

KEYNOTE:

JONATHAN WESTIN, GRIDH

A THOUSAND YEARS OF GRAFFITI
WAR, CULTURAL HERITAGE IN
PERIL, AND DIGITAL HUMANITIES

DHKO.NO

KONFERANSE 2024

Digitale og digitaliserte spor

21 - 22 OKTOBER
TRONDHEIM
NTNU



PROGRAM 21. & 22. OKTOBER

NO: DHKOs årlige konferanse, i samarbeid med Universitetsbiblioteket ved NTNU, forskergruppa **Verden på norsk**, og forskergruppen **Media, Data, Museums** (HF, NTNU).

EN: DHKO's annual conference, in cooperation with the university library of NTNU, the research group **The World in Norwegian**, and the research group **Media, Data, Museums** (HF, NTNU).

(21.10) DAG 1: VERKSTED (Workshops, only in Norwegian)

12:30 - 13:00	Registration	Sted: Dragvoll, NTNU		
13:00 - 14:30	parallel-sesjon	NO	<p>Innføring for humanister (og samfunnsvitere) i programmeringsspråket Python og bruk av Jupyter Notebooks for databehandling og analyse</p> <p>Dette er et verksted for de som har ingen kunnskap om programmering, men har lyst til å få en innføring. Verkstedet tar utgangspunkt i Versjon 1.0 av den nye norske heldigitale og åpent tilgjengelige læreboka "Databehandling og analyse med Python on OpenRefine".</p> <p>Verkstedet gir først en kort presentasjon av boka, og kort og grunnleggende om hvordan et programmeringsspråk fungerer. Deretter vil hovedtyngden av verkstedet være å praktisk komme i gang med programmering og prøve ut enkle analyser på et datasett. Et av datasettene som vi vil ta utgangspunkt i er "Top 500 "greatest" novels" (amerikansk datasett).</p> <p>Målgruppen er alle humanister (og samfunnsvitere) som ønsker en innføring i hvordan bruke programmeringsspråk for å gjøre enkle analyser. Det kreves ingen forkunnskaper. Her kan du stille så grønn som bare det.</p> <p>ROM: Kursrom 6575 (bygg 6b) på biblioteket Dragvoll</p>	Camilla H Soelseth, OsloMet
13:00 - 14:30	parallel-sesjon	NO	<p>Introduksjon til SAMLA.no og innføring i verktøyet Transkribus (maskinassistert tekstgjenkjenning)</p> <p>Mange institusjoner digitaliserer og publiserer arkivmateriale. Et av prosjektene er SAMLA – digitalisering av norske tradisjonsarkiv, der materiale fra tre kulturhistoriske arkiv er publisert på arkivportalen samla.no: Etno-folkloristisk arkiv (UiB), Norsk Folkeminnesamling (UiO), og Norsk etnologisk granskning (Norsk Folkemuseum). Mye arkivmateriale består av håndskrevne tekster som kan være vanskelig tilgjengelig både for publikum og forskere, rett og slett fordi</p>	Therese Foldvik, Universitetet i Oslo

			<p>håndskriften er vanskelig å lese. Ett viktig steg mellom det fysiske og det digitale er å gjøre det digitale lesbart for brukere, og med det klargjøre materialet for videre spørsmål og analyser. SAMLA bruker maskinassistert tekstgjenkjenning ved hjelp av verktøyet Transkribus for å gjøre materialet mer tilgjengelig. Verktøyet kan brukes på flere måter, enten man trenger hjelp til å lese håndskriften, ønske om en digital tekst som lesestøtte til bildefilene, eller å lage en korrekt transkripsjon som kan fungere som et selvstendig digitalt objekt.</p> <p>Verkstedet starter med en presentasjon av SAMLA, og hvordan Transkribus brukes for å transkribere materialet, både enkeltdokumenter og større mengder materiale. Så følger en kort gjennomgang av de grunnleggende funksjonene, etterfulgt av en arbeidsøkt der deltakere kan jobbe med eget materiale.</p> <p>Målgruppen er de som jobber med tekst, og som vil ha en innføring i Transkribus som verktøy. Verkstedet krever at deltakerne på forhånd har opprettet en konto i Transkribus, og helst har noen sider med tekstmateriale i PDF eller JPEG tilgjengelig for opplastning.</p> <p>ROM: D-lab (rom 8563, bygg 8) på biblioteket Dragvoll</p>	
14:30 - 14:45	PAUSE			
14:45 - 16:15	parallel-sesjon	NO	<p>Å ta i bruk Nasjonalbiblioteket sine byggesteiner for tekstanalyse</p> <p>Dette verkstedet gir en kort introduksjon til Nasjonalbiblioteket sin API, og hva det er mulig å hente ut både tekst, metadata og datapakker for analyse fra Nasjonalbiblioteket sitt system gjennom dette brukergrensesnittet.</p> <p>API'et til Nasjonalbiblioteket inneholder også ulike byggesteiner for gjøre forskjellige typer analyse, som tekstlikhet, LIX-score, datagenerering for word2vec, samforekomstanalyse (tematisk analyse).</p> <p>Verkstedet går praktisk til verks i hvordan lage en Jupyter Notebook som</p>	Lars G Bagøien Johnsen, Nasjonalbiblioteket

			<p>kombinerer data og metadata med byggestener for å utføre ulike typer analyser.</p> <p>Målgruppen er de som ønsker å jobbe med tekstkorpset fra Nasjonalbiblioteket sin samling. Forkunnskaper er grunnleggende kunnskaper om Python og bruk av Jupyter Notebooks. Det er mulig å kombinere innføringsverkstedet i Python tidligere denne dagen med dette kurset.</p> <p>ROM: D-lab (rom 8563, bygg 8) på biblioteket Dragvoll</p>	
14:45 - 16:15	parallellesjon	NO	<p>Innføring i analyseverktøyet 4CAT og innsamlingsverktøyet Zeeschuimer for undersøkelser av materiale fra Instagram og TikTok spesielt, og sosiale medier og nettfora generelt.</p> <p>Verkstedet gir en kort introduksjon til et open source brukergrensesnitt for analyse av materiale fra sosiale medier (4CAT). Etter dette følger en gjennomgang av hvordan man går frem for å bestemme hva man vil samle inn hvis man ønsker å undersøke et fenomen på sosiale medier.</p> <p>Videre vil verksteddeltagere få mulighet til å ta i bruk nettleserapplikasjonen Zeeschuimer for å samle inn et korpus fra sosiale medier som de så enten kan få utforske å analysere i 4CAT, eller se nærmere på i Jupyter Notebook (med bruk av datapakken Pandas), eventuelt (for mindre korpus) i excel.</p> <p>Målgruppen er de som ønsker å samle inn korpus fra sosiale medier og gjøre utforskende analyser av aggregerte mønstre av datamaterialet på en forsvarlig måte. Det kreves ingen forkunnskaper, med mindre man ønsker å jobbe med datasettet i Jupyter Notebook.</p> <p>ROM: Kursrom 6575 (bygg 6b) på biblioteket Dragvoll</p>	Camilla H Soelseth, OsloMet

_DIGITAL KULTUR LIVE | DIGITAL CULTURE LIVE

16:30 - 17:30	Artistic Keynote + 15 min Q&A	<p>Code Poetry: The constraints of programming and making art with code</p> <p>Chris Kerr (UK) will explore how code poetry can be seen as a form of highly-constrained writing. He will look back at the history of French collective Oulipo and their engagement with computer science, focusing on work by Oulipo members Noël Arnaud and Valérie Beaudouin. Chris will also explore how programmers play with constraints today, from "code golf" to esoteric programming languages. Meanwhile, he will present the constraints of his collection <code>./code --poetry</code>, co-authored with Daniel Holden</p> <p>Bio: Chris Kerr is a poet who lives in London. He co-wrote <code>./code --poetry</code> with Daniel Holden. This work, published by Broken Sleep Books in 2023, is a collection of code poems each written in the source code of a different programming language. Every poem is also a valid program which produces a visual representation of itself when compiled and run (see code-poetry.com). Some code poems from <code>./code --poetry</code> are being exhibited in the <i>Maschinenpoesie</i> (Machine Poetry) exhibition at Museum Strauhof in Zurich, Switzerland until January 2025.</p> <p>ROM: D-lab (rom 8563, bygg 8) på biblioteket Dragvoll</p>
---------------	-------------------------------	--

_KONFERANSEMIDDAG | CONFERENCE DINNER

17:30	MIDDAG	<i>Information is sent out to everyone who registered for the dinner in a separate e-mail.</i>
-------	--------	--

(22.10) DAG 2: Presentasjoner & keynote | Conference presentations & keynote

08:30 - 09:00	Registration	Sted: Kalvskinnet, rom: LY32, Lysholmbygget, 3.etasje/3rd floor.		
09:00 - 09:10	Velkommen	NO & EN	Velkommen og praktisk informasjon Welcome and practical information	Camilla H. Soelseth, DHKO, og Inger H. Schmidt-Melbye, UB NTNU

09:10 - 10:05	(Academic) Keynote + short Q&A	EN	<p>A Thousand Years of Graffiti War, Cultural Heritage in Peril, and Digital Humanities</p> <p>Saint Sophia Cathedral in Kyiv is a world heritage a thousand years old and a source of knowledge that is now threatened by the ongoing war. Carved into its frail frescoes are more than seven thousand inscriptions in multiple languages from those that have visited the cathedral over the centuries. Observations and marks from travellers, merchants, and princesses mix with crude drawings and messages of yearnings and pleads for salvation. Jonathan Westin, Director of GRIDH, Gothenburg Research Infrastructure in Digital Humanities, presents how he together with his colleagues strive to help Ukraine to preserve and analyse these threatened historical sources digitally before they are destroyed, and what role Digital Humanities can play in contingency planning and securing cultural heritage for the future.</p> <p><i>BIO: Jonathan Westin is an associate professor in Conservation and current Director of Gothenburg Research Infrastructure in Digital Humanities (GRIDH) at the University of Gothenburg. His research focuses on how we form our perception of culture through representations, and how these representations become part of our cultural heritage. By focusing on the communicative aspects of culture management, I approach the creation of visual representations as a negotiation process between new research and established images.</i></p>	Jonathan Westin, Göteborg universitet, Sverige
10:05 - 10:15	Pause	Short break		
		PRESENTASJONER / PRESENTATIONS rom: LY32, Lysholmbygget, 3.etasje/3rd floor.		

		TEMA: Kartlegging og klassifisering <i>Chair: Per-Olav Broback Rasch</i> 10 -15 min presentations, Q&A for all at the end	
10:15 - 11:15		<p>NO The geography of Norwegian literary history 1814-1849</p> <p>Can digital methods help us to dig out a better knowledge of literary history? If we use digital methods to identify, count and map out the place names mentioned in literary history, what is the geographical horizon(s) organizing it? Which places stand out as the literary centers and peripheries, and why? My paper will discuss and try to answer these questions.</p>	Frode Boasson, NTNU
		<p>NO Å bruke LLMer til geotagging av tekst</p> <p>En standard prosess for å finne steder i en tekst, og den som er brukt i NFR-prosjektet ImagiNation ved NTNU, er 1) finn alle tekstuelle kandidater til et sted, såkalt NER (Named Entity Recognition) og 2) fra listen over mulige steder konsulter en database over lokasjoner (for eksempel GeoNames) for å finne beste plassering med lengde- og breddegrad.</p> <p>Et problem for slike system ligger både i selve NER, men også i disambiguering, for eksempel er Gaula ikke bare en elv utenfor Trondheim, Gaula er også en elv i India.</p> <p>Som en del av å anvende LLMer (Store Språkmodeller) skal vi se på hvordan språkmodeller kan instrueres til å analysere store mengder tekst og gjøre (1) og (2) over i en operasjon. Altså at språkmodellen både finner stedsnavn og kobler navnet til en lokasjon som en output. Her er et eksempel på output fra boken "Tre i Norge ved to av dem". Her er LLMen (ChatGPT) instruert til å antyde nogenlunde hvor plassen er, om den ikke finner en nøyaktig posisjon.</p> <p>Kristiania (now Oslo): 59.9139° N, 10.7522° E Fjorden - Place: Likely refers to the Oslofjord or a similar fjord in Norway (Oslofjord coordinates: 59.2° N, 10.7° E) Victoria Hotell - Organization, Likely located in Oslo, Norway Grand Hotell: Coordinates: 59.9145° N, 10.7376° E Area: Oslo, Norway</p>	Lars G Bagøien Johnsen, Nasjonalbiblioteket & Anders Skare Malvik, NTNU

		I foredraget vil vi komme inn på forskjellige måter å instruere på, og variasjoner i output, som markup eller andre format.	
	NO	<p>Nasjonalbibliotekets prototype på en bildesøkapplikasjon: Bildeklassifisering basert på bildemodellen SigLIP</p> <p>Digitale verktøy for tekstprosessering og -analyse har blitt en viktig del av å tilgjengeliggjøre digitaliserte samlinger. Fremskritt innen maskinsyn, spesielt de basert på dyp læring, har nylig introdusert lignende muligheter for visuelle materialer. Siden mange bøker i Nasjonalbibliotekets samling har visuelt innhold som illustrasjoner, diagrammer og bilder kan det å utnytte disse teknologiene åpne nye innganger til samlingen og gjøre visuelle materialer mer tilgjengelige.</p> <p>I denne presentasjonen viser vi frem en prototype på en bildesøkapplikasjon, utviklet for å utforske bilder i Nasjonalbibliotekets bøker fra før år 1900. Vi har hentet 422 589 utsnitt fra sider som tekstgjenkjenningsdelen av Nasjonalbibliotekets digitaliseringsløype markerte som visuelle elementer. Data som er hentet ut av en slik prosess har helt unike utfordringer. For eksempel segmenterer digitaliseringsløyna noen sider feil, slik at blanke sider, deler av tabeller o.l. markeres som visuelle elementer og blir med i datasettet.</p> <p>Vi vil vise resultater fra analyse basert på boksideutsnittene hvor vi undersøkte hvor godt tre populære bildemodeller – Vision Transformers (ViT), Contrastive Language-Image Pre-Training (CLIP) og Sigmoid loss for Language-Image Pre-training (SigLIP) – fungerte som grunnlag for søk og klassifisering av visuelle element i bøker. Resultatene viser at SigLIP gjorde det noe bedre enn CLIP og ViT for både søk og klassifisering. I tillegg viser vi at bildeklassifisering basert på SigLIP kan brukes for å effektivt rense bildesamlinger hentet fra bibliotekets digitaliseringsløype.</p>	Marie Roald, Nasjonalbiblioteket
11:15 - 11:30	Pause	Coffee Break	
		<p>THEME: Digital Culture <i>Chair: Camilla H. Soelseth</i></p> <p>10 -15 min presentations, Q&A for all at the end</p>	

11:30 - 12:30		EN	<p>Traces of AI Generation, 'Wild Forensics,' and the Hermeneutics of Suspicion</p> <p>While discussions about prompt-to-image AI-platforms like DALL·E, Stable Diffusion, or Midjourney are overshadowed for a while now by the rapid development of prompt-to-text-applications generative imagery continues to flood social media platforms. AI platforms' abilities to generate images that very closely resemble photographs has received a lot of attention and stirred up new questions about credibility and deception, especially in the form of meme series like the "Balenciaga Pope," Donald Trump's "AI-imagined arrest," or a series called "Time Travel Selfies." Roland Meyer recently observed an interesting mutual relation between AI-memes and AI-fakes through wild forensics, a "collective speculations about whether a shared pictures was manipulated or not, the search for revealing details and traces of manipulation."</p> <p>My proposed contribution presents some current observations on the integration – and resistance – of AI-generated imagery within processes of media/image convergence and circulation. My assumption is that there can be no talk of a seamless "integration" into the wider media landscape at the moment (within films, comic books, or video games, for example) as one of countless other image production techniques. The medial "site" of AI image circulation thus appears to remain anchored to social media communication and fan cultures, insofar as the assumption that *any* image now *could* be AI-generated is increasingly present as a generalized "hermeneutics of suspicion."</p>	Lukas R.A. Wilde, NTNU
		EN	<p>Echoes of History? The Impact of AI-Generated Historical Figures on Informal Education</p> <p>Imagine having a text conversation about philosophy with Aristotle, discussing capitalism with Karl Marx through voice chat, or asking a holographic Marie Curie about nuclear physics. These interactions used to be confined to science fiction narratives. Now, they are seemingly becoming a reality due to advancements in generative artificial intelligence (AI) and large language models.</p> <p>Histobots — a term I have coined for chatbots that blend historical reenactment with generative AI — are emerging as significant tools for informal education in museums and online. This study critically reflects on AI portrayals within museums, such as an avatar of Vincent van Gogh in the Musée d'Orsay and representations of Marie Curie on the chatbot applications Hello History and Character.Ai. While these emulations of historical figures offer new and seemingly interactive educational experiences, they also raise substantial concerns. Histobots reflect the challenges of and with artificial intelligence, such as data privacy, misinformation, and other ethical issues. This study evaluates</p>	Lina Harder, UiB

			histobots through case study reviews and autoethnographic approaches, highlighting the risks of inaccurate information, potential biases, and the ethical implications of anthropomorphising AI. Additionally, it examines the impact of AI as a cultural phenomenon through the imaginaries created around these tools, how they influence our perception of AI, and the potential consequences for educational practices.	
	EN	A distant reading of the r/wallstreetbets and the Gamestock craze The Reddit forum r/wallstreetbets was founded in April 2012 and for years was a relatively small forum dedicated to discussion of options trading strategies (both orthodox and not) and betting on corporate earnings calls. In January 2021 it became almost a household name as it became the organizational home for thousands or even millions of stock day-traders who bought massive amounts of a failing mall-based video game store known as Gamestop. The loosely coordinated buying of Gamestop resulted in its price exploding and caught out some major Wall Street firms who had shorted the stock and thus lost sizeable amounts of money. As r/wallstreetbets grew it attracted some academic and journalistic interest and once it became (in)famous for the Gamestock craze explanations and exegesis of the forum came from all corners. However, these analyses were based on participant interviews or close readings, invariably of the most popular posts. However, given that discussion board comprises some 650 million words from 2012 to mid-2021, close readings and interviews can only offer a partial picture. This size makes it an object where computational text analysis, distant reading as Franco Moretti has called it, provides a way to visualize and map the corpus as a whole. This presentation will look at the results of topic modelling of the r/wallstreetbets corpus and try to tease out the major narratives and ask just what were the forum's millions of users talking about in January 2021.	Gregory Ferguson-Cradler, HINN	
12:30 - 13:15	Lunsj	Lunch		
		THEME: Digital Cultural heritage / Digital kulturarv <i>Chair: Nina Lager Vestberg</i>		

			10 -15 min presentations, Q&A for all at the end		
13:15 - 14:15		EN	<p>Henrik Ibsen's Writings: Exploring Ibsen's Legacy through Digital Innovation</p> <p>This presentation will introduce the unique digital resources related to Henrik Ibsen at the Center for Ibsen Studies (CIS), University of Oslo. The main focus will be on the Digital Scholarly Edition of Ibsen's complete works, a pioneer project in digital humanities and the most ambitious project in digital scholarly editing in Norway. It offers open access to high-quality facsimiles and transcriptions of all Ibsen's works in TEI XML, along with annotations, editorial notes and historical introductions. The project was a collaboration between universities and the archives, libraries, and museums sector, and the edition is now managed by the Centre for Ibsen Studies at the University of Oslo. The other resources at CIS include IbsenStage: a relational database on international performances of Ibsen's works, an international Ibsen bibliography, a text archive with numerous full text translations and reviews, and a media archive with performance documentation, photographs and paintings of Ibsen, and Ibsen's own art. Together these resources form a unique entrance to explore Ibsen's legacy from all angles and perspectives.</p>	Nina Marie Evensen, Senter for Ibsen-studier, UiO	
			<p>Portal Propagations: The Rediscovery of the Urnes Portal</p> <p>The northern portal of Urnes Stave Church stands spoliated and preserved within the 12th-century building as the seemingly frozen memory of an earlier past. Yet this portal continues to be extended through the circulation of its hyper-reproduced image with fluid and shifting receptions. This paper will explore this heritage icon through Digital Humanities methods as an archival and historiographical study of its paintings, drawings, plaster casts, carvings, tattoos, coins, and commercial packaging. It asks questions of museum and preservation practices and how those responsibilities might extend to the communication of monuments, and the reception of their extended image.</p> <p><i>Portal Propagations</i> also bases its exploration on an ethnographic study of digital reproduction, as part of a residency at <i>Factum Foundation for Digital Technology in Preservation</i>. The Madrid-based practice pioneers an approach to facsimile production through new imaging techniques and in the return of the digital recordings to a physical state, exploring a closing gap between the object and its material reproduction. This</p>	Nick Walkley, Arkitektur- og designhøgskolen	

		<p>study provides a foundation to explore how the Urnes Portal might be further reproduced, as rediscovered in line with its lost 11th-century architectural context as the western entrance to an earlier church. Working with digital scan data, this paper presents new images of the portal within a visualisation of the earlier building.</p> <p>Summarising from a concluding PhD study, as the scrutiny of a single artefact, this paper will explore the reproduction of the Urnes Portal through both past and future perspectives. Through understanding how the portal's past reproduction has contributed to the making of the monument, the paper explores how the continuation of its circulated image might be made in respect of its considered future growth.</p>	
	EN	<p>Remembering the cultural heritage of Jewish life before the Holocaust: Digitizing and translating historical Yiddish press sources</p> <p>This project is a born-digital translation of Yiddish press annotated and contextualised as a scholarly digital edition. The edition and supporting data are displayed on a static, open-source website. The focus lies in transforming scans of the newspaper in Five-Star Open-Source and Linked Data in a universally understood language without losing the essence of Yiddish press sources.</p> <p>דער אַמְתִּיר יִז – The Genuine Jew was published from February to October 1904 in Lviv, Ukraine. The scans are provided by the Austrian National Library already including the automated transcription in Hebrew letters in a separate text document. Taking those digitised sources, the text is manually translated and interpreted, as existing translation tools are not proficient (enough) for Yiddish. The translation is then followed by the annotation in TEI. Using the Cookiecutter library and XSLT stylesheets, the data is transformed and processed into HTML files which are then deployed on GitHub pages. On the website, the newspapers are published as text and viewable as HTML sources. People, places, Geopolitical Entities, important events, and Jewish concepts / traditions are linked into indices and to external databases like Wikidata or the YIVO encyclopaedia.</p> <p>Special focus is put on creating open, findable, accessible, interoperable, and reusable data by applying the FAIR principles. The project creates Yiddish sources available to non-Yiddish speakers to ensure remembrance of the cultural heritage of Jewish life before the Holocaust.</p>	Robin Luger, University of Vienna
14:15 - 14:30	Pause	Coffee Break	

		TEMA: Metodologiske veier videre <i>Chair: Alexandra Angeletaki</i>	10 -15 min presentations, Q&A for all at the end ONLY TWO PRESENTATIONS! EXTRA TIME FOR Q&A!	
14:30 - 15:30		NO	<p>Datametoder i folkloristikken: innovasjon og kontinuitet i fagets metodologiske utvikling.</p> <p>Datametoder i folkloristikken utgjør både innovasjon og kontinuitet i fagets metodologiske utvikling. For et fag som har vært dominert av kvalitative forskningsmetoder de siste femti årene, er introduksjonen og utviklingen av dataverktøy for folkloristisk forskning en betydelig utvidelse av fagets metodologiske arsenal. Dersom vi løfter blikket og skuer tilbake før 1970-tallet ser vi imidlertid at bruk av store datasett og beregningsmetoder var sentrale i folkloristiske studier, og at det er tydelige paralleller til datametodene av i dag. Når beregningsmetoder nå blir tatt i bruk i et paradigme dominert av kvalitativ metodologi, er det verdt å reflektere over hvordan disse metodene påvirker forutsetningene for folkloristisk kunnskapsproduksjon.</p> <p>I presentasjonen vil jeg anskueliggjøre aspekter ved de epistemologiske implikasjonene datametoder har for folkloristikken som fagfelt. For å diskutere datametodenes muligheter og begrensninger for kunnskapsproduksjon tar jeg utgangspunkt i sammenligning mellom beregningsmetoders anvendelse i to ulike paradigmer innen samme fagfelt. Hensikten med denne sammenligningen er å bidra til en dypere forståelse av hvordan vi kan navigere de metodologiske og epistemologiske utfordringene som følger med innføringen av datametoder i et fagfelt som tradisjonelt har verdsatt kvalitativ dybde. Ved å kontekstualisere dagens metodologiske innovasjoner innenfor en bredere faghistorisk ramme, håper jeg å bidra til en mer nyansert diskusjon om hvordan vi kan videreutvikle metodologiske strategier som er både faglig robuste og epistemologisk reflekerte.</p>	Marte Færevaag Wetteland, UiB
		NO	<p>Hvordan AI og spesielt LLM'er kan brukes som forskningsassistenter.</p> <p>I presentasjonen skal jeg ta for meg hvordan AI og spesielt LLM'er kan brukes som forskningsassistenter. I forlengelsen av temaet digitale spor, skal jeg se på to tilfeller. Det første handler om kartotek-kort (spesielt fra historiske kartotek) og digitalisering av dem,</p>	Lars G Bagøien Johnsen, Nasjonalbiblioteket

			<p>det andre om å behandle data som genereres fra tekst, som for eksempel samforekomstanalyser (kollokasjoner, temamodeller, word2vec med mer).</p> <p>For begge dreier det seg dels om støy i data, dels om modernisering og historiske former, og også hvordan finne geolokasjoner for termer som er knyttet til sted.</p> <p>Første case om katalog-kort, eller metadata fra ustrukturert tekst. Etter en første OCR står man med noe tekst, og den teksten inneholder informasjon om utgivelser. I den her delen skal jeg diskutere hvordan LLM'er kan instrueres til å hente ut og strukturere informasjonen på kortene, finne forfatter, tittel samt spesifikke koder, utgivelsår og -sted, samt utgave. Her kan de gi støtte til modernisering av stedsnavn. Med geolokasjon kan de gjøre en katalog-base direkte analyserbar med tanke på hvor dokumentene som beskrives har sitt opphav, og kan benyttes i visualisering av aspekter av katalogen.</p> <p>Andre case, om data fra samforekomstanalyse skal jeg ta for meg hvordan LLM'er kan instrueres til å si noe om spesielt historisk materiale, og se nærmere på modernisering og geolokasjon. Her tar jeg utgangspunkt i folkeslag, og hvordan de er beskrevet i 1800-tallslitteraturen. Det å geolokalisere folk er et annet problem enn å lokalisere et stedsnavn, noe LLM'er fungerer utmerket til.</p>	
15:15 - 15:15	Pause		Coffee Break	
15:15 - 15:45	Avslutning	NO & EN	<ul style="list-style-type: none"> - Takk for i år! - Om DHKO 2025 (Hamar) og utvidelse av konferanse til å fra neste år ha lengre presentasjoner som vil kvalifisere til CFP til spesialutgave av DHNB Publications journal. - Invitasjon til å bli medlem av DHKO 2025 arbeidsgruppe - Tilbakemeldinger for å forbedre konferansen og nettverket 	Camilla H. Soelseth, DHKO
			POST-KONFERANSE-ARRANGEMENT, onsdag 23. okt	
11:30 - 12:00			<u>Kunnskap: Kaffe</u> - forskergruppa Verden på norsk holder et relatert DH-lunchforedrag. Sted: Biblioteket, Dragvoll	Forskergruppa Verden på norsk