencies keep struggling with implementing mechanisms to fund a (generic) national e-Infrastructure and meanwhile keeping a balance with (other) research infrastructure priorities. Taken to the extremes, there are two mechanisms:

1. 'market approach': fund researchers and research communities and let them develop the services or buy them from whatever provider they want
2. "public provider approach": fund a (generic) national e-Infrastructure and offer these to researchers and research communities for free at the point of use;

Clearly both approaches have their pros and cons. Intermediate and mixed models (e-Infrastructure service development co-funded by users and national e-Infrastructure funds, including the commitment to a long lasting relationship between both) are also frequently advocated. A distinct variety is funding the development of a discipline specific e-Infrastructure on top of a generic (national) e-Infrastructure.

e-IRG observes that in various countries funding agencies are using different models but clearly there is no systematic approach as to assessing their effects and there is no forum to exchange best practices. This lack of clarity on who funds what and assigns what to whom also hinders the resource provisioning on the European level (i.e. for the benefit of European organised user communities).

In this context e-IRG refers to one of the conclusions of the Competiveness Council May 28-29, 2015, of the European Council conclusions on open, data-intensive and networked research: "ESFRI is invited to explore mechanisms for better coordination of Member States' investment strategies in eInfrastructures, covering also HPC, distributed computing, scientific data and networks." For this purpose ESFRI has established a working group with e-IRG representation. The main goal of the group is to analyse the investment strategies of the Member States in e-Infrastructures for research and innovation and formulate recommendations for how mechanisms for coordinating these strategies could be implemented. Important issues to be discussed in connection to national investment strategies include coordinated governance of service delivery, life cycle management for services, and coordinated funding streams for investments and operations.

e-IRG believes, that the funding system must facilitate the right incentive structure to reach the ideal situation, believed to be consisting at least of a balanced mix of:

1. base funding for the innovation of the (national) e-Infrastructure;
2. funding by users derived from service delivery by the providers;
3. top up funding based on (national) priorities for (demanding and well organised) research communities;

e-IRG supports strongly that national e-Infrastructure funding and governance mechanisms are analysed, so that best practices can be identified, which can contribute to more specific recommendations to national governments and funding bodies.

