

# INFRASTRUKTUR & GRENSESNITT

# INFRASTRUCTURE & INTERFACE

# DHKO 2025



15.-16. OKTOBER 2025  
UNIVERSITETET I INNLANDET  
HAMAR

## DHKO: Norsk nettverk for digital humaniora og kulturorganisasjon

DHKO er en løst organisert gruppe av akademikere som har interesse for og/eller befinner seg i digital humaniora, både i bred og smal forstand. Dette innebefatter:

- Humanistiske perspektiver på det digitale, digital teknologi og på digitalt materiale.
- Bruk av software og digitale metoder for analyse og tolkning i humanistisk orientert forskning.

Siden humanistisk materiale ofte har en tilknytning til arkiver, biblioteker og museer har dette nettverket til hensikt å situere seg i ABM-feltet og knytte forskningen opp mot, og med, kultur- og informasjonsinstitusjoner. Nettverket er derfor åpent for akademikere i bred forstand, og det kreves ingen tilknytning til en bestemt institusjon.

Med begrepet kulturorganisasjon mener vi det som innebefatter organisering, lagring, og tilgjengeliggjøring av kulturelle artifakter og kultur som data. Med andre ord; infrastrukturelle forhold.

Nettverkets formål er å skape rom for utveksling av faglig erfaringer, samt bidra til en styrkelse av digital humaniora i Norge - og særlig opp mot ABM-feltet. Formålene skal fremmes ved å holde konferanser, seminarer og webinarer. Samt å jobbe i samarbeid med forskningsstøtte ved institusjonsbibliotekene i Norge.

På lang sikt vil nettverket jobbe for å få til phd-kurs, sommerskolekurs for masterstudenter, og tematiske utgivelser ved ulike tidsskrift. Nettverket håper også å kunne være en katalysator for utvikling av forskningsprosjekter på tvers av institusjoner, både norske og internasjonale samarbeid. Vi har derfor som mål at nettverket vil lede til nye og fruktbare forskningssamarbeid, samt å bidra med forskning fra, til og for ABM-sektoren.

For å sikre utviklingen vil vi samarbeide med lignende nettverk i andre land i Europa, med særlig fokus på Norden. Her er særlig [DHNB](#) en organisasjon vi ser som viktig for å dekke den nordiske og baltiske delen.

Se [dhko.no](http://dhko.no) for mer informasjon, og mulighet for å melde seg på e-postliste.

### DHKOs arbeidsgruppe per oktober 2025:

Leder: Camilla Holm Soelseth | [camil@oslomet.no](mailto:camil@oslomet.no) (OsloMet)

Dianna Walla (NTNU)

Lina Ruth Harder (CDN, UiB)

Emma Aadland (UB, UiB)

Nick Walkley (AHO)

Rådgiver: Annika Rockenberger (UB, UiO)

Webmaster: Henrik Askjer (UiB)

### Konferansegruppa 2025

Konferansestyret består av et utvalg fra arbeidsgruppa og et utvalg fra institusjonen som er vert for årets konferanse. I 2025 er det Universitetet Innlandet.

Fra DHKO: Dianna Walla, Nick Walkley og Camilla Holm Soelseth

Fra Universitetet i Innlandet:

- Siri Fürst Skogmo (representant for forskningsprosjektet [Verden på norsk](#))
- Ida Jahr (DigKom – representant for forskergruppa [Digital kommunikasjon og kultur](#) og studieprogramansvarlig for [Digital Kommunikasjon og kultur](#))
- Karianne Dæhlin Hagen ([Universitetsbiblioteket](#))

Studentassister fra masterprogrammet [Digital Kommunikasjon og kultur](#) ved Universitetet i Innlandet.

## DHKO 2025:

# Infrastruktur og grensesnitt

Årets konferanse har temaet «infrastruktur og grensesnitt». Med dette ønsker vi å favne all humanistisk orientert forskning på, med og om infrastrukturer og grensesnitt. Fra visuell epistemologi, via digitale medier og materialer, til forskningsinfrastrukturer og organisering av kultur. Ikke minst arkiver, museer og biblioteker som infrastrukturer og grensesnitt for digital humanioraforskning for fortid, nåtid, og fremtid.

Infrastructure is both relational and ecological—it means different things to different groups and it is part of the balance of action, tools, and the built environment, inseparable from them. It also is frequently mundane to the point of boredom, involving things such as plugs, standards, and bureaucratic forms.

-- Susan Leigh Star (1999)

Alan Liu definerer videre digital humaniora som et felt som er unikt i sin evne til å konseptualisere, tolke og kritisere infrastruktur (Liu, 2018). Humaniora har i de siste årene hatt sin «infrastructural turn», med en gjenoppdagelse, men også fornyelse av, sin strukturtenkning. Innen digital humaniora og digital kulturforskning betyr dette et fokus på materialitet, forbindelser og de fysiske forutsetningene for digital kommunikasjon. Shannon Mattern (2014) definerer videre infrastruktur som: «The casting, staging and direction that determine what happens on the stage, and that allow it to function as a stage».

Digital forsknings og kulturinfrastruktur har i løpet av det siste året også fått fornyet politisk relevans. For bare to år siden kunne Poell, Nieborg og Duffy (2022) hevde at «an interest in the material structure of how apps and platforms operate in the background» var tilsynelatende langt viktigere å se på når denne infrastrukturen befant seg i Kina «than in discussions of Silicon Valley-born platforms». Dette er ikke lenger tilfelle, og DHKO ønsker med denne konferansen å åpne for nye tanker om både hvordan infrastrukturen strukturerer kultur og forskning (det teoretiske perspektivet) og hvordan best organisere kultur og forskning i et endret politisk landskap (det praktiske perspektivet). Vi lever i en samtid med infomediaries (Morris, 2015), og hvor det er mer enn noen gang viktig å forstå hvordan noe blir til, hva vi får se, hvordan vi ser det og når vi får tilgang. Samt hva slags data det gir oss tilgang til, og hvordan vi kan ta i bruk disse dataene og hva de kan gi oss av humanistisk innsikt. Mens infrastrukturer skaper og opprettholder systemene, er grensesnitt det som gir inngang til interaksjon med slike systemer.

Vi er i «a situation where the computer's interface seemingly becomes omnipresent and invisible, and where it at once is embedded in everyday objects and characterized by hidden exchanges of information between objects...» (Andersen and Pold 2018). I den andre enden av digital kulturs materielle forutsetninger finner vi nemlig grensesnittet (og infrastrukturer skaper også grensesnitt), og i grensesnittet ligger det mye makt. Her har teknografiske mediestudier og human-computer-interactions (HCI) vært felt som bidrar til å forstå teknologien diskursivt gjennom studier av stedene menneske møter maskin.

Som Grant Bollmer (2018) påpeker: «Often, the goal of interface design is to make technologies invisible, blending into the everyday architecture of our lives. Yet we only interact with digital media



based on the physical qualities of our technologies». De fysiske forutsetningene for digital humaniora og kulturorganisasjoner er dermed hovedfokuset for dette årets konferanse. Vi inviterer derfor årets deltakere til å studere noe som infrastruktur eller grensesnitt eller som del av en infrastruktur eller påvirket av grensesnitt, eller infrastrukturen eller grensesnittet i seg selv.

Vi inviterer til bidrag enten i form av presentasjon eller verksted. Overordnet er fokuset på infrastrukturer og grensesnitt, og særlig bruk av infrastrukturelle tilnærninger til studie av det menneskeskapte (kulturelle uttrykk og teknologi) og/eller metodologiske betraktninger knyttet til en eller flere av temaer relatert til dette.

### Tema for konferansen

Overordnet er fokuset på infrastrukturer og grensesnitt, og særlig bruk av infrastrukturelle tilnærninger til studie av det menneskeskapte (kulturelle uttrykk og teknologi) og/eller metodologiske betraktninger knyttet til en eller flere av temaer relatert til dette. Dette innebefatter organisering, lagring, og tilgjengeliggjøring av kulturelle artefakter og kultur som data, og bruken av disse dataene og hvilke innsikter de gir oss. Noen eksempler på hva dette kan være:

- Utvikling og bruk av verktøy for data-drevne studier av kultur
- Data-drevne studier av kultur som infrastruktur eller del av infrastruktur
- Bruk av digitale verktøy og/eller metoder for å studere humaniora, eller underliggende fagfelt
- ABM-institusjoner som formidlere av data og verktøy for DH-forskning (særlig tilgang, integrering og gjenbruk av data)
- Visuelle grensesnitt og visuelle epistemologier: Visualisering av data (inkl. bruk av GIS-teknologi)
- Datafeministiske eller teknografiske studier (f.eks. situerte data analyser) av teknologi og grensesnitt, særlig AI-teknologier
- Studier av infrastrukturelle vilkår for kulturelle fenomener
- Digital edisjonsfilologi
- Utfordringer og muligheter for bevaring, vedlikehold og kuratering av digital kultur (f.eks. spill og interaktiv litteratur) eller digitalisert kultur(arv).

For verkstedsbidrag er vi særlig interessert i verksted som tar for seg ulike programvarer og andre forskningsrettede brukergrensesnitt som kan tas i bruk av forskere og andre fagpersoner for å jobbe med digital eller digitalisert materiale.



## Keynote infrastructure: Organizing Information, Technological Change and Ideological Contestation

Jack Andersen (Copenhagen University)

Jack Andersen is an Associate Professor in the Department of Communication at the Copenhagen University. His research focuses on how classification practices, particularly in digital media, influence the categorization, understanding, and treatment of information, individuals, and groups in society. Andersen investigates the social, cultural, and political implications of these practices, which range from algorithms to digital archiving. His work incorporates concepts from rhetorical genre theory, classification theory, cultural techniques, communication and media theory, as well as cultural and social theory.

In 2025 he co-edited a new book called *Exploring Contemporary Classification Practices*. Based on this book, he will give a keynote on systems of classification, media technologies, and cultural institutions (such as libraries, archives, and museums) and how they respond to challenges, including classificatory bias, truth, neutrality, institutional tradition, and technological innovation. The scholarly discussion on the classification and organization of knowledge has developed with digital technologies from a bibliographic paradigm into something much wider, as the need for metadata and classification has become critical for usability and legitimacy. Raising awareness of classification practices in modern culture serves to emphasize how sorting things into categories is both an everyday accomplishment and a highly cultural and political activity with consequences for those who are classified and for those who classify.



## Keynote interface: Game Poems and Born-Videogame Poetics

Jordan Magnuson (University of Southampton)

Jordan Magnuson is a game designer and media artist whose work explores how videogames can express human experience in poetic and culturally resonant ways. For over fifteen years, he has created award-winning independent games, "game poems," and other interactive works. He is the editor of *Game Poems*, an online literary magazine dedicated to exploring the artistic and poetic potential of short-form videogames.

To suggest that videogames can be poems feels like an "outrage against decorum," to quote game scholar Piotr Kubiński. Poems, after all, are made up of words, where games are made up of 1s and 0s and flashing lights. Dig deeper, and the differences only seem to grow: games are associated with fast action and object-oriented outer experience, while poetry is often bound to slowness, ambiguity, and the exploration of inner life. Yet I want to suggest that these apparent contradictions stem less from the inherent properties of the media themselves than from the rigid taxonomies and preconceived notions of family resemblance that tend to frame our encounters with them. When we encounter something labelled a "poem," we pay one kind of attention; when we encounter something labelled a "videogame," we pay another--and it becomes difficult to imagine a space between.

I propose that thinking of poems and games as interfaces can help us bridge this gap. Poems can be understood not simply as collections of lyrical words, but as interfaces between language and subjective human experience. Videogames, similarly, can be understood as interfaces between code and human experience. If we approach them in this way, there is no reason why games cannot also serve as poetic interfaces, mediating subjective experience, metaphor, rhythm, ambiguity, and ritual through gameplay and interaction.

Drawing on frameworks from poetry studies, digital humanities, and computational media, and looking at examples from lyric poetry, digital poetry, and experimental videogames, I will explore what makes the "interface of poetry" tick, and examine how these same characteristics can be made to operate in videogames--even in games that don't feature any words. My aim is to show how interface perspectives can "make familiar objects be as if they were not familiar" (Percy Shelley) and sketch a born-videogame poetics that might be useful to digital culture scholars and practitioners alike.



## Verksted: Bit for bit, vers for vers: Et kodeverksted for lyrikkanalyse (45 minutter)

Ingerid Løyning Dale

“Bit for bit, vers for vers” er et verksted som gir en innføring i et nytt norsk Python-bibliotek som kan brukes til å annotere og analysere norsk lyrikk. Kodebiblioteket er utviklet av datalingvist Ingerid Løyning Dale ved Nasjonalbibliotekets DH-lab og doktorgradsstipendiat Ranveig Kvinnslund fra Senter for Ibsen-studier. Koden brukes i første omgang i Kvinnslands avhandling for å gi innsikt i sammenhenger og trekk i den nye norske lyrikken fra 1890-tallet, men den er skrevet med et grensesnitt som tillater å kjøre de mørnstergjenkjennende algoritmene på en hvilken som helst tekst der linjeskiftet er meningsbærende. Kodebiblioteket gjør det mulig å hente ut informasjon om stemmen i diktet, tekstlengde og versstruktur, samt gjentakelsesmønstrene allitterasjon, anafor og enderim.

Vi vil demonstrere hvordan man kan bruke kodebiblioteket med tekster fra ulike kilder, inkludert fra ulike grensesnitt i Nasjonalbibliotekets infrastruktur. Nettbiblioteket tillater nedlastning av avbildning og digital tekst for publikasjonene som er falt i det fri, og DH-labens API tillater uthenting av fulltekst. Vi vil gjennomgå hvordan vi har jobbet iterativt med litteraturvitenskapelige forskningsspørsmål i undersøkelsen av tekstmaterialet. I løpet av verkstedet viser vi noen enkle eksempler som deltakere kan følge og som illustrerer hvordan vi kan gjøre en form for digital nærlæring i stor skala. Vi ønsker en diskusjon om hvilke muligheter digitale verktøy og metoder gir litteratutforskningen, og hva slags spørsmål som nå blir interessante å stille.

Verkstedet holdes på norsk. Hvis deltakerne vil følge eksemplene, må de ha installert Python på egen maskin og kunne åpne og kjøre en Jupyter Notebook selvstendig. Det er derimot ikke et krav å følge eksemplene, hvis man kun vil få med seg presentasjonen og delta i diskusjonen.

## Verksted: Emitoner i praksis – semantisk annotasjon med Nasjonalbibliotekets dhlab API (2 x 45 minutter)

Lars G. B. Johnsen

Denne workshoppen gir deltakerne hands-on erfaring med å bruke *emitoner* – minimale tolkningsenheter generert med språkmodeller – for å utforske store tekstkorpus. Vi tar utgangspunkt i Nasjonalbibliotekets DH-lab for å hente ut konkordanser fra bøker, aviser og tidsskrifter.

Emitoner oppstår når et tekstfragment (f.eks. en setning) knyttes til en konseptuell tolkning, som *Ola* → *empatisk* eller *fisker* → *substantiv*. Metoden kombinerer tekstfragmenter med tolkningsbare, småskala annotasjoner, og kan brukes både kvalitativt og kvantitativt.

I løpet av verkstedet vil deltakerne:

1. Lære hvordan man henter ut konkordanser med DH-lab enten via app eller programering.
2. Velge et analyse tema (for eksempel egennavn i skjønnlitteratur, eller ord og begreper).
3. Formulere presise instruksjoner (*prompts*) som danner grunnlaget for å kjøre språkmodeller via API – med mulighet til å bruke frie modeller.
4. Generere emitoner, tilpasset konseptuelle kategorier underveis, og teste lingvistiske bruksforskjeller (f.eks. om et ord brukes som verb eller substantiv, *fisker*).

Målet er at deltakerne går fra tekstuddrag til et strukturert, tolkningsbart datasett som kan brukes til videre analyse, konseptkartlegging eller visualisering. Metoden viser hvordan språkmodeller kan fungere som grensesnitt mellom menneskelig fortolkning og storskala tekstinfrastruktur, og gir en lav terskel for å koble fortolkende humanistisk metode med digitale verktøy.

## Verksted: Digital humaniora på norsk- bidrag til termportalen (2 x 45minutter)

Camilla Holm Soelseth

Dette verkstedet er en kombinasjon av infomøte om en ny termgruppe for digital humaniora i Norge (del av Termportalen ved Universitetet i Bergen), og praktisk arbeid med oversettelse av sentrale termer. Verkstedet er åpent for alle interesserte.

Verkstedet gir en introduksjon til Termportalen, hva termgruppa for digital humaniora vil gjøre, og innføring i begrepsarbeid. Siste del av verkstedet vil gå til praktisk arbeid med å jobbe med forslag til oversettelse av noen aktuelle termer i digital humaniora.

Det er ingen forpliktelse om å bli med i termgruppa ved å møte på dette verkstedet, men verkstedet kan forhåpentligvis gi «blod på tann» i det å jobbe med oversettelsesarbeid.

På verkstedet vil vi:

- Fastsette formål og målgruppe
- Avgrense fagområdet
- Komme med forslag til datakategorier, med utgangspunkt i presentasjon av *TADiRAH – Taxonomy of Digital Research Activities in the Humanities*
- Jobbe med forslag til oversettelser av termer fra TADiRAH

Målet med etablering av termgruppen er å styrke det norske fagspråket for termer relatert til the digitale, samt etablere en nasjonal arbeidsgruppe som kan ha et fast årlig verksted på konferansen, med presentasjon av arbeid for konferansedeltagere i tillegg.

Aktivitetsnivået i termgruppen utenom konferansen kan justeres opp eller ned etter man har kapasitet, men det er viktig at det jobbes kontinuerlig.

Det vil være naturlig at begreper vil overlappe med både term-gruppa for IT og term-gruppa for bibliotek- og informasjonsvitenskap, og derfor kan det i fremtiden bli aktuelt med samarbeid når det gir mening.



## Verksted: Introduksjon til databehandling med OpenRefine (45 minutter)

Camilla Holm Soelseth

OpenRefine er en gratis open-source programvare beskrevet som “a power tool for working with messy data” ([David Huynh](#)). Det er et program som er særlig nyttig for både forskere og andre som jobber i forskningsbibliotek og som er tett på forskningsdatahåndtering. Dette gjør også at OpenRefine inngår som en sentral programvare i Library Carpentry, som er den delen av bibliotek- og informasjonsvitenskaplig praksis som dekker grunnleggende ferdigheter med datahåndtering.

OpenRefine er laget for å hjelpe med oppdeling, vasking og beriking av data, og gjør dette gjennom et oversiktlig grensesnitt. Verkstedet er særlig egnet for forskere og biblioteksansatte som ønsker å få en innføring i gode praksiser for enkel forskningsdatahåndtering av «messy data».

Camilla H. Soelseth er sertifisert *carpentry instructor*, og vil med utgangspunkt i en tilpasset norsk versjon av det fulle Carpentry-kurset gi en kort introduksjon og hands-on innføring på 45 minutter.



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

Jon Tønnesen

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, datauthenting og analyse.<sup>11</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 (Giellatekno) 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>



# Presentasjoner



# Open Access is just the Beginning: Multimodal Data Representations for Disability-Inclusive Datasets

Hannah Ackermans (Universitetet i Bergen)

This presentation will address the media-specific affordances of databases through the lens of accessibility to disabled users. Accessibility considerations make the modalities of databases obvious in ways that would otherwise remain obscured or taken for granted.

I will explore the intersection between databases and accessibility on three levels: metadata, metainterface, and multimedia. In each section, I outline accessibility considerations and its methodological implications. To appeal to the cross-disciplinary audience, I will use examples from a range of databases/datasets in various fields.

## Metadata

Although a lot of attention is given to Open Access of data using the FAIR principles (findability, accessibility, interoperability, and reusability), they are rarely combined with Web Content Accessibility Guidelines' (WCAG) POUR principles (Perceivability, Operability, Understandability, Robustness). How do we combine the two to make data truly accessible?

## Metainterface

Communication does not only happen between people, but also between interface layers. Attention to accessibility reveals these layers because it demonstrates how people are accessing the same data using different displays, assistive technologies, and settings.

## Multimedia

Making multimedia data accessible tends to mean adding a media modality, such as adding captions or translations to audio. Adding these layers of data can be a valuable exercise to understand one's dataset, as well as create a richer dataset for people using it.

Bringing together these levels, I reflect in the conclusion on ways of accessing knowledge through disability on a technical and a human level.

## Showcase of the Consortium on Electronic Literature

Hannah Ackermans (Universitetet i Bergen)

As a literary genre, electronic literature (E-Lit) is characterized by creative experimentation and subversive use of digital media. Functioning outside the confines of traditional publishing, divergent practices of collecting, documenting, and archiving E-Lit have emerged within the field's community.

The Consortium on Electronic Literature is an open access, non-commercial resource offering qualified access to literary databases, archives, and institutional programs in the literary arts and humanities scholarship. The recent revitalization of CELL has led to the work-in-process The CELL Index, which has a new interface, infrastructure, and data model that will ensure longevity and lessen the burden on individual scholars.

With rise of linked open data, we have chosen to use the collaborative, open-data knowledge graph Wikidata to communicate our data. This migration provides new opportunities to make metadata about electronic literature available to more people, as well as promoting the individual electronic literature databases who have been doing foundational work over the decades.

At the same time, we are creating our own focused Wikibase environment with taxonomy relevant to the electronic literature community in particular. Each record in the Wikibase contains links to the various places in which a work is documented or archived by other databases. In some cases the work itself has become unavailable, but has either been archived for preservation in its entirety, or documentation exists. In those cases, screenshots and walkthroughs created by individual working groups are a replacement for the original work. For this reason it's especially useful to have the overarching Consortium Wikibase linking to all these efforts to piece together what it left of the work.

In partnership with a variety of institutions archiving electronic literature, we are providing access for reuse of collections and datasets.



## En sosiopolitisk maskinlesning av norsk mellomkrigslitteratur, 1920-1945

Sofie Arneberg (Nasjonalbiblioteket), Pål C. Halvorsen (OsloMet), Lars G. B. Johnsen (Nasjonalbiblioteket)

Litteratursosiologi kan handle om å se på litteraturen som et sosialt artefakt, og som en inngang til en bestemt tids tankesett. I vår undersøkelse sammenligner vi tekster skrevet av to grupper forfattere: Forfattere som sympatiserte med okkupasjonsmakten i Norge mellom 1940 og 1945, og forfattere som var nærmere tilknyttet motstandsbevegelsen. Utvalget er basert på en studie av norske forfatteres posisjoner innenfor økonomiske, kulturelle og sosiale hierarkier (Hjellbrekke et al., 2025).

Vår tilnærming tar sikte på å hente ut sosiologiske variabler (kjønn, klasse, alder og annet) fra litterære tekster og undersøke hvordan disse reflekterer underliggende samfunnsstrukturer. På denne måten vil vi undersøke hvordan fiktive karakterer blir posisjonert, sosiale roller og allianser kommer til uttrykk, og konflikt er innvevd i språket. I stedet for å basere oss utekkende på næranalyse, kombinerer vi korpusmetoder som frekvenstelling og navngitt enhetsgjenkjenning (NER) med store språkmodeller (LLM-er).

Korpuset er hentet fra den norske nasjonalbibliografien og tilhørende digitaliserte filer. Det har vært et poeng å ikke bryte koblingen til kilden, den åpne bibliotekkatalogen, med sine internasjonalt standardiserte formater for bibliografisk beskrivelse. Et av målene med arbeidet er derfor å utvikle en metode som ivaretar sporbarhet og som kan anvendes til lignende arbeider med utgangspunkt i bibliografiske beskrivelser.

### A socio-political machine reading of Norwegian interwar literature, 1920-1945

Literary sociology can approach literature as a social artifact and as an entry point to the mindset of a given era. In this study, we compare texts written by two groups of Norwegian authors: those who sympathized with the occupying power during 1940–1945, and those who were more closely aligned with the resistance movement. The selection is based on a study of Norwegian authors' positions within economic, cultural, and social hierarchies (Hjellbrekke et al., 2025).

Our approach aims to extract sociological variables (such as gender, class, and age) from literary texts to investigate how these reflect underlying social structures. In doing so, we analyze how fictional characters are positioned, social roles and alliances are expressed, and conflict is embedded in language. Rather than relying solely on close reading, we combine corpus-based methods including frequency counts and named entity recognition (NER) with large language models (LLMs).

The corpus is derived from the Norwegian national bibliography and its associated digitized files. Preserving the link to the source, the open library catalogue with its internationally standardized formats for bibliographic description has been central to our work. One of the aims of the project is therefore to develop a method that safeguards traceability and can be applied to similar studies grounded in bibliographic metadata.



## Memes som grensesnitt og danningsarena

Jonas Christoffer Berger (Universitetet i Innlandet)

Lars Løvlie (2003) skisserer danningskonseptets historie i sitt kapittel «Teknokulturell dannelses». Her beskriver han en ny forståelse for danning, hvor danning ikke er noe som skjer med individet når selvt møter verden, men snarere noe som skjer i grensesnittet mellom individet og verden, noe verden og individet opplever felles. Et slikt grensesnitt kan være i interaksjonene vi har med andre mennesker til daglig, det kan være et maleri, et sjakkparti, eller gjennom smarttelefonen. Og nettopp i smarttelefonen finner vi det kanskje beste eksempelet, for her finner vi også et grensesnitt på et mer konkret nivå, en flate hvor verden er tilgjengelig for oss på mange forskjellige måter, men hvor vi også er tilgjengelig for verden. Her knytter vi vår identitet til det eksterne for oss, og teknologien fungerer som et medie for danning.

En parallel til dagens danningsarena er å finne i antikkens teater. Doseth skriver om hvordan imitasjon, eller mimesis, stod sentralt i paideia, da som en form for lek eller underholdning. Vegen til de riktige dyder, viste seg ofte gjennom tragediene og komedienes skuespill; mytene, tragediene og komediene kontekstualiserte det alvorlige, livet, det gode og det skjønne, i samspill med lek og spill (2011, s. 15-16). Der hvor mimesisbegrepet brukes til å beskrive en imitasjon av dannelsesidealene i antikken, kan man spørre seg om memes-begrepet nå kan brukes til å skildre deler av dannelsesidealene for dagens samfunn, eksempelvis medborgerskap. Med utgangspunkt i teknologien og medieringen som skjer via denne, ser dette paperet på hvordan memes kan fungere som grensesnitt for danning, og hvorvidt memes endrer noe ved vår forståelse av danningsbegrepet.



## Videoessayet som metode i litteraturforskningen og- undervisningen

Adriana Margareta Dancus (Universitetet i Sørøst-Norge)

Videografisk kritikk innebærer den audiovisuelle studien av audiovisuelle og skjermmedier ved hjelp av digitale redigeringsprogrammer. Det er et voksende fagfelt innen film- og medievitenskap, men foreløpig lite brukt i andre humanistiske fag, eksempelvis i litteraturvitenskap der skriften anses å være dominerende. Samtidig har litteraturbegrepet vært utvidet betydelig siden 2000-tallet for å omfatte multimodale, digitale tekster som dataspill, litterære apper og Instaserier. Lesekulturen blant unge generasjoner har også endret seg, med nordiske og europeiske studier som viser en dalende leselyst særlig blant unge gutter, og ungdommer som står for stadig større mengder audiovisuelt innhold som de selv produserer og distribuerer via sosiale medier. Mot denne bakgrunnen skal jeg i denne presentasjonen utforske hvordan videografisk kritikk kan være en metodisk inngang til studien av litteratur, både i forsknings- og undervisningssammenheng. Jeg reflekterer med utgangspunkt i erfaringer fra forskningsprosjektet Litteratur og videoessayet og vil vise eksempler på videoessay utviklet under prosjektet. Presentasjonen vil fokusere på videoessayet som en forskningsmetode, men vil også fremme noen litteraturdidaktiske tanker om hvordan praktisk arbeid med videoessayet kan styrke elevenes og studentenes digitale og litterære kompetanse.

### The Video Essay as Method in the Research and Teaching of Literature

Videographic criticism entails the audiovisual analysis of audiovisual and screen media using digital editing software. It is a growing field within film and media studies, and for now little deployed in other humanistic subjects, for example in literary studies where prose is seen to have a dominating position. At the same time, the concept of literature has been significantly expanded since the 2000s to include multimodal, digital texts such as video games, literary apps, and Instacomics. The reading culture among young generations has also changed, with Nordic and European studies showing a decreasing desire to read for pleasure, particularly among young boys, and youth producing large amounts of audiovisual content which they distribute through social media. Against this background, I will discuss in this presentation how videographic criticism can be a method in literary studies, both in the context of literature research and teaching. My reflections are based on experiences with the research project Literature and the Video Essay, and I will share some examples of video essays developed through this project. The presentation focuses on the video essay as research method, but I will also propose some literary didactic ideas about how practical, digital work with the video essay can strengthen pupils' and students' literary and digital literacy.



# Color as A Physical Description of the Dead Sea Fragments: An explanation of a misconception and a future recommendation

Hilda Deborah (Colorlab, NTNU)

Color is often used as the physical description of Dead Sea scrolls and fragments. This is especially true throughout the volumes of the Discoveries in the Judaean Desert (DJD) where we can often find variations of brown or tan. Such description of color is, however, a subjective and qualitative way to describe color and it is far from being an objective and accurate description of surface color. In DJD 36, these color terms are usually accompanied by the Munsell color category, e.g., very pale brown 10YR 7.5/4. Despite the attempt to standardize the color description, this particular use and documentation using the Munsell color chart is still not an objective use of color. Color is a perceptual property of materials and objects, and not physical. This means that its perception is always affected by the characteristics of our visual system and the illumination under which a visual observation is carried out. To use color as an objective measure, we must understand the concept and follow the procedure laid out in colorimetry, allowing to quantify and describe physically the human color perception. Finally, the understanding that color is a perceptual property will hopefully make us more careful in deriving interpretations or making theories based on the observation of color images of the fragments, even if the images are of very high quality. We need to constantly remind ourselves that state of the art imaging techniques and very high-quality images do not automatically give us objective color measurements.



## Digital slip narratives as Cultural infrastructure

Cathrine Felix (Universitetet i Innlandet)

A slip is an error made by an agent when she performs an unintended action (Norman, 1981) such as buying margarine instead of butter (Anscombe, 2000) or pouring tea into a tomato can instead of a cup (Norman, 1981). While often trivial, slips can also be disastrous, as in medical errors or failures in high-risk domains such as aviation or nuclear power (Reason, 2007). Beyond their consequences, however, slips also generate narratives. They are retold, shared, and commented on, becoming cultural material in their own right.

The digital world provides a unique infrastructure for such narratives. Everyday mistakes are recorded, uploaded, and circulated across social media platforms, blogs, and video-sharing sites. Interfaces like TikTok, YouTube, and Instagram not only host these narratives but actively shape their form and dissemination through templates, comment functions, and algorithms that structure visibility. What might otherwise remain private errors are transformed into collective stories, memes, and sometimes even global spectacles, such as the widely publicized “Nipplegate” incident.

This presentation explores how slip narratives may be seen as cultural practices shaped by the infrastructures and interfaces that mediate them. Platforms operate as infrastructures that determine how, when, and to whom slip stories become accessible, while interfaces frame the aesthetic and discursive possibilities of narration. At the same time, slips invite creativity, empathy, and humor; they become icebreakers, bonding devices, and sites of self-reflection. By studying digital slip narratives as infrastructurally conditioned cultural expressions, we can better understand how digital culture organizes error, visibility, and affect.

In contemporary digital culture, infrastructures such as YouTube, TikTok, Instagram, blogs, and memes have become key sites where everyday errors are rendered visible, shareable, and narratable. These platforms function simultaneously as infrastructures that organize circulation and accessibility, and as interfaces that afford specific modes of representation and engagement. Studying how slips are transformed into digital narratives within these infrastructural and interface conditions allows us to understand processes of cultural production, affective exchange, and meaning-making that are central to digital humanities.



## From Archive to Interface: Participatory Methods for Digital Catalogue Raisonnés

Elizabeth Ashley Fox-Jensen (Malmö University)

This presentation examines how participatory design methodologies can transform archival art documentation into accessible digital interfaces, drawing from my interdisciplinary doctoral research and the ongoing Ted Stamm Catalogue Raisonné (TSCR) project. Positioning digital catalogue raisonnés (CRs) at the intersection of art history, digital humanities (DH), design, textural studies, and this research discusses the connection between archival materials and Art Scholarly Databases (ASD).

Building on Bearman's ASD concept as "information sources containing art historical data and imagery" and incorporating Drucker's DH theories on democratisation and shared sustainable resources, this work advances participatory approaches to digital art documentation. Bearman discusses "Considerations in the Design of Art Scholarly Databases" (1988) established ASD as a repository for collectors, conservators, curators, historians, researchers, and GLAMs, yet their effectiveness depends on collective resources for successful CR projects.

Digital CRs function as authentication, documentation, and contextualisation tools for artists' oeuvres or specific series. Despite substantial growth (3,750 published CRs; 371 in development, IFAR), issues of public accessibility, technical obsolescence, and sustainable preservation persist when traditional top-down development approaches dominate ASD integration.

This research investigates:

- How can participatory design methods advance Bearman's ASD framework and connect the archive-to-interface transformation to catalogue raisonné and their resources?
- What collaborative frameworks emerge when diverse stakeholder communities actively inform shared scholarly database development?

Using the TSCR as a case study, the presentation demonstrates:

- Collaborative ASD methodology: Action research with user communities, contextual design, and stakeholder involving estates, archives, GLAMs, and technical partners
- Archive-to-interface translation: Transforming ephemera, correspondence, and documentation into coherent digital narratives through UI/UX/GUI design
- Sustainable participation: Building long-term collaborative workflows that honor archival integrity while embracing contemporary DH practices and ASD standards

By sharing implementation successes and challenges, this presentation offers replicable methodologies for CR developers, GLAM professionals, and DH practitioners creating collaborative cultural heritage infrastructure through participatory ASD development.



## Exploring Visual Identity and Social Influence in the Picturemarks Project

Ole Goethe (Høyskolen Kristiania)

The Picturemarks project, initially developed in collaboration with a research lab at Harvard University and supported by Getty Images, began as an exploratory digital platform for aggregating and visualizing personal image collections. Building on prior quantitative analyses—including usability testing, A/B testing, and cluster analysis, this new phase shifts toward a qualitative investigation of how individuals construct identity and interpret social influence through curated visual assemblages. This research explores how users engage with their aggregated images, particularly those drawn from social media platforms such as Instagram, and how the recontextualization of these images on a shared digital canvas affects self-perception, memory, and interpersonal communication. Through in-depth interviews, participatory design sessions, and visual ethnography, we examine user experiences with the Picturemarks interface, focusing on emotional resonance, narrative formation, and the perceived agency in visual self-representation. Preliminary observations, such as a student's astonishment at seeing 2,000 personal images, many of them selfies, reorganized on a single canvas, suggest that visual aggregation can provoke reflective self-assessment and new modes of digital storytelling. This project investigates these moments of insight, analyzing how visual patterns, repetition, and arrangement influence users' understanding of their digital identities over time. By centering user voices and interpretive practices, this study contributes to humanistic and pedagogical discussions in digital humanities, particularly around visual literacy, data ownership, and the emotional dimensions of digital archives. Findings will inform the design of more reflective, ethically grounded tools for personal data visualization in educational and creative contexts.



# Evaluating Collocation Measures: Which Best Support Lexicographic Choice

Aleksandra Ignateva (Universitetet i Bergen)

This study investigates the effectiveness of statistical association measures in identifying strong Norwegian collocations and multiword expressions (MWEs), with a focus on their relevance to lexicographic practice. A set of 20 high-frequency Norwegian words -- covering adjectives, nouns, adverbs, and verbs -- was selected as collocation nodes. Collocations containing these target words were extracted from two annotated corpora: the Norwegian Newspaper Corpus (NNC) and the Lexicographic Corpus of Bokmål (LBK), both available through the Corpuscle interface.

After sorting the collocations and MWEs according to various association measures, they were compared to their occurrences in the online dictionary ordbokene.no for Bokmål. The dictionary collocations and MWEs were assigned scores based on how highly each association measure prioritized them. This ranking process aims to determine which association measures align the most closely with lexicographic judgment, thereby providing insights into the most suitable tools for dictionary work.

The following statistical measures were tested: collocation frequency, Log-likelihood, Mutual Information (MI), t-score, MI3 (cubic mutual information), MI\*log(frequency), z-score, corrected z-score, logDice, and Pearson's chi-squared.

Among these, the collocation frequency consistently produced the highest scores across all experiments, suggesting it is the most reliable indicator of lexicographically relevant collocations. T-score, LogDice, MI3, and Pearson also showed strong alignment with dictionary choices. In contrast, the corrected z-score and Log-likelihood yielded the lowest scores. However, it is worth noting that Log-likelihood exhibited a distinct bias toward MWEs, more so than the other measures.

This research introduces a novel method that analyzes collocations and MWEs selected directly from dictionary entries and compares them with results obtained from two corpora — the Norwegian Newspaper Corpus and the Lexicographic Bokmål Corpus. The dictionary examples represent standardized usage in contemporary Norwegian. This study provides useful insights for lexicographers into which statistical association measures are most effective for identifying dictionary-relevant collocations and MWEs, while also revealing how well the measures align with their choices.

## Evaluering av kollokasjonsmål: Hvilke støtter leksikografiske valg best

Denne studien undersøker hvor effektive statistiske assosiasjonsmål er for å identifisere sterke norske kollokasjoner og flerordsuttrykk (MWEs), med særlig fokus på deres relevans for leksikografisk arbeid. Et utvalg på 20 høyfrekvente norske ord – inkludert adjektiver, substantiver, adverb og verb – ble valgt ut som kollokasjonsnoder. Kollokasjoner som inneholder disse ordene ble hentet ut fra to annoterte korpus: Norsk Aviskorpus (NNC) og Leksikografisk Bokmålkorpus (LBK), begge tilgjengelige gjennom Corpuscle.

Etter å ha sortert kolloasjonene og uttrykkene etter ulike assosiasjonsmål, ble de sammenlignet med forekomster i nettordboken ordbokene.no for Bokmål. Kolloasjonene og uttrykkene fra ordboken ble gitt poeng ut fra hvor høyt hvert assosiasjonsmål rangerte dem. Denne rangeringsprosessen har som



mål å fastslå hvilke assosiasjonsmål som best samsvarer med leksikografisk vurdering, og dermed gi innsikt i hvilke verktøy som er mest hensiktsmessige for ordboksarbeid.

Følgende statistiske mål ble testet: kollokasjonsfrekvens, Log-likelihood, mutual information (MI), t-score, MI3 (kubisk mutual information), MI\*log(frekvens), z-score, korrigert z-score, logDice og Pearsons kjikvadrattest.

Blant disse ble kollokasjonsfrekvens rangert jevnt over høyest i alle eksperimentene, noe som antyder at dette målet er den mest pålitelige indikatoren for kollokasjoner av leksikografisk relevans. T-score, LogDice, MI3 og Pearson viste også god overensstemmelse med leksikografiske valg. Derimot ga korrigert z-score og Log-likelihood de laveste poengene. Det er likevel verdt å merke seg at Log-likelihood hadde en tydelig tendens mot flerordsuttrykk, mer enn de andre målene.

Denne masteroppgaven introduserer en ny metode for å analysere kollokasjoner og flerordsuttrykk hentet direkte fra ordbokartikler og sammenligne dem med resultater fra to korpus – Norsk Aviskorpus og Leksikografisk Bokmålkorpus. Eksemplene fra ordboken representerer standardisert bruk av moderne norsk. Studien gir nyttige innsikter for leksikografer hvilke statistiske assosiasjonsmål som er mest effektive for å identifisere ordboksrelevante kollokasjoner og flerordsuttrykk, og viser samtidig hvor godt målene samsvarer med leksikografenes valg.



## AI language/ language on AI: language as infrastructure

Ida Jahr

One could argue that underlying the whole linguistic turn in cultural studies and history is an idea of language as cultural infrastructure. Today, that seems more relevant than ever. As Jani Marjanen and Pasi Ihälainen point out, “The meanings of key terms through which political, social, cultural and scientific debates are conducted are often multi-layered, ambiguous, contradictory and contested.” According to the linguist Emily Bender, ‘artificial intelligence’ is a marketing term, used to “make certain kinds of automation sound sophisticated, powerful, or magical” and created in order to attract funding to a particular field of information technology in 1956. Through the successful marketing of OpenAI, ‘artificial intelligence’ has today come to be used interchangeably with ‘generative AI’, and there seems to be a concerted effort to create a slippage between the terms AI and AGI. In this paper, I present a work in progress, in which I would like to use the methods of conceptual history to interrogate the actual uses of the term ‘artificial intelligence’ within the field of AI research, historically and today.

### KI-språk/KI i språket: Språk som infrastruktur

Man kan argumentere for at en ide om *språk som kulturell infrastruktur* ligger under den språklige vendingen i felt som kulturstudier og historie. I dag virker dette viktigere enn noen gang. «Viktige begreper som blir brukt til å føre politiske, sosiale, kulturelle og vitenskapelige debatter har ofte flere lag, og de er flertydige, motstridende og omstridte», påpeker Jani Marjanen og Pasi Ihälainen. Emily Bender hevder at begrepet ‘kunstig intelligens’ er et markedsføringsbegrep, som i dag blir brukt for å “få enkelte former for automatisering til virke sofistikerte, spesielt kraftige eller magiske”, og i 1956 ble skapt for å trekke finansiering til et bestemt felt innen informasjonsteknologi. OpenAIs markedsføring har ført til at «kunstig intelligens» i dag ofte blir forstått som «generativ KI», og det arbeides for å skape en glidende overgang mellom begrepene KI og KGI. Jeg vil presentere et pågående prosjekt, der jeg ønsker å bruke metoder fra begrepshistorie for å undersøke hvordan begrepet «kunstig intelligens» faktisk har blitt og blir brukt, innenfeltet KI-forskning.



## Emitonet: en minimal tolkningsenhet for humanistisk tekstanalyse med språkmodeller

Lars Gunnarsønn Bagøien Johnsen (Nasjonalbiblioteket)

Forskere møter ofte et tilbakevendende problem med digitale metoder: hvordan forene fortolkende dybde med storskala beregning. Statistiske teknikker som kollokasjonsanalyse eller topic modeling er kraftige verktøy, men kan fremstå som «svarte bokser» for dem uten teknisk bakgrunn.

Språkmodeller (LLM) tilbyr mer intuitive tekstoperasjoner – oppsummering, parafrasering, klassifisering – men hva finnes mellom fulltekstoppsummeringen og de enkelte datapunktene?

Vi introduserer emitonet som en mikroenhet for fortolkning. Emitonet oppstår når en språkmodell knytter et tekstfragment til en konseptuell tolkning, som at *Ola* → *intellettuell* fra fragmentet *Ola leser bøker*, eller *spise* → *verb* fra *De forsøkte å spise kake*. Et emiton består av (1) et tekstfragment, (2) et fokuselement (navn, entitet, nøkkelord) og (3) en tolkning.

Emitonet skiller seg fra oppsummeringer, embedding-representasjoner og klassifikasjoner: det er lite, tolkningsbart og kan telles, noe som gjør det egnet for semantisk annotasjon, konseptkartlegging og kombinasjon av kvalitative og kvantitative analyser. Vi viser også hvordan et lite, representativt utvalg av konkordanser – ofte et par tusen linjer selv ved millioner av treff – kan gi et solid grunnlag for analyse. Kombinert med LLM gir dette rask, etterprøvbar innsikt som i mange tilfeller kan supplere eller erstatte fullskala kollokasjonsanalyser.

Vi viser hvordan emitonet kan brukes til å kartlegge meningsfelt i historisk litteratur, og hvordan det kan fungere som et lett, generativt format for semantisk kartlegging i digital humaniora.

### Emitons: a Minimal Unit of Interpretation for Humanistic Text Analysis with Language Models

Researchers often face a recurring challenge in digital methods: how to reconcile interpretive depth with large-scale computation. Statistical techniques such as collocation analysis or topic modeling are powerful tools, but they can appear as “black boxes” to those without technical backgrounds. Language models (LLMs) offer more intuitive text operations – summarization, paraphrasing, classification – but what exists between full-text summaries and individual data points?

We introduce the emiton as a micro-unit of interpretation. An emiton arises when a language model links a text fragment to a conceptual interpretation, such as *Ola* → *intellectual* from the fragment *Ola reads books*, or *eat* → *verb* from *They tried to eat cake*. An emiton consists of (1) a text fragment, (2) a focus element (name, entity, keyword), and (3) an interpretation.

The emiton differs from summaries, embedding representations, and classifications: it is small, interpretable, and countable, making it suitable for semantic annotation, concept mapping, and the combination of qualitative and quantitative analyses. We also demonstrate how a small, representative sample of concordances – often just a few thousand lines even when millions of hits are available – can provide a solid foundation for analysis. Combined with LLMs, this yields rapid, verifiable insights that in many cases can complement or even replace full-scale collocation analyses.

We show how the emiton can be used to map semantic fields in historical literature, and how it can function as a lightweight, generative format for semantic mapping in the digital humanities.

# Fra katalog til kunnskapsgraf: Lars Dahles boksamling som forskningsobjekt

Lars G Bagøien Johnsen (Nasjonalbiblioteket), Jennifer Thøgersen (VID vitenskapelige høgskole)

Hvilke historier kan en samling fortelle – om sin samler og verdenen rundt ham? Og kan AI-teknologier hjelpe til å synliggjøre disse historiene?

Midt på 1900-tallet fikk misjonsskolen i Stavanger (nå VID) boksamlingen til den kjente norske misjonæren og akademikeren Lars Dahle. Vi digitaliserte nylig kortkatalogen for denne samlingen og jobber med å importere deler av samlingen til Alma. Dette er viktig for å gjøre samlingen søkbar, men samtidig ønsker vi å løfte fram katalogens verdi som forskningsressurs i seg selv. Potensialet strekker seg langt utover bibliografisk gjenfinning. Hvert kort bærer spor av tid, sted og tanke: forfatternavn, titler, utgivelsessted og -år danner et historisk, geografisk og tematisk kart over en intellektuell horisont i bevegelse – med misjonærrens samling som ett eller flere punkt i dette bildet. Katalogen er ikke bare en inngang til bøker – den er en bok i seg selv, skrevet i metadata.

Når vi leser katalogen som data, åpner det seg nye forskningsoppgaver. Titlene kan tematiseres, språk identifiseres, og digitale versjoner spores opp i nasjonale og internasjonale samlinger.

Utgivelsesstedene kan geolokaliseres og danne grunnlag for kartbaserte analyser, mens forfatternavn gir mulighet for kjønnsstudier og nettverksanalyser basert på publiseringsmønstre.

For alt dette, utforsket vi bruk av språkmodeller for å rengjøre, berike og strukturere dataene. Vi skal presentere våre suksesser og utfordringer med denne fremgangsmåten og hvordan vi håndterte modellenes mangler. Når det arbeidet er fullført, kan vi løfte katalogen videre – fra digitalisering til strukturering. Resultatet kan bli en kunnskapsgraf som både gir grunnlag for Alma-oppføringer og åpner for digitale humaniorastudier: tematisk, språklig og historisk.

## From catalog to knowledge graph: The Lars Dahle book collection as a research object

What stories can a collection tell – about its collector and the world around him? And can AI help reveal these stories?

In the mid-1900's the mission school in Stavanger (now VID) was gifted the book collection of Lars Dahle, a well-known Norwegian missionary and scholar. We recently digitized the card catalog for this collection and are working toward importing portions of the collection into Alma. This is important to make the collection searchable, but we also want to highlight the value of the catalogue as a research resource itself. The potential extends far beyond bibliographic retrieval. Each card carries traces of time, place and thought. Author name, title, place and year of publication form a historical, geographical and thematic map of an intellectual horizon in motion – with the missionary's collection as one or more points in this picture. The catalogue is not just an entrance to books – it is a book itself, written in metadata.

When we view the catalogue as data, new research possibilities emerge. Titles can be thematized, languages identified, and digital versions tracked down in national and international collections. The



places of publication can be geolocated and form the basis for map-based analyses, while author names provide the opportunity for gender studies and network analyses based on publication patterns.

To achieve this, we explored the use of language models to clean, enrich, and structure the data. In this presentation, we will discuss the successes and challenges we had with this approach and how we addressed shortcomings in the language models' output. When that work is complete, we can take the catalogue further – from digitization to structuring. The result can be a knowledge graph that both provides the basis for Alma entries and opens up for digital humanities studies: thematic, linguistic, and historical.

A complex network graph with numerous small, colorful dots (yellow, orange, red) connected by thin, translucent blue lines, representing a digital or scholarly infrastructure.

## Ad fontes! Ad fragmenta!

## Reinventing the Dead Sea Scrolls

Årstein Justnes (Universitetet i Agder) & Hilda Deborah  
(Colorlab, NTNU)

The inventory system for the Dead Sea Scrolls was created in the early 1950s by the Polish priest Józef T. Milik. The system's primary goal was to compile all Dead Sea Scroll fragments into manuscripts. Strictly speaking, the designation "Dead Sea Scrolls" is misleading; there are only a few complete scrolls, but there are tens of thousands of fragments. Because of the fragmentary nature of the material, most Dead Sea Scroll manuscripts are de facto modern scholarly constructs. They were largely created in the 1950s and early 60s based on more than 1200 museum plates of fragments stored in the Palestinian Archaeological Museum (later: Rockefeller Museum). Some manuscripts that came into being these years were later deconstructed and ceased to exist. Others were split up and divided. In 2003, the Norwegian scholar Torleif Elgvin, for instance, argued that the manuscript 4Q413 (4QComposition concerning Divine Providence) should be split into two manuscripts of two distinct works (4Q413 [4QExhortation] and 4Q413a [4QApocryphal Psalm B]).

In this paper, we will propose a new way of approaching the Dead Sea Scrolls using knowledge graph. Our point of departure is that the Dead Sea Scrolls are, first and foremost, physical fragments, and that any scholarly construct must respect the physical nature of these objects. This is a radical shift in focus from the current system, where objects are often arranged to fit one's idea of a manuscript.



## Poetry of/as interface: language and resistance in creative encounters with corporate AI

Anne Karhio (Universitetet i Innlandet)

This presentation focuses on creative projects engaging with digital interfaces through the idea of language itself as interface between human and machine exchanges in the age of AI and the digital platform economy. I will particularly consider recent work on poetry, and how it as a language-oriented art can both demonstrate and resist the appropriation and commodification of verbal expression for economic and corporate profit. The work of artist-practitioners and scholars including Pip Thornton, Søren Pold, and Ben Gosser highlights how critical creative practice can make visible and challenge many concerning aspects of current AI tools and the platform economy, and address the often problematic encounters between visible user interfaces and the infrastructures supporting them. At the same time, poetry or "creativity" should not be seen as easy answers to difficult questions concerning AI. As Pip Thornton suggests, poetry (and other art forms) can nevertheless help develop "a range of language-based creative refusal tactics and interventions", and allow authors, citizens, and communities to "reclaim some agency in the midst of AI's crushing inevitability".

### Poesi av/som grensesnitt: språk og motstand i kreative møter med teknologiselskapers KI

Denne presentasjonen ser på kreative prosjekter som befatter seg med digitale grensesnitt fra perspektivet at språk er et grensesnitt mellom menneskelige og maskinelle utvekslinger i KIs tidsalder og i økonomien rundt digitale plattformer. Jeg er spesielt interessert i nyere arbeid som tilnærmer seg poesi som en språkorientert kunstform som kan både vise og gjøre motstand mot at språklige uttrykk blir annektert og gjort til en handelsvare for profit. Kunstnere, forfattere og forskere som Pip Thornton, Søren Pold og Ben Gosser fremhever hvordan kritisk kreativ praksis kan synliggjøre og utfordre mange bekymringsfulle aspekter ved nåværende KI-verktøy og økonomien rundt, og adressere de ofte problematiske møtene mellom synlige brukergrensesnitt og infrastrukturene som støtter dem. Samtidig bør ikke poesi eller «kreativitet» bli sett som enkle svar på vanskelige spørsmål om KI. Som Pip Thornton foreslår, kan poesi (og andre kunstformer) likevel bidra til å utvikle «a range of language-based creative refusal tactics and interventions», og la forfattere, borgere og lokalsamfunn «reclaim some agency in the midst of AI's crushing inevitability».

# Envisioning a National Digital Platform: Interconnecting Music Archives, Data, and Discovery

Olivier Lartillot (Nasjonalbiblioteket, RITMO/Uo, MishMash) & Ingrid Romarheim Haugen (Nasjonalbiblioteket, MishMash)

The National Library of Norway hosts a range of digital music collections that are currently not readily accessible and are dispersed across various databases, encompassing formats such as music recordings, sheet music, and related materials like texts, videos, and interviews. We envision creating a centralised infrastructure that brings these collections together on a high-standard platform, adhering to the latest metadata standards to enhance search capabilities and facilitate integration with international networks. This would allow users, even those querying from other systems abroad, to access music data and related information seamlessly.

Additionally, we aim to extract musical information from diverse formats—such as audio recordings and sheet music images—promoting interconnectivity between these modalities. This foundational work will also pave the way for incorporating emerging AI technologies for music analysis and search in the future.

To inform our vision, we are surveying music researchers in Norway to gain insights into the needs of potential users. This infrastructure represents the initial step toward developing more user-friendly interfaces targeting a broader audience, ultimately extending the impact of music research

## Mot en nasjonal digital plattform: sammenkobling av musikkarkiver, data og utforskning

Nasjonalbiblioteket forvalter Norges største digitale musikktsamling, som består av lydopptak, noter, så vel som tekster, intervjuer og video. Deler av den er imidlertid vanskelig tilgjengelig og fordelt på ulike databaser. Vi ønsker å etablere en sentralisert infrastruktur som kobler sammen disse delene på en samlet plattform. Ved å bruke de nyeste metadata-standardene, er målet å legge til rette for utvidede søkemuligheter og integrasjon med internasjonale databaser og musikktsamlinger. Dette vil gi forskere, også de som søker fra utenlandske systemer, sømløs tilgang til musikkdata og relatert informasjon. I tillegg ønsker vi å trekke ut musikalsk informasjon fra ulike formater – som lydopptak og notebilder – for å gjøre det mulig å koble de ulike modalitetene sammen. Dette grunnarbeidet vil også legge til rette for å ta i bruk nye AI-teknologier for musikkanalyse og søk i fremtiden.

Som en del av arbeidet mot en nasjonal forskningsinfrastruktur, vil vi gjennomføre en undersøkelse blant musikkforskere i Norge, for å få innsikt i behovene til potensielle brukere.

Den foreslalte infrastrukturen representerer det første steget mot å utvikle mer brukervennlige grensesnitt rettet mot et bredere publikum, og dermed styrke musikkforskningens gjennomslagskraft.



## "No Slave to the Machine!"

# Why Digital Scholarly Editions Ought to be Datasets, not Interfaces An Interference

Annika Rockenberger (Universitetsbiblioteket, Universitetet i Oslo)

In this intentionally provocative paper, I investigate the consequences of an ‘interface first’ approach, a strategy often observed in modern digital scholarly editions. With ‘interface first’, I mean that modes and qualities of digital visual presentation via an end-user interface dictate the choice of data format and encoding, instead of philological principles, rigorous argumentation, encoding standards and domain-specific, empirically founded user expectations. I will draw from an in-depth survey of digital scholarly editions that I conducted in 2024 and 2025 at the University of Oslo, as well as discussions of the ‘interface first’ paradigm in the international academic community. Instead of ‘interface first’, I will argue that by employing a ‘dataset first’ approach, many of the current and long-term challenges DSEs face are avoidable. While I do not pose that researchers can access, analyse and manipulate DSEs without any interface, I claim that the primacy of the end-user interface must be brought to an end. In conclusion, I will demonstrate that the preservation and archiving issues DSEs currently face can be avoided under a ‘dataset first’ paradigm, resulting in significantly lower and more sustainable maintainability and hosting demands on technical, administrative, and human infrastructure at higher education and cultural heritage institutions.

I dette bevisst provoserende foredraget undersøker jeg konsekvensene av en «interface first»-tilnærming, en strategi som ofte observeres i moderne digitale vitenskapelige utgaver. Med «interface first» mener jeg at modi og kvaliteter av visuell presentasjon via et brukergrensesnitt dikterer valget av dataformat og koding, i stedet for filologiske prinsipper, argumentasjon, kodingsstandarder og domene-spesifikke, empirisk funderte brukerforventninger. Jeg vil trekke fra en grundig undersøkelse av digitale faglige utgaver som jeg gjennomførte i 2024 og 2025 ved Universitetet i Oslo, samt diskusjoner om «interface first»-paradigmet i det internasjonale akademiske miljøet. I stedet for «interface first» vil jeg argumentere for at ved å bruke en «dataset first»-tilnærming, er mange av de nåværende og langsiktige utfordringene digitale utgaver står overfor unngåelige. Selv om jeg ikke påstår at forskere kan få tilgang til, analysere og manipulere digitale utgaver uten noe som helst grensesnitt, hevder jeg at forrangsstilling av brukergrensesnittet må settes en stopper for. Avslutningsvis vil jeg demonstrere at bevarings- og arkiveringsproblemene digitale utgaver møter i dag kan unngås under et «dataset first»-paradigme, noe som fører til betydelig lavere og mer bærekraftige vedlikeholds krav og vertsdrift knyttet til teknisk, administrativ og menneskelig infrastruktur ved høyere utdannelses- og kulturarvinstitusjoner.



## Reading Food: presenting its infrastructure and user interface(s)

Diana Santos (Universitetet i Oslo), Elizaveta Khachaturyan (Universitetet i Oslo), Michael Preminger (OsloMet), Åse Kristine Tveit (OsloMet) & Eckhard Bick (Syddansk Universitet)

In the Reading Food project, we investigate what four literatures (Norwegian, Italian, Portuguese and Brazilian) can tell us about food and drink in the 1840-1920 period. The project encompasses several different aims: a lexical taxonomy of F&D in the three languages, a classification of metaphors and idiomatic expressions related to this domain, a novel annotation of uses of F&D in literary works (after all, a significant number of mentions of food and drink in the texts analysed are not literal), an investigation of pragmatic issues associated with F&D, and the study of food-related episodes and their purposes in literary text. In a nutshell, both linguistic and literary problems associated with food and drink.

In this presentation, we motivate and explain our choice to use CorpusEye, <https://corp.visl.dk>, as the infrastructure to make browseable and queriable the syntactically and semantically annotated corpora, which have been initially created and annotated in the European COST action “Distant Reading for European Literary History”, <https://www.distant-reading.net> and are being enriched by the Reading Food project.

We will also discuss our choice of organising user workshops with literature scholars in order to design the interface for food-related episodes (e.g. meals and food preparation), something that as far as we know is a unique feature of our project.

### Reading Food: Infrastruktur og brukergrensesnitt

I prosjektet *Reading Food* gransker vi hva skjønnlitteraturen i fire land (Norge, Italia, Portugal og Brasil) kan fortelle oss om mat og drikke i tidsrommet 1840-1920. Vi har fem mål for prosjektet: Utvikle en leksikalsk taksonomi for mat og drikke i tre språk; Klassifisere metaforer og idiomatiske uttrykk som gjelder mat og drikke; Lage en ny måte å annotere hvordan mat og drikke brukes i litterære verk. (For mange ganger nevnes ikke mat og drikke i bokstavelig mening.); Undersøke pragmatiske forhold som gjelder mat og drikke, og granske avsnitt som gjelder mat - hvilke formål de har i litterære tekster.

Kort sagt både språklige og litterære problemer i samband med mat og drikke. I denne presentasjonen forklarer vi valget av CorpusEye, <https://corp.visl.dk>, som infrastruktur for å søke i de syntaktisk og semantisk annoterte korporaene som har blitt lagd og annotert i det europeiske prosjektet *COST action “Distant Reading for European Literary History”*, <https://www.distant-reading.net>, og som vi holder på å videreutvikle innen prosjektet *Reading Food*. Vi kommer også til å drøfte hvordan vi har valgt å legge opp arbeidsseminar med litteraturvitere for å lage et nytt grensesnitt for å finne og kategorisere avsnitt der mat og drikke omtales (for eksempel måltider og matlaging). Så vidt vi veit har dette ikke blitt gjort tidligere.

# Forsvunnet, men bevart: Lenkeråte som utfordring for replikerbar forskning

Jon Carlstedt Tønnessen (Nasjonalbiblioteket)

Verdensveven er et sentralt informasjonssystem for både forskning og samfunnsliv.

Forskningsresultater publiseres primært på nett, og en betydelig andel av litteraturen og kildene som ligger til grunn finnes utelukkende på nett. Samtidig har nettinnhold en iboende flyktighet som skaper en rekke utfordringer. Domenenavn oppstår, forsvinner og endrer eierskap. Innholdet blir stadig endret, flyttet eller slettet (Chapekis et al, 2024a). Når sentrale kilder for en studie kan være endret eller utilgjengelig før resultatene har blitt publisert utfordres et grunnleggende vitenskapelig prinsipp om etterprøvbarhet.

Dette paperet adresserer problemet gjennom å undersøke forekomsten av lenkeråte på .no-domenet. Basert på nettsider arkivert av Nasjonalbiblioteket (NB) i 2005 og 2020, har man besøkt disse URLene på nytt i 2025 og sammenliknet HTTP statuskoder (Chapekis et al, 2024b). Resultatene viser en signifikant andel «råtne» URLer, særlig for den eldre delen av datasettet. For et subsett av data fra nettavisar har man også undersøkt utgående lenker og funnet et betydelig antall artikler med én eller flere brutte lenker. Dette bekrefter at innhold på .no-domenet er like flyktig og ustabilt som det som er observert globalt (Chapekis et al, 2024a; Alam, 2025).

I lys av denne erkjennelsen diskuteres utfordringer med gjenbruk av nettbaserte forskningsdata, og hvordan det NFR-finansierte prosjektet «Forskningsinfrastruktur for nettdata» kan bidra til mer etterprøvbar forskning (Brügger, 2021; Forskningsrådet, 2025). Presentasjonen konkluderer med at kulturarvsinstitusjoner og forskere kan bidra til mer persistente forskningsresultater, blant annet gjennom:

- a) Bruk av tidssituerte data fra nettarkiv
- b) Aktive henvisninger med persistente identifikatorer

## Bibliografi og kilder

Alam, S. (2025). «Lost, but Preserved: A Web Archiving Perspective on the Ephemeral Web». [Foredrag]. *IIPC Web Archiving Conference 2025*. Oslo, Norway.

Ben-David, A. (2021). «Critical Web Archive Research». In: Gomes, D., Demidova, E., Winters, J., Risse, T. (eds) *The Past Web*. Springer, Cham. [DOI:10.1007/978-3-030-63291-5\\_14](https://doi.org/10.1007/978-3-030-63291-5_14)

Brügger, N. (2021). «The Need for Research Infrastructures for the Study of Web Archives». I Gomes et al. (Red.), *The Past Web: Exploring Web Archives* (s. 217–224). Springer International Publishing. [DOI:10.1007/978-3-030-63291-5\\_17](https://doi.org/10.1007/978-3-030-63291-5_17)

Chapekis et al. (2024a). «When Online Content Disappears». Pew Research Centre. Arkivert av Internet Archive, 17.05.2024. [URN:PWID:archive.org:2024-05-17T18:31:42Z:page:https://www.pewresearch.org/data-labs/2024/05/17/when-online-content-disappears/](https://www.pewresearch.org/data-labs/2024/05/17/when-online-content-disappears/)

Chapekis et al (2024b). «Methodology». Pew Research Centre. Arkivert av Internet Archive, 18.05.2024. [URN:PWID:archive.org:2024-05-](https://www.pewresearch.org/data-labs/2024/05/17/when-online-content-disappears/)



[18T23:05:53Z:page:https://www.pewresearch.org/data-labs/2024/05/17/when-online-content-disappears/](https://www.pewresearch.org/data-labs/2024/05/17/when-online-content-disappears/)

Frew et al. (2025). «Temporally Extending Existing Web Archive Collections for Longitudinal Analysis». *arXiv preprint*. [DOI:10.48550/arXiv.2505.24091](https://doi.org/10.48550/arXiv.2505.24091).

Forskningsrådet (2025). «Nasjonal forskningsinfrastruktur for nettdata». Arkivert av Internet Archive, 25.05.2025. [URN:NW-25T09:02:18Z:page:https://prosjektbanken.forskningsradet.no/project/FORISS/350342](https://urn.nbn.no/URN:NW-25T09:02:18Z:page:https://prosjektbanken.forskningsradet.no/project/FORISS/350342)

van Horik, R. et al. (2025). *FAIR Research Data Management - A deeper dive into putting FAIR RDM into practice. Part 1.* EU Open Research: Medical and Health Science. Belgium. [COI: 20.500.12592/8ngfddh](https://doi.org/10.500.12592/8ngfddh)

Singer, E. (2025): «Thousands of U.S. Government Web Pages Have Been Taken Down Since Friday». *New York Times*. Arkivert av Internet Archive, 02.02.2025. [URN:NW-02T20:27:22Z:page:https://www.nytimes.com/2025/02/02/upshot/trump-government-websites-missing-pages.html](https://urn.nbn.no/URN:NW-02T20:27:22Z:page:https://www.nytimes.com/2025/02/02/upshot/trump-government-websites-missing-pages.html).



## Reel me in – Exploring the Dichotomy of Storytelling and Spectacle in Reels on Social Media

Håvard Vibeto (Universitetet i Innlandet)

Reels on platforms such as Instagram, TikTok, and Facebook have become one of the most popular audiovisual formats of our time. While brief and ephemeral, they blend elements of narrative and spectacle that can be traced back to earlier media histories, from Tom Gunning's "cinema of attractions" to the aesthetics of blockbuster cinema. My article explores this dual identity of Reels: as fragmented stories that engage audiences emotionally, and as spectacles that rely on technological novelty and visual intensity.

Yet in order to study Reels within a humanities framework, one must confront a central challenge: how to treat Reels as research data. Unlike films in an archive or artifacts in a museum, Reels are embedded in opaque infrastructures of platforms, algorithms, and interfaces. They are difficult to collect systematically, subject to constant deletion, remixing, and recommodification, and highly dependent on personalized feeds. This raises methodological questions: Which Reels should be chosen? According to what parameters—virality, hashtags, themes, or algorithmic visibility? How can reproducibility be ensured when the dataset itself is unstable?

My article depends on finding ways to handle these methodological and infrastructural dilemmas in order to answer its research questions. I therefore use this presentation to articulate the challenges of identifying, collecting, and selecting Reels in a way that is methodologically rigorous, rather than based on arbitrary or personal preference. By opening up the "messy middle" of data selection, I hope to spark discussion and gather feedback on possible approaches to working with such volatile materials.



## Glemte Stemmer i norsk musikkarkiv – digital infrastruktur for fremtidens musikkarkiv

Daniel Henry Øvrebø (Universitetet i Bergen) & Arnulf Christian Mattes (Senter for Griegforskning)

Prosjektet Glemte Stemmer, støttet av UiBs humaniora-satsning, utvikler en digital infrastruktur for kuratering og formidling av glemte norske komponister, med utgangspunkt i vestnorsk musikkliv. På tvers av musikkvitenskap, digital humaniora, og arkiv- og bibliotekarbeid, bygger Senter for Griegforskning et framtidsrettet musikkarkiv basert på det åpne XML-baserte verktøyet MEI (Music Encoding Initiative).

I samarbeid med Zentrum Musik – Edition – Medien, Universität Paderborn og Centre for Digital Music Documentation ved Akademie der Wissenschaften und der Literatur, Mainz, har vi i løpet av våren 2025 implementert merMEId (Metadata Editor and Repository for MEI Data) til bruk i arbeidet med å organisere og formidle digitale verkkataloger til norske komponister tilknyttet den nasjonale musikkarkiv-satsingen, fra Edvard Grieg til Bergenskomponisten Anne-Marie Ørbeck. Sistnevntes arkiv med brev og manuskripter ble donert til Bergen Offentlige Bibliotek, og er utlånt til Universitetsbiblioteket i Bergen.

UiB ved Senter for Griegforskning og Griegakademiet er også partner av det nye MishMash-senterets satsing på KI for Kulturarv (WP6). I denne sammenheng vil prosjektet bidra med innsikt i hvordan AI kan styrke tilgjengelighet, representasjon og etisk kuratering av musikkarkiv i arkiver og biblioteker. Det åpner for en større satsing på å gjøre musikkpartiturer og manuskripter maskinlesbare gjennom verktøyet merMEId.

I presentasjonen vil vi fokusere på potensialet for denne typen maskinlesbar musikktranskripsjon til å adressere MishMash-senterets forskningsspørsmål knyttet til mangfold i kulturarven. Mer spesifikt drøfter vi senterets første forskningsspørsmål “How can we significantly advance AI to effectively discover, organise, and showcase the vast richness of cultural heritage and creative works in all forms?”

## Forgotten Voices in Norwegian Musical Heritage – Digital Infrastructure for the Music Archive of the Future

The project Forgotten Voices, supported by the University of Bergen’s Humanities Strategy, is developing a digital infrastructure for the curation and dissemination of forgotten Norwegian composers, with a focus on musical life in Western Norway. Bridging musicology and digital humanities together with archival and library work, the Grieg Research Centre is building a future-oriented music archive based on the open XML-based tool MEI (Music Encoding Initiative).

In collaboration with Zentrum Musik – Edition – Medien at Universität Paderborn and the Centre for Digital Music Documentation at the Academy of Sciences and Literature in Mainz, we implemented merMEId (Metadata Editor and Repository for MEI Data) during spring 2025. This tool is used to organize and present digital work catalogues of Norwegian composers associated with the national musical heritage initiative, ranging from Edvard Grieg to Bergen composer Anne-Marie Ørbeck. The



latter's archive of letters and manuscripts was donated to Bergen Public Library and is currently on loan to the University Library in Bergen.

UiB, through the Grieg Research Centre and the Grieg Academy, is also a partner in the new MishMash Centre's initiative on AI for Cultural Heritage (WP6). In this context, the project contributes insights into how AI can enhance accessibility, representation, and ethical curation of musical heritage in archives and libraries. This opens up broader efforts to make musical scores and manuscripts machine-readable through the merMEId tool.

The presentation will focus on the potential of this type of machine-readable music transcription to address the MishMash Centre's research questions related to diversity in cultural heritage. More specifically, we will discuss the centre's first research question: "How can we significantly advance AI to effectively discover, organise, and showcase the vast richness of cultural heritage and creative works in all forms?"