

Hur söker du efter forskningsdata?

Iris Alfredsson (SND/Göteborgs universitet) och
Olivia Ekman (Umeå universitet)
SND Nätverksträff 8/11 2023



Diskussionsfrågor

Dela upp er i grupper om 4-5, skriv ned en sammanfattning av er diskussion.

- Hur gör ni på ditt lärosäte, söker ni data? Vill ni göra det?
- Vad behövs för att du ska kunna göra det?
- Har du testat söka i repositorier?
- Något repositorium som funkar bättre än andra?



Sammanfattning av diskussionen



Skillnader och likheter mellan litteratursökning och forskningsdatasökning

Skillnader

- Data - inte bara text
- Kan utgöra STORA mängder data
- Saknar uniform datastandard

Likheter

- Metadata
- Samma söklogik
- Leverera sökresultat som EndNote/Zotero-fil

Varför använder forskare annans data?

Table 2. Background purposes for seeking data.

Discipline	Comparison	Evaluate systems/ algorithms	Overview of new topic/method	Learning data-centric skills	Supporting others	Verifying own results	Making predictions	Teaching/ student learning	Preparing proposals
Acoustical engineering	15					15			
Business and finance			2						
Ecology – industrial and evolutionary			16					4	4
Information/ computer science	9, 10	5, 7, 8, 14		9		8			
Librarians/literature manager					20, 21, 22				
Molecular biology	13						13		
Psychiatry, medicine, neuroscience	12		12						
Psychology, cognitive psychology, neuroscience				11					
Water resources						6			
Astronomy (reported)	6								
Students (reported)								20	

Blue indicates researcher, green indicates support, orange indicates industry and purple indicates student (For interpretation of the references to colours in the table, refer to the online version of this article).

Table 3. Foreground purposes for seeking data.

Discipline	Model/ system inputs	New questions of existing data	Research using social media data	Identify new substance, organism	Replication
Chemistry				19	
Ecology – industrial and evolutionary	4, 16	4, 16			
Information/computer science	7, 8, 14	9	2, 9		
Medical statistics					1
Palaeontology				17	
Popular culture			18		
Psychology, cognitive psychology, neuroscience		11 (planned)			11
Water resources	6				
Climate change (reported)		21			
Sociologists (reported)		21			

Blue indicates researcher and yellow indicates citizen (For interpretation of the references to colours in the table, refer to the online version of this article).

Steg i en systematisk översikt



CC-BY-SA
Illustration skapad av Karolinska Institutet Universitetsbiblioteket.

Välja ämne

Avgränsa frågan &
upprätta ett protokoll

Litteratursökning



Olika strategier kan
behövas beroende på
ämne.

- Varför söker du data? (Behöver du öppen data eller har du tid att begära ut data?)
- Vad är din forskarfråga?
- Metod, andra avgränsningar
- Kända forskare/institut inom fältet
- Gör sedan första provsökning i zenodo/figshare, efter förfining sök i relevant ämnesspecifikt repository (samt i databaser för att hitta artiklar)

Sök bland:

- Databaser
- Repositorier
- DMP
- Patent
- Artiklar
- Fråga kollegor



Att tänka på

Först efter nedladdning av data vet forskaren om data är användbart.
Metadata kan vara bristfällig.

Utgå inte från antal nedladdningar som aktualitet markör.

Återvinningssystem för data bygger på nyckelordssökningar lämpade för artikelsökning - ej anpassad efter data.



Hur går en sökning till?

Boolsk logik fungerar oftast (AND, OR, “”, NOT, ()), wildcard)

Läs hjälptexten

FigShare exempel: `:title: Line balancing for improving production AND :keyword: cancer cell`

Zenodo exempel: `+metadata.title:"open science" -metadata.title:policy` or e.g. `metadata.title:(-open +science)`



Några exempel

Var börjar vi leta?

Vet vi vad vi vill ha eller behöver vi starta med en översikt?

Kan EOSC hjälpa mig?

Några exempel från samhällsvetenskap - CESSDA & ICPSR



Översikt

re3data.org

Globalt register över repositorer från olika discipliner

<https://www.re3data.org/search>



Några EOSC-relaterade initiativ

EOSC Marketplace

<https://marketplace.eosc-portal.eu/>

SSH Open Marketplace

<https://marketplace.sshopencloud.eu/>

GoTriple

<https://www.gotriple.eu/>



CDC - CESSDA Data Catalogue

<https://datacatalogue.cessda.eu/>

Innehåller dataset från 14 av CESSDA:s 22 Service Providers. Fler på väg att ansluta sig.

[User Guide](#)

[Introduktionsvideo](#)



CDC - byggstenar

[ELSST](#) - European Social Science Thesaurus

En flerspråkig thesaurus som består av över 3 300 begrepp översatta till 16 språk. Ägs av CESSDA och dess SP:s.

[Vocabulary Service](#)

Service för kontrollerade vokabulärer

[Metadata Validator](#)



Hitta data hos ICPSR

<https://www.icpsr.umich.edu/web/pages/ICPSR/index.html>

- Search/Compare Variables
- Discover Data via the Literature
- Thematic Collections



Hur går vi vidare?

Vårt förslag:

Börja med studenter. Ta fram ett material för metodkurser där studenter lär sig söka data. Ta fram en MOCC som alla kan använda och bearbeta.

Men vad mer: Nätverk? Arbeta tillsammans? Hur?