

Bruxelles, den 19.4.2016
COM(2016) 178 final

**MEDDELELSE FRA KOMMISSIONEN TIL EUROPA-PARLAMENTET, RÅDET,
DET EUROPÆISKE ØKONOMISKE OG SOCIALE UDVALG OG
REGIONSUDVALGET**

**Det europæiske cloudinitiativ - Opbygning af en konkurrencedygtig data- og
vidensøkonomi i Europa**

{ SWD(2016) 106 final}
{ SWD(2016) 107 final}

Indledning

Verden er vidne til en dramatisk stigning i den mængde og type af data, der bliver genereret. Samtidig med at der genereres data af milliarder af mennesker, som anvender digitale apparater og tjenester af personlige og erhvervsmæssige årsager, og data af et stigende antal netforbundne apparater, genereres der også data fra forskning, fra digitaliseret litteratur og digitale arkiver, og fra offentlige tjenester som hospitaler og ejendomsregistre. Fænomenet "big data" giver nye muligheder for at dele viden, forske og udvikle og gennemføre offentlige politikker.

Det bliver også stadig lettere at udnytte disse data takket være cloudtjenester. Cloudtjenester kan forstås som en kombination af tre af hinanden afhængige elementer: den datainfrastruktur, hvor data lagres og administreres, de højhastighedsnet, der transporterer data, og de stadig kraftigere computere, der kan bruges til at behandle data. Mulighederne for at analysere og udnytte big data har konsekvenser for den globale økonomi og det globale samfund, da den giver mulighed for en større grad af industriel og social innovation. En væsentlig del af disse konsekvenser er en ændring af den måde, hvorpå videnskabelig forskning udføres, fordi vi hastigt bevæger os hen imod [Åben forskning](#) (Open Science).

Cloudtjenester gør det muligt uden problemer at flytte, dele og genbruge data på tværs af grænser og globale markeder, og mellem institutioner og forskningsområder. Med den kapacitet, der er tilgængelig i Europa på nuværende tidspunkt, behandles de data, der produceres af EU's forskermiljø og erhvervsliv, ofte andetsteds, og blandt europæiske forskere og innovatorer er der en tendens til at flytte til de steder, hvor der er bedre adgang til data- og databehandlingskapacitet. Eftersom Europa genererer den største mængde videnskabelig viden i verden, ville det samtidig være på sin plads, at Europa tog den globale føring inden for udviklingen af en forskningscloud.

For at udnytte datas potentiale som en nøgelfaktor inden for åben forskning og den [4. industrielle revolution](#) skal Europa finde svar på adskillige specifikke spørgsmål:

- Hvordan får man mest muligt ud af incitamenterne til at dele data, og hvordan øger man evnen til at udnytte dataene?
- Hvordan sikrer man, at data kan bruges så bredt som muligt på tværs af videnskabelige fagområder, samt i både den offentlige og den private sektor?
- Hvordan forbinder man de eksisterende og nye datainfrastrukturer på tværs af Europa bedre?
- Hvordan koordinerer man bedst den støtte, der er tilgængelig for europæiske datainfrastrukturer, efterhånden som de bevæger sig hen imod databehandling i exaskala¹?

De potentielle gevinster for forskning, teknologi og innovation ved at imødegå disse udfordringer blev understreget af [forsker miljøet selv](#), men også af [regeringerne i OECD-landene](#). EU's medlemsstater bekræftede i 2015², hvor vigtigt det er for samfundet og hele økonomien at imødegå disse udfordringer. Som direkte svar på udfordringerne fremsættes der

¹ Databehandling i exaskala henviser til computersystemer med kapacitet på mindst én exaFLOPS – 10^{18} beregninger pr. sekund – hvilket er ca. 1 000 gange hurtigere end, hvad vore dages maskiner kan.

² Se [konklusionerne](#) fra samlingen i Rådet (konkurrenceevne) fra 2015.

i denne meddelelse forslag om at oprette et europæisk cloudinitiativ, der kan sikre Europas plads i den globale datadrevne økonomi³.

Det europæiske cloudinitiativ bygger på strategien for et digitalt indre marked, som bl.a. sigter mod at få mest muligt ud af vækspotentialt i Europas digitale økonomi⁴. Det sigter mod at udvikle et sikkert, åbent miljø til forskere, hvor de kan lagre, dele og genbruge videnskabelige data og resultater, den **åbne europæiske forskningscloud**⁵. Ligeledes sigter det mod at udrulle den underliggende superberegningskapacitet, de hurtige tilslutningsmuligheder og de cloudløsninger med høj kapacitet, der er nødvendige, via en **europæisk datainfrastruktur**⁶. 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. At nå dette mål vil kræve en samarbejdsindsats, der er åben for alle, der er interesseret i at drage nytte af Europas datarevolution som en væsentlig komponent i den globale vækst.

Det europæiske cloudinitiativ bygger på resultaterne af den europæiske strategi for cloudcomputing⁷ og strategien for højtydende databehandling⁸. Det bygger på initiativer som f.eks. det nyligt annoncerede vigtige projekt af fælleseuropæisk interesse (IPCEI) om anvendelsen af højtydende databehandling og big data⁹. Initiativet videreudvikler den politik, der er skitseret i meddelelsen om big data¹⁰, og støtter den europæiske politiske dagsorden for åben forskning, som sigter mod at øge kvaliteten og virkningen af forskning¹¹, samt bygger videre på resultaterne inden for fri adgang¹². Denne meddelelse markerer begyndelsen på en proces, hvorigennem Kommissionen vil gå sammen med medlemsstaterne og alle relevante interessenter for at sikre, at det europæiske cloudinitiativ kan nå sine mål.

Det europæiske cloudinitiativ vil blive understøttet af yderligere tiltag i forbindelse med strategien for et digitalt indre marked, som omfatter cloudkontrakter for virksomhedsbrugere og muligheden for at skifte leverandør af cloudtjenester, samt initiativet om frie datastrømme¹³.

Fem grunde til, at Europa ikke udnytter datapotentialet fuldt ud endnu

For det første mangler mange europæiske virksomheder, forskningssamfund og offentlige enheder at udnytte det fulde **datapotentiale** og dets potentielt transformerende virkning på

³ Tale af Kommissionsformand Jean-Claude Juncker, oktober 2015, <http://bit.ly/1Y52pGi>.

⁴ COM(2015) 192 final.

⁵ Det forberedende arbejde blev igangsat via en ekspertgruppe på højt niveau fra Kommissionen, som fik til opgave at fremsætte anbefalinger om infrastrukturens oprettelse: <http://bit.ly/1RK7lhh>.

⁶ Forberedende arbejde udført, dvs. gennem rådgivende grupper som f.eks. refleksionsgruppen vedrørende e-infrastruktur.

⁷ COM(2012) 529 final og resultaterne af arbejdsgrupperne <http://bit.ly/1QVrvIb>.

⁸ COM(2012) 45 final.

⁹ Formålet er at støtte ny industriel anvendelse af højtydende databehandling og garantere offentlig og privat forskning adgang til højtydende databehandlingsfaciliteter, <http://bit.ly/1RMFq0i>.

¹⁰ COM(2014) 442 final.

¹¹ Rådets orienterende debat (9385/15); Rådets konklusioner (8970/15).

¹² COM(2012) 401 final.

¹³ Eventuelle lovforslag vil blive underlagt Kommissionens krav om bedre lovgivning i overensstemmelse med Kommissionens retningslinjer for bedre regulering, SWD(2015) 111.

traditionelle sektorer og den måde, hvorpå forskning bedrives¹⁴. **Data fra offentligt finansieret forskning er ikke altid frit tilgængelige**, og data, der genereres eller indsamles af virksomheder, deles ofte heller ikke, og det er ikke altid af kommercielle årsager. Nogle ser stadig data som et aktiv, der skal beskyttes, men mange i virksomhederne (særligt SMV'er), akademiske kredse og den offentlige sektor simpelthen er ikke bevidst om værdien af at dele data. Årsagerne hertil er bl.a. **mangel på en klar incitamentsstruktur** og manglende belønning for at dele data (primært i akademiske kredse), mangel på et klart retsgrundlag¹⁵ (primært i den offentlige sektor) og mangel på datafærdigheder samt manglende anerkendelse af datas værdi (i alle sektorer). EU's databeskyttelsesramme forhindrer, at personlige oplysningers frie bevægelighed begrænses på grundlag af privatlivets fred og beskyttelse af personoplysninger. Andre juridiske og tekniske barrierer for datas fri bevægelighed skal reguleres via det kommende initiativ om frie datastrømme i forbindelse med strategien for et digitalt indre marked.

For det andet forhindrer manglen på interoperabilitet, at der kan tages hånd om store samfundsmæssige udfordringer, som forudsætter effektiv datadeling og en tilgang, hvor mange fagområder og aktører er involveret, som f.eks. klimaforandringer, der ikke kan afhjælpes af klimatologer alene. Selv om problemerne med interoperabilitet og datadeling er blevet løst i visse sektorer (f.eks. placering af data i INSPIRE-direktivet¹⁶, sundhedsoplysninger i direktivet om patientrettigheder¹⁷), er der stadig mange datasæt, der ikke er tilgængelige for forskere, erhvervslivet, den offentlige forvaltning og politiske beslutningstagere. Selv om interoperabilitet af administrative data primært forudsætter minimumsstandards, retssikkerhed for så vidt angår adgang og anvendelse samt praktisk support¹⁸, hæmmes deling af forskningsdata også af datasættenes størrelse, deres forskellige formater og kompleksiteten i den software, som er nødvendig for at analysere dem, og de dybt rodfastede skel mellem fagområderne. Der er behov for simple "metadata"¹⁹ til at identificere data og specifikationer for datadeling for at gøre dem tilgængelige i videre omfang og for at muliggøre, at de kan behandles via fælles open source-dataanalyseværktøjer. Problemer med langtidsopbevaring og forvaltning af data skal også løses. Der findes allerede globale græsrodsinitiativer²⁰, og nogle medlemsstater gør fremskridt på dette område, men den europæiske deltagelse i disse initiativer er begrænset, og indsatsen er i høj grad fragmenteret.

¹⁴ Det er f.eks. tilfældet inden for sundhedsvæsenet <http://bit.ly/1XEeaTN> (og EFR-projekterne BIOTENSORS, DIOCLES, SMAC), astronomi (f.eks. SparseAstro); klimaforandring, migration eller internettet (f.eks. DIADEM, MIGRANT, RAPID, THINKBIG).

¹⁵ I INSPIRE-direktivet 2007/2/EF fastsættes reglerne for deling af europæiske lokaliseringsdata. Anvendelsesområdet for disse regler er imidlertid begrænset til specifikke data og tjenester vedrørende miljø, naturkatastrofer og sundhedspolitik, og ikke alle forhindringer som vedrører datapolitikker, er blevet fjernet effektivt.

¹⁶ Forordning 1089/2010 om gennemførelse af direktiv 2007/2/EF.

¹⁷ Med arbejdet på e-sundhedsnetværket, som er oprettet i medfør af direktiv 2011/24 om patientrettigheder, e-sundhed, digital serviceinfrastruktur, på e-recepter og på patientdatatjenester om udveksling af sundhedsoplysninger og den fælles aktion, der støtter e-sundhedsnetværkets rapport om "brug af cloudcomputing i sundhedsvæsenet", understøttes anvendelsen af data, der ikke vedrører den enkelte patients behandling.

¹⁸ Håndteres gennem Kommissionens ISA-program: <http://bit.ly/24DxWUs>.

¹⁹ Disse kan omfatte statistiske metadata af høj kvalitet fra officielle statistikker med henblik på at øge muligheden for søgning i, interoperabilitet af og integration af data.

²⁰ Der findes flere globale initiativer om emnet: Principperne om fair data, G8-landenes principper om åbne forskningsdata, RDA-retningslinjerne, Belmont Forums anbefalinger, OECD's principper og fagspecifikke retningslinjer.

For det tredje hæmmer **fragmentering** datadreven forskning²¹. Datainfrastrukturer er opdelt i det videnskabelige og det økonomiske domæne, efter land og efter styringsmodeller. Politikkerne for adgang til net, dataopbevaring og it-anvendelse er forskellige. Afkoblede og langsomme data- og databehandlingsinfrastrukturer er en hindring for videnskabelige opdagelser, de skaber siloer og gør udbredelsen af viden langsommere. Forskningsdata, der kan deles, værktøjer til analyse af frit tilgængelige data og netforbundne databehandlingsfaciliteter skal være tilgængelige for det store flertal af forskere²² i Europa, og ikke kun for forskere inden for de førende fagområder på vigtige forskningsinstitutioner. Endvidere arbejder europæiske universiteter og forskningscentre normalt inden for nationale grænser, og de mangler et miljø på europæisk plan til it-anvendelse, lagring og dataanalyse. Dette gør forskningssamarbejde i EU sværere, særligt databaseret tværfagligt samarbejde²³. I en nylig offentliggjort høring²⁴ svarede langt størstedelen af respondenterne, at en åben europæisk forskningscloud ville gøre forskningen mere effektiv, fordi man bedre ville kunne dele ressourcer på nationalt og internationalt niveau.

For det fjerde er der i Europa en stigende efterspørgsel efter **højtydende databehandlingsinfrastruktur i verdensklasse til at behandle data**²⁵ inden for forskning og ingeniørfag. Simulering af den kommende generation af fly, klimamodellering, forbindelsen mellem genomet og sundhed, forståelse af den menneskelige hjerne og in silico-test af kosmetik for at reducere antallet af test på dyr er alle opgaver, der kræver databehandling i exaskala. Selv om kvantedatabehandling på langt sigt giver løfte om at løse databehandlingsproblemer, der kræver mere, end hvad nutidens supercomputere kan yde²⁶, afhænger EU's konkurrenceevne også af højtydende databehandling for at kunne oprette paneuropæiske datainfrastrukturer.

På verdensplan gør USA, Kina, Japan, Rusland og Indien hurtige fremskridt. De har erklæret, at højtydende databehandling er en strategisk prioritet, de finansierer programmer til udvikling af højtydende databehandlingsøkosystemer (hardware, software, anvendelsesområder, færdigheder, tjenester og forbindelser) og arbejder på at idriftsætte supercomputere i exaskala²⁷. Europa er ikke med i kapløbet om at udvikle højtydende databehandling i overensstemmelse med vores økonomiske og videnskabelige potentiale. Vi sakker bagud i forhold til andre regioner, eftersom vi ikke investerer i højtydende databehandlingsøkosystemer, og vi høster derfor heller ikke gevinsterne af intellektuelle ejendomsrettigheder på dette område. På udbudssiden står EU's industri for ca. 5 % af verdens højtydende databehandlingsressourcer, mens vi forbruger en tredjedel af disse ressourcer. Da Europa i stadig større grad er afhængig af, at andre regioner leverer den kritiske teknologi, er der en risiko for, at Europa bliver teknologisk fastlåst, forsinket eller frataget strategisk viden. Europa sakker også bagud i forhold til rå samlet regnekraft: kun en ud af de ti førende højtydende databehandlingsinfrastrukturer ligger i EU, hvor Tysklands Höchstleistungsrechenzentrum i Stuttgart ligger på 8.-pladsen. USA har fem, og Kina har haft verdens hurtigste supercomputer siden 2013.

²¹ Høringen om forskning 2.0 understregede den manglende integration af eksisterende infrastruktur som en af forhindringerne for forskeres arbejde.

²² Forskere er enten ikke klar over, hvordan de skal opbevare og vedligeholde deres data (54 %), eller har ikke faciliteterne til at gøre det (37 %) (bit.ly/206u6hm)

²³ <http://bit.ly/1SkL9wm>

²⁴ <http://bit.ly/1JEymCY>

²⁵ Anmodninger om databehandlingscykluser vedrørende dobbelt tilgængelighed af PRACE (Partnerskab for avanceret databehandling i Europa): <http://bit.ly/1So2sgc>

²⁶ SWD(2016) 107.

²⁷ SWD(2016) 106.

Der er ikke en eneste medlemsstat, der selv har de finansielle ressourcer til at udvikle **det nødvendige højtydende databehandlingsøkosystem** i konkurrence med USA, Japan og Kina²⁸. Der er imidlertid ikke indtil videre taget nogen fælles skridt til at lukke hullet mellem den interne efterspørgsel og EU's udbud²⁹. EU har oprettet et kontraktuelt offentlig-privat partnerskab om højtydende databehandling for at udvikle teknologi i exaskala, men der findes ikke nogen europæisk ramme, som skal integrere det i computersystemer i stor skala.

Endelig skal ophavsmænd til og brugere af forskningsdata være i stand til at genbruge data og til at anvende avancerede analyseteknikker, som f.eks. tekst- og data-mining, i et miljø, der er mindst lige så driftsikkert, som deres egne faciliteter. Medlemsstaterne har i høj grad henvist til, hvor vigtigt det er at have EU-forskningsdata og sikre, at datadreven forskning giver det europæiske samfund en fordel³⁰. Enhver brug eller genbrug af forskningsdata skal sikre, at personoplysninger er behørigt beskyttet i henhold til EU's regler om databeskyttelse³¹. Disse og fremtidige revisioner af EU's lovgivning om ophavsret³² skal opstille generelle regler, som er relevante i den forbindelse.

Hvilke løsninger er der?

1. En åben europæisk forskningscloud

Den åbne europæiske forskningscloud 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 gratis, åbne og problemfri tjenester inden for lagring, forvaltning, analyse og genbrug af forskningsdata på tværs af landegrænser og videnskabelige faggrænser. Udviklingen af det skal drives af forskermiljøet, som er de mest avancerede brugere og de største ophavsmænd til videnskabelig forskning i verden. Den åbne europæiske forskningscloud vil også være tilgængelig med henblik på uddannelsesformål for videregående uddannelser og med tiden blive tilgængelig for regerings- og erhvervsbrugere, når de udviklede teknologier bliver tilgængelige for bredere anvendelsesområder.

Den åbne europæiske forskningscloud skal indledningsvis samle eksisterende videnskabelige datainfrastrukturer, som i dag er spredt på tværs af fagområder og medlemsstater. Den vil gøre adgangen til forskningsdata lettere, billigere og mere effektiv. Den vil skabe nye muligheder på markedet og nye løsninger på nøgleområder, som f.eks. sundhed, miljø og transport. Den åbne europæiske forskningscloud vil skabe et sikkert miljø, hvor beskyttelsen af privatlivets fred og data er garanteret via dens udformning på grundlag af anerkendte standarder, og hvor brugere kan vide sig sikre på datasikkerheden og ansvarsrisici. Den vil gavne andre af Kommissionens tiltag vedrørende fremme af åben forskning i Europa, som f.eks. fri adgang til videnskabelige publikationer og data i forbindelse med Horisont 2020, og samle nøgleinteressenter, der skal bidrage til at udforme de næste tiltag. Forvaltningen af en

²⁸ USA's forsvarsministerium vil investere 525 mio. USD i at opnå 3 præ-exaskala-systemer i 2017/2018 ("CORAL"). Japan planlægger at investere 1,38 mia. USD i at installere et nær-exaskala-system i 2019.

²⁹ Selv om PRACE giver mulighed for at dele visse medlemsstaters databehandlingsressourcer, er fremskaffelse af højtydende databehandlingssystemer et nationalt valg uden EU-koordinering eller -finansiering.

³⁰ Rådets konklusioner (8970/15).

³¹ COM(2012) 9 final

³² COM(2015) 626 final

åben europæisk forskningscloud vil blive fastlagt efter gennemførelsen af en grundig forberedelsesproces, som allerede er i gang.

Med henblik på at udvikle den åbne europæiske forskningscloud vil det i særdeleshed være nødvendigt at:

- **Gøre alle forskningsdata frembragt under Horisont 2020-programmet frit tilgængelige pr. definition.** Dette vil udvide det igangværende pilotprojekt³³, der betyder projekter indfører dataforvaltningsplaner, som gør data søgbare, tilgængelige, interoperable og genanvendelige³⁴.
- **Øge bevidstheden og ændre incitamentsstrukturerne** for akademikere samt tjenester i erhvervslivet og den offentlige sektor med henblik på at få dem til at dele deres data og forbedre uddannelsen inden for dataforvaltning, datakompetencer og datahåndteringsfærdigheder. Samtidig vil principperne og retningslinjerne for adgang til forskningsdata i Europa³⁵ blive revideret med henblik på at fremme og koordinere deres gennemførelse.
- Udvikle specifikationer for **interoperabilitet og datadeling** på tværs af fagområder og infrastrukturer på grundlag af eksisterende initiativer som f.eks. Forskningsdataalliancen (RDA) og Belmont Forum og lovbestemmelser som f.eks. INSPIRE-direktivet. Med tiden vil der blive taget hånd om alle standardiseringsbehov gennem prioriteterne for IKT-standardisering i forbindelse med det digitale indre marked.
- Skabe en **formålstjenlig paneuropæisk forvaltningsstruktur** for sammensluttede videnskabelige datainfrastrukturer og undgå fragmentering. Den institutionelle opbygning vil varetage den langsigtede finansiering, bæredygtighed, databevarelse og -forvaltning. Den skal bygge på eksisterende strukturer og involvere forskere, forskningsfinansierende og -gennemførende aktører³⁶.
- Udvikle **cloudbaserede tjenester for åben forskning**. Tjenesterne skal med støtte fra den europæiske datainfrastruktur give forskere mulighed for at finde og tilgå delte forskningsdata, anvende avanceret analysesoftware, bruge højtydende databehandlingsressourcer og lære om bedste praksis inden for digital forskning fra de førende fagområder.
- **Udvide det videnskabelige brugergrundlag** for den åbne europæiske forskningscloud med forskere og innovatorer fra alle fagområder og medlemsstater

³³ Pilotprojektet om fri adgang til forskningsdata under Horisont 2020 omfatter: Fremtidige og nye teknologier, forskningsinfrastruktur, informations- og kommunikationsteknologier, emnerne "nanosikkerhed" og "modellering" inden for nanoteknologi, avancerede materialer, avanceret produktion og forarbejdning samt bioteknologi, udvalgte emner inden for samfundsmæssige udfordringer: Fødevaresikkerhed, bæredygtigt landbrug og skovbrug, havforskning og forskning i indre vandveje samt bioøkonomien; klimatiltag, miljø, ressourceeffektivitet og råmaterialer – undtagen råmaterialer; Europa i en verden i forandring – rummelige, innovative og reflekterende samfund; Videnskab med og for samfundet, samt den tværgående aktivitet og fokusområdet smarte og bæredygtige byer. Bemærk, at projekter, som ikke er en del af disse "kerneområder", kan tilslutte sig på frivillig basis.

³⁴ De eksisterende muligheder for at fravælge datadeling, hvis fri adgang til data forhindrer en fremtidig kommerciel anvendelse, eller hvis deling strider mod reglerne for databeskyttelse og persondatabeskyttelse, sikkerheden eller beskyttelse af fortrolige EU-oplysninger, vil blive bevaret. Analysen af pilotprojekterne viste, at de fleste projekter anvender frit tilgængelige data, men at det er vigtigt at have en fravalgsmulighed.

³⁵ C(2012) 4890 final

³⁶ Blandt andre ESFRI, INSPIRE, Refleksionsgruppen vedrørende e-infrastruktur (eIRG), GEANT, PRACE, den europæiske biovidenskabsinfrastruktur for biologiske oplysninger (ELIXIR), Belmont Forum og lignende samlende initiativer.

samt fra partnerlande og globale initiativer, således at de kan bidrage til kvaliteten og få del i initiativets fordele³⁷.

Initiativet skal underbygge andre tiltag inden for åben forskning, som Rådet³⁸ og Europa-Parlamentet³⁹ har opfordret til at gennemføre, og tiltag i forbindelse med Kommissionens kommende politiske dagsorden for åben forskning. Det skal fremme bedste praksis inden for søgemuligheder i og adgang til data og hjælpe forskere med at få deres datafærdigheder anerkendt og belønnet, det skal bidrage til at gøre det lettere at efterprøve resultater og begrænse dataspild, f.eks. ved kliniske forsøgsdata (forskningsintegritet), det skal bidrage til at klarlægge finansieringsmodellen for generering og bevaring af data, begrænse profitsøgende adfærd ("rent-seeking") og klargøre markedet for innovative forskningstjenester (f.eks. avanceret tekst- og data-mining). Initiativet kan også bidrage til at afhjælpe problemer med dataclearing og beskyttelse af personoplysninger⁴⁰. Kommissionen vil høre interessenter og samarbejde med leverandører af FoU om behovet for at indføre retningslinjer for det videnskabelige område i forbindelse med EU's politik og lovgivning inden for databeskyttelse og om behovet for at sikre, at initiativet fra begyndelsen udformes på grundlag af de relevante retsprincipper.

Tiltag	Tidsplan
Kommissionen vil sammen med globale politiske partnere og forskningspartnere arbejde på at fremme samarbejde og skabe lige vilkår inden for deling af forskningsdata og datadreven forskning.	Fra 2016
Kommissionen vil bruge Horisont 2020-arbejdsprogrammerne til at skaffe finansiering til at integrere og konsolidere e-infrastrukturplatforme, til at samle eksisterende forskningsinfrastrukturer og cloudløsninger til forskere og til at understøtte cloudbaserede tjenester til åben forskning.	Fra 2016
Kommissionen vil gøre åbne forskningsdata til standardvalget, og samtidig sikre mulighed for at fravælge denne løsning, for alle nye projekter i Horisont 2020-programmet.	Fra 2017
Kommissionen vil revidere Kommissionens henstilling om adgang til og bevaring af videnskabelig information fra 2012 ⁴¹ med henblik på at fremme deling af forskningsdata og skabe incitamentsordninger, belønningssystemer og uddannelsesprogrammer til forskere og virksomheder, så de kan dele data, og så en tæt forbindelse til initiativet om frie datastrømme i forbindelse med	Fra 2017

³⁷ Nye initiativer fra Kommissionen kan finansieres gennem de europæiske struktur- og investeringsfonde (ESIF), forudsat at medlemsstaterne er enige i at finansiere dem og ændre deres operationelle programmer i overensstemmelse hermed.

³⁸ Rådets konklusioner (8970/15).

³⁹ Europa-Parlamentets betænkning 2015/2147(INI).

⁴⁰ Under overholdelse af artikel 7 og 8 i Den Europæiske Unions charter om grundlæggende rettigheder og gældende og fremtidige bestemmelser om brug af data til forskning kan initiativet f.eks. udvikle tjenester til tekst- og data-mining, der identificerer intellektuelle immaterielle rettigheder, adgangskontrol for forskellige brugere, uigenkaldelig anonymisering af følsomme data inden datafusioner, "personoplysningsrum", der skal værne om privatlivets fred og fremme udbredelsen af innovative anvendelser, eller lægge vægt på licenserings- og privatlivsmetadata, der kan læses af maskiner og vedhæftes datasæt, som er tilgængelige via cloudløsninger, samt indføre retningslinjer og bedste praksis for organisationsprocesser, der lever op til disse og understøtter initiativet. Selv om disse værktøjer og processer er tekniske, indbyggede og standardiserede, kan de bidrage til at nedbringe forekomsten af uredelighed og manglende overholdelse af lovbestemmelser.

⁴¹ C(2012) 4890 final.

det digitale indre marked sikres.	
Kommissionen vil samarbejde med medlemsstaterne om at forbinde de højtprioriterede europæiske forskningsinfrastrukturer ⁴² til den åbne europæiske forskningscloud.	Fra 2017
Sammen med interessenter og de relevante globale initiativer vil Kommissionen arbejde hen imod en handlingsplan for forskningsdatas interoperabilitet, herunder "metadata", specifikationer og certificering.	Inden udgangen af 2017

2. Den europæiske datainfrastruktur

Den **europæiske datainfrastruktur** vil, når den er indført fuldt ud, understøtte den åbne europæiske forskningscloud. Europa har behov for integreret højtydende databehandlingskapacitet i verdensklasse, højhastighedsforbindelser og førende data- og software-tjenester⁴³ for forskere og andre af sektorens førende brugere (herunder SMV'er) og den offentlige sektor. Denne infrastruktur vil give mulighed for at udnytte værdien af big data og digitalisering som standard⁴⁴. Den europæiske datainfrastruktur vil også hjælpe EU op blandt verdens førende magter inden for højtydende databehandling ved at **anvende supercomputere i exaskala omkring 2022 på baggrund af EU-teknologi, som skal bringe EU op blandt verdens tre førende**. Europa bør sigte mod at have mindst to kilder til denne teknologi.

Selv om den eksisterende strategi for højtydende databehandling⁴⁵ understøtter forskning og udvikling inden for højtydende databehandlingsteknologier, der kan markedsføres, omfatter den ikke indførelsen af supercomputere i exaskala. Den europæiske datainfrastruktur vil samle de nødvendige ressourcer og kapaciteter til at gennemføre overgangen fra forskning og udvikling til levering og drift af højtydende databehandlingssystemer i exaskala, som brugere og leverandører har udarbejdet i fællesskab. Dette vil omfatte dataforbindelser og opbevaring af big data for at sikre, at højtydende databehandlingstjenester er tilgængelige i hele EU, uanset hvor supercomputerne er placeret. Luxembourg, Frankrig, Italien og Spanien tog for nylig et skridt i denne retning med et **vigtigt projekt af fælles europæisk interesse (IPCEI) om anvendelsen af højtydende databehandling og big data**⁴⁶.

På grundlag af den paneuropæiske højtydende databehandlingsinfrastruktur og tjenester (PRACE), det transeuropæiske højhastighedsnet (GÉANT), det kontraktuelle offentlig-private partnerskab om højtydende databehandling⁴⁷, fællesforetagendet ECSEL⁴⁸ og IPCEI om anvendelsen af højtydende databehandling og big data vil Kommissionen og de deltagende medlemsstater:

- fremme et højtydende databehandlingsøkosystem, der er i stand til at udvikle ny

⁴² Som udpeget af Det Europæiske Strategiforum for Forskningsinfrastrukturer (ESFRI) <http://bit.ly/1pfqOe7>

⁴³ Herunder eksisterende tjenester OpenAIRE, EUDAT, Det europæiske vækstinitiativ (EGI), IndigoDataCloud, HelixNebula, PRACE, GÉANT.

⁴⁴ "Digitalisering som standard" henviser til tjenester og processer, der som udgangspunkt gøres tilgængelige online eller i digital form.

⁴⁵ COM(2012) 45 final.

⁴⁶ <http://bit.ly/1QxERan>.

⁴⁷ <http://bit.ly/1WZH8wF>.

⁴⁸ <http://www.ecsel-ju.eu>.

europæisk teknologi, som f.eks. **lavenergichips til højtydende databehandlingssystemer**⁴⁹

- integrere teknologier i systemprototyper, udvikle⁵⁰ løsninger i fællesskab og skaffe højtydende databehandlingssystemer. Den resulterende højtydende databehandlingsinfrastruktur skal fokusere på **de allerkraftigste supercomputere**, som er forbundet med mellemstore nationale databehandlingscentre i EU og den paneuropæiske data- og softwareinfrastruktur, så højtydende databehandling kan tilbydes som en tjeneste
- levere **problemfri, pålidelige og sikre højhastighedsforbindelser**, som skal gøre højtydende databehandling tilgængelig på tværs af EU. Det transeuropæiske højhastighedsnet (GÉANT) og National Research and Education Networks (NREN) forbinder allerede 50 millioner forskere og studerende. Disse infrastrukturer vil blive opgraderet, så de kan klare de stigende datamængder, der skal overføres, og brugergrundlaget kan udvides.

Den europæiske datainfrastruktur vil bidrage til digitalisering af industrien, til at udvikle europæiske platforme til nye, strategiske anvendelsesområder (f.eks. medicinsk forskning, luftfart og energi) og til at fremme industriel innovation. Den vil **udvide brugergrundlaget for højtydende databehandling** og give lettere adgang via cloudtjenester til både forskere inden for videnskabelige nøglefagområder og andre videnskabelige medarbejdere. Industrien, herunder navnlig SMV'er uden egne forudsætninger og offentlige myndigheder (inden for f.eks. intelligente byer og intelligent transport) vil kunne drage fordel af cloudbaserede og let anvendelige højtydende databehandlingsressourcer, anvendelsesmuligheder og analyseværktøjer⁵¹. I den forbindelse vil Kommissionen fremme indførelsen af behandlings- og udnyttelseskapaleteten for data fra Sentinel-satellitter, information om Copernicus-tjenester og andre jordobservationsdata, således at forskellige datasæt kan understøtte udvekslingen af ideer, at udviklingen af innovative produkter og tjenester kan fremmes, og at de socioøkonomiske fordele ved jordobservationsdata i Europa kan øges.

Den europæiske datainfrastruktur vil fungere i samspil med nationale og regionale, videnskabelige og offentlige datacentre. Den vil udvikle og indføre gode praksisser på grundlag af certificeringsordninger, fælles europæiske og globale standarder og specifikationer⁵², således at den nuværende mangel på interoperabilitet mellem nationale og fagspecifikke datacentre kan afhjælpes⁵³.

⁴⁹ Energieffektive maskiner i exaskala vil påvirke hele databehandlingsspektret og give Europa tekniske, økonomiske og sociale fordele. På nuværende tidspunkt ville en enkelt maskine i exaskala have behov for et kraftværk på 700 MW for at fungere, hvilket er nok strøm til 140 000 husholdninger i et år. Derfor er der behov for lavenergichips.

⁵⁰ Udvikling i samarbejde er en tilgang til at udvikle, hvor man aktivt involverer kunder og brugere i designprocessen for at sikre, at resultaterne lever op til deres behov og er anvendelige.

⁵¹ <http://bit.ly/1pqny20>.

⁵² RDA-Europe har sat gang i dialogen med den europæiske gruppe for flere interessenter om IKT-standardisering med henblik på at forelægge den bedste praksis for interoperabilitet i datainfrastrukturer, som er udviklet af forskningsdataalliancen (RDA- Research Data Alliance).

⁵³ Som f.eks. INSPIRE's interoperable geodata og -tjenestespecifikationer.

Den europæiske datainfrastruktur vil omfatte en forvaltningsstruktur for ledelse og udvikling af datainfrastrukturer og -tjenester⁵⁴ samt beslutningstagning inden for finansiering, bæredygtighed og sikkerhed på langt sigt. Forvaltningen af den bør involvere brugere (af den åbne europæiske forskningscloud og andre, der vil bruge systemet i længere tid, samt den offentlige sektor), gennemførende aktører (PRACE, GEANT) og investorer, og den bør bygge på eksisterende forvaltningsstrukturer.

Tiltag	Tidsplan
Kommissionen og de deltagende medlemsstater bør udvikle og indføre et højtydende europæisk databehandlingssystem og en højtydende data- og netværksinfrastruktur, herunder:	2016-2020
– Indkøb af to i samarbejde udviklede prototyper af supercomputere i exaskala og to driftssystemer, som skal være blandt de tre største i verden	fra 2018
– Etablering af et europæisk center for big data ⁵⁵	fra 2016
– Opgradering af backbonenettet for forskning og innovation (GEANT) og integrering af det europæiske net for offentlige tjenester.	fra 2016

Udnyttelse af kvanteteknologiens potentiale

Det næste gennembrud inden for højtydende databehandling og sikre net kan være baseret på kvanteteknologier. Førende virksomheder i Europa, Asien-Stillehavsområdet og Nordamerika er begyndt at investere i kvanteteknologi, men det er nødvendigt med et højere investeringsniveau, hvis der skal skabes produkter, som kan markedsføres. Europa skal være med i front i den kommende udvikling⁵⁶. Den europæiske datainfrastruktur bør komplementeres af et ambitiøst og langsigtet flagskibsinitiativ i stor skala, der skal give adgang til kvanteteknologiernes fulde potentiale, accelerere deres udvikling og skaffe offentlige og private brugere kommercielle produkter. Europa-Kommissionen vil påbegynde de indledende skridt til flagskibsinitiativet, herunder høring af interessenter og udarbejdelse af konsekvensvurdering, under hensyntagen til resultaterne af den foreløbige evaluering af Horisont 2020-programmet i slutningen af 2017⁵⁷.

Tiltag	Tidsplan
Europa-Kommissionen vil tage de indledende skridt til flagskibsinitiativet, herunder høring af interessenter og udarbejdelse af konsekvensvurdering ⁵⁸ ,	2016-2019

⁵⁴ Disse bygger på eksisterende tjenester, der udbydes af OpenAIRE, EUDAT, Det europæiske vækstinitiativ (EGI), IndigoDataCloud, HelixNebula, PRACE, GÉANT.

⁵⁵ Der eksempelvis hostes af Det Fælles Forskningscenter for data fra flere fagområder, men som har fokus på INSPIRE-/GEOSS-/Copernicus-geodata.

⁵⁶ <https://goo.gl/zBVi8N>.

⁵⁷ SWD(2016) 107.

⁵⁸ Konsekvensvurderingen vil være en del af forberedelsesprocessen af relevante finansieringsprogrammer efter 2020. Yderligere gennemførelsestiltag, som kan have væsentlig indvirkning, kan kræve en separat, individuel konsekvensvurdering.

under hensyntagen til resultaterne af den foreløbige evaluering af Horisont 2020-programmet i slutningen af 2017 ⁵⁹ med henblik på at igangsætte opstartsfasen i 2018 ⁶⁰ .	
--	--

3. Bredere adgang og opbygning af tillid

Udbredelsen af cloudtjenester i den offentlige sektor er ujævn og langsom⁶¹. Dette skyldes manglende tillid til og begrænsede synergier mellem den offentlige sektor og akademiske kredse. Fragmentering af datainfrastrukturer er en hindring for at opbygge kritisk masse og fælles løsninger for forskellige brugergrupper. **Brugergrundlaget for den åbne europæiske forskningscloud og for den europæiske datainfrastruktur vil blive udvidet til den offentlige sektor**, eksempelvis gennem pilotprojekter i stor skala, der omfatter e-forvaltning⁶² og interessenter fra den offentlige sektor, og gennem en progressiv åbning af den europæiske datainfrastruktur for **brugere fra erhvervslivet** og den offentlige sektor, indtil der nås en europæisk dimension. Den åbne europæiske forskningscloud vil med tiden sikre, at offentlige data er fuldt ud søgbare, tilgængelige og at forskere, politiske beslutningstagere og virksomheder kan udnytte dem. Erfaringerne vil give praktisk vejledning til de offentlige forvaltnings indførelse af cloudbaserede tjenester i hele Europa.

Eftersom den offentlige sektor genererer massive mængder data (f.eks. jordobservationsdata fra Copernicus og lokaliseringsdata fra INSPIRE) og har behov for større databehandlingskapacitet (f.eks. til trafikinformations- og rejseinformationssystemer, til anvendelse i intelligente byer eller udformning af politikker), vil den kunne få gavn af stordriftsfordele, fleksibilitet og kontinuitet. Offentligheden vil således drage fordel af billigere, hurtigere og bedre netforbundne offentlige tjenester og af bedre politikudformning på grundlag af prismæssigt overkommelige og sikre databehandlingstjenester og dataintensive tjenester.

Den åbne europæiske forskningscloud og den europæiske datainfrastruktur vil ligeledes være til fordel for virksomheder, herunder SMV'er, der mangler en omkostningseffektiv og let adgang til lagring af data, datatjenester og avanceret databehandling. Der vil løbende blive igangsat tiltag for at udvide brugergrundlaget til innovative SMV'er og erhvervslivet via data- og softwareekspertisecentre samt innovationsknudepunkter for datatjenester til SMV'er. Disse tiltag vil kræve et tæt samarbejde med den private sektor: SMV'er, store forsknings- og erhvervsbrugere af højtydende databehandling og sektoren for cloudtjenester, som skal være involveret fra begyndelsen.

Endvidere vil det europæiske cloudinitiativ skulle leve op til **høje standarder for kvalitet, pålidelighed og fortrolighed** for at sikre beskyttelse af personoplysninger og intellektuelle ejendomsrettigheder **samt sikkerhed** for så vidt angår modstandsdygtighed og beskyttelse mod uautoriseret adgang. Eksisterende faciliteter i den offentlige sektor – navnlig Connecting Europe-faciliteten (CEF) og den digitale infrastruktur for digitale tjenester, som er vigtige byggeklodser for at opnå tillid og sikkerhed – kan genanvendes og udrulles til forskermiljøet med henblik på at opnå besparelser, få let adgang og sikre overordnet sammenhæng. De overordnede regler fastsættes i de almindelige databeskyttelsesregler, i direktivet om net- og

⁵⁹ SWD(2016) 107.

⁶⁰ FET-flagskibe som beskrevet i referencedokumenterne til Horisont 2020.

⁶¹ SMART 2013/0043: Organisationer i den offentlige sektor var 10 % bagud i 2013 i forhold til den private sektor for så vidt angår brugen af cloudcomputingtjenester.

⁶² EU-handlingsplan for e-forvaltning 2016-2020 – fremskyndelse af forvaltningernes digitale transformation.

informationssikkerhed (NIS)⁶³ og i den reviderede udgave af EU's lovgivning om ophavsret. Eftersom cloudcomputing er karakteriseret ved at være verdensomspændende, er det vigtigt, at den europæiske dataøkonomi er forbundet med resten af verden, og at de globale standarder for databeskyttelse hæves til et højt niveau, der svarer til det europæiske.

Udarbejdelsen af passende standarder er en del af prioriteterne i IKT-standardiseringsplanen under det digitale indre marked⁶⁴. Der vil blive lavet en passende certificeringsordning på EU-niveau med henblik på at garantere sikkerhed, dataportabilitet og interoperabilitet i overensstemmelse med de juridiske krav⁶⁵, herunder den certificeringsordning, der allerede er beskrevet i den generelle databeskyttelsesforordning om sikker behandling af personoplysninger. Selv om der allerede findes en række certificeringsordninger⁶⁶, varierer deres anvendelsesområder i væsentlig grad, og der er ikke nogen fælles tilgang til minimumskrav inden for indkøb og forvaltning af cloudressourcer i den offentlige sektor. I den forbindelse vil samarbejdet med erhvervslivet og offentlige myndigheder matche erhvervslivets kapacitet med forskningsverdenens og den offentlige sektors krav.

Udvidelsen af adgangen til den åbne europæiske forskningscloud og den europæiske datainfrastruktur vil blive gennemført i takt med den relevante lovgivning, navnlig i forbindelse med genanvendelse af data til andre formål.

Tiltag	Tidsplan
I samarbejde med erhvervslivet og den offentlige sektor vil Kommissionen: - tilpasse løsninger inden for højtydende databehandling og big data til cloudmiljøer med henblik på at give bredere adgang, især til SMV'er - udvikle et økosystem, der skal styrke cloudindustrien i Europa, og som anvender den åbne europæiske forskningscloud som prøvesten for innovative løsninger inden for cloudteknologi - skabe en platform for de offentlige myndigheder, hvor de kan give fri adgang til deres data og tjenester, ved at skabe et udgangspunkt for offentlige forvaltningers rolle som leverandør af tjenester ("Government as a Service" - GaaS) i EU.	2016-20
For at lette udbredelsen af big data-teknologier vil Kommissionen stille et testmiljø til big data (pilotprojekter i stor skala) til rådighed for offentlige forvaltninger, bl.a. inden for rammerne af den foreslåede IPCEI.	Fra 2016
Kommissionen vil i samarbejde med erhvervslivet og medlemsstaterne fremme brugen af de relevante, eksisterende certificeringsmuligheder og -standarder samt indførelse af certificering og mærkning på EU-niveau, hvor det er relevant, især med henblik på at understøtte offentlige indkøb af cloudtjenester.	Fra 2016

⁶³ COM(2013) 48.

⁶⁴ COM(2016) 176.

⁶⁵ Forordning 765/2008.

⁶⁶ <https://resilience.enisa.europa.eu/cloud-computing-certification>.

Finansielle virkninger

Den digitale transformation i Europa forudsætter aktiviteter i stor skala. Adskillige kilder til finansiering i EU kan anvendes til det europæiske cloudinitiativ:

- Horisont 2020-rammeprogrammet for forskning og innovation (Horisont 2020).
- Connecting Europe-faciliteten (CEF)
- De Europæiske Struktur- og Investeringsfonde (ESIF)
- Den Europæiske Fond for Strategiske Investeringer (EFSI)

Det er nødvendigt med forskellige finansieringskilder for at understøtte hele investeringscyklussen. Store infrastrukturprojekter får i begyndelsen støtte fra offentlige tilskud, og efterhånden som de modnes, af markedsbaserede risikodelingsinstrumenter. Eftersom sådanne initiativer kræver en vedvarende og koordineret indsats, er fragmenteringen af tilgængelige budgetkilder imidlertid en klar ulempe.

Eksisterende finansieringsmuligheder under Horisont 2020-programmet vil give mulighed for at støtte den åbne europæiske forskningscloud og kickstarte den europæiske datainfrastruktur. En indledende vurdering af de yderligere midler, der er behov for i form af offentlige og private investeringer, lyder på 4,7 mia. EUR over en 5-årig periode. Dette omfatter 3,5 mia. EUR til datainfrastruktur⁶⁷, 1 mia. EUR til et flagskibsinitiativ i stor skala inden for kvanteteknologi i EU og 0,2 mia. EUR til tiltag, der skal give bredere adgang og skabe tillid. Supplerende bestemmelser vil blive drøftet med medlemsstaterne med henblik på at fremme støtten til den åbne europæiske forskningscloud ud over Horisont 2020-programmet. Initiativet vil med tiden skabe sin egen indtægt, når forskermiljøets, innovative opstartsvirksomheders og den offentlige sektors brug af det tager fart.

Kommissionen har til hensigt at fremlægge et forslag om, hvordan forskellige finansieringskilder i EU og på nationalt niveau kan kombineres med henblik på fuldt ud at nå målene i denne meddelelse. Kommissionen vil drøfte forslaget med medlemsstaterne efter behørig evaluering, konsekvensvurdering og høring. Infrastruktur med dette ambitionsniveau vil kræve et stort engagement fra medlemsstaternes side, navnlig ved at benytte strukturfonde og EFSI-garantier⁶⁸, men også fra store investeringer fra den private sektor og egnede koordineringsmekanismer. I den forbindelse viser det foreslåede vigtige projekt af fælleseuropæisk interesse (IPCEI) om højtydende databehandling og big data, hvilke muligheder og positive virkninger medlemsstaternes engagement har.

Tiltag	Tidsplan
Kommissionen vil i samarbejde med medlemsstaterne og interessenter undersøge, hvilke forvaltnings- og finansieringsmekanismer, der er egnede til den åbne forskningscloud og den europæiske datainfrastruktur, og fastlægge en køreplan for gennemførelse.	Fra 2016
Kommissionen vil foreslå metoder til at blande forskellige finansieringskilder, som skal drøftes med medlemsstaterne og interessenter, med henblik på at nå målene i denne meddelelse.	2016

⁶⁷ SWD(2016) 106.

⁶⁸ EIB's rådgivningstjenester under det europæiske centrum for investeringsrådgivning vil også være involveret.

KONKLUSIONER

Det europæiske cloudinitiativ er designet til at hjælpe forskere, erhvervslivet og offentlige myndigheder i Europa med at få tilgang til datainfrastruktur og cloudbaserede tjenester i verdensklasse, eftersom disse vil blive afgørende faktorer for succes i den digitale økonomi.

Et europæisk cloudinitiativ skal give ethvert forskningscenter, ethvert forskningsprojekt og enhver forsker i Europa adgang til højtydende databehandling, datalagring og analysekapacitet i verdensklasse, som er nødvendig, hvis de skal have succes i det globale, digitale innovationssystem.

Initiativet vil gøre det muligt at udvide brugergrundlaget for infrastrukturer og tjenester til den offentlige sektor og erhvervslivet, herunder SMV'er, at garantere et hensigtsmæssigt niveau for sikkerhed, dataportabilitet og interoperabilitet samt at leve op til EU's retlige krav.

Initiativets succes vil blive afgjort af, i hvor høj grad medlemsstaterne og den private sektor udnytter de fordele, der kan opnås ved at tackle denne udfordring, og i hvor høj grad de engagerer sig i at samarbejde om at få mest muligt ud af dem.