

# **STATUSRAPPORT NATIONAL STRATEGI FOR FORSKNINGS DATAMANAGEMENT**

**NOVEMBER 2016**

Version 1.1

## INDHOLDSFORTEGNELSE

1	LEDELSESRÉSUMÉ.....	3
1.1	KONKLUSION.....	4
2	BAGGRUND – DET NATIONALE SAMARBEJDE OM FORSKNINGSDATAMANAGEMENT.....	4
2.1	DET STRATEGISKE UDGANGSPUNKT.....	4
3	ORGANISATION OG STYRING AF DET NATIONALE SAMARBEJDE INDEN FOR FORSKNINGSDATA MANAGEMENT.....	5
3.1	DEIC, DEFF OG DM LEDELSESCAB.....	5
3.2	UNIVERSITETERNE OG BEVARINGSINSTITUTIONERNE.....	6
3.3	DM FORUM.....	6
3.4	DM TEKREF.....	7
3.5	DM OPERATØROPGAVER.....	8
4	BEHANDLEDE INDSTILLINGER FRA ESCIENCE KOMITEEN, DM TEKNISK REFERENCEGRUPPE OG NATIONALT FORUM FOR FORSKNINGSDATA MANAGEMENT.....	8
4.1	INDSTILLINGER FRA DM FORUM OM DM PILOTPROJEKTER.....	8
5	TILSYNET MED DEN NATIONALE INDSATS.....	9
5.1	DM PILOTPROJEKTER - STATUS OG FREMDRIFT.....	9
5.2	DM FORUMS AKTIVITETER – OPNÅEDE RESULTATER.....	9
5.3	TRAIN-THE-TRAINERS KURSUS.....	10
5.4	DATAMANAGEMENT I PRAKSIS.....	11
5.5	DEIC'S VIRKE PÅ OMRÅDET "AKADEMISK DROPBOX" OG "ARBEJDSGRUPPE OM KRAV TIL LAGRING AF FORSKNINGSDATA I EN SKY".....	11
6	DM LEDELSESCAB'S ØVRIGE ARBEJDSOPGAVER.....	11
6.1	POLITIKUDVIKLINGEN PÅ OMRÅDET "OPEN DATA" OG FAIR.....	11
6.2	DATA MANAGEMENT PLANS ONLINE (DMP-ONLINE).....	12
6.3	REGISTERFORSKNINGEN OG FORSKNINGSDATA.....	12
7	VURDERING AF STRATEGIENS GENNEMFØRELSE.....	13
7.1	PERSPEKTIVERING.....	13
7.1.1	ET KIG UD I VERDEN.....	13
7.1.2	EN MULIG TILGANG I DANMARK TIL STORAGE-INFRASTRUKTUREN ("HÅRD DATA MANAGEMENT").....	14
7.1.3	DATAMANAGEMENT UD OVER INFRASTRUKTUREN ("BLØD DATAMANAGEMENT").....	14
7.1.4	ET INDSATSOMRÅDE MED RELEVANS FOR ERHVERVET, SUNDHEDEN OG VIDENSKABEN: DATAMANAGEMENT I LIFE SCIENCE.....	15
7.1.5	REGISTERFORSKNING.....	16
8	SLUTNOTE.....	16
8.1	EKSTERNE KONTAKTER.....	16
8.2	EUROPEAN OPEN SCIENCE CLOUD (EOSC) – PRESENTATION OG DISKUSSION.....	17

## 1 Ledelsesresumé

Denne statusrapport er udarbejdet af DM LedelsesCAB'en i overensstemmelse med dens kommissorium, der bl.a. pålægger den at føre tilsyn med den nationale strategi for forskningsdatamanagement og at rapportere årligt om dette tilsyn.

Rapporten beskriver de mangeartede aktiviteter, der er iværksat for at realisere strategiens mål. Det drejer sig om et antal pilotprojekter og flere andre aktiviteter, der er forankrede i det bredt sammensatte nationale Datamanagement Forum, samt om aktiviteter, der gennemføres i andet selvstændigt regi.

Rapporten beskriver også, hvordan den i strategien forudsatte organisering af arbejdet er gennemført, og at organiseringen opfylder ønsket om en bred forankring, ligesom det er lykkedes at få tilgodeset det nationale perspektiv i det udførte arbejde. Strategien er meget hurtigt blevet accepteret og har opnået legitimitet blandt samtlige involverede organisationer.

Endelig perspektiveres datamanagement i et afsluttende afsnit, hvor der peges på, hvilke rige og i stort omfang uudnyttede muligheder der er for at styrke forskningen i Danmark gennem en kraftig målrettet indsats også efter strategiperiodens udløb.

Rapporten indeholder en række delkonklusioner, og de kan resumeres som følger:

- 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 kunne have bidraget i meget væsentligt omfang, ligesom alle de valgte pilotprojekter ikke har fået fuld medfinansiering, hvilket peger på, at der med fordel vil 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.
- Alle de øvrige aktiviteter, DM Forum har iværksat og indstillet medfinansieret, bidrager i meget væsentligt omfang til realisering af den nationale strategis mål.
- 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 at sikre yderligere videndeling.
- Hvad den separate Train-the-Trainers aktivitet angår, skal der bygges videre på de indvundne erfaringer, og det internationale udsyn skal styrkes yderligere. Det bør derfor overvejes, hvordan der kan skaffes midler til at forsætte denne aktivitet eller en lignende ud over foråret 2017.
- Den samlede indsats på området åbne data er ikke tilstrækkelig, og det bør drøftes, hvordan den kan øges og især målrettes understøttelsen af de helt generelle principper og politikker, der udspringer af det internationale samarbejde
- Initiativet vedr. værktøjet DMPOnline til udarbejdelse af datamanagement-planer er helt afgørende for at kunne understøtte de stigende krav til og behov for systematisk planlægning af datamanagement i forskningsprojekter. Hvis ansøgningen til DEFF om finansiering til implementering af værktøjet ikke bevilges, bør der anvises alternative finansieringsmuligheder.
- Registerforskningen i Danmark er af en sådan strategisk vigtighed, at der bør være et særligt fokus på de datamanagement-discipliner, der har relevans for registerforskningen. Det bør gennem strategiperioden løbende specifikt vurderes, om der iværksættes og gennemføres tilstrækkeligt med initiativer på området.

- Med den valgte organisering af det strategiske arbejde kan det være vanskeligt at gennemskue, hvor beslutninger træffes, og at bevare et overblik, hvorfor det bør overvejes, om organisationen omkring strategien kan forenkles.
- Operatøropgaven er nødvendig og udføres tilfredsstillende.
- 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.

## 1.1 Konklusion

Det er DM LedelsesCAB'ens samlede vurdering, at den nationale strategi efterleves og udfoldes gennem de i denne rapport beskrevne aktiviteter, men det er samtidig vurderingen, at der ikke investeres nok i området, hvis strategiens overordnede mål skal nås inden for en rimelig tidshorisont.

I rapportens afsnit 7.1 Perspektivering er der angivet helt konkrete eksempler på, hvor indsatsen kunne styrkes, og hvad det medfører af investeringsbehov.

I forlængelse af denne perspektivering opfordrer DM LedelsesCAB'en til, at DeiC's bestyrelse og DEFF Styregruppe iværksætter udarbejdelsen af en strategi hhv. en udvidelse af den eksisterende omfattende de områder (bl.a. EOSC og "open data"), der p.t. ikke er adresserede.

## 2 Baggrund – Det nationale samarbejde om Forskningsdatamanagement

Ledelses Change Advisory Board for Forskningsdata Management (DM LedelsesCAB) er, ved sit kommissorium<sup>1</sup>, etableret som en del af *Den Nationale Strategi for Forsknings Datamanagement*<sup>2</sup>, der blev vedtaget i december 2014 af DeiC's bestyrelse og DEFF styregruppen.

DM LedelsesCAB'en har ifølge kommissoriet ansvar for at føre tilsyn med den nationale indsats og rapportere til DeiC bestyrelsen og DEFF Styregruppen. Denne statusrapport udgør den første regulære afrapportering i strategiperioden.

### 2.1 Det strategiske udgangspunkt

Den nationale strategi tager udgangspunkt i formålet "at sikre Danmark bedre og mere konkurrencedygtig forskning gennem effektiv opsamling, sikring, formidling og genbrug af relevante forskningsdata".

Strategiens ramme og de konkrete indsatser tager udgangspunkt i det aktuelle forhold, at forskningsdata ofte lever en isoleret og ikke fremtidssikret tilværelse privat hos forskeren, i IT-afdelinger eller på kommercielle skytjenester. Forskningsdata er desuden sjældent entydigt linket til forskere, projekter og publikationer og er derfor ofte svære at dele, genfinde for verifikationsformål eller genbruge til nye forskningsformål.

Værdien af forskningsdata øges imidlertid signifikant, når de er forvaltede, forbundne, søgbare og genbrugbare, og derved muliggør nye typer af forskning med udgangspunkt i genbrug af datasæt – det såkaldte "4. forskningsparadigme". Bedre data-management vil derfor medføre et markant løft til dansk forskning i den offentlige og private sektor og dermed medvirke til nationens konkurrenceevne. De indsatser, der udspringer af strategiens gennemførelse, skal derfor vurderes i forhold til deres bidrag til dette løft.

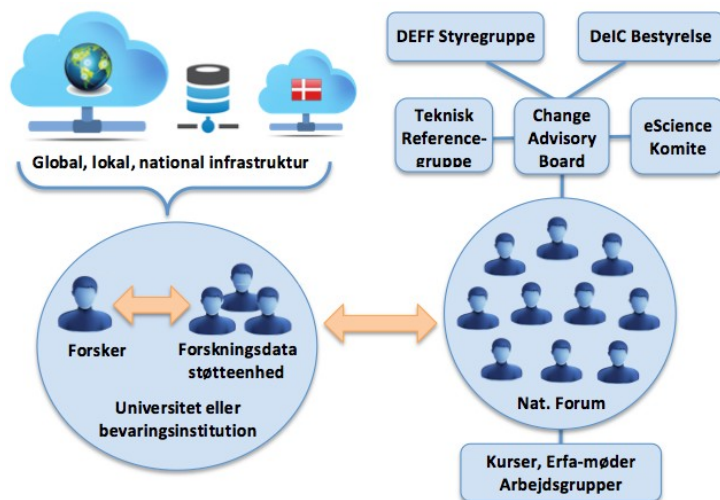
Det er desuden et væsentligt træk ved den nationale strategi, at den netop er national og samler alle interessenter om arbejdet med strategiens mål. Indsatserne skal derfor i anden række også vurderes i

---

<sup>1</sup> <https://www.DeiC.dk/sites/default/files/uploads/PDF/Kommissorium%20for%20Ledelses%20CAB%20NDM%201.0.pdf>

<sup>2</sup> [https://www.DeiC.dk/datamanagement/national\\_strategi\\_med\\_bilag](https://www.DeiC.dk/datamanagement/national_strategi_med_bilag)

forhold til deres bidrag til at fastholde det nationale perspektiv, samtidig med at det ikke må skygge for, at der også kan hentes inspiration og samarbejdes internationalt.



### 3 Organisation og styring af det nationale samarbejde inden for Forskningsdata Management

DM LedelsesCAB'en indgår i den samlede organisation, der støtter gennemførelsen af den nationale

Figur 1 Organisation og styring

strategi (se figur 1 nedenfor).

#### 3.1 DeiC, DEFF og DM LedelsesCAB

DM LedelsesCAB har til opgave at behandle indstillinger fra eScience Komiteen, DM Teknisk Reference-gruppe og Nationalt Forum for Forskningsdata Management og at anbefale beslutninger i henholdsvis DeiC Bestyrelse og DEFF Styregruppe.

DM LedelsesCAB har derudover, som nævnt ovenfor, ansvar for at føre tilsyn med den nationale indsats og rapportere til DeiC Bestyrelsen og DEFF Styregruppen.

Områdets samlede økonomi på nationalt niveau er i størrelsesordenen 40 millioner kroner over tre år (13,3 MDKK/år). Af dette beløb er godt 8 millioner kroner imidlertid allokeret inden etableringen af samarbejdet og styres alene af DEFF, men anvendes til formål der er relevante for og i en vis udstrækning supplerer det nationale samarbejde.

Midlerne er disponeret som anført i figuren herunder.

(Pr.2016 Q3)											
Strategi Ref.	DM Forum (Økonomisk estimat 2015-17; DKK)	DEFF	DeiC	Uni.	Bib.	Uni+Bib	Σ DKK	Disponeret	Disponeret %total budget	Σ Total Forbrugt	Σ Total Forbrugt %total budget
A1	Frikøb af formand	1,800,000					1,800,000	750,000	42		0
A2	Deltids frikøb af universiteternes nøglepersoner	5,235,000		5,235,000			10,470,000	6,757,800	65	2,949,800	28
A3	Deltids frikøb af bevaringsinstitutioners nøglepersoner	1,965,000			1,965,000		3,930,000	0	0	602,200	15
A4	Driftsomkostninger (Operatør v/Sandfær)	500,000					500,000	500,000	100	500,000	100
B	Train the trainer kursus	500,000					500,000	0	0	187,605	38
C	Fagspecifikke pilotprojekter		5,000,000			5,000,000	10,000,000	10,000,000	100	10,000,000	100
D	Akademisk "dropbox"		1,000,000				1,000,000	1,000,000	100	1,000,000	100
E	DeiC sekretariatsbetjening (Løn + Direkte udgifter)		3,600,000				3,600,000	1,380,402	38	517,000	14
F	"Data management i praksis" projekt	4,522,000				3708000	8,230,000	8,230,000	100	8,230,000	100
	<b>Total DKK</b>	<b>14,522,000</b>	<b>9,600,000</b>	<b>5,235,000</b>	<b>1,965,000</b>	<b>8,708,000</b>	<b>40,030,000</b>	<b>28,618,202</b>	<b>71</b>	<b>23,986,605</b>	<b>60</b>

Tilskud afregnes på basis af en fast gennemsnitlig timeløn på DKK.: 350

Disponeret beløb til frikøb (DKK):	Uni	Bib	Uni+Bib
1 DM Forum Deltagelse	1,050,000	250,000	1,300,000
2 DM Forum Aktiviteter	424,900	51,100	476,000
3 DM Forum Lokal støttefunktion (Antagelse, maks)	1,474,900	301,100	1,776,000
<b>Total forbrug til dato DKK</b>	<b>2,949,800</b>	<b>602,200</b>	<b>3,552,000</b>

Figur 2 Økonomi pr. 3. kvartal 2016

### 3.2 Universiteterne og Bevaringsinstitutionerne

Det nationale samarbejde om Forschungsdata Management er forankret hos Danske Universiteter (Rektorkollegiet), der udpegede 8 af de i alt 11 medlemmer af den styregruppe, der forestod udarbejdelsen af den nationale strategi. De 3 resterende medlemmer blev udpeget af bevaringsinstitutionerne, der nu er hhv. Nationalbiblioteket og Rigsarkivet.

Det er en del af samarbejdsgrundlaget og dermed strategien, at universiteterne og bevaringsinstitutionerne medfinansierer aktiviteterne under strategien i forholdet 1:1.

### 3.3 DM Forum

DM Forum samler ca. 30 nøglepersoner delvist fra de støttefunktioner for forskningsdata, som enten er under etablering eller udbygning på de danske universiteter, Nationalbiblioteket og Rigsarkivet (jf. Figur 1 Organisation og styring).

Formanden har bestræbt sig på, at universiteterne repræsenteres både ved forskere, forskerstøttemedarbejdere, biblioteksfolk, systemadministratorer, jurister og andre tæt på ledelserne. DM Forum er det operative organ, ansvarlig for implementering af den nationale strategi ved dets aktiviteter inden for forskningsdatamanagement i universitetsregi. Disse aktiviteter tager afsæt i en fælles vision<sup>3</sup> for DM i Danmark pr. juni 2018, og falder i disse kategorier:

1. Tværgående vidensdeling (forums møder mv.). DM Forum har afholdt 6 møder inklusive opstart august 2015. Her har man fundet frem til en fælles forståelse af genstandsfeltet, DM, og aftalt en fælles vision for status for DM i Danmark pr. juni 2018. DM Forummedlemmernes institutioner kompenseres for den tid medlemmerne bruger i DM Forum, med et beløb på DKK 50.000 årligt per deltagende person, hvoraf halvdelen dækkes af frikøbsmidlerne fra DEFF
2. Tværgående aktiviteter, som fx forskellige former for kortlægninger, afklaringer og udvikling af bedste praksis samt vidensdeling om aktiviteterne i projektet DM i Praksis. DM Forum arbejder med aktiviteter indenfor: Internationalt samarbejde og relationer; Politik skabelon; eLabBook; PhD undervisning; Digital DM guide; Politikimplementering og modeller for opbygning af DM support

<sup>3</sup> Common Ground - Vision for juni 2018: [https://www.DeiC.dk/dalamanagement/DM\\_forum/common\\_ground](https://www.DeiC.dk/dalamanagement/DM_forum/common_ground)

3. Forankring af DM i lokale universitetsenheder ved etablering eller udbygning af forskningsdatastøttefunktioner. DM Forum videndeler om lokale tiltag på institutionerne samtidigt med, at dette arbejde støttes med nationale midler i en størrelsesorden, der maksimalt er summen af de ovenstående to andre støtteformer
4. Kompetenceudvikling - Train-the-Trainers – blandt andet for DM Forum medlemmer. Udgangspunktet har indledningsvis været at opnå en fælles forståelse i DM Forum af DM problemstillingen. En "Future Search" metode har været anvendt til problemafløring; internationale eksperter har givet forelæsninger; Flipped Classroom og anden blended learning har været i spil
5. Udvælgelse af og vidensdeling om og tilsyn med fire pilotprojekter omhandlende e-Infrastruktur (se nedenfor), samfinansieret mellem DeiC og universiteterne

Alle virkemidler fra økonomien i DM Forum er disponerede dog ofte med afløb helt frem til sommeren 2018. Dette afspejler naturligvis et meget positivt engagement i DM. Samtidig udfordres Danmark imidlertid af det stærke fokus, som EU har givet til dagsordenen Open Data gennem EOSC. DM Forum vil fremadrettet tage denne udvikling op på møderne, men der er ikke midler til at iværksætte nye aktiviteter. DM Forum vil naturligvis arbejde på at opdatere igangværende aktiviteter på EOSC-dagsordenen.

Formanden ønsker intelligent åbenhed om aktiviteterne i DM Forum. Dette betyder, at den deling af informationer, som allerede er iværksat på DeiC.dk under fanebladet Datamanagement, senest 1. december 2016 vil blive udbygget til at indeholde en beskrivelse af både de delvist DEFF-finansierede lokale aktiviteter og DM Forums aktiviteter inklusive løbende skriftlige rapporter om opfølgning af aktiviteter; beskrivelser af Train-the-Trainers-aktiviteter inklusive evalueringer og eventuelle kursusmateriale; beskrivelser af de fire Pilotprojekter om e-Infrastruktur inklusive løbende skriftlige rapporter på opfølgning.

Derudover gives DM LedelsesCAB adgang til alt mødemateriale vedr. de fremtidige og de allerede afholdte 6 møder i Forum inklusive referater, ligesom den løbende opfølgning på udviklingen i økonomien gøres tilgængelig. Dette materiale gøres tilgængelig på data.DeiC.dk (senest 1. december) sammen med formandens afrapporteringer til DEFF-DeiC Styregruppe-Bestyrelse – pt. har formanden udarbejdet 5 sådanne afrapporteringer plus bidraget til denne rapport.

Ud over at forstå arbejdet som beskrevet ovenfor deltager formanden løbende i DM-aktiviteter nationalt og internationalt<sup>4</sup>.

Formanden er medlem af

- Reference Group for Nordic Data Services, NeIC
- Rigsarkivets følgegruppe vedrørende bevaring af forskningsdata (indstillet af SDU og udpeget af Danske Universiteter)
- Advisory Board for Den Danske Forskningsdatabase
- DM LedelsesCAB
- Knowledge Exchange
- SDU DM Forum, formand
- Arkivudvalget, foreslået af Ministeriet for Videnskab, Teknologi og Udvikling og udpeget af Kulturministeren

### 3.4 DM TekRef

I organiseringen af den nationale strategi indgår en teknisk referencegruppe (TekRef), der gennem et tværinstitutionelt samarbejde fremmer kendskabet til, opbygger viden om, tester og evaluerer tekniske datamanagement værktøjer. Der formidles bredere forståelse for de tekniske aspekter i Forskningsdata

---

<sup>4</sup> Se 8 Slutnote.

Management blandt teknikere, forskere og de administrative ledelser på universiteter og bevaringsinstitutioner.

DM TekRef understøtter Nationalt Forum for Forskningsdata Management og DM LedelsesCAB gennem teknisk rådgivning og anbefalinger.

### 3.5 DM Operatøropgaven

Operatørens rolle er at varetage de administrative, økonomiske og juridiske aspekter ved gennemførelsen af den nationale strategi. Operatøren administrerer således det samlede projekt i løbende dialog med såvel forums formand og sekretariat som bevillingsgiveren DEFF. I opstartsfasen havde operatøren tillige til opgave at sikre fremdrift i planlægningen af forums Train-the-Trainer kursusrække.

Operatørens opgaver er således:

- Aflæggelse af årligt regnskab indeholdende projektets udgifter og indtægter/finansiering, underskrevet af den juridisk/økonomisk ansvarlige
- Udarbejdelse og indgåelse af frikøbskontrakt for den udnævnte formand for Nationalt Forum for Data Management.
- Udarbejdelse og indgåelse af "consortium agreement" eller lignende rammeaftale, der regulerer det nationale forums virksomhed og danner baggrund for udbetaling af frikøbsmidler til de deltagende institutioner.
- Sammen med DeiC's projektleder for forskningsdata management at igangsætte planlægningen af Train-the-trainer kurserne
- Deltagelse i aktiviteterne i DM LedelsesCAB i hele perioden
- Administration og udbetaling af frikøbsmidler til de institutioner, der indgår i Nationalt Forum for Forskningsdata Management, vedrørende:
  - Formanden – i henhold til ovennævnte frikøbskontrakt
  - Nøglepersoner fra universiteterne – som indstillet af formanden og i overensstemmelse med ovennævnte rammeaftale
  - Nøglepersoner fra bevaringsinstitutionerne – som indstillet af formanden og i overensstemmelse med ovennævnte rammeaftale
- Indrapportering af evt. principielle ændringer af opgaverne eller forudsætningerne for deres løsning til DEFF Sekretariatet

## 4 Behandlede indstillinger fra eScience Komiteen, DM Teknisk Referencegruppe og Nationalt Forum for Forskningsdata Management

DM LedelsesCAB'en har behandlet en række indstillinger fra de tilknyttede organer, herunder særligt om et resurseforbrug på 10 mio. kr. til projekter og piloter<sup>5</sup>, der skal katalysere udviklingen på data-managementområdet.

### 4.1 Indstillinger fra DM Forum om DM pilotprojekter

Som i aftale med interessenterne og efter godkendelse i DeiC's bestyrelse og DEFF's styregruppe, har DM LedelsesCAB'en allokeret MDKK 5 til i alt 4 projekter under forudsætning af mindst samme medfinansiering fra institutionerne:

- **Video Life Cycle Data Management:** Projektet skal udvide den eksisterende multimedia-platform Edumedia med datamanagement-funktioner. Det omfatter blandt andet mere avancerede metadata og mulighed for at dele fortrolige data mellem forskere. Endvidere skal Edumedia integreres med DanPID, så mediefiler kan udstyres med langtidsholdbare identifikatorer (Persistent IDentifiers)
- **National Science App Store:** Projektet er et bud på en national open science cloud, der skal udvikle en generisk app-container til dataanalyseapplikationer. Der skal også udvikles integrati-

---

<sup>5</sup> Infrastruktur pilotprojekter: <https://www.DeiC.dk/datamanagement/pilotprojekter>



on mellem data.DeiC.dk og Abacus 2.0, så data frit kan flyde begge veje. Endvidere skal systemet integreres med Zenodo

- **International Infrastructure Integration of DK Science Cloud for Smart Cities:** Det strategiske forskningscenter CITIES (Centre for IT–Intelligent Energy System in Cities) vil udvikle et koncept for en national infrastruktur til near realtime-opsamling af big data. Det omfatter en prototype til datalagring baseret på OwnCloud og et system til behandling af de opsamlede data. Integration med DataCite og Zenodo skal tilføje langtidsholdbare identifikatorer (DOI'er, Digital Object Identifier)
- **Danish Infrastructure for Storing and Updating Actionable Biomarkers in Human Disease:** Biomarkører er data, der kan være med til at diagnosticere en sygdom. Projektet ser på de biomarkører, der kan hentes ud fra genomer. Det skal hjælpe med at give overblik over biomarkører, der i dag er spredt over mange databaser rundt om i verden. Det bygger på DeiCs nationale supercomputere

Det er DM LedelsesCAB'ens opfattelse, at de pågældende projekter alle er væsentlige for at udbrede og kvalificere datamanagement i Danmark, og at de derfor medvirker til opfyldelsen af den nationale strategi.

Det bemærkes dog samtidigt, at der måtte fravælges flere pilotprojektkandidater, der også ville kunne have bidraget i meget væsentligt omfang, ligesom alle de valgte pilotprojekter ikke har fået fuld medfinansiering. Dette understreger, at der med fordel vil kunne anvendes flere midler på området.

## 5 Tilsynet med den nationale indsats

### 5.1 DM pilotprojekter - status og fremdrift

DM pilotprojekterne er nu alle iværksat. Formanden har planlagt besøgsrunde i uge 49 2016 med henblik på sikring af fremdrift og på vidensdeling og læring i DM Forum. På nuværende tidspunkt kan det konstateres, at alle fire projekter har påbegyndt deres arbejde med et godkendt budget og godkendt indhold.

Det er DM LedelsesCAB'ens opfattelse, at der er en fornuftig fremdrift i pilotprojekterne, men det må også samtidig konstateres, at projekterne vil kunne lide under, at de involverede ressourcer typisk kun kan allokere tid i begrænset omfang. Det bør derfor overvejes, om en øget medfinansiering fra centralt hold vil kunne modvirke dette forhold.

### 5.2 DM Forums aktiviteter – opnåede resultater

Afsluttede og igangværende aktiviteter, der er besluttet i og følges af DM Forum, omfatter:

- **Skabelon for udarbejdelse af en institutionel DM-politik:** En oversigt over afsnit og emner, der bør indgå i en institutionel politik samt hvilke interessenter, der bør inddrages i politikudviklingen.
- **Digital Data Management Guide:** En guide, der skal understøtte forskere inden for alle discipliner i at håndtere forskningsdata korrekt og effektivt – og på sigt kunne føre til rådgivning af forskere med en velbeskrevet forskningsdatatype om en passende lagring af aktive og låste data
- **eLabBook:** Guidelines og bedste praksis for integration af elektroniske laboriebøger inden for et bredere spekter af fagområder og arbejdsgange, end sådanne systemer hidtil har været anvendt i - dette med henblik på at understøtte dokumentation af dataindsamling, processering og analyse og dermed understøtte forskningsintegritet
- **Internationalt samarbejde – Videndeling og synergi omkring den danske indsats:** At få kortlagt, prioriteret, udvalgt, formidlet, samordnet og operationaliseret det danske engagement i internationale DM aktiviteter - på såvel et politisk som teknisk og praktisk niveau
- **Udvikling af skalerbare læringsforløb og læringsmål i datamanagement for Phd-studerende (og deres vejledere):** Kortlægger eksisterende forskningsdatamanagement kompetenceinitiativer, vidensressourcer og erfaringer indenfor humaniora og sundhedsvidenskab mhp. at kravspecifisere en didaktisk model for et ph.d.-kursus og udforme udbudsmateriale til udvikling heraf

- **Politikimplementering:** Kortlægger infrastruktur relateret til RDM, services og supportfunktioner og aktiviteter. Identificerer ansvarsfordelingen og organiseringen samt informationsflowet ift. datamanagement mhp. at understøtte implementeringen af en datamanagementpolitik
- **Research Data Management behov og infrastrukturer – modeller for RDM:** Afdækker udvalgte eksisterende modeller for håndtering af forskningsdata mhp. at foreslå, hvilke modeller der med fordel kan anvendes i en dansk kontekst og være fælles referenceramme

Aktiviteten vedr. skabelon for udarbejdelse af en institutionel DM-politik er afsluttet og en skabelon er rundsendt til institutionerne, som lader den indgå i det videre arbejde med udvikling af en institutionel DM Politik. Skabelonen kan findes på DeiC.dk under fanebladet Datamanagement under Aktiviteter. CBS har haft brug for en engelsk version af denne skabelon, og har efterfølgende givet en oversættelse tilbage til DM Forum.

De andre aktiviteter er alle i god fremdrift undtagen dele af aktiviteten vedr. Internationalt samarbejde, der er udfordret på blandt andet vidensdeling, da en passende blog til formålet har ladet vente på sig. Alle aktiviteter afrapporteres i DM Forum både mundtligt og fra mødet i oktober måned også skriftligt.

De lokale aktiviteter omhandlende implementering af DM kan, som tidligere nævnt, også delfinansieres af DM-budgettet. Sekretariatet har netop modtaget to detaljerede beskrivelser fra DTU hhv. KU. Det forventes, at disse vil modtage finansiering, og at udviklingen fortløbende vil blive delt med DM Forum ligesom alle DM Forums aktiviteter og de delfinansierede lokale aktiviteter vil blive delt på DeiC.dk under fanebladet Datamanagement fra senest 1. december.

Det er DM LedelsesCAB'ens opfattelse, at alle de nævnte aktiviteter i meget væsentligt omfang bidrager til realisering af den nationale strategis mål, og at de også til fulde illustrerer, hvordan et nationalt perspektiv kan anlægges i aktiviteter, der formentlig ellers ville være blevet gennemført udelukkende med et institutionsspecifikt sigte. Det er imidlertid også DM LedelsesCAB'ens opfattelse, at koordinationen af og udvekslingen mellem de forskellige aktiviteter i nationalt regi og lokale aktiviteter med fordel vil kunne styrkes for både at undgå overlappende aktiviteter og at sikre yderligere videndeling. Det bør derfor overvejes, hvordan dette vil kunne ske.

### 5.3 Train-the-trainers kursus

Der har været afholdt en serie Train-the-Trainers-aktiviteter, som i nogen grad låner fra de lande, der er ca. 10 år foran Danmark på området. Det drejer sig om følgende 11 aktiviteter:

- August 2015 kick-off
  - Future Search
  - RDM experiences & challenges – talks by Elixir, 3TU.Datacentrum and OPEN
- Januar 2016
  - Data Management Planning - talk by Kevin Ashley, Digital Curation Centre, UK
- Marts 2016
  - The making of a domain specific RDM policy – talk by Nicole Schmitt, KU Health
  - Legal aspects of RDM – talks by Experts from CPH University
- April 2016
  - Essentials 4 Data Support - a course series with a flipped classroom flavour adapted for a Danish context <http://datasupport.researchdata.nl>
- Oktober 2016
  - RDM as a wicked problem – course by Andrew Cox, Sheffield
  - RDM for Grant Office People (SDU workshop offered as an inspiration for similar workshops at all universities)
- November 2016
  - Q&A with Barend Mons, Chair of HLEG for EOOSC
- Marts 2017
  - Research Data and Meta Data for Experts
- Primo 2017

- A series of workshops on Legal aspects of Open Data (open access, sensitive and personal data, ownership)

Train-the-Trainers-aktiviteterne er blevet godt modtaget. Der er tilbud til både novicer og eksperter. Beskrivelser, materialer og evalueringer kan findes på DeiC.dk under fanen Datamanagement under Train-the-Trainers. Der er dog ikke lige god dokumentation af alle tilbud.

Det fremgår, at der er planlagt tilbud også i foråret 2017, men derefter forventes budgettet at være forbrugt. Dette kan blive en udfordring for Danmarks tilpasning til European Open Science Cloud (EOSC).

Over to dage har EOSC har været forlagt for og diskuteret af henholdsvis DeiC's ledelse og repræsentanter fra Styrelsen for Forskning og Innovation, samt DM Forum og eScience komiteen.

Det er DM LedelsesCAB'ens opfattelse, at Train-the-Trainer aktiviteterne i væsentlig grad har medvirket til, at det internationale perspektiv, der også er omfattet af den nationale strategi, er blevet tilgodeset. Det er vigtigt, at der bygges videre på de indvundne erfaringer, og at det internationale udsyn styrkes yderligere. Det bør derfor overvejes, hvordan der kan skaffes midler til at forsætte denne aktivitet eller en lignende ud over foråret 2017.

#### 5.4 Datamanagement i Praksis

Data Management i Praksis (DMiP) er et DEFF-finansieret projekt der blev påbegyndt 1. marts 2015, altså umiddelbart før etableringen af det nationale forum. Projektet har sin egen organisering med egen styregruppe mv. men da arbejdet er så tæt relateret til det øvrige nationale datamanagement-arbejde har projektet været præsenteret og diskuteret ved flere lejligheder i flere af de øvrige nationale fora.

Projektet arbejder i 2 dimensioner, henholdsvis faktiske cases hvor de respektive forskningsprojekter hjælpes hele vejen i gennem en livscyklus for datamanagement og sideløbende i en række temaer der netop er udvalgt ud fra en livscyklus-model for at få udtrykt de erfaringer de enkelte cases kan bidrage med ift. disse temaer.

De konkrete cases skal bidrage til at de deltagende universitetsbiblioteker og Rigsarkivet ved Dansk Data Arkiv får praktisk erfaring med datamanagement og herigennem får et kompetenceløft ligesom projektet afdækker mulige fremtidige roller for netop bibliotekerne.

Temaarbejdet bidrager til at opsamle erfaringer som kan deles og formidles som mere generel viden til fremtidige projekter, som en slags best practices.

Projektet afsluttes medio 2017 og omkring marts 2017 planlægges et 1-dags afslutningsarrangement hvor projektets resultater vil blive præsenteret og evalueret af et eksternt ekspertpanel.

#### 5.5 DeiC's virke på området "Akademisk Dropbox" og "Arbejdsgruppe om krav til lagring af forskningsdata i en sky"

Der er i samarbejde med bl.a. Danske Universiteters CIO Forum blevet etableret en "Arbejdsgruppe om krav til lagring af forskningsdata i en sky", der har haft til opgave at opstille de krav man som universitet eller bevaringsinstitution bør og kan have til en national lagringsinfrastruktur, der i funktionalitet er at sammenligne med Dropbox og lignende produkter. Arbejdsgruppen har afrapporteret i form af en overordnet kravspecifikation ultimo oktober 2016, og der vil nu blive taget stilling til, hvordan arbejdet kan fortsætte, så der kan tages en beslutning på området.

## 6 DM LedelsesCAB's øvrige arbejdsopgaver

### 6.1 Politikudviklingen på området "Open Data" og FAIR

DM LedelsesCAB'en har diskuteret betydningen af det europæiske samarbejde på området "Open Data" og FAIR-principperne (Findability, Accessibility, Interoperability, and Reusability) for Danmark.

DM LedelsesCAB'en er af den opfattelse, at det europæiske samarbejde og generelt set arbejdet med datamanagement ud fra principper som dem, der er formuleret i FAIR, kun vil få stigende betydning i

de kommende år, og DM LedelsesCAB'en er overbevist om, at udviklingen vil gå meget stærkt (jf. de skærpede krav om datamanagement-planer og åbenhed, der gælder fra 1. januar 2017 i H2020-programmet).

Der bør på nationalt plan sættes hurtigt ind med målrettede initiativer, så kravene om åbenhed og samarbejde kan imødekommes ud fra en konkret vurdering dels i det enkelte forskningsprojekt dels af de specifikke datasæt.

De initiativer, der er omtalt i denne statusrapport, medvirker for en stor dels vedkommende til, at åbenhed og samarbejde kan fremmes, men den samlede indsats er efter DM LedelsesCAB'ens vurdering ikke tilstrækkelig, og det bør drøftes, hvordan den kan øges og især målrettes understøttelsen af de helt generelle principper og politikker, der udspringer af det internationale samarbejde.

## 6.2 Data Management Plans online (DMP-online)

Datamanagement-planlægning er det første vigtige skridt ifm. datamanagement. Dette arbejde gennemføres allerede ifm. planlægningen af et forskningsprojekt der skal anvende og/eller producere digitale data og planen revideres løbende efter behov.

Til støtte for dette arbejde eksisterer der flere web-baserede værktøjer og et af disse er DMPOnline som udvikles og vedligeholdes af det britiske DCC (Digital Curation Center).

I Danmark har DMPOnline været testet i DEFF-projektet FIF og siden implementeret i en pilot-installation i DMiP-projektet. Denne pilot-installation er brugt i DMiP's arbejde med de konkrete cases i projektet men installationen er i stigende grad også blevet inddraget af andre som har hørt om værktøjet og fået lov at bruge det. Værktøjet er desuden nu allerede en del af ph.d.-kurser på et par universiteter.

DMiP-projektet har følt behov for at få løftet installationen fra en pilot-installation til at blive et system i faktisk produktions-drift og for at løfte denne overgang og en efterfølgende drift er der i oktober 2016 afsendt en ny ansøgning til DEFF om finansiering af dette. Der er søgt penge til at få flyttet installationen fra DTU til en operatør og få etableret en national struktur omkring systemet med ansvar for helpdesk og lokal hjælp til systemet gennem de lokale universitetsbiblioteker.

DM LedelsesCAB'en finder initiativet helt afgørende for at kunne understøtte de stigende krav til og behov for systematisk planlægning af datamanagement i forskningsprojekter, så hvis ansøgningen til DEFF ikke bevilges, skal der efter DM LedelsesCAB'en mening tænkes i alternative finansieringsmuligheder.

## 6.3 Registerforskningen og forskningsdata

DM LedelsesCAB har diskuteret aktiviteterne i regi af *Det Koordinerende Organ for Registerforskning (KOR)*. Danmark har unikke muligheder indenfor registerforskning på grund af vores lovgivning, brugen af unikke CPR-numre og de mange registre med lang opfølgningstid. Forskere frembringer på basis heraf unbiased resultater, der ikke kan skaffes andre steder. Eksempler herpå er: bevis på at tvillinger klarer sig lige så godt i skolen som enkeltfødte børn<sup>6</sup>, og at mæslingevaccinationer ikke giver en øget risiko for autisme<sup>7</sup>.

KOR arbejder aktuelt på flere fronter: Oparbejdning af høj kvalitetsdokumentation finansieres af KOR's bevilling. Der arbejdes på et web-baseret overblik over danske registre. Der er startet et pilotprojekt i forhold til hvordan der i Danmark kan skabes et multigenerationsregister. Et sådant vil kunne være af interesse for forskere fra mange faglige discipliner, f.eks. sundhed, historie, sociologi osv.

---

<sup>6</sup> <http://www.registerforskning.dk/registerforskning/tvillinger/>

<sup>7</sup> <http://www.sum.dk/Sundhedsprofessionelle/Forskning/STARS-Strategisk-alliance-for-register-og-sundhedsdata/-/media/Files/20-%20dokumenter/STARS/Eksempelsamling/01-Eksempelsamling-STARS-20160610.ashx>

Herudover har der været kontakt mellem KOR og DEIC med hensyn til sikkerhedsspørgsmål; discovery-tjenester; tools; meta-tagging; jura; AALog data-værdiforøgelse. Muligheden for en fælles teknisk løsning omfattende både Danmarks Statistik og Sundhedsdatastyrelsen er blevet drøftet. Sundhedsdatastyrelsens opgaver blev gennemgået, herunder at styrelsen udvikler digitale løsninger til sundhedsvæsenet og samler data om befolkningens sundhed. Der bliver opdelt mellem administrative stabsdata og forskningsdata.

Det er DM LedelsesCAB'ens opfattelse, at registerforskningen i Danmark er af en sådan strategisk vigtighed, at der bør være et særligt fokus på de datamanagement-discipliner, der har særlig relevans for registerforskningen f.eks. med henblik på at kunne koble flere datasæt sammen. Det bør derfor gennem strategiperioden løbende specifikt vurderes, om der iværksættes og gennemføres tilstrækkeligt med initiativer på området. DM LedelsesCAB vil foretage denne vurdering og rapportere om den.

## 7 Vurdering af strategiens gennemførelse

DM LedelsesCAB'en vurderer samlet set på baggrund af det første års arbejde, at den nationale strategi efterleves og tilstræbes realiseret i de indsatser, der er igangsat. Det er således vurderingen, at indsatserne repræsenterer en tilstrækkelig bredde, både hvad angår de involverede institutioner, og hvad angår strategiens hovedpunkter og dens indsatsområder.

Det er i særlig grad lykkedes at skabe et samarbejde mellem de mange interessenter og institutioner, og det er også i pæn grad lykkedes at få koordineret og delt en række indsatser, der formentlig ellers udelukkende ville have haft lokal forankring og betydning.

De iværksatte pilotprojekter og andre aktiviteter er efter DM LedelsesCAB'ens opfattelse alle relevante og vil kunne få positiv indflydelse på opfyldelsen af strategiens overordnede mål, og det afsatte budget er med det for øje allokeret hensigtsmæssigt.

Der kan imidlertid herske tvivl, om det er muligt med de givne rammer at gennemføre tilstrækkeligt med indsatser på strategiens hovedområder, når det tages i betragtning, hvor stærkt udviklingen går, og på hvilket niveau datamanagement generelt set befinder sig i Danmark. Det første års arbejde med den nationale strategi har afdækket et stort uopfyldt behov for initiativer og har afsløret, at vi i Danmark har endnu et stykke vej at gå, før vi kan sammenligne os med de bedste, hvad datamanagement angår. I afsnittet 7.1 Perspektivering below er denne betragtning uddybet.

Hertil kommer, at det er ressourcekrævende at koordinere de mange interessenter og initiativer og dermed sikre de åbenlyse fordele, der på lidt længeres sigt er forbundet med at tænke og handle på tværs. I den sammenhæng er det også vigtigt at understrege behovet for, at det er tydeligt for interessenterne, hvordan beslutninger træffes, initiativer iværksættes og bevillinger gives i relation til strategien. Med de relativt mange organer, der er involveret (jf. Figur 1 Organisation og styring), så kan det være vanskeligt at gennemskue og bevare et overblik, hvorfor det bør overvejes, om organisationen kan forenkles, uden at interessenterne mister indflydelse eller den opnåede faglige bredde i involveringen sættes over styr.

Det er DM LedelsesCAB'ens samlede vurdering, at den nationale strategi efterleves og udfoldes gennem de i denne rapport beskrevne aktiviteter, men det er samtidig vurderingen, at der ikke investeres nok i området, hvis strategiens overordnede mål skal nås inden for en rimelig tidshorisont, og at der skal være særlig opmærksomhed på organisering og styring.

### 7.1 Perspektivering

Blandt e-Infrastrukturer er datalagring, -håndtering og -management ligeså vigtig som HPC og forventes at blive vigtigere i fremtiden.

#### 7.1.1 Et kig ud i verden

DeiC har peget på flere forskellige internationale HPC-løsninger inklusive ideer fra Norge og EPSRC i UK. Lad os se nærmere på vores partner i NeIC og NORDUnet, Norge, som tidligere end mange har erkendt, at forskningsdata er centrum for eScience. De to nationale HPC løsninger i Norge har tilkøbt datalagringsinfrastruktur. Disse to lagringsinfrastrukturer er i virkeligheden én og samme på to adres-

ser (HPC værtskaberne), idet det er tale om asynkron spejling (natlige kørsler, der skaber gensidige backups).

Lagringsinfrastrukturer tager afsæt i den eksisterende, noget ældre NorStore infrastruktur, der primært er til forholdsvis statiske data, men er en blandingsinfrastruktur, idet der til forskere uddeles "Project Storage Areas" til dynamiske data (på samme fysiske infrastruktur).

NorStore bygger på samme teknologier som EUDAT: B2SAFE bygger på iRODS teknologier; B2DROP bygger på OwnCloud og NextCloud.

EUDAT er ikke den eneste europæiske lagringsinfrastruktur, men EUDAT er den lagringsinfrastruktur, som vores nordiske naboer har valgt at satse på og samarbejder omkring i regi af NeIC.

De nationale lagringsinfrastrukturer er horisontalt integreret med nationale HPC installationer. Vertikalt er det tanken, at lagringsinfrastrukturen skal integreres både med de lokale lagringsinfrastrukturer og med European Science Cloud (EOSC) på nordisk og europæisk niveau.

Der arbejdes med begrebet "Data Centric Architecture" hvilket betyder, at det bliver data (ikke HPC), der er i centrum med tilknyttede services, herunder beregningskapacitet.

Om de mere bevaringsinstitutionelle sider af forskningsdatamanagement (blød datamanagement) er der ikke så meget at lære af Norge, idet man der ikke er kommet så langt med andre dele af Data Life Cycle, eksempelvis tilknytning af "digital object identifier (DOI)" og systemer til meta-tagging, indeksering og genfindning af data.

Det bliver op til Sigma2-bestyrelsen at afdække og udarbejde strategi for DM, men der syntes enighed om, at det også er et universitetsanliggende og ikke alene et arkiv- og biblioteksanliggende.

Lagringsinfrastrukturer bliver 100 % økonomisk integreret i Sigma2, idet det bliver en del af den samlede økonomi (ca. MNOK/år 110 på nationalt niveau), som beskrevet i den norske Sigma2 analyse.

Til hardware (CAPEX) forventes der heraf afsat ca. MNOK 7 årligt til det fremtidige NorStore, der forventes drevet (OPEX) med yderligere MNOK 7 årligt fra brugerne eller lokale institutioner og indledningsvis med subsidier fra det nationale budget. Dvs. et samlet årsbudget på ca. MNOK 14, alene til lagringsinfrastrukturer og dets drift. Afskrivningsmodellen er som for HPC 3 til 4 år, ligesom geografisk lokation ikke forventes at være permanent, men derimod følger HPC installationerne.

### **7.1.2 En mulig tilgang i Danmark til storage-infrastrukturen ("hård data management")**

Ministeriet har tidligere ladet sig inspirere af Norge, fx med hensyn til den Bibliometriske Forskningsindikator. Givet Danmarks forventede ophæng til EOSC, er der næppe grund til at gentage analysen i Norge. Svaret vil blive det samme: Forudsætningen for datamanagement er infrastruktur, og det koster ca. 15 mio. kr. pr år til generisk drift og generisk infrastruktur.

Hertil kommer særlige tiltag, hvis Danmark fortsat vil udvikle relevansen af forskningen gennem samarbejde mellem vidensinstitutioner og erhverv omkring præcisionsmedicin og mere generelt Life Science.

Hvorledes universiteterne og ministeriet sammen skal løfte finansieringen på 15 mio. kr. pr. år, skal afklares.

### **7.1.3 Datamanagement ud over infrastrukturen ("blød datamanagement")**

Danmark har etableret et godt samarbejde mellem de 10 institutioner, repræsenteret i Det Nationale Forum for Forskningsdata Management, og der er igangsat samarbejde om både infrastruktur og aktiviteter til fremme af god datamanagement.

Imidlertid vil der fortsat være behov for, at de 2 bevaringsinstitutioner (nu en) i samarbejde med de 8 universiteter kommer endnu tættere på en national understøttelse af

- Den europæiske Open Science vision
- Præcis metadatering af alle slags forskningsdata
- Håndtering af management af Big Data

- Arkivering af alle typer interessante data (i dag tager Rigsarkivet sig af en mindre del; Det nye Arkivudvalg under Kulturministeriet kan måske skrue et samarbejde sammen, der kan håndtere 100% af de bevaringsværdige forskningsdata)
- Samarbejde med Danmarks Statistik, Sundhedsdatastyrelsen og andre offentlige myndigheder med ansvar for relevante registre om datamanagement, herunder applikationer til at arbejde på data.

Det nationale samarbejde mellem de 8 universiteter og de 2 bevaringsinstitutioner (nu 1) bør fortsætte i yderligere 4 år begyndende medio 2018 –med et skærpet fokus på koordinering af fysisk infrastruktur til aktive og statiske data, og på muligheder for at genfinde, dele og genanvende data.

Finansieringsmodellen for denne del kunne bestå af et bidrag fra DEFF på 15 mio. kr. med matchende finansiering fra universiteter og FI således, at der er etableret et budget på 30 mio. kr. sammenlagt over 4 år.

Ovenstående er generisk infrastruktur, som løftes nationalt, baseret på samme rationaler som for HPC. Nedenfor sættes fokus på en fagspecifik infrastruktur, der organiseres for den aktuelle disciplin. Tiltaget skal indpasses med internationale samarbejder, fx ELIXIR, EMBL, BBMRI-ERIC, EMBRIC. Den mest oplagte kandidat for en national fagspecifik satsning på datalagringsområdet er af mange grunde, Life science.

#### **7.1.4 Et indsatsområde med relevans for erhvervet, sundheden og videnskaben: Datamanagement i Life science**

De biologiske, bioteknologiske og medicinske forskningsområder har i det sidste 20 år skiftet karakter. En lang række massivt parallelle, eksperimentelle teknikker genererer store mængder heterogene data, hvis omfang gør det både muligt og nødvendigt at tage ny informationsteknologi i brug. På kort tid er behovet for kompetence indenfor data og knowledge-management, herunder infrastrukturer der kan bruges både af eksperter og ikke- eksperter, blevet meget større.

I forhold til andre datatunge, videnskabelige discipliner er en helt afgørende komplicerende faktor, at mange af disse nye data er person-sensitive. Data produceres både i forbindelse med patientbehandling og i forskningsprojekter, og de er omfattet af forskellige former for lovgivning, der også ændrer sig hurtigt. Helt generelt er det således vigtigt, at der i udviklingen af løsninger og tilrettelæggelsen af regler og retningslinjer skelnes mellem adgang til data til kliniske formål og til forskningsformål. F.eks. har man til forskningsformål i sagens natur brug for adgang til store mængder data, mens der til kliniske formål fortrinsvis vil være brug for adgang til helt specifikke data. Dette forhold har betydning for, hvordan man designer en infrastruktur, og hvordan regelsæt og vejledninger disponeres.

Data indgår samtidigt i arbejder med et meget bredt spektrum af formål, f.eks. kvalitetssikring, metadataharmonisering, statistisk analyse, maskinindlæring, og integration på tværs af regionale og nationale landegrænser.

Danmark har med sin stærke akademiske og industrielle life-science profil brug for nye tiltag, der sikrer, at de nyeste data management teknologier kan udvikles og tages i brug, således at data kan udnyttes optimalt samtidig med at de i praksis kan beskyttes og benyttes korrekt indenfor lovgivningsmæssige rammer. Området er i Danmark præget af duplikeret arbejde, hvor lokale løsninger ofte genskabes grundet fraværet af modulære, generiske løsninger, der kan indgå i databehandlingsprocesser, der er opbygget effektivt og økonomisk.

Den internationale udvikling har paralleller til mobiltelefoniområdet, hvor lande, der før var tilbagestående og nærmest uden infrastruktur, bringer sig i front, fordi nye løsninger implementeres på meget kort tid, ofte under medvirken af private aktører. Danmarks konkurrencemæssige situation er derfor udfordret.

På grund af de unikke tekniske og juridiske forhold inden for life-science området er der brug for tiltag, der sikrer nye muligheder på tværs af de offentlige og private sektorer i Danmark gerne som samfundspartnerskaber, der også har fokus på præ-kompetitive aktiviteter.

Samarbejde med de andre nordiske lande er også en oplagt mulighed.

Afhængigt af ambitionsniveau vil gensekventering kræve massive investeringer i tiered storage-infrastruktur. Der er ligeledes behov for anvendelse og videreudvikling af kompressions algoritmer for at sikre economy of scale. På baggrund af disse punkter er der et estimeret behov på 15-20 mio. kr. per år.

### 7.1.5 Registerforskning

Registerforskningen kan blive en betydelig kilde til forskning og innovation, hvis udfordring med datas følsomhed takles rigtigt. Dette potentiale bør undersøges.

Danmark har globalt set et helt unikt forskningspotentiale ved de data, som i generationer er samlet i de danske offentlige registre. Få andre lande har en så lang og grundig tradition for systematisk at indsamle sådanne indsigtssrige data. Flere og flere forskningsområder har fået interesse for disse data i takt med, at lagrings- og beregningsteknologierne er modnet og er blevet mere udbredte. Men disse data adskiller sig fra mange andre forskningsdata ved deres personfølsomme natur.

Det Koordinerende Organ for Registerforskning (KOR) har vurderet det store potentiale for forskning og innovation såvel som de udfordringer det personfølsomme element udgør.

Konklusionen er, at der er behov for nye infrastrukturtyper, der på en og samme tid muliggør en ensartet, brugervenlig og gnidningsfri adgang til registrene for forskerne samtidigt med, at der ikke hersker tvivl om sikkerheden i forhold til anonymisering og adgangsregulering. Således er det blevet afgørende at få samlet parterne om en nærmere afgrænsning og specifikation af teknologibehovet samt sikre, at infrastrukturen finansieres og tilvejebringes. Endvidere skal en sådan infrastruktur give mulighed for, at data fra registrene kan kobles til andre datatyper og store datamængder (fx biomarkører), samt at de store datamængder kan analyseres tilstrækkeligt hurtigt.

De involverede parter vurderes indledningsvis at være Sundhedsdatastyrelsen, Danmarks Statistik og den nationale forskningsinfrastrukturleverandør, DeIC. Der er således behov for samarbejde mellem de tre ministerier (Sundheds- og Ældreministeriet; Social- og Indenrigsministeriet; Uddannelses- og Forskningsministeriet), hvor det er naturligt at Forskningsministeriet er den drivende part.

Et estimat for omkostninger og afkast ved denne indsats må afvente en sådan udredning.

## 8 Slutnote

### 8.1 Eksterne kontakter

Formanden for DM Forum har besøgt følgende institutioner i forbindelse med nedsættelse af DM Forum: SB, AU, KU, DTU, RA og har deltaget i følgende møder:

DeiC All Hands, kort oplæg "DM Forums kommende opgaver", 22.05.15

SDU Bibliotek, Datamanagement, med deltagere fra SDU-Grant Office, 15.06.15

NeIC DM-workshop i Arlanda, kort oplæg "Nationalt F-DM Forum", 30.08-01.09.15

Norske UniNet besøger DeiC, kort oplæg "Tilgangen til DM i Danmark", 15.09.15

Statusmøde, DEFF-projektet DM i Praxis, kort oplæg "Synergi med Forum", 01.10.15

DeiC konference, kort oplæg "DM Forum, Status og Strategi", 06-07.10.15

Forskningsbibliotekernes Chef-forum, kort oplæg "Hvordan bidrager bibliotekerne til og får indflydelse på det nationale samarbejde?", 28.10.15

DeiC Årsmøde, kort oplæg "Dansk datamanagement på vej", 03.11-04.11.15

The Knowledge Initiative, Helsinki "Pathways to Open Scholarship", 30.11-01.12.15

The 7th. European Data Documentation Initiative Conference, København, kort oplæg "The Danish National Forum for Research Data Management", 03.12.15



DeiC møder NeIC, kort powerpoint med vision 2018 for Datamanagement inkl. storage og Compute, 14.12.15

Deltagelse i "Jura i forskningen", Workshop arrangeret af DeiC Kompetence Center, DTU, 18.01.16

Deltagelse i ERA workshop, Styrelsen for Forsknings og Innovation, 27.01.16

Knowledge Exchange workshop i London om RDM Training and Skills, kort oplæg "RDM: Working with the Researchers", 09.02-10.02.16

NeIC Workshop om Datamanagement, Stockholm, 18-19/5.16

E-Science Komitemøde med gennemgang af DM Forums arbejde siden etablering maj 2015, 11.05.16

Møder med FI om EU Open Science Cloud den 07.03 og 02.05, 2016

Statsbiblioteket var vært for et dagsmøde med Canadian Research Libraries den 27.06 2016, hvor blandt andre formanden holdt oplæg: "Research Datamanagement in Denmark, Status June 2016"

Formanden deltog den 16.09 2016 sammen med Rene Belsø i møde mellem Videnskabsministeriet og CIO, Kimmo Koski fra CSC, Finland. Det fremgik blandt andet, at CSC er interesserede i ideen med et Nationalt DM Forum

Formanden deltog i Konferencen "Digital Infrastructures for Research, 2016", 28-30.09 2016, Krakow, Polen.

Formanden deltog i Knowledge Exchange Telco møde, 07.10, 2016

Formanden deltog i Knowledge Exchange Møde i Bonn, 18.10 2016, Research Data Expert Group Meeting

Formanden interviewede DeiC-direktør Steen Pedersen om data.deic.dk til brug for et Knowledge Exchange projekt om Federated e-Infrastructures, 26.10, 2016

Formanden var medarrangør, som formand for SDU DM Forum, af: Workshop on Research Data Management for 30 deltagere, 27.10, 2016. Arrangementet deles som et Train-the-Trainers-element med det Nationale DM Forum

Formanden deltog via Scopia i DM-delen af dialogmødet mellem med NeIC og DeiC, 28.10, 2016

Formanden deltager i konferencen, National Seminar on Open Science in a Horizon 2020 context, med oplæg: New developments in RDM in Europe and in Denmark, 16.11, 2016, SDU

Formanden deltager i Nordic Open Science and Research Forum 2016, med oplæg: Implementation of the Danish RDM Strategy, 21-23.11, 2016, Helsinki

Formanden har ansvar for en workshop om datamanagement ved NeIC 2017 Konferencen, Umeå, 29.05 – 01.06.

## 8.2 European Open Science Cloud (EOSC) – Presentation og diskussion

Over to dage den 18. og 19. november 2016 har EOSC har været forlagt for og diskuteret af henholdsvis DeiC's ledelse og repræsentanter fra Styrelsen for Forskning og Innovation, samt DM Forum og eScience komiteen.

### Deltagere fra Styrelsen og DeiC's ledelse i møde med Barend Mons torsdag den 17. november 2016

Barend Mons

European High Level Expert Group

Lars Christensen	UFM
Troels Rasmussen	UFM
Jonas Bak	UFM
Hanne Louise Kirkegaard	UFM
Steen Pedersen	DeIC
Henrik Pedersen	DeIC
Rene Belsø	DeIC
Martin Bech	DeIC
Lene Krøl Andersen	DeIC
Kurt Gammelgaard Nielsen	DeIC
Gitte Julin Kudsk	DeIC
Rene Buch	NORDUnet
Lars Fischer	NORDUnet
Erik-Jan Bos	NORDUnet
Josva Kleist	NORDUnet

**Deltagere fra DeiC DM Forum og eSciencekomite i møde med Barend Mons fredag den 18. december 2016**

Barend Mons	European High Level Expert Group
Bente Maegaard	Københavns Universitet
Jeppe Olsen	Aarhus Universitet
Christian Hagen Thomasen	Kulturministeriet
Mikael Elbæk	DTU
Paula Martinez Lavancy	DTU
Falco Hüser	DTU
Mogens Sandfær	DTU
Karsten Kryger	Aalborg Universitet
Lars Nondal	Copenhagen Business School
Jonas Langeland Pedersen	IT Universitetet
Bodil Stenvig	Statens Arkiver
Anders Sparre Conrad	Det Kongelige Bibliotek
Mareike Buss	Copenhagen Business School
Kirsten Villadsen Kristmar	Statens Arkiver
Bo Öhrström	DEFF
Rene Belsø	DeIC

Diba Markus	DeiC
Lene Krøl Andersen	DeiC
Torben B. Sørensen	DeiC
Helle Meldgaard	DeiC