

## Verksted: Metadata og verktøy for nettarkiv-forskning (2 x 45 minutter)

Jon Tønnessen

*Målgruppe:* Forskere, bibliotekarer i UH

*Forutsetninger:* Deltakerne trenger egen datamaskin

Nasjonalbiblioteket (NB) har siden 1990-tallet samlet inn og bevart store mengder innhold fra internett. Dette nettarkivet er etterspurt blant forskere og har stort potensial for akademisk kunnskapsproduksjon, men har inntil nylig vært utilgjengelig for forskere. Gjennom WebData-prosjektet skal det bygges en nasjonal forskningsinfrastruktur for nettdata for å muliggjøre søk og utforskning, dataauthenting og analyse.<sup>1</sup>

Infrastrukturen skal utvikles i dialog med forskere og baseres på forskernes behov. Vi inviterer derfor deltakere på DHKO2025 til å være med og vurdere hvilke typer av metadata som er viktige å utvinne, og hvilke metoder og teknikker de anser som viktige for forskning på nettdata, både i dag og den nærmeste framtida.<sup>2</sup>

Etter en kort introduksjon til nettarkiv og infrastrukturens målsetninger vil deltakerne få:

1. Vurdere viktigheten av ulike typer metadata for egen forskning,
2. Dele sine behov for digitale metoder og verktøy,
3. Diskutere behov for metadata og verktøy med andre forskere.

Resultatene fra verkstedet vil inngå i WebData-prosjektets behovsstudie. Denne studien brukes til å identifisere hvilke behov forskere har knyttet til dataplattform, metadata-modell, samt ulike verktøy og tjenester.

WebData-prosjektet ledes av Nasjonalbiblioteket, med Norsk regnesentral, Universitetet i Oslo (Humit og Språkteknologigruppa) og Norges arktiske universitet (Gjellatekno) som partnere. Prosjektet er finansiert av Norges forskningsråd. Verktøy og tjenester utvikles i tråd med FAIR-prinsippene og vil bli gjort tilgjengelig som åpen kildekode.<sup>3</sup>

<sup>1</sup> Mer informasjon om WebData-prosjektet finnes på Forskningsrådets sider, <https://prosjektbanken.forskningsradet.no/project/FORISS/350342>.

<sup>2</sup> Vi vil benytte en metode kalt MoSCoW-prioritering, der deltakerne får rangere viktigheten av ulike metadata-felt og analytiske verktøy. For metodisk beskrivelse, se Agile Alliance (2017): *Agile Extension to the BABOK Guide*, s. 84-86. For eksempel på praktisk applikasjon, se Kravchenko et al. (2022): "Ranking Requirements Using MoSCoW Methodology in Practice", in Silhavy, R. (ed) *Cybernetics Perspectives in Systems. CSOC 2022. Lecture Notes in Networks and Systems*, vol 503, (Cham: Springer), [DOI:10.1007/978-3-031-09073-8\\_18](https://doi.org/10.1007/978-3-031-09073-8_18)

<sup>3</sup> Wilkinson et al. (2016): "The FAIR guiding principles for scientific data management and stewardship", *Scientific Data* 3, 160018, [DOI:10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18); "The Open Source Definition", *Open Source Initiative*, arkivert i Internet Archive 18.04.2025, <http://web.archive.org/web/20250418004844/https://opensource.org/osd>.