

Bibliotekarstudentens nettleksikon om litteratur og medier

Av Helge Ridderstrøm (førsteamanuensis ved OsloMet – storbyuniversitetet)

Sist oppdatert 18.02.25

Om leksikonet: https://www.litteraturogmedieleksikon.no/gallery/om_leksikonet.pdf

Fjernlesing

“In 2000, Franco Moretti coined the term “distant reading” to refer to the process of “understanding literature not by studying particular texts, but by aggregating and analyzing massive amounts of data.” [...] in a world where documents such as novels, speeches, song lyrics, poems, newspaper articles, movie and television scripts, and courtroom proceedings are increasingly available online to the public, distant reading has become popular and possible. Text-mining software can expand our ability to analyze a variety of patterns.” (<https://uta.pressbooks.pub/datanotebook/chapter/1-4-distant-reading/>; lesedato 12.12.24)

“Pioneren Franco Moretti mener at vi går fra “nærlesing” av noen få bøker (litteraturvitere forholder seg stort sett til bare 1 prosent av verdenslitteraturen) til “fjernlesing” av store mengder. Fjernlesing eller datagraving (*data mining*) åpner detaljer ved et korpus som ikke er tilgjengelig for nærlesing – og *vice versa*. Kvantitative og kvalitative metoder forenes på nye måter. [...] Matthew L. Jockers ved universitetet i Nebraska-Lincoln frembringer et statistisk basert bilde av hvordan litterære genrer sprer seg over epoker og regioner.” (*Morgenbladet* 16.–22. mai 2014 s. 26) Jockers har publisert boka *Macroanalysis: Digital Methods and Literary History* (2013).

Fjernlesing er en “quantitative, large-scale, computer-assisted approach to the study of literature. As Ted Underwood stresses in “A Genealogy of Distant Reading,” such a quantitative turn enables distant reading to “aspire to the methods of social science: it is defined not only by a commitment to historical breadth, but by a version of the scientific method appropriate for a historical discipline”” (Antonija Primorac, Rosario Arias m.fl. i <https://www.digitalstudies.org/article/id/8855/>; lesedato 24.09.24).

“Distant reading is the computational processing of textual information in digital form. It relies on automated procedures whose design involves strategic human decisions about what to search for, count, match, analyze, and then represent as outcomes in numeric or visual form. For texts or collections of texts to be read computationally, they have to be divided into meaningful units that an automatic process can recognize. This process is known as tokenization, the decision about

how to break a string (or set) of elements (or phenomena) into meaningful units. What is considered meaningful will vary depending on the model of research. A model is a general scheme, or template, that is independent of the individual cases in the study: research on a set of novels has a different model than work on census data [= demographic, economic and population data from the U.S. Census Bureau], for instance. The models are argument structures that represent components of the information (features of texts, images, numbers, etc.) and their relations to each other. If your project is about tracking spelling changes, then you might need to design a model of analysis that recognizes individual letters and sequences. If your research is on social relations in a narrative, then proper names and other nouns might be what you want to search for in the text. If your work focuses on sentiment analysis, then keywords in context provide a better model of how to sort and count.” (Johanna Drucker i <https://www.jstor.org/stable/pdf/27037376.pdf>; lesedato 18.01.25)

“*Data mining* includes any activity of abstracting information to create or detect patterns. In digital processes, this activity uses algorithms that follow instructions about what to find, match, or count according to the parameters set by the model. The step-by-step instructions are simply ways of making explicit what the computer is to find – just as if a horde of readers were hired to search the entire corpus of Anthony Trollope novels to count the instances of feminine and masculine nouns and pronouns. The results can be listed in a table, spreadsheet, graph, chart, or other mode of display. This data aggregation and display can be embedded in an interactive environment, where filters are applied to search selectively – by date, title, or other relevant information. The difference between a computational version of this process and a human one is largely scale. The mathematical and computational complexity of algorithms might make them hard for an individual to imitate, but, conversely, intuitive decisions that are part of human reading are difficult to specify algorithmically. When these decision-making processes are automated, they can be done at speeds, volumes, and scales (nano-, micro-, macro-) impossible for a human reader to match. Using algorithms to make decisions is, in essence, how the automated text analysis known as distant reading works. Still, no matter how sophisticated the algorithms, they are all based on models designed as interpretative acts.” (Johanna Drucker i <https://www.jstor.org/stable/pdf/27037376.pdf>; lesedato 18.01.25)

Litteratur studeres fra et “fugleperspektiv” (Schweighauser m.fl. 2014), fra et makroperspektiv i motsetning til et mikroperspektiv på enkelttekster. Data-innsamlingen skal skaffe nytt tolkningsgrunnlag, og innebærer altså ingen slutt på hermeneutisk arbeid (Schweighauser m.fl. 2014). Moretti mener at fjernlesing genererer nye spørsmålsstillinger som så kan besvares gjennom ulike metoder.

“[T]echnology might be best suited for helping us read at scale. Critics like Franco Moretti refer to this kind of analysis, when we use technology to get a bird’s eye

view of a corpus, as distant reading. If close reading [...] gives careful attention to every word in a text, distant reading assumes that we can get new insight from thinking more broadly, by using computers to take in more texts than would otherwise be possible. Thus, we might have a computer give us schematic representations of thousands or even hundreds of thousands of texts. [...] On our own, we would never be able to read all 19th century British novels. But computers can help us to at least get some sense of this great body of work. Reading at such a great scale can also offer us a chance to chip away at what Margaret Cohen has called the “great unread”, all that writing that has gone unnoticed because it never became part of the literary canon. [...] The work only begins once you have some results and a graph. You then have to figure out what elements are meaningful and what they might indicate. And the exploration very often takes you back to particular parts of the corpus that you want to read in more detail.” (Alberto Cairo i <https://walshbr.com/textanalysiscoursebook/book/reading-at-scale/distant-reading/>; lesedato 12.12.24)

Moretti bruker “kvantitativ-algoritmiske analyser” av enorme, digitaliserte tekstkorpus i stedet for å nærlese noen få verk i en kanon (Schweighauser m.fl. 2014). Han ønsker å visualisere resultatene av undersøkelsene, slik at mønstre blir tydeligere. Dette gjør han med kart, grafer, stamtrær m.m.

Moretti har formulert sitt program polemisk: “What we really need is a little pact with the devil: we know how to read texts, now let’s learn how not to read them.” (sisert fra Schweighauser m.fl. 2014)

Moretti oppfatter ikke bøker som resultat av individuell skaperkraft, men som resultat av en kollektiv prosess (Schweighauser m.fl. 2014). Han flytter fokus fra individuelle tilfeller til serier, mønstre, statistiske strukturer, sykluser, sjangrer, stamtrær og andre store utviklingslinjer. Mengden materiale omfatter for lengst glemte bøker. De aller fleste bokutgivelser blir med tiden til glemte og tilsynelatende ubetydelige bøker (Schweighauser m.fl. 2014).

“Nå kan “fjernlesing” komme til å utfordre nærlesningen som humanisters foretrukne praksis. Det er utviklingen av stadig mer raffinerte digitale verktøy og metoder, brukt på et raskt voksende digitalisert materiale, som er årsaken. Betegnelsen “fjernlesing” ble introdusert av den amerikanske litteraturviteren Franco Moretti, som polemisk hevder at for å forstå litteratur i sin bredde og natur, må vi slutte å lese bøker. For uansett hvilket svimlende antall enkelttekster en forsker kan håpe på å få lest i jakten på la oss si medisinsk metaforikk i det seine 1800-tallets franske romaner, så vil antallet bøker som det er mulig for et menneske å lese alltid bare utgjøre en ørliten del av det materialet som ble publisert i perioden, og som kunne vært relevant for undersøkelsen. Å øke antall tekster på leselisten hjelper fortvilende lite. Skal en gå vitenskapelig til verks, må selve strategien endres. Avanserte søkemotorer og nøye uttenkte framgangsmåter (som kombinerer

kvalitativ modellering med kvantitative søk) gjør det nemlig mulig å bedrive nye former for “tekstlig gruvedrift” (*text mining*) og/eller hypotese-drevet “tråling” i et stort hav av tekster. Takket være private og offentlige aktørers satsning på digitalisering er dette “havet” ikke begrenset til en velkjent kanon, men omfatter stadig større deler av det som sjelden har vært lest; “the great unread” [...] Et raskt søk i Google N-gram Viewer (basert på Google books’ tekstkorpus) kan for eksempel vise hvordan forekomsten av sammenstillingen mellom “menneske” og “rettigheter” eksploderte på slutten av 1780-tallet, parallelt med at sammenstillingen av “rettigheter” og “privilegier” falt, en solid bekreftelse på menneskerettighetstenkningens gjennombrudd i perioden. [...] er søkenes nytte avhengig av at man på forhånd har stilt gode forskningsspørsmål, slik at resultatene kan settes inn i relevante tolkningsrammer. Det samme gjelder selve visualiseringen av store data. Kart og grafer kan hjelpe til å identifisere idéhistoriske trender, også over lengre tidsspenn.” (Ellen Krefting i *Morgenbladet* 30. april–7. mai 2015 s. 30) “[T]he data pool underlying Google Books is already one of the richest document collections on human history and culture ever compiled. [...] According to its own affirmations, Google had scanned 15 million books out of an estimated 130 million in October 2010” (Kircz og Weel 2013 s. 65 og 69).

“Ifølge Moretti utgjorde tradisjonell litteraturhistoriskrivning usikre vitenskapelige byggverk basert på lite empiri. Sannsynligvis husker og dokumenterer vi bare 1 % av det som faktisk er blitt skrevet, mens 99 % går i glemmeboken (Moretti 2000b). Morettis mål var å skrive en verdenslitteraturhistorie som representerte en større del av det skrevne materialet, og med slike ambisjoner kunne ikke forskeren forventes å ha førstehåndskjennskap til alle tekster. Tvert imot måtte man konstruere “a patchwork of other people’s research, without a single direct textual reading” (Moretti 2000a, 57). Morettis begrep om fjernlesning (“distant reading”) innebar således en lesning av nasjonallitteraturer som baserte seg på annenhåndskunnskap, samtidig som det innebar en oppskalering av antallet tekster og forfatterskap som skulle representeres. Ideen om fjernlesning oppstod altså før massedigitaliseringen av arkivene og hadde i utgangspunktet lite med digitalisering å gjøre. Det var imidlertid først da digitaliseringen skjøt fart at ideen om en “fjernlesning” virkelig fikk gjennomslag, men nå som betegnelse for det å utforske større tekstmengder i digitale arkiver.” (Frode Boasson og Anders Skare Malvik i <https://www.idunn.no/doi/10.18261/issn.1504-288X-2019-02-03>; lesedato 22.01.25)

Moretti etablerte i 2010, sammen med Matthew L. Jockers, Stanford Literary Lab. “Jockers har videreført Morettis kritikk av den tradisjonelle litteraturvitenskapen, og i boken *Macronalysis: Digital Methods and Literary History* (2013) mobiliserer han et solid angrep på tradisjonell nærlesning som metode. Her sammenligner Jockers studiet av litteratur med studiet av økonomi og viser til hvordan differensieringen mellom et makro- og et mikronivå endret økonomifaget etter J. Maynard Keynes’ *The General Theory of Employment, Interest and Money* (1936). Jockers benytter analogien til å peke på hvordan digitaliseringen åpner et makronivå i litteraturstudiet

der man kan studere litteraturens overliggende strukturer og utvikling ut fra et større materiale enn tidligere. Ifølge Jockers er nærlesning som metode uegnet for å analysere dette nye makronivået. For å gjøre det, må vi omfavne digitale metoder for innhenting og databehandling: “Like it or not, today’s literary-historical scholar can no longer risk being just a close reader: the sheer quantity of available data makes the traditional practice of close reading untenable as an exhaustive or definitive method of evidence gathering. Something important will inevitably be missed.” (Jockers 2013, 9) Datamaskinen *fortolker* ikke, den prosesserer, og for å understreke denne forskjellen investerer Jockers i begrepet *makroanalyse* fremfor Morettis mer fortolkende *fjernlesning* (Jockers 2013, 25). For Jockers representerer nemlig “analyse” også et steg mot høyere grad av objektivitet. Dette fordi nærlesning, ifølge ham, innebærer anekdotisk og spekulativ fortolkning, mens makroanalysen frembringer “comprehensive and definitive” faktakunnskap (Jockers 2013, 31).” (Frode Boasson og Anders Skare Malvik i <https://www.idunn.no/doi/10.18261/issn.1504-288X-2019-02-03>; lesedato 22.01.25)

“[D]istant reading is conducted at a large scale and results are also often wide-ranging, telling stories about the progress of literature over the course of a century or longer. Some topics that have been addressed include: how the language of novels becomes less abstract as the 19th century progresses; why the titles of novels are shorter in the 19th century than in the 18th and who (and what) is behind the Elena Ferrante pseudonym. Nor is the method restricted to literary texts: Moretti and Dominique Pestre have examined why the language of World Bank Reports becomes more abstract towards the turn of the millennium.” (Justin Tonra i <https://www.rte.ie/brainstorm/2019/1114/1090846-what-is-distant-reading/>; lesedato 12.12.24)

“[D]istant readers have grappled with social questions about money, gender, race, geography, and literary circulation, as well as formal questions about genre, plot, emotion, and time [...] Many important patterns in literary history are still poorly understood, because they weren’t easily grasped at the scale of individual reading. [...] large-scale literary history is now suffused with ideas drawn from corpus linguistics, information retrieval, and machine learning.” (Ted Underwood i <https://www.digitalhumanities.org/dhq/vol/11/2/000317/000317.html>; lesedato 12.12.24)

Ved fjernlesing kan det estetiske ved litterære tekster lett gli ut av synsfeltet (Schweighauser 2014). Det er en fare for at det ikke-reduserbare ved litteratur forsvinner, til fordel for begrepsinnhold. Stilistisk flertydighet blir kun et problem, ikke en god kvalitet. Tekstenes historiske kontekst, hvordan de henger sammen med konkrete samfunnsforhold, er heller ikke vektlagt i Morettis metoder.

“Både Jockers og Moretti argumenterer tungt for digitale metodens høye grad av objektivitet og vitenskapelighet og deler en langt på vei positivistisk tiltro til datamaskinens analytiske overlegenhet med flere andre i DH-feltet [dvs. digital

humaniora]. Ikke sjeldent kombineres denne tiltroen også med en sterk entusiasme over at digitale metoder visker ut de tradisjonelle skillene mellom natur- og åndsvitenskapene (Ganascia 2015). Mens DH-feltets forskere på den ene siden dermed har hatt en tendens til å redusere betydningen av og gjennomslagskraften til den tradisjonelt nærlesende litteraturvitenskapen, kan de på den andre siden selv anklages for å ha en vel positivistisk holdning til humanistisk forskning. Det er imidlertid nødvendig å overvinne denne (ofte svært polemiske) dikotomiseringen om vi skal nå frem til en mer *kritisk* digital litteraturvitenskap. Det motsatte av nærlesning er, som Michael Hancher har minnet om, ikke fjernlesning eller makroanalyse, men en “loose, casual, and careless reading” (Hancher 2016, 125). Formulering og testing av hypoteser går alltid begge veier og de to tilnærmingene bør derfor ikke ses i et motsetningsforhold (Olsen 2014).” (Frode Boasson og Anders Skare Malvik i <https://www.idunn.no/doi/10.18261/issn.1504-288X-2019-02-03>; lesedato 22.01.25)

“Maskinell fjernlesing brukes gjerne for hurtig å gå gjennom store mengder materiale for å finne det som senere kan studeres nærmest under lupen gjennom mer nærsynt tradisjonell lesing. Her har den rivende utviklingen av det som kalles OCR-teknologi vært spesielt interessant. OCR står for Optical Character Recognition, eller automatisk tegngjenkjenning på godt norsk. Dette er et redskap for søk i store mengder data og materiale, som for eksempel gamle aviser.” (Gunnar Iversen i <https://rushprint.no/2022/11/sensasjonelt-nytt-om-fiskerlivets-farer/>; lesedato 22.01.25)

Undersøkelsene kan gjelde korte eller lange historiske perioder, og noen ganger på tvers av nasjonale grenser. Når empirien omfatter tekster på ulike språk (og når det dessuten gjennom tidene er ulike skrivemåter for et bestemt ord i et språk), oppstår det problemer med å lage statistikk over hyppigheten av bestemte ord og lignende.

“Distant reading today is able to investigate larger and larger collections of literary texts that are becoming available, and to use appropriate software for text processing, machine learning, visualization, and statistics (e.g., Python, R, Java with dedicated packages or libraries such as scikit-learn, stylo, or MALLETT). Using these datasets and tools, distant reading of style today focuses on the occurrence of textual patterns defined at various levels, from characters and single tokens to n-grams [dvs. sekvenser av ord] and larger, more complex, and more abstract linguistic structures or features. It often focuses on the various ways that these patterns interact with contextual factors, such as authorship, author gender, period of publication, or literary genre at the document level (and as available to the analysis via document-level metadata) as well as dialogue and narrative or stage direction vs. character speech at the textual level (available to the analysis via inner-textual annotations or pattern finding methodology). It also frequently focuses on the relationship of such patterns with a range of potential functions of literature that can be aesthetic, communicative, social, as well as cognitive. As these multiple avenues of research

indicate, analysis of style in the distant reading paradigm has become a wide-ranging, multi-faceted research activity.” (Antonija Primorac, Rosario Arias m.fl. i <https://www.digitalstudies.org/article/id/8855/>; lesedato 24.09.24)

“Stylometry is the quantitative study of literary style through computational distant reading methods. It is based on the observation that authors tend to write in relatively consistent, recognizable and unique ways. [...] Scholars have used stylometry as a tool to study a variety of cultural questions. For example, a considerable amount of research has studied the differences between the ways in which men and women write or are written about.” (François D. Laramée i <https://programminghistorian.org/en/lessons/introduction-to-stylometry-with-python>; lesedato 18.01.25)

“Kollokasjonsanalyse er ett eksempel på en digital forskningsmetode som gir innblikk i hvilke ord som statistisk sett er nært assosiert med et bestemt ord i et større digitalt korpus. For å identifisere hvilke ord som statistisk sett er sterkt assosiert, eller *kollokert*, med et bestemt ord, er vi ute etter relativfrekvensen. Da må ordfrekvensen i kollokasjonene divideres på ordfrekvensen i et egnet referansekorpus. Nasjonalbiblioteket har tilrettelagt for at vi kan utføre kollokasjonsanalyser av det digitaliserte materialet som er tilgjengeliggjort i deres DH-Lab via den åpne kildekodeapplikasjonen Jupyter Notebook. [...] Temamodellering, eller emnemodellering (“topic modeling”), er en metode for automatisk generering av “temaer”. “Temaer” er i denne sammenheng mengder av ord – ordkløstre – med høy samfrekvens i et korpus. Teknisk sett er et “tema” en sannsynlighetsdistribusjon over alle ordene i korpuset. Det finnes en rekke forskjellige algoritmer som kan benyttes, men som gjør mer eller mindre det samme. Basert på algoritmens statistiske utregning angis en liste med ord som ofte opptrer sammen i korpuset man analyserer. [...] Felles for disse metodene er at de på effektivt vis genererer unike, formelle objekter som kan lede til kunnskap om for eksempel tematiske eller stilistiske mønstre i store tekstmengder.” (Heidi Karlsen i https://boygen.net/wp-content/uploads/2021/09/boygen_02_2021_v05.pdf; lesedato 22.01.25)

“En amerikansk studie som ofte trekkes frem i diskusjoner om mangfold i barne- og ungdomslitteratur viser at 50 % av hovedpersonene er hvite, 10 % afrikanske eller afro-amerikanere, 7 % asiater, 5 % latin-amerikanere, 1 % urbefolkning/indianere og 27 % dyr. En tilsvarende undersøkelse er det ikke så lett å gjøre på norsk fordi det ikke er tradisjon for å angi etnisitet i litteraturen. Etnisitet kan, men behøver ikke fremkomme av illustrasjoner og navn på personer og steder. I Norge angis ikke etnisitet på virkelige personer heller. Statistisk sentralbyrå opererer for eksempel ikke med tall på antall samer. [...] I stedet angis befolkningstallet for det som defineres som samiske, geografiske områder. På samme måte fremkommer det av befolkningsstatistikken hvilke land innvandrere kommer fra. [...] Der Kulturrådet opererer med kategorien ‘Synlige minoriteter og grupper som representerer flerkulturelt og etnisk mangfold’, undersøkte vi i stedet representasjonene av land

og språk. [...] For å kunne gjennomføre opptellingen har vi måtte gjennomgå så store mengder tekst at det ikke lar seg gjøre av oss som enkeltpersoner, derfor har vi benyttet maskinell lesing av et definert materiale, såkalt fjernlesing (Moretti, 2005, Giddens, 2017). Utgangspunktet for kartleggingen skulle være Nasjonalbibliotekets digitaliserte samling, som er basert på pliktavlevering av alle trykte norske utgivelser, og i senere år også digitaliserte versjoner av disse. Nasjonalbibliotekets digitaliserte korpus er registrert med metadata og søkbart helt ned på ord-nivå. “Det gjør det mulig å hente ut trekk ved tekstene som frekvenslister, kollokasjoner og konkordanser for ord” (Johnsen, 2019, s. 193). Frekvenslistene gir informasjon om hvor mange ganger et bestemt ord opptrer innenfor én og samme tekst eller et definert sett av tekster. Vi kan, for eksempel, få ut tall på hvor mange ganger ordet muslim opptrer, eller ordet hørselshemmet, eller ordet transseksuell. Kollokasjoner viser hvilke ord som opptrer sammen, som en enhet. Det gir informasjon om hvilke andre ord som markerer den identiteten vi er ute etter å finne. Konkordans er en liste som viser hvilke setninger markøren opptrer i. Det gir oss anledning til å nærlese på setningsnivå for å kvalitetssikre at markøren er benyttet slik vi forutsetter, eller få innsikt om andre betydninger og mulige feilslutninger.” (Kristin Ørjasæter og Lars Johnsen i <https://barnebokinstituttet.no/mangfold/kan-vi-male-graden-av-litteraert-mangfold/>; lesedato 22.01.25)

Ørjasæter og Johnsen tok “utgangspunkt i kategoriene land, språk, religion, seksualitet og kjønnsidentitet, funksjonsnedsettelse og døve. Deretter har vi definert de ulike identitetene innenfor hver kategori. Vi har også lagt til noen identiteter som angår forflytning mellom land (som emigrasjon, flyktning, innvandring etc.). Så har vi benyttet maskinell kollokasjon for å avgjøre hvilke markører vi skal søke på. Og så har vi benyttet konkordans som kvalitetssikring. Fjernlesing gjør det mulig å basere slutningene på et stort materiale, men innsikten det gir er overflatisk, poengterer Alsvik og Glad (2020). De minner om at tradisjonell humaniora har vært basert på nærlesning av et svært lite empirisk materiale, som man har tillatt seg å trekke omfattende slutninger fra. Derfor kombinerer de kvantitativ fjernlesing, som egentlig handler om å telle opp antall forekomster, med kvalitativ fortolkning. [...] Vi har måttet definere visse identiteter innenfor hver kategori, og finne relevante markør-ord som peker på identitetene. Kategorien religion har identitetene kristendom, islam, hinduisme, etc. Kristendom vil for eksempel kunne markeres gjennom søkeordet kirke, som da vil kunne bli talt opp fra en setning der hovedpersonen går i kirken. Men selv om vi utarbeider en liste på ytterst relevante markører vil ikke alle eksempler på fenomenet nødvendigvis dukke opp. I litteratur blir en muslim ikke nødvendigvis nevnt hverken ved religionens navn eller ved tilhørigheten til en moské. Kanskje fremgår det bare av at hun ikke kjøper pølser på bensinstasjonen når de andre gjør det. Men det å ikke kjøpe pølse er et usikkert tegn. Kanskje hun er mett, eller vegetar? Derfor er dybdelesing og mer tradisjonell fortolkning av utvalgte eksempler nødvendig, i tillegg til fjernlesing og telling. Da barne- og ungdomslitteratur dessuten utmerker seg ved visuell narrasjon i tillegg til verbal, vil vår ord-baserte undesøkelse uansett være ufullkommen.” (Kristin

Ørjasæter og Lars Johnsen i <https://barnebokinstituttet.no/mangfold/kan-vi-male-graden-av-litteraert-mangfold/>; lesedato 22.01.25)

“Text analysis and distant reading embody the peculiar aspiration of the humanities to achieve the conditions of authority perceived to inhere in the natural sciences and their empirical methods. But human perception is characterized by infinite variation and a high degree of specificity, alongside a capacity for ambiguity, ambivalence, contradiction, and association. Reading is an act of self-production, subject enunciation, and, as such, is an emergent and shifting process in human cognition.” (Johanna Drucker i <https://www.jstor.org/stable/pdf/27037376.pdf>; lesedato 18.01.25)

“Many distant reading projects have produced disappointing results because they have been more interested in validating their tools – showing that their computational methods are able to confirm existing stereotypes – than in pursuing genuine discoveries. Many others, meanwhile, produce provocative results that cannot be meaningfully validated” (Adam Hammond sitert fra https://www.researchgate.net/publication/327954876_A_New_Research_Programme_for_Reading_Research_Analysing_Comments_in_the_Margins_on_Wattpad; lesedato 19.01.23).

Litteraturliste (for hele leksikonet): <https://www.litteraturogmedieleksikon.no/gallery/litteraturliste.pdf>

Alle artiklene i leksikonet er tilgjengelig på <https://www.litteraturogmedieleksikon.no>