



VIDEN- SKABEN SKAL SÆTTES FRI

EU's konkurrenceråd vil gøre offentligt finansierede forskningsartikler- og data frit tilgængelig for alle, og det vil både kunne demokratisere videnskaben og højne kvaliteten af forskningsarbejdet, vurderer dansk forsker. Men begrænsede ressourcer, datarettigheder og betalingsmure er blot nogle af de forhindringer, der kan spænde ben for drømmen om at sætte videnskaben fri.

AF HENNING DUE, JOURNALIST
ILLUSTRATION BY DOMINIC KESTERTON

Al offentlig finansieret forskning skal være frit tilgængeligt inden 2020. Det var budskabet fra EU's videnskabs- og forskningskommissær Carlos Moedas i maj måned efter et møde i Rådet for Konkurrenceevne med både videnskabsministre og toppen af erhvervslivets interesseorganisationer fra EU-landene.

Den portugisiske EU-kommissær kaldte selv beslutningen en "skelsættende reform", og på det principielle plan er det et stort skridt nærmere målet for fortalene for såkaldt Open Science ('åben videnskab', red.) - en bevægelse, der vil gøre videnskabelig forskning- og data gratis tilgængelig for alle borgere.

Konkurrencerådets udmelding betyder dels, at alle publicerede videnskabelige artikler, som er delvist eller helt finansieret af offentlige kroner, ikke længere må låses

inde bag betalingsmure. I dag er det typisk kun forskningspersonalet på universiteter og andre forskningsinstitutioner, der har adgang til videnskabelige artikler og kun mod betaling.

Men udmeldingen betyder også, at al den forskningsdata, som offentligt finansierede videnskabsartikler bygger på, skal lægges ud til fri afbenyttelse for alle, medmindre der ifølge Rådet "er velbegrundede argumenter for ikke at gøre det". Det kunne fx være "intellektuelle ejendomsrettigheder, sikkerheds- eller privatlivsproblemer", som rådet formulerer det i en pressemeddelelse.

Først og fremmest handler EU's udgave af Open Science om at skabe flere arbejdspladser og hælde nyt brændstof på EU-landenes vækstmotor, der har kørt i slæbegearet siden den internationale finanskrisen. Tanken

er at gøre Europa så attraktivt som muligt for forskere, iværksættere og investeringslystne virksomheder ved at dele viden gratis.

Men vejen til et sådant EU kan blive besværlig og dyr. Det vurderer to forskere med speciale i indsamling og behandling af videnskabsdata, som *Magasinet P* har talt med. De påpeger, at der foreløbigt er en verden til forskel på, hvad EU-politikerne drømmer om, og hvad der rent praktisk kan lade sig gøre.

FLERE BENSPÆND

Christine Borgman er professor og leder af Information Studies ved University of California, USA, og udgav sidste år en bog om emnet, *Big Data, Little Data, No Data: Scholarship in the Networked World*.

- Open Science handler om at give skatteborgerne adgang til forskning og om at gøre forskerne ansvarlige for deres arbejde i forhold til dem, der betaler gildet. Så hvis EU's medlemslande betaler for at få produceret forskning, skal deres borgere også have adgang til resultaterne. Det er svært at være uenig i. Men her er det vigtigt at sondre mellem at give fri adgang til de videnskabelige artikler, der udgives, og den data, som artiklerne bygger på, siger hun.

Christine Borgman forklarer, at forskningsdata er for komplicerede til, at man bare kan lægge dem ud på en server til fri afbenyttelse som fx en tv-serie, mp3-sang eller en spillefilm. Skal forskere eller andre almindeligt interesserede mennesker kunne bruge forskningsdata til noget som helst, kræver det, at dataene rengøres, gøres delbare og læsbare, og det kræver et stort stykke ekstraarbejde, forklarer hun.

- Og så er der spørgsmålet om, hvad der overhovedet menes med forskningsdata? Er det tal i et Excel-ark? Er det 1'er og 0'er på en harddisk? Er det ubehandlet rådata? Det mangler EU også at afklare, siger hun.

Irene Pasquetto, ph.d.-studerende ved UCLA Graduate School of Education & Information Studies, USA, ser også flere problemer i EU's målsætning. Hun har speciale i kritiske datastudier med fokus på politiske og økonomiske konsekvenser af åben data og videnskab og undersøger i sit ph.d.-projekt, hvordan ekstremt komplekse genom-data (undersøgelse af alle gener på én gang, red.) kan gøres genbrugelige, selvom computerdata-kraften endnu ikke er helt tilstede.

- Det er uklart, hvordan beslutningen skal udføres i praksis. Hvem er det, der skal opgive deres rettigheder? Er det forskerne bag artiklerne, der skal opgive deres datarettigheder, eller er det udgiverne af artiklerne?, spørger Irene Pasquetto.

Irene Pasquetto har brugt det seneste års tid på at undersøge, hvad frit tilgængelig data fra seks forskellige forskningsområder reelt er blevet brugt til, og hun

vurderer, at det vil kræve mellem 30 og 50 procents mere arbejde for et forskerhold at gøre de indsamlede data klar til, at andre mennesker kan få nytte af dem.

Et andet problem er ifølge Christine Borgman, at det i praksis er svært at adskille dataindsamlingen fra selve databearbejdningen og dataanalysen, og hun efterlyser også en beskrivelse af, hvem der skal have nytte af forskningsdataene.

- Skal dataene bruges af andre forskere ti år ude i fremtiden eller af skolebørn om 50 år? At gøre data videnskabeligt brugbare til kommende generationer kræver en bestemt type arkiver, der ikke bliver forældede, og derfor er tidsperspektivet og målgrupperne for dataene afgørende, siger hun.

Den amerikanske professor peger desuden på, at det heller ikke fremgår hvem, der skal samle regningen op for ekstraarbejdet. Og vil dataene overhovedet blive brugt af andre, eller om de ender med at dø en langsom død i internettets såkaldte mørkearkiver?

- Det er 50.000 euro-spørgsmålet, siger Christine Borgman.

Da *Magasinet P* stiller Irene Pasquetto spørgsmålet, smiler hun, inden hun svarer via Skype-forbindelsen.

- Paradoksalt nok findes der i dag næsten ingen data, der viser, hvor meget frit tilgængelig data der anvendes af andre forskere eller andre mennesker. Det, jeg foreløbigt kan se i mit forskningsarbejde, tyder på, at dataene ikke anvendes særligt meget, siger hun.

Alligevel har mange politikere i årevis drømt om at sætte forskningen fri. Fx søsatte EU-Kommissionen for nylig den foreløbigt største politiske satsning på området i form af European Open Science Cloud, en digital infrastruktur, som skal gøre det muligt for Europas cirka 1,7 millioner forskere at uploade, dele og genbruge forskningsdata.

Men det internationale forskningsmiljø er mere splittet i spørgsmålet, og det er der flere grunde til.

SKABER SPLID

For det første vil fri adgang til forskningsartikler og forskningsdata kunne frarøve de førende internationale videnskabelige tidsskrifter en milliardfortjeneste, som i dag hentes hjem ved at sælge videnskabelige artikler stykvis og via dyre abonnementer til offentlige og private forskningsinstitutioner i hele verden.

En anden bekymring så dagens lys i januar måned i en leder i det anerkendte videnskabelige tidsskrift *The New England Journal of Medicine*, hvori forfatterne luftede deres bekymring for, at en ny type forsker - såkaldte "forskningsparasitter" - vil dukke op i strømmen af frit tilgængelig data og "overtager systemet". "Mennesker, der intet har haft at gøre med at designe eller udføre en undersøgelse, men bruger en anden

gruppes data til deres eget formål og endda måske bruger dataene til at underkende det, de oprindelige forskere havde foreslået”, hed det i lederen.

Uenigheden var til at tage og føle på, da chefredaktøren fra selvsamme tidsskrift få dage senere gik kontra sammen med en gruppe redaktører fra andre medicinske tidsskrifter i en leder i tidsskriftet *Annals of Internal Medicine*. I lederen understregede redaktørerne, at forskere ”er etisk forpligtede” til at dele data, som er blevet indsamlet via kliniske forsøg, da deltagerne i forsøgene ”har sat deres helbred på spil”.

Godt nok har medicinalgiganter som GlaxoSmithKline, Johnson & Johnson og Medtronic grebet i egen barm og valgt at dele udvalgt data fra medicinske forsøg via Yale University’s Yale Open Data Access Project, men opfordringen rammer stadig et ømt punkt hos de dele af erhvervslivet og forskermiljøet, der ønsker at holde fast i den lukkede model.

- Som udgangspunkt deler forskere helst ikke data med nogen, heller ikke med de tidsskrifter, der publicerer forskernes arbejde. Her taler vi især om data fra kliniske medicinske forsøg, som kan være meget profitable, fordi de kan vise sig at bane vejen for fremstilling af ny medicin, siger Irene Pasquetto.

Men bekymringerne handler ikke kun om udsigten til tabt profit og tabet af intellektuelle rettigheder til nye og potentielt guldrandede medicinprodukter, men også om at indsamling af forskningsdata ofte tager tid, og at belønningen for arbejdet ikke er særlig stor.

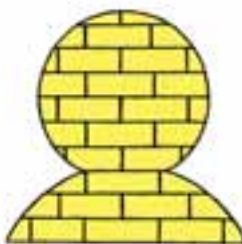
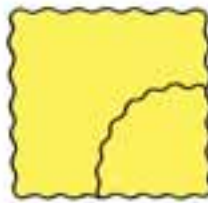
POLITISK MISBRUG

Endelig kan dataene også misbruges til politiske formål, påpeger Gretchen Goldman, der er analysechef hos amerikanske Center for Science and Democracy ved Union of Concerned Scientists (UCS).

Hun peger på et eksempel fra 2013, da det republikanske medlem af Repræsentanternes Hus, Lamar Smith, krævede adgang til rådata fra en skelsættende undersøgelse af luftforurening, den såkaldte Harvard Six Cities-undersøgelse, der viste en sammenhæng mellem luftforurening og sygdomme og dødelighed hos 8.000 amerikanere, og som det amerikanske miljøagentur, the Environmental Protection Agency, har fremhævet adskillige gange i forsøg på at stramme lovgivningskravene til luftforurening fra kraftværker.

Lamar Smith antydede, at dataene fra undersøgelsen var blevet manipuleret, og den slags eksempler bør ifølge Gretchen Goldman tjene som en advarsel.

- Mine forbehold handler om, hvordan vi bevarer den integritet, som forskningsdata- og artikler har opnået, især på grund af peer review-systemet. Hvis vi helt fjerner kontrolmulighederne, bliver vi mere sårbare over for misinformationer, og den opgave har de videnskabelige



tidsskrifter, godt nok mod betaling, løftet historisk set, siger Gretchen Goldman.

Set i det lys varsler EU-politikernes drøm om fri adgang til forskningsartikler og forskningsdata også et opgør med peer review-processen, altså forsker-bedømte videnskabsartikler, der har været praksis i den videnskabelige verden i mere end hundrede år, og hvor en videnskabelig artikel, som et tidsskrift ønsker at udgive, på papiret kvalitetssikres ved at blive sendt til faglig bedømmelse hos to eller tre eksperter inden for området, som kigger artiklen efter i sømmene for eventuelle fejl eller mangler.

En proces, der ifølge fortalere som Gretchen Goldman er med til at sikre, at udgivne forskningsartikler hviler på tilstrækkeligt solide data og forskningslitteratur.

Men vendes blikket mod den brutale virkelighed, er billedet mere mudret, påpeger David Budtz Pedersen. Han er lektor og forskningsleder ved Humanomics Research Centre på Aalborg Universitet København og har i årevis arbejdet med at omsætte Open Science-principperne til praksis i ind- og udland.

- Peer review vil i lang tid fremover være den gyldne standard for kvalitetssikring af videnskab, men i praksis er systemet under et stort pres, fordi forskerne har svært ved at finde tid til arbejdet. Det betyder, at de

videnskabelige tidsskrifter alligevel ender med at give spalteplass til dårlig forskning, og tidspresset bliver kun større af, at alt muligt skal peer reviews i dag - lige fra forskningsansøgninger til videnskabelige priser og udmærkelser og hele universitetsafdelinger, siger David Budtz Pedersen.

Hans største problem med peer review-processen er, at den foregår i det skjulte, og det er Open Science-bevægelsen netop i fuld gang med at revolutionere, vurderer han.

- Udfordringen er så at sikre, at den mest interessante forskning finder vej til læserne. En løsning er at bruge antallet af downloads af en artikel som en kvalitetsindikator, siger han.

Men Gretchen Goldman mener ikke, at den type parametre siger ret meget om en forskningsartikels og de bagvedliggende datas kvalitet.

- Fri adgang til data og artikler kan medføre flere synspunkter og meninger og kan betyde, at et stykke forskningsarbejdes skæbne ikke kun overlades til en lille gruppe tidsskriftsredaktører. Men almindelige mennesker er ikke i stand til at vurdere kvaliteten af arbejdet, og så ender kvalitetskontrollen måske alligevel i hænderne på en lille gruppe forskere, siger Gretchen Goldman.

David Budtz Pedersen peger også på såkaldt åbne evalueringer som et bud på, hvordan forskning kan kvalitetssikres i en åben proces fremover.

- EU har allerede udtrykt ønske om officielt at anerkende det åbne evalueringsarbejde af videnskabelige artikler, og det giver flere forskere et incitament til at gå ind i processen, siger han.

David Budtz Pedersen er ikke i tvivl om, at forskningskvaliteten på den lange bane øges ved mere åben vidensdeling.

- En åben videnskabelig kultur vil - alt andet lige - gøre det vanskeligere for forskere at skjule bias og manipulere data.

Hans rationale lyder, at når forskningen udføres i det åbne, er det nemmere at få adgang til baggrundsdata, som betyder, at forskningsresultater lettere kan testes.

Af samme grund ser mange forskere ifølge David Budtz Pedersen fri dataadgang som en del af løsningen på replikationskrisen (at videnskabelige forsøg ofte ikke kan gentages med samme resultat, red.) inden for bl.a. det biomedicinske felt.

Set gennem hans sorte hornbriller har Open Science-bevægelsen potentiale til både et radikalt opgør med den dominerende videnskabelige arbejdsproces og forskernes arbejds motivation.

- Belønningssystemet i forskningsverdenen er meget individualiseret, og den enkelte forsker får anerkendelse og status ved at hjemtage bevillinger eller udgive artikler. Den model kalder David Budtz Pedersen "uholdbar",

fordi laboratorieprofessoren typisk løber med æren for sine medarbejdes arbejde.

- Hvis Open Science skal blive til virkelighed, kræver det nye incitament, der opmuntrer til samarbejde, og fælles problemløsning er nødvendige, siger han.

Skal forskningsverdenen vende ryggen til konkurrence-logikken, skal nye generationer af forskere ikke tvinges til at konkurrere med hinanden om spalteplass og sparsomme forskningsmidler. De skal i stedet have øjnene op for, at de kan dele viden og arbejde sammen om et fælles mål, og at de endda kan blive belønnet for en fælles indsats.

David Budtz Pedersen peger blandt andet på et af verdens største videnskabelige forskningscentre the European Organization for Nuclear Research (CERN), uden for Geneve, som et friskt eksempel på en mere samarbejdsorienteret forskningspraksis.

- Hos CERN bliver de videnskabelige opdagelser til i et samarbejde mellem tusindvis af forskere og tekniske medarbejdere, der alle har en aktie i opdagelserne, og på de udgivne forskningsartikler nævnes flere end 1.000 forskere.

Et andet eksempel finder han hos Open Access-tidsskriftet *Open Mind*, der udgives af MIT Press og giver forskere inden for kognitionsvidenskaben mulighed for at publicere forskningsartikler mod et gebyr.

- Artiklerne vurderes i et åbent peer review-forløb, og man kan følge hele artiklens tilblivelse. Processer, der tidligere var lukkede, åbnes her op, og det er i den retning, progressive forskere bevæger sig, siger David Budtz Pedersen.

Han vurderer, at vi skal kigge mellem 10-20 år ind i fremtiden, før Open Science-kulturen for alvor har slået rod i forskningslaboratorierne.

- Forandringen kan kun ske ved eksemplets magt og ved at skabe et politisk og folkeligt pres, der kræver en mere gennemsigtig videnskab og baserer sig på samarbejde frem for konkurrence. Det vil skabe en mere demokratisk, mindre elitær og hierarkisk videnskab.

I oktober måned tog EU et stort skridt nærmere en mere åben videnskabssektor, da Det Europæiske Medicinagentur, EMA, besluttede at give fri adgang til forskningsdata og kliniske rapporter fra al nygodkendt medicin, som får tilladelse til at blive solgt i EU. Hver rapport indeholder typisk flere tusinde siders data, og målet er at give adgang til omkring 4.500 kliniske rapporter om året - senest 60 dage efter, EMA har besluttet, om et produkt skal have tilladelse til at blive markedsført og solgt i EU. Men opgaven skal ifølge en pressemeddelelse fra EMA løses "med eksisterende ressourcer". ●