

Forskningsrådets policy for tilgjengeliggjøring av forskningsdata

Marte Qvenild, Avd. for forskningsinfrastruktur



Datarevolusjonen

"We never, ever in the history of mankind have had access to so much information so quickly and so easily" (Internet pioneer Vinton Cerf)



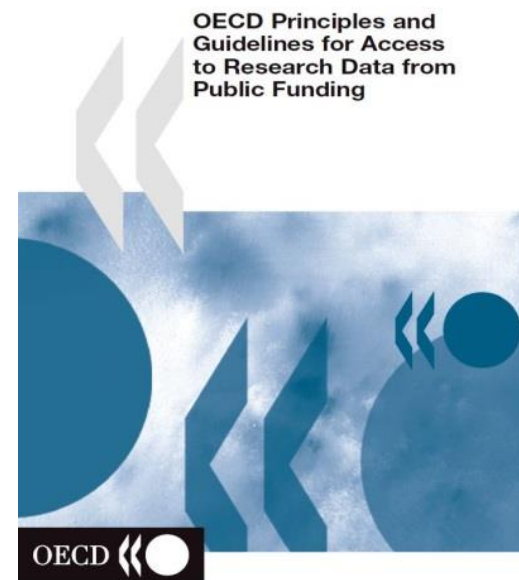
Institutt for teoretisk astrofysikk
inn 2,5 TB data pr dag ved gode
forhold

"Open access to data in agriculture is one of the key issues that we consider important if you want to combat hunger and poverty" (Johannes Keizer, FAO)

"Quite simply, knowledge is the engine of our economy. And data is its fuel." (Neelie Kroes, tidligere visepresident i EU kommisjonen)



"Verdien i data ligger i bruken. Full og åpen tilgang til vitenskapelige data bør implementeres som den internasjonale normen for utveksling av forskningsdata finansiert av offentlige midler" (OECD 2007, Principles and Guidelines for Access to Research Data from Public Funding).



- Oppfølging av foreslåtte tiltak i OECD rapporten i stortingsmeldinger:
 - *Klima for forskning* (2008-2009)
 - *Lange linjer – kunnskap gir muligheter* (2012-2013)
- Forskningsrådets policy «Tilgjengeliggjøring av forskningsdata» (2014) – anbefalinger, ikke krav



Forskningsrådets policy (2014)

- Policyen dekker data generert gjennom forskning som er støttet av Forskningsrådet
- Åpen som standard, men unntak for data
 - som må beskyttes av sikkerhetshensyn, personvern eller andre juridiske forhold
 - som skal benyttes kommersielt
 - der kostnaden ved tilgjengeliggjøring er høyere enn verdien av åpen tilgang
- Anbefalinger fremfor absolutte krav, men krav om arkivering: «forskningsgenererte data, deriblant alle data som ligger til grunn for publikasjoner, skal arkiveres i sikre, digitale arkiv. Slik arkivering skal skje snarest og senest innen to år etter at prosjektet er avsluttet.» (Forskningsrådets Generelle vilkår for FoU prosjekter (2015) punkt 8.2.)



Photo: Shutterstock



Photo: Photodisc

Anbefalinger fra policyen

Forskningsdata bør

- arkiveres på en sikker måte
- gjøres tilgjengelige for alle relevante brukere, under like vilkår
- gjøres tilgjengelig så tidlig som mulig, men en embargoperiode for å publisere er ok
- utstyres med metadata basert på internasjonale standarder
- utstyres med lisenser som legger så få begrensninger som mulig på tilgang, gjenbruk og videredistribusjon (for eksempel Creative Commons Attribution license (CC-BY))
- gjøres tilgjengelig til lavest mulig kostnad; maksimalt de faktiske kostnadene knyttet til tilgjengeliggjøring
- utstyres med en langtidsplan



Photo: Shutterstock/Marynchenko Oleksandr

```
<!DOCTYPE html PUBLIC "-//W3C//  
<html xmlns="http://www.w3.org/  
<head>  
<meta http-equiv="Content-T  
<meta http-equiv="Content-  
<meta http-equiv="Content-  
<title>Document Title</titl  
<link rev="made" href="mail  
<link rev="start" href=". /'  
<style type="text/css" medi  
@import "style/boas.css"
```

Photo: Shutterstock/Alexey V Smirnov



Illustration: Creative Commons

- Kostnader i prosjektene kan inkluderes i søknader om støtte fra Forskningsrådet
 - Kostnader ved å gi åpen tilgang til resultatdata
 - Kostnader ved å få tilgang til kilde data

Meld. St. 27 (2015-2016) – «Digital agenda for Norge - IKT for en enklere hverdag og økt produktivitet»

Tilgjengeliggjøringen av offentlige data
skal styrkes gjennom
strategier/handlingsplaner for fem
områder:

- Kultur
- Geodata
- Statlige utgifter
- Transport
- **Forskningsdata**



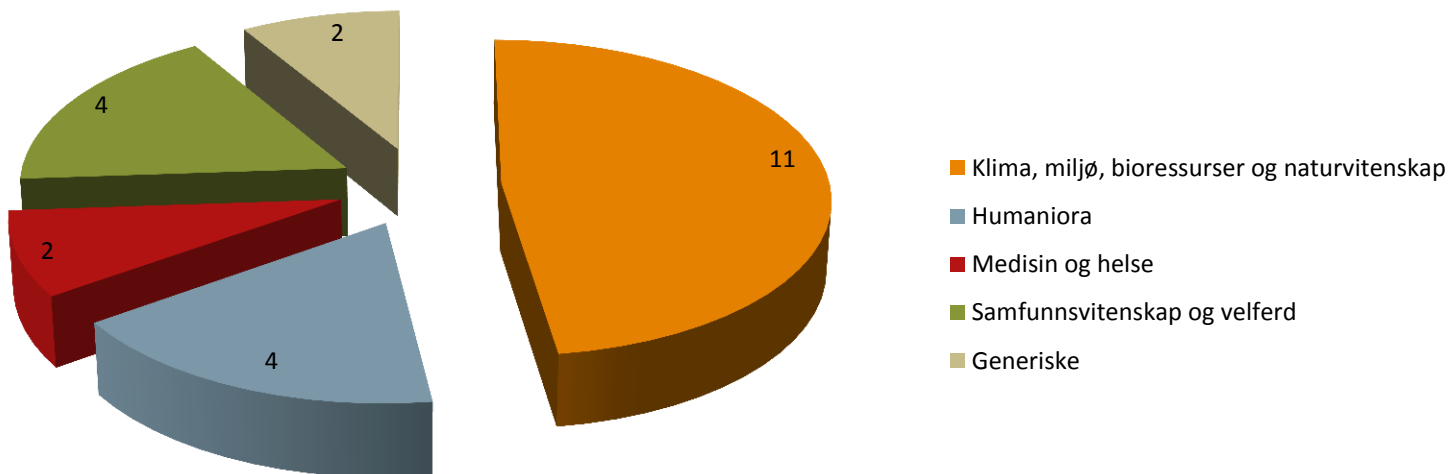
Hva Forskningsrådet skal gjøre:

- Kartlegge Forskningsrådets finansiering av infrastrukturer som støtter tilgjengeliggjøring av forskningsdata
- Kartlegge hvilke tjenester for datalagring som finnes i dag
- Gi innspill til arbeidsdeling og rolleavklaring mellom tjenesteleverandører
- Gi innspill til mulige finansieringsmodeller for nasjonalt viktige tjenester knyttet til lagring og deling av data

Avgrensning og fremgangsmåte

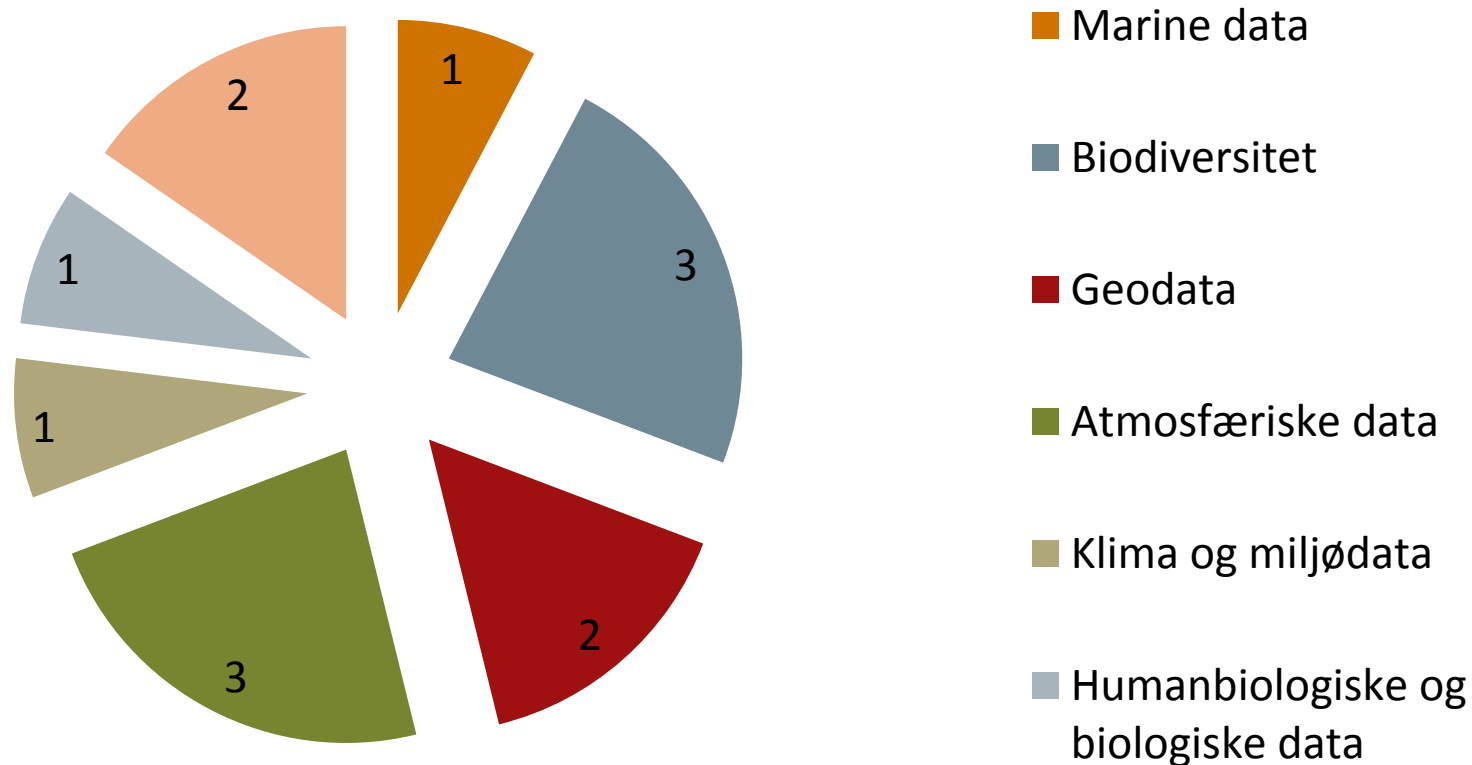
- Systematisert alle bevilgninger til *datainfrastrukturer* fra Nasjonal satsing på forskningsinfrastruktur
 - Fire utlysninger hittil: 2009, 2010, 2012 og 2014 (og 2016)
- 25-30 intervjuer (pågår fremdeles)
 - Universiteter
 - Forskningsinfrastrukturer
 - Offentlige institusjoner
 - Forskere
 - Interesseorganisasjoner
- Deltar i arbeidsgruppe nedsatt av OECD Global Science Forum som utreder hvordan data-infrastrukturer kan gjøres økonomisk bærekraftige

Oversikt over Forskningsrådets investeringer i datainfrastrukturer



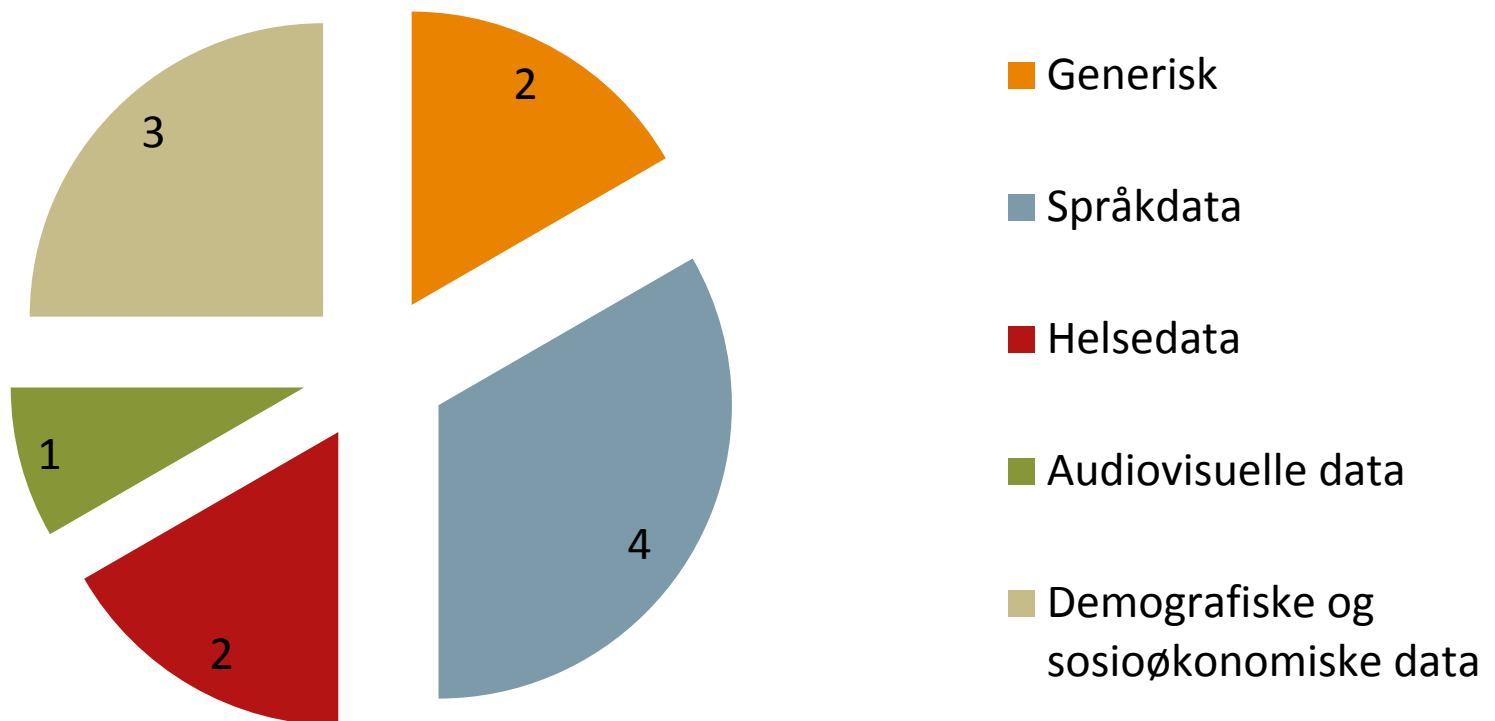
Klima, miljø, bioressurser og naturvitenskap

Antall datainfrastrukturer



Øvrige datainfrastrukturer (samfunnsvitenskap, humaniora, medisin og helse, og generiske tjenester)

Datainfrastrukturer pr fagområde



Hvor kan forskere få tilgang til data?

Hvor kan forskere lagre data?

- Stor variasjon i type data forskere har tilgang til
- Hvem som har tilgang til disse dataene varierer
- Ulike autoriseringsløsninger
 - Tilgang uten brukerid
 - Tilgang med FEIDE/eduGAIN
 - Egen løsning for tilgang basert på brukerid/passord
 - Tilgang fra godkjente institusjoner
 - Åpen søke- og metadataportal



Ikke mange norske, generiske "repositorier" pr i dag

- Norsk senter for forskningsdata (NSD)
- NorStore (Uninett Sigma2)
- Tromsø Open Research Data (UiT)
- BIRD (BIBSYS)



Noen problemstillinger som de ulike aktørene har løftet fram:

- Finansiering av arkiver/repositorier for forskningsdata må være tilstrekkelig langsiktig
- Finansiering av tilgang til arkiver opprettet for forvaltningsformål
- Bruk av helsedata i forskning
- Incentiver for publisering og deling av data
- Harmonisering med internasjonale løsninger
- Avklaring av hvilke data som skal tas vare på