

# Massedigitalisering og edisjonsfilologi

Nordisk nettverk  
for edisjonsfilologer  
Skrifter 15

Nordisk nettverk for edisjonsfilologer. Skrifter 15

# Massedigitalisering og edisjonsfilologi

Bidrag til en konferanse arrangert av  
Nordisk nettverk for edisjonsfilologer

Voksenåsen, Oslo  
2. til 4. september 2022

Utgitt ved  
Annika Rockenberger  
Aasta Marie Bjorvand Bjørkøy  
Nina Marie Evensen  
og Ellen Nessheim Wiger

Universitetsbiblioteket i Oslo  
Oslo 2025

ISBN 978-82-8037-203-1  
ISSN 1601-1562

# Innhold

<b>Forord</b>	<b>2</b>
Massedigitalisering og edisjonsfilologi ( <i>Annika Rockenberger, Aasta Marie Bjorvand Bjørkøy, Nina Marie Evensen, Ellen Nessheim Wiger</i> ) . . . . .	3
Bidragene . . . . .	4
Presentasjon av konferansebidrag som ikke er blitt omgjort til artikkel	5
Takk . . . . .	10
<b>Essay</b>	<b>13</b>
Nya vägar till det förflutna: Två editionsfilologiska frågor till den storskaliga digitaliseringen ( <i>Mats Malm, Ljubica Miočević</i> ) . . . . .	14
<b>Artikler</b>	<b>21</b>
Bibliography, database, and beyond: Towards queer scholarly editing ( <i>Jenny Bergenmar</i> ) . . . . .	22
«Op ad Midten, krydser og falder af» Dansemanuskriptet <i>XII Engels Danse</i> som kildeutgivelse ( <i>Elizabeth Svarstad</i> ) . . . . .	39
Digitale udgaver som HITL-garanter. Om editionsfaglig domæneeksportberigelse og datakvalitetens betydning for integrationen af computationelle tilgange i humanvidenskaben ( <i>Katrine Frøkjær Baunvig, Krista Stinne Greve     Rasmussen</i> ) . . . . .	50
<b>Biografier</b>	<b>67</b>
Redaktører . . . . .	68
Forfattere . . . . .	69
<b>Personregister</b>	<b>72</b>



# Forord

## Massedigitalisering og edisjonsfilologi

Annika Rockenberger<sup>1</sup>, Aasta Marie Bjorvand Bjørkøy<sup>2</sup>, Nina Marie Evensen<sup>3</sup>,  
Ellen Nessheim Wiger<sup>4</sup>

<sup>1</sup> Universitetsbiblioteket, Universitetet i Oslo

<sup>2</sup> Institutt for lingvistiske og nordiske studier, Universitetet i Oslo

<sup>3</sup> Senter for Ibsenstudier, Universitetet i Oslo

<sup>4</sup> Nasjonalbiblioteket

Nordisk nettverk for edisjonsfilologer (NNE) ble startet i 1995. Formålet med nettverket er todelt: på den ene siden å skape rom for utveksling av faglige erfaringer, på den andre siden å bidra til en styrking av edisjonsfilologi som fagdisiplin i Norden. NNE har i årene 1996–2024 arrangert 18 konferanser:

- 1996 – Edb-teknikkens muligheter for editionsfilologien; København
- 1997 – Brev- og dagbogsudgaver; Oslo
- 1998 – Valget af grundtekst; Stockholm
- 1999 – Realkommentering; København
- 2000 – Forholdet mellem bogudgaver og elektroniske udgaver; Oslo
- 2001 – Text och tradition – om textedering och kanonbildning; Stockholm
- 2002 – Varia – om variantapparat och bibliografisk beskrivning; Helsingfors
- 2003 – Udgavetyper og målgrupper; Sønderborg
- 2005 – Hermeneutikk og filologi; Oslo
- 2007 – Bokens materialitet; Uppsala
- 2009 – Tryckta och elektroniska utgåvor – erfarenheter, planering och teknik i förändring; Helsingfors
- 2011 – Editionshistorie og editionsfilologiens faglige status; Sønderborg
- 2013 – Sensur og edisjonsfilologi; Oslo
- 2015 – Textkritik som analysmetod; Göteborg
- 2017 – Utgåvor i användning; Helsingfors
- 2019 – Kildeudgivelse og tekstkritik i forskellige traditioner; København
- 2022 – Massedigitalisering og edisjonsfilologi; Oslo
- 2024 – Recirkulering: Textkritik och (åter)användbarhet; Sigtuna

Tidligere konferansebidrag finnes publisert i serien *Skrifter*, som er tilgjengelig digitalt på nettverkets nettsider, [www.nnedit.org](http://www.nnedit.org). De fleste konferansebindene er også tilgjengelig i trykt format.

Konferansen *Massedigitalisering og edisjonsfilologi* var egentlig planlagt til september 2021, men måtte på grunn av covid-pandemien utsettes to ganger, først til februar 2022 og deretter til september 2022. Men 2.–4. september 2022 kunne endelig 59 deltakere fra Norge, Sverige, Danmark, Finland, Tyskland og Italia møtes på Voksenåsen hotell i Oslo.

Konferansen behandlet temaer knyttet til spørsmålet om – og i så fall hvordan – massedigitaliseringen endrer hvordan vi arbeider med tekstutgivelse og tekstkorpus. Stadig mer materiale blir digitalisert, og i Norden har



massedigitalisering bidratt til at store mengder tekst blir gratis tilgjengelig for alle. Konferansedeltakerne diskuterte hvilke muligheter og hvilke problemer dette åpner for, deriblant: Hvilke problemstillinger oppstår når vi digitaliserer eldre verk? Hvilke konsekvenser får massedigitaliseringen for etterlivet til edisjonsfilologiske utgaver? Representerer massedigitaliseringen en ny filologisk tradisjon? Massedigitaliseringen bidrar dessuten til økt demokratisering, siden innholdet blir tilgjengelig for alle. Men hvordan forholder dette seg når tilgjengeligheten er begrenset for verk som er vernet av opphavsretten?

Disse temaene åpnet for en rekke interessante bidrag fra deltakerne på konferansen, både i form av foredrag og som del av diskusjonene i etterkant. Den foreliggende publikasjonen gjengir noen av disse bidragene i artikkelform. Noen av bidragsyterne har dessverre ikke hatt mulighet til å omgjøre bidragene sine til artikler, og disse presenteres kort i slutten av forordet.

Alle bidrag i dette konferansebindet har blitt fagfellevurdert av to fagfeller i enkeltblind vurdering. I tillegg har alle bidrag blitt vurdert av redaktørene før publikasjon.

## **Bidragene**

*Mats Malm og Ljubica Miočević*

Konferansepublikasjonen innledes av essayet «Nya vägar till det förflutna: två editionsfilologiska frågor till den storskaliga digitaliseringen» av Mats Malm og Ljubica Miočević (Litteraturbanken, Göteborgs universitet). I forbindelse med at den svenske digitale tekstbasen vokser og digitalisering generelt øker, vurderer Litteraturbanken behovet for tekstkritiske perspektiver på to områder. Den første knytter seg til vitenskapelig publiserings kontekstualisering og kommentarer, og angår de mange forestillingene om kjønn, etnisitet, det nasjonale og det fremmede som er nedfelt i vår kulturarv. Det andre gjelder framtiden til de store tekstkritiske forfatterprosjektene. Det blir stadig vanskeligere å finne midler til slike prosjekter, og spørsmålet er hvordan vi kan bevare det edisjonsfilologiske perspektivet når digitaliseringen blir mer storstilt.

*Jenny Bergenmar*

I artikkelen «Bibliography, database and beyond: Towards queer scholarly editing» presenterer Jenny Bergenmar (Göteborgs universitet) prosjektet Queerlit og tar opp viktige spørsmål rundt LHBTQ-litteratur, bibliografisk metadata og edisjonsfilologi. Hun peker på hvordan den tradisjonelle kanon innebærer en kanonisering av et fåtall forfattere basert på et selektert utvalg, der smalere forfatterskap faller utenfor. Innenfor dette perspektivet representerer edisjonsfilologien en kanonopprettelse, der særlig litteratur skrevet av kvinner er marginalisert. Massedigitalisering kan ses som en løsning på dette problemet, ved at materiale ikke velges ut og dermed heller ikke bort. Likevel utgjør datasett og digitalt materiale alltid et utvalg, en kuratering, bearbeidet av et subjekt som kurerer og velger ut. Massedigitaliseringen er dermed heller ikke nødvendigvis objektiv, men et utvalg basert på samlingen som er valgt. Siden heller ikke

bibliotekataloger er nøytrale, men reflekterer dominerende verdsett, har det vært viktig for skeivt litteraturarbeid å lage nye metadatasett og kategorier, *Information activism*, der valg av nye emneord gjør det mulig å speile virkeligheten uten fordommer.

*Elizabeth Svarstad*

I artikkelen «‘Op ad Midten, krydser og falder af’. Dansemanuskriptet *XII Engels Danse* som kildeutgivelse» viser Elizabeth Svarstad hvordan man ved hjelp av digitalisering og kildeutgivelser av dansenotasjoner kan gi ny innsikt i tidligere tiders omgangsformer. Danseforskning og informasjon knyttet til en dansebok kan fortelle mye utover den konkrete informasjonen dokumentet inneholder, og studier av fortidig dans kan si noe om tidens estetikk, måte å oppføre seg på og omgang med mennesker.

*Katrine Frøkjær Baunvig og Krista Stine Greve Rasmussen*

Katrine Frøkjær Baunvig og Krista Stine Greve Rasmussen (Aarhus Universitet) diskuterer i sin artikkel «Digitale udgaver som HITL-garanter. Om editionsfaglig domæneeksportberigelse og datakvalitetens betydning for integrationen af computationelle tilgange i humanvidenskaben» i hvor stor grad den automatiserte massedigitaliseringen krever en bevisst bruker og menneskelig gjennomføring. De gir eksempler på og reflekterer rundt konkrete utfordringer de møter som del av edisjonsarbeidet, og som ingen kunstig intelligens eller forbedret algoritme ser ut til å kunne løse. Deres konklusjon er at edisjonsfilologien fortsatt spiller en avgjørende rolle i humaniora generelt og digital humaniora spesielt, og at den edisjonsfilologiske kompetanse er særlig verdifull nå som alle tekster er blitt digitale.

## **Presentasjon av konferansebidrag som ikke er blitt omgjort til artikkel**

*Gioele Barabucci*

I åpningsforedraget «Let’s agree to disagree: Challenges and recent developments in recording multiple points of view in Humanities data» reflekterte Gioele Barabucci (da NTNU - Norges teknisk-naturvitenskapelige universitet, nå Universität zu Köln) rundt metadata og utfordringene ved å registrere humanistiske data i dataformater som er lite fleksible. Mye av det humanistiske forskningsarbeidet består av forarbeid og refleksjoner rundt metadataregistreringen, men disse refleksjonene går ofte tapt når dataene publiseres fordi dataformatene ikke åpner for fleksibilitet og ambivalens. Det kuraterte resultatet skjuler dermed ofte ambivalensen som ligger i materialet. De sju syndene ved dataregistrering identifiserte han som:

1. utelatelse: at kjent informasjon ikke registreres fordi den ikke passer inn i formatene som er valgt, eller er problematisk av andre grunner,
2. utflating: at man taper informasjon ved at derivasjonskjeden samles i ett trinn (middels spesifikk informasjon),

3. tvang: at man tvinger informasjon inn i et felt som ikke passer helt,
4. dumping: at man legger inn strukturert informasjon i fritekstfelt,
5. overse variasjon: at man forenkler komplisert informasjon for at det skal passe formatet,
6. underforståtte fakta: at man blander fakta og personlige oppfatninger,
7. antatt sikkerhet: at man ikke skiller mellom ulike grader av sikkerhet (bekreftet, mulig, spekulasjon).

Kollisjonen mellom behovet for konsistente metadata og den ambivalensen og usikkerheten som ofte finnes i dataene, bør i større grad kunne gjenspeiles i registreringen. Barabuccis forslag var å utvide RDF-modellen (Resource Description Framework) til å kunne uttrykke noe uten å slå det fast som fakta, og kunne gjengi ulike standpunkter ved uenighet eller usikkerhet. Dette ville innebære å åpne for ambivalens og usikkerhet ved å gjøre det mulig å uttrykke tvil og refleksjoner i et strukturert format.

### *Annika Rockenberger*

I sitt foredrag «Hva snakker vi egentlig om når vi snakker om edisjon (i massedigitaliseringens tidsalder)?» reflekterte Annika Rockenberger (Universitetsbiblioteket i Oslo) rundt hvordan massedigitaliseringen av dokumentbasert kulturarv utfordrer den tradisjonelle edisjonsfilologien. Mye av det edisjonsfilologiske arbeidet – faksimilering, transkripsjon og transgrafering av historiske kilder, sammenstilling av alle verk samt alle utgaver i et forfatterskap (œuvre) og ikke minst tilgjengeliggjøring til et bredt publikum – gjøres i dag ved hjelp av automatiserte prosesser i regi av de store kulturarvsinstitusjonenes digitaliseringssatsninger. Historiske dokumenter blir dermed tilgjengelige på en måte og i et omfang som var utenkelig for kort tid siden. Samtidig har det blitt mulig å bruke tekstanalysemetoder (n-grams, konkordanser, ord- og begrepsnettverk, topic modeling, named-entity recognition, osv.) som egner seg for evidensbasert forskning uten at det er nødvendig å manuelt annotere eller ekstrahere data fra tekstene selv. Kanskje vil noen si at det ikke lenger er nødvendig – og ikke lenger verdt – å satse på noen få, men store, historisk-kritiske utgivelser av et fåtall forfattere, slik at edisjonsfilologi som yrke egentlig er blitt overflødig?

Med utgangspunkt i Nasjonalbibliotekets og Arkivverkets massedigitalisering av dokumentbasert kulturarv diskuterte Rockenberger hvilke deler av det edisjonsfilologiske arbeidet som allerede kan gjøres fullautomatisk av (ikke-filologiske) algoritmer, hvilke deler av dette arbeidet som kommer til å automatiseres i nær framtid, og hva som forblir menneskearbeidet som ikke kan integreres i slike prosesser. Verktøyene, særlig OCR (Optical Character Recognition) og HTR (Handwritten Text Recognition), blir stadig bedre, og løypene er laget for å håndtere store mengder materiale. Men her er det et skille

mellom målrettet digitalisering ved forskningsbehov og massedigitalisering av den kulturelle dokumentarven, der feil ikke anses som forstyrrende. Ut fra dette formulerte Rockenberger to forutsigelser: 1) at edisjonsfilologien vil dø ut, og 2) at edisjonsfilologien vil gjenoppstå i andre former.

Den tradisjonelle edisjonsfilologien, som har resultert i flere store forfatterbaserte, historisk-kritiske utgaveprosjekter, er utdatert. Prosjektene blir færre, og det blir stadig vanskeligere å finansiere nye prosjekter av samme størrelsesorden. Men samtidig som massedigitaliseringen er vår største konkurrent, åpner den for nye arbeidsområder for framtidens edisjonsfilologer. I møte med digitalisert materiale er det behov for kunnskapsbasert redigering og forbedring av tekstkorpus, der teknologien tas til hjelp i stedet for å ses som en trussel. Edisjonsfilologien kan blant annet bidra til utvikling av gode verktøy og algoritmer for å forbedre de eksisterende rutinene for massedigitalisering, og for å overkomme hindringene som oppstår ved maskinelt skapte tekstkorpus med høy andel OCR-feil. På denne måten kan edisjonsfilologien forbli relevant. Men for å oppnå dette, må edisjonsfilologisk arbeid bli mer synlig for ledelsen av kulturinstitusjoner og universiteter og høyskoler. Det krever at vi ikke er digitalignorante og hindres av teknofobi, men derimot inntar en aktiv rolle i møte med massedigitaliseringens utfordringer og muligheter.

#### *Andre Kåsen og Mette Witting*

Hovedtemaet for foredraget «Erfaringer ved bruk av Transkribus på Collett-brev» var å vise hvordan maskinell håndskriftgjenkjenning brukes både som støttetekster i det digitale nettbiblioteket ved Nasjonalbiblioteket (NB) og som grunnlag for vitenskapelige kildeutgaver. André Kåsen og Mette Witting (Nasjonalbiblioteket, Oslo) presenterte et nylig gjennomført prosjekt der webapplikasjonen for håndskriftgjenkjenning, Transkribus, ble brukt for å skape maskinlesbar tekst av brev fra Camilla Collett. Brevutgaven blir publisert (på [www.bokselskap.no](http://www.bokselskap.no)) som del av den digitale kildekriftserien *NB kilder*, der tekstene er tilgjengelige som kvalitetssikrede transkripsjoner med kommentarer, kodet i XML (eXtensible Markup Language) etter standarden TEI (Text Encoding Initiative).

Bakgrunnen for prosjektet var at Nasjonalbiblioteket i 2019 etablerte et laboratorium for digital humaniora (DH-lab) for å styrke arbeidet med digitale samlinger og metadata, og for å ta i bruk flere digitale verktøy innenfor dette feltet. De gjenopptok da forsøk med Transkribus, som var blitt testet på Colletts håndskrift i 2016/17 med for dårlig resultat, og gjennom å kombinere flere modeller for håndskriftgjenkjenning klarte de å forbedre gjenkjenningen og få et brukbart transkripsjonsresultat. Transkripsjoner av brev som allerede var utgitt, ble også knyttet til faksimiler av brevene ved såkalt Text2Image i Transkribus som grunnlag for videre forbedringer.

Collett-prosjektets fremgangsmåte var å beholde transkripsjonen inne i Transkribus til studenter hadde lest grovkorrektur og rettet åpenbare feil, for deretter å hente ut tekstene, slik at erfarne filologer kunne lese detaljkorrektur og rette dem i tråd med utgaveprosjektets retningslinjer. En fordel var at mye tid ble spart ved å slippe inntasting. Men en ulempe var at man dermed mistet den forbedrede forståelsen av håndskriften som skjer ved manuell transkripsjon. Det

skjer en opptrening av maskinen, ikke av mennesket. Det er et paradoks at de metodene man nå i stadig større grad tar i bruk, kan resultere i at det blir vanskeligere å finne gode korrekturlesere i tiden framover.

I tillegg til de kvalitetssikrede utgavene har Nasjonalbiblioteket tilgjengeliggjort mer enn 30 000 brev og manuskripter som del av NBs digitale korpus. Disse er imidlertid ikke korrekturlest, men den HTR-etablerte teksten er søkbar og koblet til faksimilene ved hjelp av koordinater, slik at søkeordet blir optisk framhevet. Teknologien på dette feltet er i rivende utvikling, og feilprosenten nærmer seg nå 5%. Datasettet *NorHand* finnes åpent tilgjengelig i det åpne datarepositoriet Zenodo.

Kåsen presenterte også muligheter som ligger i Transkribus for navnegjenkjenning og tagging, der man kan lage grafer over brev, nevnte personer, avsendere og mottagere, og lenke navnene videre til databaser som Wikidata.

### *Mette Gismerøy Ekker*

I sitt foredrag «Kvemens manuskript til Haugianerbrev som prøvekanin for digital skriftgjenkjenning og publisering» presenterte Mette Ekker (Riksarkivet, Oslo) Arkivverkets arbeid med kildeutgivelser eksemplifisert ved et upublisert manuskript som finnes i arkivets samling. Utgangspunktet var at kildeutgivelse er svært ressurskrevende, og store arbeid kan iblant bli liggende uferdige. I hvilken grad kan kunstig intelligens og tekniske løsninger bidra til å ferdigstille slike arbeid og gjøre dem tilgjengelig for bruk?

I Riksarkivet ligger Ingolf Kvemens uferdige manuskript til en utgave i fire bind med Haugianerbrev, dvs. brev mellom haugianere, en norsk, kristen lekmannsbevegelse som ble grunnlagt i 1796–1804 av predikanten Hans Nielsen Hauge. Materialet består av fire esker med maskinskrevne transkripsjoner på papir, kronologisk ordnet i fire perioder fra 1760 til 1842, og utgjør i alt rundt 760 brev og 1300 sider. Utgaven er uferdig, Kvemen rakk ikke å utarbeide innledning, registre eller opplysning om proveniens. Riksarkivet har nå undersøkt gevinsten av dagens tekniske hjelpemidler og nye publiseringsmetoder, i forbindelse med Arkivverkets satsing på å utforske kunstig intelligens og spesielt maskinlæring med henblikk på kildene i arkivet.

De ulike tilnærmingene som ble prøvd ut på dette materialet, inkluderte skanning med publisering i Riksarkivets publiseringsløsning Digitalarkivet, skriftgjenkjenning i det skannede materialet gjennom Transkribus og andre verktøy, og publisering av haugianerbrevene i publiseringsløsningen Generic Digital Edition Platform (GDEP).

Resultatene av disse metodene ga samlet et godt resultat, men de hadde også noen begrensninger. Transkribus ble kjørt i tre runder som til slutt ga en CER (Character Error Rate) på 0,08%, og kombinert med de gode redigeringsmulighetene i Transkribus ble feilprosenten til slutt lik null. Som digital publiseringsløsning testet man først den finske plattformen GDEP, utarbeidet av Svenska Litteratursällskapet i Finland (SLS). Den er basert på en enkel TEI XML-struktur og ga et godt resultat, men løsningen er ikke fullført og det mangler driftsmidler, så man valgte å ikke basere seg på denne alene. I stedet ble Riksarkivets løsning Digitalarkivet valgt, som gir en bedre inngang til materialet og inkluderer metadata (personregister, roller) og globalt søk, og som er mer langsiktig. Riksarkivet har for tiden flere HTR-prosjekter på gang og samarbeider med institusjoner i hele Norden.

Både avslutningsvis i foredraget og i diskusjonen etter bidraget ble det pekt på det paradoksale ved at så mange prosjekter har egne visninger av digitalt materiale til tross for at mange bruker felles verktøy og har overlappende behov. Det savnes en felles løsning for visninger, og en levedyktig driftsløsning for digitale prosjekter.

*Alessandra Molinari*

Alessandra Molinari (Università degli Studi di Urbino Carlo Bo, Urbino) presenterte i sitt foredrag «Editing medieval texts between ‘subjective’ interpretation and ‘objective’ FAIR Principles: A case-study from digital fragmentology» et pågående digitalt forskningsprosjekt med tittelen *Textus invisibilis*, der forskningsobjektet er fragmenter av resirkulerte middelaldermanuskripter. Slike manuskripter ble i tidlig moderne tid i Europa ofte resirkulert på ulike måter og gjenbrukt for eksempel som bokbindingsmateriale eller arkivemballasje. Molinaris prosjekt går ut på å sette opp en open access multimedia database med bilder og metadata for ca. 4000 middelalderfragmenter som er bevart i biblioteker og arkiver i Italia (se [www.textusinvisibilis.it](http://www.textusinvisibilis.it)).

Som bakgrunn for prosjektpresentasjonen tematiserte Molinari utfordringer knyttet til FAIR-prinsippene, som ethvert forskningsprosjekt som arbeider med store mengder data i dag må forholde seg til. Disse prinsippene er basert på vitenskapelige standarder fra naturvitenskapelige og «harde vitenskaper» og søker å fremme allmenn tilgang til kunnskap og vitenskap. Dette bringer nye utfordringer for humaniora-forskning. På den ene siden støtter FAIR-prinsippene ideen om universell filologi og utdanning i humaniora. På den annen side kolliderer de med noen grunnleggende prinsipper innen humaniora-epistemologi, som anerkjenner forskerens subjektivitet og bruk av kvalitative og fortolkende metoder. Spørsmålet blir hvordan man kan finne en fruktbar tilnærming som både respekterer FAIR-prinsippene og opprettholder humanioras mangfold.

Prosjektet har møtt stor velvilje og interesse både hos arkiver og studenter, som har samarbeidet om å identifisere tekstfragmenter og gjøre dem digitalt tilgjengelige uten å demontere dem fra nåværende innbinding. Formålet er å digitalt rekonstruere enhetene og lenke til digitale utgaver der det er mulig, og gjøre databasen interoperabel med eksisterende databaser (f.eks. Books within books, Fragmentarium). En drøm som foreløpig ikke er innenfor rekkevidde, er å kunne etablere alle tekstfragmentene som søkbar tekst og skape et fragment-google, der man kan gjenfinne konteksten basert på universelle søk.

Imidlertid har prosjektet støtt på utfordringer knyttet til finansiering, drift og langtidssikring, og det er uklart hvordan dette vil løses i tiden som kommer. Diskusjonen etter foredraget gjorde det tydelig at langsiktig bevaring av databaser og digitale utgaver er en utfordring for mange prosjekter, og at det er vanskelig å få de store institusjonene til å forplikte seg på lang sikt. Dette bør adresseres i fellesskap for å få større gjennomslagskraft.

*Christian Smidt Alenkjær*

I foredraget «Digitaliseringens litteraturdidaktiske potentiale – om brugen af H.C. Andersens eventyrmanuskripter i gymnasieundervisningen» presenterte Christian

Smidt Alenkjær (Syddansk Universitet) doktorgradsprosjektet «Undersøgende litteraturhistorie i gymnasiet – et interventionsstudie med utgangspunkt i den digitale udgave af H.C. Andersens eventyrmanuskripter». Prosjektet knytter an til utgivelsen av H.C. Andersens eventyr og historier ved Syddansk Universitet. Alenkjær viste hvordan (masse)digitalisering kan brukes i undervisningssammenhenger og hvordan litterære tilblivelsesprosesser synes å være særdeles egnet til det han kaller «undersøgende undervisning».

## **Takk**

Konferansen *Massedigitalisering og edisjonsfilologi* ble gjennomført med støtte fra Fritt Ord, UiO: Norden og Institutt for lingvistiske og nordiske studier ved Universitetet i Oslo.

Vi takker våre kollegaer Krista Stine Greve Rasmussen, Janicke Stensvaag Kaasa, Federico Aurora, Heidi Rohde Rafto, Margrethe Støkken Bue, Sebastian Köhler, Nils Holger Németh Berg og Hilde Bøe for deres viktige arbeid med fagfellevurderinger.

Takk også til Johanne Emilie Christensen for hjelp med referansehåndteringen.

---







# Essay

# Nya vägar till det förflutna: Två editionsfilologiska frågor till den storskaliga digitaliseringen

Mats Malm<sup>1</sup>, Ljubica Miočević<sup>2</sup>

<sup>1</sup>Svenska Akademien

<sup>2</sup>Stockholms universitet

## Introduksjon

Litteraturbanken intar en rätt egenartad position mellan editionsfilologiska projekt och storskalig digitalisering. Verksamheten började år 2004 genom ett anslag från Riksbankens Jubileumsfond och riktar sig mot forskningen, undervisningen, biblioteken och den breda allmänheten. Webbplatsen litteraturbanken.se förenar så ambitioner om storskalig digitalisering med tydlig uppmärksamhet på vetenskapliga utgåvor. Å ena sidan innebär det att den fungerar som ett arkiv där de som vet vad de söker kan finna material för läsning eller hämtning, inte minst för datorstödd analys – å andra sidan vårdar sig Litteraturbanken om att skapa ingångar och presentationer så att de som inte är så bekanta med det litterära kulturarvet kan orientera sig och finna fram till nya bekantskaper.

Den (förhållandevis) storskaliga digitaliseringen sker för närvarande inom ett projekt finansierat av Riksbankens Jubileumsfond, som går ut på att digitalisera all svenskskriven, separat publicerad skönlitteratur från 1800-talet och ta fram nya redskap för att utforska materialet. En bärande ambition är att erbjuda forskningen ett konsistent material som inte styrts av någon annan urvalsprincip än fullständighet. Projektet omfattar omkring 9 000 böcker och avslutas under 2024. Allmänt tillgängliggör Litteraturbanken mer och mer av det svenska litterära kulturarvet, med utgångspunkt i äldre material där upphovsrätten gått ut. Däri ingår moderna vetenskapliga utgåvor genom samarbete med en rad större och mindre utgivningsprojekt. Tillgänglighet är nyckelordet och det innebär alltså inte största möjliga antal digitaliseringar, utan i mycket hög grad kurering och anpassning för att öppna materialet för så många användare som möjligt: det finns en skolavdelning som erbjuder lärare material och pedagogiska redskap, en avdelning som erbjuder inläst litteratur samt musik och film, en utställningsavdelning för större tematiska presentationer, en litteraturkarta som gradvis täcker allt större delar av Sverige, en lång räckta specialskrivna presentationer till författare och teman, det stora Svenskt översättarlexikon, Dramawebben som lyfter fram den i hög grad okända dramatiken och en särskild ingång för bibliotekarier. Hösten 2022 var antalet verk omkring 8500, 1200 av dem tillgängliga som korrekturlästa epubfiler för läsplattor och mobiltelefoner. Vid slutet av 2024 beräknas antalet verk närma sig 14000.

I samband med att webbplatsen växer och digitaliseringen över lag tilltar har Litteraturbanken haft anledning att särskilt överväga behovet av textkritiska

perspektiv inom två områden. Det första anknyter till den vetenskapliga utgivningens kontextualisering och kommentar, och rör de många föreställningar om kön, etnicitet, det nationella och det främmande som ligger inbäddade i vårt kulturarv. Det andra angår de stora utgivningsprojektens framtid: det blir allt svårare att finna finansiering för författarskapsutgivning och frågan är hur vi kan bevara det editionsfilologiska perspektivet och helst till och med vidareföra det när digitaliseringen blir mer storskalig.

### **Att umgås med kulturarvet**

Vår historia erbjuder ovärderliga estetiska verk, filosofiska idéer och samhällsliga utvecklingslinjer, men den är också fylld av fördomar, ideologiska manipulationer och ansatser att bygga gemenskap genom exkludering. Vi är inte befriade från sådana fördomar och strukturer i dag, men ju längre tiden går, desto tydligare framträder det tidsbundna och orimliga i många av historiens föreställningar. I Litteraturbankens samlingar finns stora mängder av strukturella fördomar formulerade och dokumenterade. De är inte sällan stötande, men dokumentationen och bearbetningen av historiska föreställningar är viktig. De är en del av vårt förflutna och måste utgöra en del av vår förståelse av vårt kulturarv. Med den breda inriktning som Litteraturbanken har är det självklart att inte utöva censur genom att undertrycka verk som rymmer vad som i vår tid ses som föråldrade fördomar, eller att förse enskilda verk med olika former av varningstexter.

Litteraturbanken eftersträvar i tilltagande grad fullständighet i urvalen, till exempel inom projektet Nya vägar till det förflutna som avser att samla all svensk separattruckt skönlitteratur från 1800-talet för att öppna upp historien för forskare och allmänt intresserade. Ambitionen är att bidra till analys och problematisering inte minst av hur föreställningar formas, speglas, förmedlas och förändras i det litterära kulturarvet. Forskningen får tillgång till underlag för fördjupad analys av hur vårt samhälle och vår kultur genom olika processer och inflytanden formats fram till dagens situation. I andra länder som kommit längre med digitaliseringen har redan förtjänsterna av sådana material och perspektiv börjat synas – nu tänker jag inte minst på Norge, där Nasjonalbibliotekets välorganiserade och alldeles föredömliga digitalisering av det tryckta materialet öppnat för studier inte minst av kvinnors och nationella minoriteters villkor i historien.

Men samtidigt som forskningen får redskap att utvidga förståelsen av hur fördomar och föreställningar utvecklats och spridits i det litterära kulturarvet, blir ju materialet tillgängligt för många som inte besitter bakgrundskunskapen. För mottagare utanför forskarkretsar finns alltså påtagligt behov av att kommentera och klargöra. Frågan om sådana föreställningar kommer ofta upp i debatten, men inget samlat grepp har tagits från minnesinstitutionernas håll, åtminstone inte i Sverige. Litteraturbanken avser att rama in sitt material med en samlad diskussion av problem och möjligheter på området. Det görs genom presentationer i första hand av:

- Misogyni

- Antisemitism
- Rasistiska föreställningar generellt i det litterära kulturarvet
- Antisemitism och rasism ur mer institutionellt perspektiv
- Queerperspektiv
- Nationella minoriteter
- Den samiska litteraturen
- Kvinnliga författares villkor
- Funktionshinder
- Föreställningar om det svenska
- Religiös intolerans

Ett antal av presentationerna finns redan på hemsidan, och när de blivit tillräckligt många kommer de att ramas in under rubriken «Att umgås med kulturarvet». Det bör bli något som liknande organisationer och minnesinstitutioner kan hänvisa till. Det kan komma att ge upphov till vidare diskussioner, vilket vore en konstruktiv utveckling på ett svårt ämne.

### **Ett nytt slags författarskapsutgivning**

Litteraturbanken rymmer, inte minst genom samarbetet med Svenska Vitterhetssamfundet, ett stort antal samlade utgåvor av klassiska svenska författarskap, ofta gjorda med stor textkritisk medvetenhet och med utförliga kommentarer: Hedvig Charlotta Nordenflycht, Erik Johan Stagnelius, C.J.L. Almqvist och August Strindberg för att nämna några. Till sådana utgåvor kan framöver läggas en annan typ av författarskapsutgivning, gjord i Litteraturbankens egen regi.

Det är nämligen tydligt att traditionella författarskapsutgivningar av ovan nämnda slag kommer att ha än svårare att få finansiering i framtiden. Efter Strindberg och Almqvist har det till exempel visat sig mycket svårt att få finansiering för Selma Lagerlöfs författarskap. Det har att göra med att stora digitaliserade material delvis ersätter samlade verk i tryckt form, men också och inte minst med att digitaliseringen som kostnadskrävande humanistisk infrastruktur gör anspråk på ekonomiska resurser som riskerar att konkurrera ut den omsorgsfulla editionsfilologin. Litteraturbanken avser att utveckla ett annat slags författarskapsutgivning, en som fyller flera behov men kan åstadkommas med mindre resurser och som kan förmedla den textkritiska medvetenheten i en tid då sådan inte prioriteras men – i samtidens stormflod av information – är viktigare än någonsin. Det är ett slags studieutgåva men för hela författarskap, som ger underlag för den som vill gripa sig an en författare eller ett verk med textkritiska perspektiv – eller, förstås, digitalhumanistiska perspektiv. Helst ska textkritisk och digitalhumanistisk kompetens förenas. Författarskapsutgåvan omfattar idealt:

- Alla förstaupplagor kollationerade
- Övriga upplagor OCRade (dvs. inte kollationerade men sökbara)
- Övergripande introduktion till författarskapet
- Textkritisk introduktion till författarskapet
- Länkar in i ordböcker (Svensk ordbok, Svenska Akademiens historiska ordbok, Dalins Ordbok Öfver Svenska Språket)
- Möjlighet till spårning av olikheter mellan olika upplagor
- Möjlighet till spårning av likheter (parallella textsekvenser inom författarskapet och mellan författarskap)
- Nerladdningsbarhet (så att forskare kan använda egna redskap för att utforska materialet)
- Översättningar till andra språk (i den mån rättigheter medger)
- Kringmaterial i form av faksimil av manuskript, korrektur, annoterade exemplar med mera.

I stället för den traditionella vetenskapliga utgåvans ord- och sakförklaringar får användaren alltså direkt tillgång till ett antal ordböcker. I stället för den traditionella vetenskapliga variantapparaten får användaren möjlighet att själv jämföra de olika upplagorna med varandra. Prioriterade författarskap är i igångsättningsfasen: Sophie von Knorring, Fredrika Bremer, Emilie Flygare-Carlén, Victoria Benedictsson (inklusive manuskripten som visar textläget före Axel Lundegårds redigeringar: textläget här är utomordentligt intressant), Selma Lagerlöf, Agnes von Krusenstjerna, Birger Sjöberg, Karin Boye.

Författarskapsutgåvan riktas inte bara mot forskare och studenter utan även mot allmänheten:

- förstaupplagorna tillgängliggörs som epub-filer – även i moderniserad version som bland annat kan utgöra underlag för bokcirkelar via webbplatsen bokcirkelar.se
- särskilda inledningar för en bredare publik, där det textkritiska perspektivet varsamt introduceras
- förlag får fri tillgång till kollationerade och moderniserade texter och anmodas att använda dem för pappersutgåvor.

Genom detta slags författarskapsutgivning är ambitionen att kombinera massdigitaliseringens möjligheter med textkritisk medvetenhet, samtidigt som Litteraturbanken kan nå ut till nya användargrupper med moderniserade texter. Det riskerar förstås att konkurrera med traditionell författarskapsutgivning, men kan också fungera som ett idealiskt underlag för en mer traditionell utgåva då en

sådan får finansiering. Ljubica Miočević samordnar och är ansvarig för detta arbete.

Genom det kontextualiserande ramverket «Att umgås med kulturarvet» är avsikten att ge hela den stora korpusen av digitaliserade verk en bred inplacering och kommentar som är mycket generell, men ändå fyller vissa av den historisk-kritiska utgåvans funktioner.

Litteraturbanken avser att under de kommande åren prova ut de möjligheter dessa två inriktningar ger, parallellt med den ordinarie verksamheten. Detta kan betraktas som en startpunkt och arbetet med dem kommer att avpassas efter vilka behov som visar sig viktigast och vilka samarbeten som blir möjliga.

---







# Artikler

# Bibliography, database, and beyond: Towards queer scholarly editing

Jenny Bergenmar<sup>1</sup>

<sup>1</sup> Department of Literature, History of Ideas, and Religion | University of Gothenburg

## Introduction

How can we capture the production and sociology of queer texts and make it easier for readers to understand them? Although this question might seem to be intimately connected with scholarly editing, it has mainly been explored within other fields. Librarians and scholars within library and information studies have taken pains to compile subject guides and bibliographies of gay, lesbian or trans literature. Efforts have also been made to develop terms that can better index and describe such literature. Literary scholars have delved into queer readership and the transmission of queer texts, while carefully unlocking the queer meaning potentialities, metaphors, language use and intertexts of literary texts. However, in the fields of book history and philology, and especially scholarly editing, queer perspectives have remained marginal. Texts can be said to be queer on many different levels. Here, «queer texts» is understood as texts that has been read and received as queer in different contexts, or in other words seen to destabilise or undermine normative gender identities and sexualities. Queer book history is found in a variety of contexts: in discussions of special collections, queer publishing and bookshops, and queer archives; in research on individual authorships and works; and in investigations of queer books in translation.<sup>1</sup> There are also important outliers of philology aimed at developing queer perspectives, such as Jeffrey Masten's *Queer Philologies. Sex, Language, and Affect in Shakespeare's Time*[5]. Masten defines queer philology both as

a philology of the queer (that is, of sexual practices, bodies, affects, and identities that seem nonnormative, whether in their own time or from this historical distance) and a queer philology (the traditional and once-hegemonic discipline of philology read and practiced in a way that will highlight its own normativizing categories and elisions). This requires attending to queer philologies that have been deemphasized or submerged in more traditional accounts of language – a process that will continue here.[5, p. 214] (...)

But what about queer scholarly editing? How would such an endeavour work, and what principles should guide it? Based on experiences from the project «Queerlit: Metadata and searchability for queer literary heritage», I will outline some suggestions of what might characterise a queer scholarly editing practice in relation to mass digitisation, feminist scholarly editing, and queer philology.

---

<sup>1</sup>For special collections, queer publishing and queer bookshops, see [1]. For queer archives, see [2]. For individual works and authors, see [3]. For translation, see [4].

The Queerlit project ties into international efforts to make LGBTQI heritage accessible for the public, especially for queer communities. It involves the creation of a bibliographic database containing Swedish LGBTQI literature (prose fiction, autobiographies, poetry, drama, graphic novels, children's literature, and picture books), the development of a thesaurus (QLIT) to better describe the material, and the creation of a user-friendly interface that supports searching by themes and topics.<sup>2</sup> The Queerlit bibliography is also accessible in Libris, the Swedish National Union Catalogue. QLIT was developed for subject indexing in Queerlit specifically, but may be used by all libraries indexing their material in Libris.

Bibliography and indexing serve as forms of LGBTQ+ information and archival activism practiced both today and in the past.[6, 7] These practices are ultimately aimed at the same goal as scholarly editing: to enrich historical materials (e.g. literary works) with descriptions that might add to an understanding of them. Like indexing, which is intended to describe content in a way that is independent of subjective interpretation, claims of objectivity have been made for scholarly editing, such as in the description of manuscripts or editions as material objects and in the careful registering of variants. Needless to say, the principles practiced in scholarly editing in general are also useful in queer scholarly editing, even though the conditions for the production and reception of LGBTQ+ texts are different, similar to the case with other marginalised groups. However, as queer theory inherently contests the entrenched claims of objectivity within the broader scholarly editing landscape, queer scholarly editing deviates from the traditional approach that originates from such a standpoint. Considering the theme of this volume, namely, mass digitisation and scholarly editing, I begin by describing the contested relationship between mass digitisation and scholarly editing and how discussion within the digital humanities ties into these issues. I then provide an overview of previous work in feminist digital bibliography and scholarly editing. Finally, I propose principles for queer scholarly editing based on queer philology and examples from the Queerlit project.

### **Mass digitisation, digital humanities and scholarly editing**

The extensive digitisation of literature and historical materials during the past two decades has radically changed conditions for scholarship. Calls for mass digitisation projects, which would make all books available in a digital format – as seen in Norway and the United States, for example – have created a need to navigate a new abundance of data in various academic disciplines.<sup>3</sup> Mass digitisation has come to mean the transformation of the book from a material object bound by sociological relations and contexts – or what Robert Darnton has named «communication circuits» – into information unbound by these that can be processed and analysed by computers.[10] The possibility of developing methods for using large text corpuses as empirical material for literary studies is being explored within the digital humanities, with a focus on different forms of

---

<sup>2</sup>When referring to Queerlit, I use LGBTQI, since the bibliography covers lesbian, gay, bisexual, trans, queer, and intersex in Swedish literature. In all other instances, I use LGBTQ+, indicating the inclusion of a broad variety of non-normative sexualities and gender identities.

<sup>3</sup>For useful reflections on mass digitisation in American and Norwegian contexts, see [8] and [9]

large-scale computer-assisted analysis. Data-driven textual analysis has been criticised for not sufficiently supporting social justice and diversity.[11] However, a reorientation of quantitative literary studies has been proposed, which enhance their utility in addressing these concerns.

Queer readings, – that is, readings of texts that lack explicit representation of LGBTQ+ motifs or characters but may have been read as conveying queer meaning, might seem to be the very opposite of a large-scale analysis. As Elsi Hyttinen phrases it in a discussion on queer readings of early 20th century prose, «There is no formal or morphological feature that could be claimed as always signaling the queer».[12] The difficulty of identifying stable elements that signal queer meaning in texts is one likely explanation for the absence of computer-assisted readings of literature from queer perspectives; that queer meanings are conveyed abstractly, covertly or elusively is another. Moreover, the ill fit between empirically directed computer-assisted readings and theoretically informed queer readings has also to do with epistemological standpoints: while data-driven methods rest on the idea that literature can be transformed into data points, queer theory insists on the idea of fluid meanings and subjective readings. This does not mean that queer digital humanities is an impossibility; merely that it needs to take other forms and apply other methods.

Queer scholarly editing could learn from existing reconceptualisations of what distant reading might entail. Lauren F. Klein argues that quantitative methods for exploring literary history may become more useful if other analytical facets are added to the close/distant dichotomy, such as the visible/invisible. «These perspectives come into focus by considering the context that surrounds the production of a particular body of work, as well as the bodies of those who labored to produce it.»[13] Interestingly, Klein contends that though quantitative methods are often connected with a shift toward a distant view, obscuring non-dominant perspectives, this doesn't have to be the case. By detaching quantitative methods from the axis close and distant, she seeks to shift focus to other analytical axes, using visible and invisible as an example specifically highlighting gendered textual practices. Such a reconceptualization of the framework for quantitative analyses may assist in uncovering aspects that may be obscured in the distant and generalising view. Klein places significant importance on the material conditions influencing text production, thereby bridging divides between material and digital realms, on one side, and between small and large scale perspectives, on the other – divides that have been particularly acute for scholarly editing in the age of mass digitisation. Klein argues that, although we must be wary of how the concept of data ties into oppressive classification models, the question is how quantitative methods may be used to challenge these models and to uncover marginalised histories. One key element in such a recalibration of quantitative methods is to be able to better account for context and to actively seek to identify a lack of data from different kinds of datasets.

Klein's argument aligns well with Katherine Bode's work with the National Library of Australia (NLA)'s Trove database, a mass digitised collection of historical newspapers that she used to create a «data-rich literary history» by employing computational methods while paying close attention to the gaps in the

data and to historical contexts.[14, p. 3] Bode redefines the concept of a scholarly edition by making it encompass a broader section of the literary system than single works or authors and attempts to move beyond «the polemics of both ‘distant reading’ and ‘macroanalysis’ on the one hand, and McGann’s ‘new philology’ on the other, and thus beyond the view of digital literary history as defined by either the multiplication of data points (in Moretti’s and Jockers’ work) or the elaboration of unique and ultimately unknowable philological objects.»[14, p. 4]

One way in which mass digitisation has affected the practices of scholarly editing is that it has called into question the micro-level that scholarly editing usually works on. Indeed, one apparent challenge for scholarly editing in this context is to argue for the value of carefully curated and edited digital resources. Scholarly editions are labour-intensive and costly endeavours that ideally have the function of making texts intellectually accessible for new readers. However, in many cases scholarly editing has functioned as a consecration of hyper-canonical authors, while one of the arguments for mass digitisation is that it expands beyond that kind of canonical gatekeeping. This does not mean that scholarly editing is an inherently elitist pursuit. It has also proven a valuable tool for digital recovery projects, devoted to marginalised and forgotten literary heritage. Amanda Gailey has described the relationship between mass digitisation and scholarly editions as a tension between «automation» and «authenticity». In scholarly editing, significant emphasis is placed on the meticulous representation of texts, employing various editorial practices like selection, commentary, and annotation. In contrast, scholars using quantitative digital methods approach texts as datasets, seeking to extract knowledge automatically or, at the very least, semi-automatically. The proliferation of computer assisted methods have also demanded of digital editions to be more accessible for computer assisted analysis. 2016 MLA Statement on the Scholarly Edition in the Digital Age identifies three opportunities offered by digital media for scholarly editing: 1) it makes editions reusable in new editions, possible to compare using collation tools and reusable as parts of larger corpuses in computer-assisted analysis; 2) it supports annotation and additional layers of user or editor information; and 3) it allows for the use of an interface as an analytical tool to explore texts.[15, p. 1] The report underlines the «interdependence of micro- and macroanalysis», thereby attempting to mitigate the juxtaposition of careful and detailed digital scholarly editions (qualitative digitisation) and mass digitisation.[15, p. 3] However, as Barbara Bordalejo has pointed out, the MLA white paper underlines a continuity between the print and digital scholarly editions. She concludes: «there is no such thing as digital scholarly editing, [...] only scholarly editing, which can be published in print or digital format, but that remains the same principle linked to meticulous historical-critical work carried out by textual scholars or under their direct supervision».[16, p. 25]

### **Feminist bibliography and editing**

In this context, it may be useful to reflect on the early feminist digitisation projects and scholarly editions that emerged a few decades ago as a response to the lack of resources on women’s literature and history. Recovery projects from the 1990s and early 2000s, such as Women Writers Online (WWO), directed by

Julia Flanders, and Orlando, created by Susan Brown, Patricia Clements and Isobel Grundy, tackled the absence of feminist perspectives in digital humanities and scholarly editing by building digital bibliographies and digital collections that functioned as infrastructures, enabling research on the publications of marginalised female authors. As Laura Mandell points out, many similar projects followed, but many perished due to a lack of resources, and «we have nothing yet on the scale of Whitman, Blake or Rossetti archives, or the sites for Shakespeare, Thomas Gray, Herman Melville, to name a few [...]»[17, p. 514] Mandell describes the dilemma of recovery projects such as these: they may write women into literary history again, but «[b]ig data threatens to eradicate the history of women’s writing altogether, given that women originally published in small print runs and via manuscript circulation.»[17, p. 521] Thus, digitisation is challenging in different ways for women writers in history, as such authors are published in formats not neatly covered in mass digitisation projects, and their works are also less likely to be the object of digital scholarly editions since they are more seldom canonised; in fact, in a Nordic context, there are still no significant examples of digital scholarly editions of female authors.

An article by Martha Nell Smith, the founder and executive editor of the Dickinson Electronic Archives, was a formative text for my own entry into both scholarly editing and the digital humanities. In this article, Smith scrutinises the inclination – in what was then called humanities computing – to strive to establish authoritative and authentic editions, extracted from the «messiness» of social relations and positions, and to see technology as a neutral, objective and rational tool to accomplish this goal. Smith argues for a collaborative, non-hierarchical model of scholarly editing that allows for non-expert input, revisions over time and an acknowledgement of difference and power structures.[18, p. 14] Anyone familiar with queer theory will feel at home in this vision of a different practice of textual scholarship, with its affirmation of community knowledge and complex and changing social practices. Smith also uses the «lesbian rule» as a metaphor for textual scholarship that is not bound to a rigid set of rules: «in a seventeenth architectural sense», a lesbian rule was a tool originating from the island of Lesbos; it was used in construction and «could be bent to fit the curves of moulding».[18, p. 1] Through this image, Smith elegantly captures both «the lesbian rule» as a practical instrument that allows for flexibility in the construction of digital archives and editions, and «the lesbian rule» as a queer methodology associated with a resistance to fixed meanings and an acknowledgement of intersectional identities.

Although resources such as the WWO, Orlando, and Dickinson Electronic Archives show that there were early attempts to ensure that women’s writing was not made invisible by the move to digital publication models, there is still a male norm within scholarly editing and book history. Kate Ozment argues that book history has revolved around male print history and «inherited a discourse where gender is not central» from «the debates among D.F, McKenzie, G. Thomas Tanselle, and Jerome McGann that focused on the bibliographic concerns of textual editing and authorship.»[19, p. 150] In these debates, the question of how to establish an edition of a literary text was central: should it be defined by

authorial intent (Tanselle), or should an edition reflect the text as a social production – that is, as a product of a process with many actors (McGann and McKenzie)? Interestingly, Ozment makes an argument about book history that has also been made concerning the digital humanities: that it functioned as a refuge from the advances of critical theory – or, at least, the critical theory that meant diversifying the subject matter.[19, p. 160] However, she expressly points to Flanders and Smith, along with contributions by Amanda Gailey, as examples of how critical theory may be employed within book history and scholarly editing, where it complicates the concept of objectivity and the ideal of the authoritative edition.[20] Ozment creates a genealogy of feminist work in bibliography, which leads her to conclude that much of the work has been carried out by women librarians, archivists, cataloguers, and indexers; that is, work outside of the boundaries of academia. Here, feminist bibliography converges with queer bibliography and archiving, engaging the same groups of practitioners and re-evaluating that work.

### **Practices of queer (digital) scholarly editing**

Feminist and queer book history and scholarly editing overlap in certain areas, such as in making visible alternative kinds of publications, practices, and collection and circulation patterns. Scholarly editing is concerned with «how the literary work emerges in a network of connections between the author, the receiving readers, and institutional actors such as publisher, printer, typesetter, and the constellations of expectations and demands these contexts entail.»<sup>4</sup>[21, p. 15] Traditionally, one important task for the scholarly editor is to study textual transformations over different media: from the manuscript, through the interventions of editors and publishers, to the text that is finally printed (and then, in many cases, reprinted, and re-edited). Being a scholarly editor requires a deep understanding of the historical contexts involved, as well as a sensitivity to the nuances of historical language use. As this description shows, there are plenty of instances in which queer textual practices may be policed, and texts can be subjected to a «symbolic annihilation» of the queer. The concept «symbolic annihilation» was used by Gaye Tuchman already in 1978 to refer to how women were represented in American mass media, by trivialization, condemnation or omission. More recently, the concept has been used for discussing absences and misrepresentation of marginalised communities in archives. Scholarly editing is a practice where queer elements may be excluded, downplayed, or misrepresented, both in individual texts and in literary works. Frequently editors may lack the requisite expertise to identify instances conveying queer meanings, or deliberate choices may be made not to address them. This could be due to considerations of perceived triviality, irrelevance, or that it would result in a deviation from what is considered an authentic representation of the text or work. Despite the ideological assumptions that have underpinned philology, be it ideas about authentic texts or

---

<sup>4</sup>My translation, quotation in original: "hur det litterära verket uppstår i ett nätverk av förbindelser mellan författaren, de mottagande läsarna och individer verksamma vid institutioner som förlag, tryckeri och sätter, samt de konstellationer av förväntningar och krav som dessa sammanhang frambringat."



heteronormative frames of interpretation, the principles in themselves can still be useful. It is precisely the scholarly editor's thorough knowledge of the material, social and historical aspects that is needed for readers to be able to understand queer texts, the conditions for their production and circulation, and their (often conflicted) relation to literary institutions.

When practicing queer philology in editorial work, it is crucial to regard it as interpretive work, rather than as an objective descriptive practice preceding interpretation. For instance, Masten points to tendencies to erase queer meanings in emendations that serve to normalise language. To practice queer philology as defined by Masten, that is both as a philology of the queer (namely non-normative bodies, actions, and identities), and as a queering of philological practices, it is necessary to pay close attention to the complex meanings of historical terms and to be able to preserve and represent this complexity, by giving justice to unfamiliar terms and to those we recognise but that might carry other meanings in their historical contexts.[5, p. 225] Glossaries that allow for multiple, complex and interconnected meanings are one tool that can be used to achieve this goal. In fact, this kind of editorial practice might be construed as the very opposite of the practice of subject indexing literature – that is, assigning fixed descriptive terms to works. Since the purpose of indexing is to support discoverability, multiple complex meanings are difficult to capture in subject indexing. A recurring criticism of general systems for subject description, such as Library of Congress Subject Headings, is precisely that such terms function as mechanisms of control, universalising identity categories and marginalising experiences not captured by them. In Adler's discussion of this problem, she emphasises that adding new terms does not suffice. Instead, she argues for subject description «based on more just conceptualisations of the richness and relations among subjects and subjectivities.»[7, p. 154]

Like Masten, Adler emphasises the complexity of language that describes sexualities and gender identities. It is easier to account for these aspects as a philologist than as an indexer. However, there are similarities between indexing and encoding in digital scholarly editing. The criticism of what can be captured by indexing terms is also similar to the feminist critique of the markup systems used in scholarly editing. Like an indexing system, markup in scholarly editing is aimed at making information about texts accessible to readers by adding different kinds of labels and relations between terms. Kathryn Holland and Susan Brown have described how the Orlando project contested the idea that markup is an objective tool, and how interpretive tags that better represented the conditions and lives of women authors were created. Some of these, such as «CulturalFormation», allow for the representation of social contexts and historical and intersectional identities through the subtags «Sexuality», «RaceAndEthnicity», and «PoliticalAffiliations».[22, p. 413] Furthermore, the tag «IntimateRelationships» can be included multiple times in a single entry or not at all. It can capture «aspects of subjects' activities and communities that involve companionship, romance, queerness, chastity, sapphism, domesticity, and separation, as well as discussions of how they resist or complicate such terms».[22, p. 415] Moreover, Gailey has argued that the drawback of having an encoding standard such as TEI

for digital scholarly editions is that it steers the scholars involved in digital editing projects to tag precisely the characteristics and structures of texts that TEI already supports. Even if it is possible to build customised tagsets, like those described by Holland and Brown, there is a higher technical and intellectual threshold for such work. Furthermore, Gailey argues that TEI works best for projects that are «primarily interested in single authors and single texts.»[20, pp. 128-129]

Having – or making sure to create – a tag system that allows for multiple definitions of intimate relations along a spectrum, instead of a heteronormative dichotomy, is a useful tool in queer scholarly editing and a deliberate design to foster interpretative diversity. In essence, such efforts are related to attempts to refute the norms of classification systems and the view that data is unambiguous and objective. How data can be modelled to capture ambiguity and complexity is an important question to handle in any work with a queer bibliography and queer scholarly editing. As the Orlando project shows, interventions into markup languages may support the understanding of queer texts and authorial practices but may need tweaking based on an analysis of what is missing in them. As in the case of the subject indexing of texts, predefined tag sets (or controlled vocabularies in indexing) may be unfit for describing complex practices, identities and meanings; however, that also depends on how these systems are used. «Perverse» markup practices may be employed, «aim[ed] to some extent at a deviant encoding, one that starts to look beyond an objective rendering of formal features of the text to interpretative structures of narrative and identity construction.»[23, p. 241]

### **Queer vocabularies, (un)queer variants and queerer versions**

Building on the work of Masten and on the interventions into markup described by Holland and Brown, we can conclude that queer scholarly editing must pay close attention to how texts are emended, how glossaries are constructed, how markup principles are employed in digital editions and what tags are allowed. The practices of explaining words in a glossary is an interpretative work with similarities to the selection and explanation of terms in a domain specific thesaurus, such as the QLIT thesaurus. All indexing terms in QLIT have a so-called «scope note» explaining the definition and intended use of the term. Of course, such definitions can support or obscure queer meanings, like the explanations of words in a glossary included in a scholarly edition. Thus, QLIT functions as a resource for queer scholarly editing, at least of Swedish literature, since it provides explanations of terms used in the literature, both current and historical. Even if every scholarly edition must consider words in a particular context (e.g. within a work or within an author's writings), important context can be provided by how the terms in QLIT are applied to other works. In particular, the category «historical terms», which contains words that are derived from works and thus represent actual use, such as «Saphists» (Sapfister) or «Tribads» (Tribader), may be useful for discovering the context of particular words.

Building a bibliography of Swedish LBGTQI literature in Queerlit has demonstrated the importance of another task for the scholarly editor: accounting for versions and variants. For example, there are interesting examples of later editions obscuring nonnormative sexualities and gender identities in works that

originally contained them. One is Alice Lyttken's novel *Flykten från vardagen* (*The escape from the everyday*; 1933), a contemporary novel with a well-educated female protagonist who breaks free from a loveless marriage and – along with primitivistic aesthetic ideals – longs for a healthy sexual life and motherhood.[24, pp. 34-35] In her dissertation, Liv Saga Bergdahl points out that there is a lesbian motif in the novel.[25, p. 37] The protagonist's friend Astrid commits suicide, and it is implied that she was lesbian. In the original edition, Astrid is described as not quite normal («inte riktigt normal») and as having a perverse inclination («perverst inställd»).[26, p. 240] In later editions (including the one now available as an e-book), Astrid is described as a questionable lady («en vidlyftig dam») and as having inappropriate heterosexual relations. It has not been determined when these changes were made, but they beg the question of to what extent representations such as these were edited out of the prose fiction of the 1930s more generally. It might also have been the author's own choice, perhaps in view of new editions and her increased popularity as a writer.

A more recent example is Inger Edelfeldt's youth novel *Duktig pojke* (1977), which is often named the first young adult coming-out story published in Sweden. The second and subsequent editions were significantly changed; among other things, they portrayed homosexuality in a more positive light.[27] In both cases, the changes in how homosexuality was treated were connected to complex relations between publishers and readers and to changing values, both aesthetic and moral. It is important to acknowledge that the need to suppress representations of nonnormative sexual practices and gender expressions is interwoven with the cultural and social contexts; thus, the scholarly editor must review such changes and declare how the scholarly edition handles them. Scholarly editors also have literary agendas, moral views and aesthetic preferences that are reflected in their editorial work. The disputes about the Sappho fragments that have haunted the image of Sappho in Sweden, inspired by nineteenth-century German philology, are a case in point.[28] It is crucial for the editor's perspective to be transparently disclosed in the introduction to the edition. Additionally, a thorough critical analysis of previous editions is imperative to prevent the perpetuation of normative assumptions from earlier versions.

These examples also highlight another aspect of queer scholarly editing to be considered: what we choose as work that is significant enough to warrant a scholarly edition. Traditionally, scholarly editions have been made for prominent – usually male – authors holding stable positions within the literary canon. Canonisation is thus one factor in scholarly editing; another is the focus on individual authors or works. A queer scholarly editor might need to challenge both factors. As the Lyttkens example shows, interesting works from a queer perspective might be found within popular or «middlebrow» literature. The Edelfeldt case demonstrates that other genres might also be of interest, such as children's and young adult literature. Most importantly, since LGBTQ+ experiences are marginalised in literature, it might be necessary to focus on clusters of works by different authors rather than reserving the philological endeavour to single authorships or works. The different conditions and circulation

---

<sup>5</sup>Here, I would like to thank Sam Holmqvist, who pointed out this erasure of the lesbian to me.

routes for female authors that Mandell identified are also relevant for literature published for LGBTQ+ audiences, which was often circulated through small publishing houses and bookshops. Finally, how to account for queer readings – that is readings of texts that lack explicit representation of LGBTQ+ motifs or characters but have been understood as queer – is a challenging question for anyone engaged in queer scholarly editing. In this instance, one basic principle of scholarly editing must be addressed: following the author’s intention. For queer scholarly editing, it is important to recover intentional queer meanings which may have been obscured by editors or publishing houses. On the other hand, texts (or parts of texts) might also accrue queer meanings for different reading communities over time, regardless of the author’s intention. Accounting for these changing understandings of words, passages or plots may be as important as restoring the original meaning of texts that have been submitted to revisions – by, for example, publishers – that obscure queer content.

### **Conclusion: beyond the politics of correction**

In *Queer/Early/Modern*, Carla Freccero states that «all textuality, when subjected to close reading, can be said to be queer.»[29, p. 5] As used by Freccero, «queer» is a broad, ultimately undefinable concept that has to do with a critique of heteronormativity, sexual and gender subjectivities, and positionalities, while being associated with the odd, strange or perverse. In many descriptions of the purpose of philology, that which is odd or strange is foregrounded. Michael Edward Moore proposes that philology involves engaging with texts that seem «wrong’» in some way. Philology is needed when a text «defies comprehension or appears to be faulty – the textual tradition seems broken into shards or has been muddied by spurious readings, interpolations, or contaminations.»[30, p. 463] In many ways, this description resonates with ideas about queer presences in texts and in the literary tradition. Queerness is often perceived by readers as wrong – something belonging to a broken tradition, or something that has been subjected to misrepresentation and misreading. The rationale for queer scholarly editing is a desire to understand the skewed, faulty, or unintelligible by means of contextualising while paying attention to language use and the conditions for queer authors and readers. As Masten puts it, queer editing philologies are about «more and different explanatory whispers in the margins and in the notes.»[5, p. 230]

In this text, I have not criticised the basic principles of scholarly editing per se; rather, I have called attention to instances in which they should be applied differently. This is the same method that was used in the *Queerlit* project regarding indexing and metadata practices. For example, we have indexed minor characters and motifs in texts, even though the National Library of Sweden states in its guidelines for indexers that, in order to index a book with a subject term, the subject must make up 20% of the book.[31] In this way, more literature can be included in the bibliography, and it is possible to visualise the marginalisation of LGBTQ+ experiences in literary history. As mentioned earlier, we have also included historical terms in the *QLIT* thesaurus that are directly derived from the texts included in the bibliography. This addition is a way of opening up texts to queer meanings for readers who are not familiar with the resonance of such terms

(e.g. the term «Stockholmare» was sometimes historically used to describe persons assumed to be homosexual).[32] Another strategy is to include texts in the bibliography that do not contain explicit representations of LGBTQI characters or motifs but do include symbolism that connects the texts to queer literary traditions or that has been read as queer by scholars or in queer communities. Although indexing and bibliographic work are usually performed by librarians rather than by scholars, they have many similarities with scholarly editing – especially digital scholarly editing. Both areas involve tools and standards, as well as indexing systems and encoding systems that set general parameters but may also be the starting point for the development of new vocabularies or tags.

The collection of works into bibliographies – such as Queerlit – and the scholarly editing of literary works have one goal in common: to make literature more accessible to readers. On a positive note, the practice of collecting and describing LGBTQI fiction, as is done in the Queerlit project, can be viewed as a way of reconstructing a marginalised literary heritage and acknowledging oppressed experiences. At the same time, library practices such as classification and indexing are tools of power that may enforce universalism, even when they strive to include marginalised topics and subjects. Similarly, scholarly editing may have a patrilinear heteronormative heritage, but I believe there is still room for queer whispers in the margin, lesbian rules and queer encoding.

# Bibliografi

- [1] Leila Kassir and Richard Espley, eds. *Queer Between the Covers. Histories of Queer Publishing and Publishing of Queer Voices*. London: University of London Press, 2021. ISBN: 1-913002-05-5. JSTOR: [j.ctv123x59r](https://www.jstor.org/stable/j.ctv123x59r). URL: <https://www.jstor.org/stable/j.ctv123x59r>.
- [2] Rebecka Taves Sheffield. *Documenting Rebellions: A Study of Four Lesbian and Gay Archives in Queer Times*. Series on Gender and Sexuality in Information Studies 11. Sacramento, California: Litwin Books, 2020. ISBN: 1-63400-113-3.
- [3] Laura L. Doan and Jay Prosser. *Palatable Poison. Critical Perspectives on The Well of Loneliness*. New York: Columbia University Press, 2002. ISBN: 0-231-11874-0.
- [4] B.J. Epstein and Robert Gillet, eds. *Queer in Translation*. London: Routledge, 2017. ISBN: 978-1-4724-5623-6.
- [5] Jeffrey Masten. *Queer Philologies. Sex, Language, and Affect in Shakespeare's Time*. Philadelphia: University of Pennsylvania Press, 2016. ISBN: 978-0-8122-4786-2.
- [6] Cait McKinney. *Information Activism. A Queer History of Lesbian Media Technologies*. Durham: Duke University Press, 2020.
- [7] Melissa Adler. *Cruising the Library. Perversities in the Organization of Knowledge*. New York: Fordham University Press, 2017. ISBN: 978-0-8232-7636-3. JSTOR: [j.ctt1xhr79m](https://www.jstor.org/stable/j.ctt1xhr79m). URL: <https://www.jstor.org/stable/j.ctt1xhr79m>.
- [8] Robert Darnton. *The Case for Books. Past, Present, and Future*. New York: Public Affairs, 2009. ISBN: 978-1-58648-826-0.

- [9] Anne-Britt Gran, Eivind Røssaak, and Linn-Birgit Kampen Kristensen. “Digital Infrastructure for Diversity—On Digital Bookshelf and Google Books”. In: *The Journal of Arts Management, Law, and Society* 49.3 (2019), pp. 171–87. DOI: 10 . 1080 / 10632921 . 2019 . 1581114. URL: <https://www.tandfonline.com/doi/full/10.1080/10632921.2019.1581114>.
- [10] Robert Darnton. “What Is the History of Books?” In: *Daedalus* 111.3 (1982), pp. 65–83. ISSN: 0011-5266. JSTOR: 20024803. URL: <https://www.jstor.org/stable/20024803>.
- [11] Elizabeth Losh and Jacqueline Wernimont, eds. *Bodies of Information. Intersectional Feminism and the Digital Humanities*. Debates in the Digital Humanities. Minneapolis: University of Minnesota Press, 2018. ISBN: 1-4529-5858-0. JSTOR: 10 . 5749 / j . ctv9hj9r9. URL: <https://www.jstor.org/stable/10.5749/j.ctv9hj9r9>.
- [12] Elsi Hyttinen. “Spectacles of Silence: The Queer in 1910’s Finnish Literature”. In: *Lambda nordica* 21.3–4 (2016), pp. 35–55. ISSN: 1100-2573. URL: <https://lambdanordica.org/index.php/lambdanordica/article/view/526>.
- [13] Lauren F. Klein. “Dimensions of Scale. Invisible Labor, Editorial Work, and the Future of Quantitative Literary Studies”. In: *PMLA* 135.1 (2020), pp. 23–39. ISSN: 0030-8129. DOI: 10.1632/pmla.2020.135.1.23. URL: <https://www.cambridge.org/core/journals/pmla/article/dimensions-of-scale-invisible-labor-editorial-work-and-the-future-of-quantitative-literary-studies/9BB30779EA9C9CF7AF176E7FEF12F484>.
- [14] Katherine Bode. *A World of Fiction. Digital Collections and the Future of Literary History*. Ann Arbor: University of Michigan Press, 2020. ISBN: 0-472-90083-8. URL: <https://muse.jhu.edu/book/59018>.
- [15] M.L.A. Committee Scholarly Editions. *MLA Statement on the Scholarly Edition in a Digital Age*. May 2016. URL: <https://www-mla-org.ezproxy.ub.gu.se/content/download/52050/file/rptCSE16.pdf>.
- [16] Barbara Bordalejo. “Digital Versus Analogue Textual Scholarship or the Revolution Is Just in the Title”. In: *Digital Philology. A Journal of Medieval Cultures* 7.1 (2018), pp. 7–28. ISSN: 2162-9544. DOI: 10.1353/dph.2018.0001. URL: <https://muse.jhu.edu/article/700506>.
- [17] Laura Mandell. “Gendering Digital Literary History. What Counts for Digital Humanities”. In: *A New Companion to the Digital Humanities*. Ed. by John Unsworth, Ray Siemens, and Susan Schreibman. Chichester: John Wiley and Sons, 2015, pp. 511–24. ISBN: 1-118-68060-X. DOI: 10 . 1002 / 9781118680605 . ch35. URL: <https://onlinelibrary.wiley.com/doi/10.1002/9781118680605.ch35>.
- [18] Martha Nell Smith. “The Human Touch of the Highest Order. Revisiting Editing as Representation.” In: *Textual Cultures* 2.1 (2007), pp. 1–15. ISSN: 1559-2936. DOI: 10.2979/TEX.2007.2.1.1.
- [19] Kate Ozment. “Rationale for Feminist Bibliography.” *Textual Cultures: Texts, Contexts*. In: *Textual Cultures: Texts, Contexts, Interpretation* 13.1 (2020), pp. 149–78.

- [20] Amanda Gailey. *Proofs of Genius. Collected Editions from the American Revolution to the Digital Age*. Ann Arbor: University of Michigan Press, 2015. ISBN: 0-472-12126-X. URL: <https://muse.jhu.edu/book/42621>.
- [21] Paula Henrikson. *Textkritisk utgivning. Råd och riktlinjer [Scholarly editing. Guidelines and principles]*. Stockholm: Svenska Vitterhetssamfundet, 2007. ISBN: 978-91-7230-139-9.
- [22] Kathryn Holland and Susan Brown. “Project|Process|Product. Feminist Scholarly Subjectivity in a Shifting Scholarly Field”. In: *A New Companion to the Digital Humanities*. Ed. by John Unsworth, Ray Siemens, and Susan Schreibman. Chichester: John Wiley and Sons, 2015, pp. 409–133. ISBN: 978-1-5179-0611-5.
- [23] Marion Thain. “Perspective: Digitizing the Diary. Experiments in Queer Encoding (A Retrospective and a Prospective).” In: *Journal of Victorian Culture* 21.2 (2016), pp. 226–41. ISSN: 1355-5502. DOI: 10.1080/13555502.2016.1156014.
- [24] Eva Heggstad. *Alice Lyttkens. I och om 1930-talet [Alice Lyttkens. Writing in and about the 1930s.]* Uppsala: Skrifter från Centrum för genusvetenskap, 2009. ISBN: 978-91-975680-9-8.
- [25] Liv Saga Bergdahl. “Kärleken utan namn. Identitet och (o)synlighet i svenska lesbiska romaner [The love without name. Identity and (in)visibility in Swedish lesbian novels]”. PhD thesis. Umeå: Umeå University, 2010. ISBN: 9789172649620.
- [26] Alice Lyttkens. *Flykten från vardagen*. Stockholm: Albert Bonniers Förlag, 1933. ISBN: 978-91-0-014793-8 91-0-014793-1. URL: [https://books.google.se/books/about/Flykten\\_fr](https://books.google.se/books/about/Flykten_fr).
- [27] Sebastian Nilsson Lindberg. “Att skriva, skriva om, och skriva om sig själv. En komparativ undersökning av Inger Edelfeldts och Bengt Martins omarbetade berättelser om homosexualitet [Writing, writing again, and writing the self. A comparative study of Inger Edelfeldt’s and Bengt Martin’s revised narratives about homosexuality]”. MA thesis. Södertörn: Södertörns högskola, Institutionen för genus, kultur och historia, 2011. 82 pp.
- [28] Sigrid Cullhed Schottenius. “När Sapfo kom ut [When Sapfo came out]”. In: *Klassisk filologi i Sverige. Reflexioner, riktningar, översättningar, öden*. Ed. by Eric Cullhed and Bo Lindberg. Konferenser / Kungl. vitterhets historie och antikvitets akademien 89. Stockholm: Kungl. vitterhets historie och antikvitets akademien, 2015, pp. 191–202. ISBN: 978-91-7402-439-5.
- [29] Carla Freccero. *Queer/Early/Modern*. Durham: Duke University Press, 2006. ISBN: 1-283-02207-9.
- [30] Michael Edward Moore. “Philology and Presence”. In: *The European legacy, toward new paradigms* 22.4 (2017), pp. 456–471. ISSN: 1084-8770. DOI: 10.1080/10848770.2017.1301106. URL: <https://www.tandfonline.com/doi/full/10.1080/10848770.2017.1301106>.



- [31] Kungliga Biblioteket. *Principer för ämnesordsindexering [Principles for subject indexing]*. Metadatabyrån. URL: <https://metadatabyran.kb.se/ammesord-och-genre-form/svenska-ammesord/principer-for-ammesordsindexering..>
- [32] *QLIT thesaurus*. Queerlit. URL: <https://queerlit.dh.gu.se/subjects..>
-





# «Op ad Midten, krydser og falder af» Dansemanuskriptet *XII Engels Danse* som kildeutgivelse

Elizabeth Svarstad<sup>1</sup>

<sup>1</sup> Fagseksjon for musikkteori, komposisjon og musikkteknologi | Norges musikkhøgskole

## Introduksjon

*XII Engels Danse* er et manuskript bevart i musikkksamlingen ved Nasjonalbiblioteket i Oslo.<sup>1</sup> Det inneholder tekstlige beskrivelser av såkalte engelskdanser med tilhørende noter til dansemelodier. I denne artikkelen skriver jeg om arbeidet med en kildeutgivelse og diskuterer spørsmål rundt hvordan tekst som representerer en fortidig kroppslig praksis kan gi mening som utgivelse.<sup>2</sup>

## Bakgrunn

I mitt doktorgradsarbeid om dans som sosial dannelse i Norge 1750–1820, valgte jeg dansebøker som er bevart i norske arkiver fra samme periode som hovedkilder [2]. Benevnelsen «dansebok» innebærer at boken inneholder beskrivelser av hvordan dansen skal utføres.<sup>3</sup> Jeg samlet 15 dansebøker fra Nasjonalbiblioteket, Universitetsbiblioteket i Bergen, Gunnerusbiblioteket i Trondheim og Statsarkivet på Hamar. De fleste er manuskripter, men det er også noen trykte dansebøker. Mange inneholder tekstlige beskrivelser av dansene. Noen inneholder dansenotasjon og andre inneholder en kombinasjon av tekstlige beskrivelser og notasjon. Mange av bøkene inneholder også noter til dansemelodiene, men det gjelder ikke alle. Seks av de 15 dansebøkene som utgjorde mine hovedkilder finnes på Nasjonalbiblioteket. Fire av dem er manuskripter, og *XII Engels Danse* er et av dem. Nasjonalbiblioteket digitaliserte bøkene, og jeg studerte også de fysiske dokumentene nøye mens jeg arbeidet med transkripsjonene.

Nasjonalbiblioteket startet tidlig med digitalisering av manuskripter, og et ønske fra Musikkarvprosjektet ble starten på en massedigitalisering av musikkmanuskripter og noter, som nå har pågått i over 10 år. En god del av manuskriptene er nedtegnelser av dansemelodier i ulike sjangre, men inneholder kun noter og ikke beskrivelser av selve dansene. Det er typisk manuskripter som har vært brukt til å fremføre musikk, og i mange tilfeller også til å spille til dans. Selve dansepraksisen kommer imidlertid ikke frem direkte i disse dokumentene.

Margrethe Støkken Bue, som arbeider ved Seksjon for musikk ved Nasjonalbiblioteket inviterte meg til en utforskning av hvordan vi kan behandle andre former for tekst og uttrykk i en tekstbasert, digital sammenheng. På NNEs konferanse i København i 2019 holdt Bue og jeg en presentasjon om dansenotasjon,

<sup>1</sup> *XII Engels Danse*. Mus.ms. 8634:657. Nasjonalbiblioteket.

<sup>2</sup> Parallelt med at jeg har skrevet denne artikkelen har jeg arbeidet med en digital kildeutgivelse av manuskriptet i Nasjonalbibliotekets kildekriftserie *NB kilder*. [1]

<sup>3</sup> «Dansebok» kan også brukes om en bok som inneholder kun dansemelodier og ikke danseinstruksjoner på samme måte som benevnelsen «dans» brukes om både dansemelodien og om dansen (som danses til dansemelodien).

ulike systemer for dansenotasjon, forholdet mellom dansenotasjon og musikknotasjon, og begge sett i sammenheng med tekst og som tekst. Problemstillingene fra konferansepresentasjonen er videre utdypet i artikkelen «Koreografi – dansekunst som tekst».[3]

Kildeutgivelse av dansebøkene som finnes i Nasjonalbibliotekets samlinger er kommet til som følge av doktorgradsarbeidet og samarbeidet med Bue, og det som særlig er interessant og utfordrende med dette utgivelsesprosjektet, er sammenkoblingen av musikk- og dansebeskrivelser og dansenotasjon tilknyttet samme verk.

### Manuskriptet *XII Engels Danse*

Manuskriptet er et tynt hefte på seks blad. Det har en forside med tittel og hvert oppslag inneholder dansebeskrivelser på venstre side og noter på høyre side. Det inneholder 12 dansemelodier, men beskrivelser av kun syv av dansene. For resten av dansene står dansens nummer angitt, men plassen til beskrivelsen står åpen. Tre av musikkstykkene har titler: *La Vertu Couronné*, *Le Printems* og *Le Triomphe*, som er typiske titler for denne typen danser. Papiret er mykt og gulnet, heftet har vært brettet i tre, og et hjørne er revet av på alle bladene. Manuskriptet har tilhørt Grevskapet Larviks arkiv før det kom til Nasjonalbiblioteket, og ved avleveringen ble det anslått å være fra ca. 1753.[4]

Som historisk kilde for dans føyer *XII Engels Danse* seg inn i en tradisjon med dansebøker som inneholder repertoar for dans og musikk, helt i tråd med europeiske trender på samme tid. Det finnes svært mange slike dansebøker i Norden, særlig i Danmark og Sverige og ellers i Europa. Et manuskript som *XII Engels Danse* kan ha tilhørt en dansemester, men mer sannsynlig er det at det er en avskrift. Manuskriptets opphavsperson har sannsynligvis skaffet seg avskriften for å praktisere dansene på ball.

### Engelskdanser

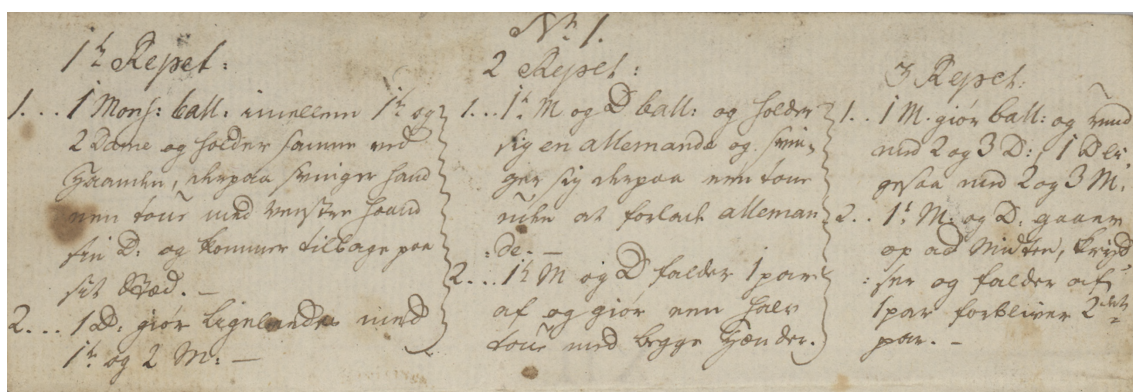
Dansene i heftet kalles engelskdanser.<sup>4</sup> Engelskdanser er en del av en større gruppe danser som på norsk kalles kontradanser. Kontradansene i det norske kildematerialet deles inn i to kategorier, nemlig danser med engelsk opphav og danser med fransk opphav. Kontradansen har røtter til engelske *country dances*, og den mest betydningsfulle kilden for denne danseformen er John Playfords samling *The English Dancing Master*.[5] Samlingen inneholder beskrivelser og noter til danser i ulike oppstillinger, for eksempel ring, firkant og rekker. Noen danser er for fire, seks eller åtte personer, og mange er for et vilkårlig antall par; for «as many as will». Dansene i *XII Engels Danse* er av typen «for alle som vil». *Country dance* var en svært populær danseform på 1600- og 1700-tallet, og Playfords samling ble utgitt i mange opplag. I tillegg finnes det et stort antall samlinger med samme type danser, både håndskrevne og trykte bøker, blant annet i den rikholdige samlingen ved Vaughan Willams Memorial Library.[6]

---

<sup>4</sup>Engelskdanser kan i det norske kildematerialet også kalles engelskdands, englois, anglaise og lignende stavemåter.[2, s. 205]

På slutten av 1600-tallet tok den franske dansmesteren André Lorin *country dances* til Frankrike. Han skrev to samlinger med danser, såkalte *contredances*, og dansene ble tilpasset den franske dansestilen med posisjoner, trinn på halvtå med avansert teknikk og eleganse[7, 8]. Fra England og Frankrike spredte danseformen seg til andre europeiske hoff og i Norden finner vi dem igjen som kontradanser. Dansene med oppstilling av parene på rekke kategoriseres som danser med engelske røtter, mens danser med oppstilling i firkant kategoriseres som danser med franske røtter. Dansene med engelske røtter er såkalte sett-danser,<sup>5</sup> og de kan ha enten småsettprogresjon eller helsettprogresjon. Småsettprogresjon vil si at dansesekvensen utføres i et sett på et færre antall dansere, for eksempel to par eller tre par, og det dannes nye sett hver gang sekvensen er avsluttet og skal begynne på nytt. På den måten kan dansene vare så lenge en vil eller til alle har danset med alle. I dansene med fransk opphav danses delene av dansen på tvers av firkantoppsettet, og når det antallet turer som dansebeskrivelsen inneholder er gjennomført, er dansen ferdig. Det er denne typen danser som utvikler seg videre til såkalte kvadriljer (*quadrille*, *lanciers*, eller *française/fransese*). Dansene i *XII Engels Danse* er engelskdanser med småsettprogresjon og tre par.

Engelskdanser består av en sekvens med turer og figurer, og parene danser med hverandre, fra hverandre og sammen med andre og bytter plasser hele tiden gjennom sekvensen. Det er gjerne et øverste par som leder dansen, de andre parene er organisert etter nummer: andre par og tredje par. Sekvensen danses om igjen og om igjen, og for hver gang den danses, flytter parene seg en plass nedover i rekken eller opp i rekken, og dermed får alle par danset med alle par. Dansen danses til alle parene er tilbake på den plassen de startet.



Figur 1. Nr. 1, *XII Engels Danse*.

## Prinsipper og transkripsjon

I transkripsjonen har jeg skrevet ut forkortelser og lagt til kursiv for å lette lesbarheten. Flere ganger i teksten angis tall når det egentlig skal være ordenstall. Jeg har løst opp og lagt til kursiv. Nedenfor følger et eksempel på transkripsjon av dans nr. 1 i manuskriptet. Ved oppløsning av forkortelser har jeg forsøkt holde meg til den språklige normen i dansebeskrivelser fra 1700-tallet. For noen ord gir

<sup>5</sup>Et sett er en formasjon av to, tre eller flere par.

manuskriptet selv svar på hva som er normen, mens andre ord, for eksempel «ball.», har jeg skrevet ut ordet «ballancere» slik det er brukt i andre dansebeskrivelser.

## Nr. 1.

### 1te Repetition

1. *1te Monsieur ballancerer* imellem 1te og 2de Dame og holder samme ved Haanden, derpaa svinge hand een tour med venstre haand sin Dame og kommer tilbage paa sit Stæd.
2. 1te Dame gjør Ligeledes med 1te og 2de Monsieur.

### 2de Repetition

1. 1te Monsieur og Dame ballancerer og holder sig en allemande og svinger sig derpaa een tour uden at forlade allemande.
2. 1te Monsieur og Dame falder 1 par af og gjør een halv tour med begge Hænder.

### 3de Repetition

1. 1te Monsieur gjør ballance og rund med 2de og 3die Dame, 1te Dame ligesaa med 2de og 3die Monsieur
2. 1te Monsieur og Dame gaaer op ad Midten, krydser og falder af 1 par forbliver 2det par.

## Dansens stammespråk

Dansebøker som *XII Engels Danse* ble skrevet som samlinger med danser for bruk. Språket er kortfattet og konsist og inneholder forkortelser og danseterminologiske uttrykk. Formen er stikkordsmessig fordi teksten skulle kunne leses raskt, memoreres og utføres i praksis. Teksten representerer en praksis. Den er skrevet som huskehjelp for elever som hadde deltatt på danseskole eller fått privatundervisning. De hadde lært praksisen som teksten beskriver og var innforstått med hva forkortelsene og terminologien betyr og hadde ferdigheter i den tekniske utførelsen av trinn og den praktiske gjennomføringen av de ulike turene.

Bestanddelene av trinn og turer i engelskdanser var av et visst repertoar, og når man kjente til dansens turer og oppbygning og progresjon, kunne man ta et raskt blick på dansebeskrivelsene og sannsynligvis nokså raskt være i stand til å gjennomføre dansen, til og med ta på seg å være det ledende paret som demonstrerte turer og trinn for de andre nedover i rekkene.

## Metodisk tilnærming

I doktorgradsarbeidet mitt brukte jeg metodiske tilnæringsmetoder som ivaretar forskeren som utøver. Jeg baserte metodeutviklingen på tre premisser av metodologisk art, nemlig hermeneutikk, taus kunnskap og praksisbasert forskning. Premissene ivaretar tolkning og meningsdanning i møte med historiske kilder, det

implisitte og udefinerbare i en dansers/dansekyndig forskers kunnskap og en anerkjennelse av praksis samt et kunstnerisk forhold til materialet som studeres i forskning på tidligere praksis[2, s. 31]. De metodologiske premissene ligger til grunn for mitt arbeid med kildeutgivelsen av *XII Engels Danse*, men en redegjørelse blir for omfattende for denne artikkelen. Jeg konsentrerer meg derfor om det mest sentrale perspektivet, nemlig hvordan mening kan trekkes ut av kildene ved hjelp av min interesse, bakgrunn og erfaring.

Når en dansebok som *XII Engels Danse* skal tolkes, oppstår en rekke spørsmål i møte med teksten, musikken og selve dokumentet som dokumentet i seg selv ikke kan gi svar på. En slik ufullstendighet kan sees på som manglende brikker, hull eller gap som må fylles med «noe» for å gi mening. En av de største utfordringene med forskning på historiske kilder for dans er at avstanden i tid er stor.

Dansebeskrivelsene er kortfattede oppramsinger av trinn og figurer som var etablert som kjent kroppslig praksis hos elever som deltok i danseundervisning. Repertoaret var godt kjent i overklassens selskapsliv. Målgruppen for dansebøkene var innforstått med hvilken praksis som lå bak ord og uttrykk. Den fortidige praksisen er imidlertid ikke tilgjengelig lenger. Dersom den implisitte praksisen skal forstås og forsøkes gjenskapt, må instruksjonene tolkes og rekonstrueres. Spørsmålene må besvares, og gapene må fylles. Eksempler på slike gap er når dansemønsteret ikke går opp, eller at uttrykk i teksten ikke gjenkjennes[2, s. 43-44].

## Mening

Historiker Knut Kjeldstadli sier at det ufullstendige i et kildemateriale må «suppleres ved at *vi* aktivt prøver å få dem til å gi mening. Videre er det å forske [...] å gjøre en serie valg». Han sier også at vi må «sammenfatte kilden og gjengi meningen med egne ord»[9, s. 40]. Det spesielle kroppslige aspektet ved å forske på fortidig dansepraksis er at deler av det å «gjengi meningen med egne ord» vil innebære en praktisk utprøving av bevegelse i kroppen. Den amerikanske danseforskeren Susan Foster sier: «Once the historian's body recognizes value and meaning in kinesthesia, it cannot dis-animate the physical action of past bodies it has begun to sense»[10, s. 7]. Når instruksjoner i kildematerialet forsøkes prøvd ut i praksis og materialet begynner å bli familiært i forskerens kropp blir utprøvningsprosessen en del av tolkningen og det å få informasjonen i kilden til å gi mening også som praksis.

Det er viktig å reflektere over hvilke svar en slik utforskningsprosess genererer, for tolkningen farges av forskerens egen kunnskap og erfaring. Slik sett vil enhver tolkning av en danseinstruksjon være unik – like unik som hver forsker er med sin kroppslige erfaring. I møte med instruksjonene i *XII Engels Danse* vil min erfaring som danser og forsker spille med på det bevisste og ubevisste plan. Innenfor feltet historiske danser er den vanligste metoden for gjenskaping av danser såkalt «rekonstruksjon». Rekonstruksjon innebærer å gjenskape danser fra tekstlige beskrivelser og dansenotasjon supplert med et så bredt omfang av så mange kontekstuelle kilder som mulig. Et bredt grunnlag legger til rette for mest mulig informerte og reflekterte valg.

For å finne ut så mye som mulig av hvordan dansene kan utføres i praksis, har jeg samlet og sammenstilt informasjon fra alle kilder jeg har hatt tilgjengelig.



Engelske, franske og tyske dansemanualer inneholder forklaringer av trinn og turer, og dansenotasjon viser mer grafisk hvordan danserne forholder seg til hverandre i rommet. Jeg har lest dansebeskrivelsene om og om igjen. Hver gang jeg har studert en kilde har jeg forstått mer og mer. Jeg har funnet «brikker» som gjør puslespillet mer komplett med færre gap.

I mitt arbeid med tolkningen av *XII Engels Danse* har jeg brukt erfaringen jeg har opparbeidet meg gjennom å utøve historiske danser i over 20 år, andres og egne tolkninger av dansebeskrivelser og andres og egne studier av dansebøker. Tilsammen danner det mitt brede – og mitt unike – grunnlag for å gi min (unike) tolkning av det implisitte innholdet i kilden.

### **Kildeutgivelse med praksis**

Kildeutgivelsen inneholder transkripsjon av tekst med forklaringer i noter og en innledning om heftet, innholdet og kontekst for dansene. Bue har transkribert melodiene og skrevet en innledning om musikken. Melodiene er gjort tilgjengelige for å kunne leses og spilles.

Men hvordan kan en kildeutgivelse inkludere tekstens implisitte dansepraksis? En mulighet som vil kunne bidra til at andre kan danse dansene er å lage mer utfyllende danseinstruksjoner, forklare alle dansens bestanddeler, trinn, figurer og turer ut fra den brede kjennskapen jeg har til andre historiske kilder for dans og den dype kunnskapen jeg har i å studere og rekonstruere historiske danser. Andre har tidligere gjort slike bearbeidinger, blant annet Reidar Warne, som har publisert en dansebok som finnes avfotografert ved Nasjonalbiblioteket, nemlig Johannes Bruus dansebok *Nogle Engelske Dantzer med sinne Tourer udi* datert 1777[11].

En annen mulighet er å inkludere lenker til videomateriale med demonstrasjoner av trinn og figurer samt dansene i sin helhet. På samme måte kan innspilt musikk akkompagnere en slik utprøving av materialet i praksis. Musikken inneholder informasjon som veileder danseutførelsen med tanke på rytme, varighet av trinn, tempo, aksenter og disposisjon av trinn innenfor takten, for eksempel. Et slikt tilleggsmateriale vil være nyttig som veiledning for lesere som ønsker å prøve ut dansene. Men det vil ligge en fare ved at et slikt demonstrasjonsmateriale kan komme til å låse innholdet i dokumentet til nettopp de tilgjengelige tolkningene. Det vil være nødvendig med refleksjoner rundt det faktum at demonstrasjoner av denne typen alltid vil være preget av tolkerens bakgrunn og tiden de er kommet i stand i. Tidens ånd og estetikk vil synes i tolkningen. Et dansemateriale tolket på 1980-tallet vil fortone seg annerledes når det tolkes i 2024 selv om begge er tolkninger av samme dansebeskrivelse.

Tilrettelegginger med utfyllende beskrivelser, videodemonstrasjoner og innspilt musikk ville vært et omfattende arbeid som krever dansere og musikere til å innstudere og fremføre dansene samt personer til å produsere filmene. Et slikt arbeid vil kreve betydelige midler, og jeg vil arbeide for at det kan realiseres i fremtiden.

## Noen avsluttende bemerkninger

En kildeutgivelse av en dansebok – hvilken relevans har det? Dans på 1700-tallet var ikke kun en forventet ferdighet for å kunne delta på ball. Dans var en omgangsform og en viktig del av dannelsen til barn og unge i velstående familier. Danseforskning og informasjon knyttet til en dansebok kan derfor fortelle mye utover den konkrete informasjonen dokumentet inneholder. Studier av fortidig dans kan si noe om datidens estetikk, måter å oppføre seg på og omgang mellom mennesker. Særlig engelskdansene var en danseform der samfunnets rangordninger ble opphevet for en stund idet alle par danser med alle par. Dans og ball var en arena for å møte en tilkommende, og en god danser var høyt ansett og beundret. Forskning på dans i selskapslivet kan vise nettverk og mobilitet og belyse andre forhold i samfunnet som ellers ikke ville komme frem på samme måte.

Det er nødvendig med en kildeutgivelse fordi kilden i seg selv er vanskelig tilgjengelig. Det fysiske manuskriptet ligger bortgjemt i arkivet, og selv om manuskriptet er tilgjengelig i skannet versjon i Nasjonalbibliotekets nettbibliotek<sup>6</sup> og kan studeres online, er selve innholdet ikke umiddelbart forståelig for oss i dag. Det er nødvendig med en fortolker som formidler innholdet gjennom kildeutgivelsen.

Kildeutgivelsen gjør også at manuskriptet, innholdet og kildens konteksts slik den er utarbeidet i utgivelsens forfattertekster blir søkbare og mer synlige.

Tilgjengeliggjøring av historiske kilder for dans bidrar også til å gjøre tidlig norsk dansehistorie mer kjent. Forhåpentligvis vil kildeutgivelsen også bidra til å dementere påstanden «Norge har en kort dansehistorie». Jeg har tatt til orde for at utsagnet, som stadig repeteres, er basert på et feilaktig premiss, nemlig at vår Nasjonalballett, som ble etablert i 1958, er ung sammenlignet med de nasjonale ballettkompaniene i Sverige og Danmark. Påstanden ekskluderer blant annet den sosiale danseformen engelskdanser og viktige kilder som *XII Engels Danse* fra norsk dansehistorie, og jeg har tatt til orde for at feltet må arbeide for tilbakevise påstanden og heller løfte frem dansepraksiser og kilder fra vår lange dansehistorie[12].

Dessuten vil en kildeutgivelse gjøre innholdet og den tilhørende konteksten tilgjengelig for personer som har lyst til å prøve ut praksisen selv. Dansene er morsomme, og 1700-tallets sosiabilitet kommer frem og nærmest synliggjør seg selv når vi trer inn i bevegelsesmaterialet. Jeg har mange ganger sett hvordan en gruppe mennesker, alt fra en skoleklasse til Den norske opera og balletts sponsorer, kan fornøye seg i engelskdans. De hilser høflig på hverandre, ser på hverandre, tar hånden og danser elegant med hverandre. I en slik praksis kan engelskdansens tidløshet komme klart frem i deltagerens fascinasjon over opplevelsen.

Det er mitt ønske at kildeutgivelsen, i tillegg til å vise fram selve dokumentet og innholdet, kan oppmuntre til bruk av engelskdansene. Dansene bør komme til sin rett i våre nåtidige kropp og ikke ligge bortgjemt i esker i arkivet.

---

<sup>6</sup>*XII Engels Danse*. Mus.ms. 8634:657. Nasjonalbiblioteket. [https://urn.nb.no/URN:NBN:no-nb\\_digimanus\\_249032](https://urn.nb.no/URN:NBN:no-nb_digimanus_249032).

# Bibliografi

- [1] Margrethe Støkken Bue and Elizabeth og Margrethe Støkken Bue Svarstad, eds. *XII Engels Danse. NB kilder 17*. Oslo: Nasjonalbiblioteket, 2024. URL: [www.bokselvskap.no/boker/dansebok..](http://www.bokselvskap.no/boker/dansebok..)
- [2] Elizabeth Svarstad. “’Aqquratesse i alt af Dands og Triin og Opførsel.’ Dans som sosial dannelse i Norge 1750–1820”. Doktoravhandling. Trondheim: Norges teknisk-naturvitenskapelige universitet, 2017.
- [3] Elizabeth Svarstad and Margrethe Støkken Bue. “Koreografi – dansekunst som tekst”. In: *Kildeudgivelse og tekstkritik i forskjellige traditioner*. Ed. by Annette Lassen, Krista Stinne Greve Rasmussen, and Kirsten Vad. NNE Skrifter 14. Odense: Syddansk Universitetsforlag, 2022.
- [4] Fin Michalsen. *Hs. 8634:657*. Maskinskrevet. Jan. 12, 1972.
- [5] John Playford. *The English Dancing Master or, Plaine and Easie Rules for the Dancing of Country Dances, with the Tune to Each Dance*. Ed. by Hugh Mellor and Leslie Bridgewater. London: Dance Books Ltd, 1984.
- [6] *Historic Dance and Tune Books*. Vaughan Williams Memorial Library. May 7, 2023. URL: [https://www.vwml.org/browse/browse-collections-dance-tune-books/..](https://www.vwml.org/browse/browse-collections-dance-tune-books/)
- [7] André Lorin. *Livre de Contredance*. Français: Bibliotheque nationale de France, Département des manuscrits, 1697;1698.
- [8] André Lorin. *Livre de la contredance du roy. Présenté à Louis XIV (1688) et retranscrit pour Louis XV (1721)*. Français: Bibliotheque nationale de France, Département des manuscrits, 1697–1698.
- [9] Knut Kjeldstadli. *Fortida er ikke hva den engang var: En innføring i historiefaget*. Oslo: Universitetsforlaget, 1999.

- [10] Susan Leigh Foster. “Choreographing History”. In: *Choreographing History, Redigert Av Susan Leigh Foster*. Ed. by Susan Leigh Foster. Bloomington & Indianapolis: Indiana University Press, 1995.
- [11] Reidar Warne. *Nogle Engelske Dantzer med sinne Tourer udi. Udskrevet af Johannes Bruu udi Tønsberg Aar Anno 1777*. Norsk musikksamling Publikasjon 10. Oslo: Universitetsbibliotekets hustrykkeri, 1982.
- [12] Hanna Järvinen et al. “Choreographing Histories: Critical Perspectives on Dance Histories in Nordic Dance Practices and Scholarship”. In: *Nordic Journal of Dance* 14.1 (Aug. 10, 2023), pp. 31–47. DOI: 10.2478/njd-2023-0004. URL: <https://sciendo.com/article/10.2478/njd-2023-0004> (visited on 12/19/2024).
-





# Digitale udgaver som HITL-garanter. Om editionsfaglig domæneeksportberigelse og datakvalitetens betydning for integrationen af computationelle tilgange i humanvidenskaben

Katrine Frøkjær Baunvig, Krista Stinne Greve Rasmussen<sup>1</sup>

<sup>1</sup>Center for Grundtvigforskning | Aarhus universitet

## Indledning

*Less is more.* I hvert fald hvis den mindre delmængde er af høj kvalitet. Sådan burde en tommelfingerregel lyde i de grene af humanvidenskaben, som gør beregningsorienteret brug af digitale data. Denne opfattelse er det vores hovedærinde at underbygge i denne artikel, som angår vigtigheden af den editionsfaglige berigelse af de kulturarvsdata, som ligger til grund for den mængde af digitale, videnskabelige udgaver (Digital Scholarly Editions, DSE), der i de seneste årtier har vokset sig omfangsrige i Norden.<sup>1</sup> Følgelig ønsker vi at pøde det programmatisk akronym HITL (Human-In-The-Loop), som i de senere år i påtrængende grad har vokset sig relevant i fagfelter, der benytter sig af maskinlæring, på den editionsfaglige selvforståelse. 'HITL' anvendes om allehånde initiativer, hvor mennesker inddrages til at kvalitetssikre eller -forbedre data, der skal ligge til grund for en given algoritmisk model.[1] Det kan være til annoterings- eller kategoriseringsopgaver, som netop hører til de editionsfilologiske kerneopgaver: genreannotering, stedopmærkning eller personnavnsregistrering etc. I takt med at den Computationelle Humanioras virkefelt udvides, bliver pålidelige højkvalitetsdata stadig mere efterspurgt. Dette skærper relevansen af DSE-projekter og af de editionsfilologiske kompetencer, som sådanne projekter er baseret på.[2]

I det følgende vil vi med andre ord argumentere for, hvorledes klassiske editionsfilologiske fagdyder vinder ny relevans i en tid, hvor generativ AI og algoritmiske metoder i det hele taget sætter dagsordenen. I denne kontekst bliver humanvidenskabelig akkuratess og kontekstforståelse attraktiv i alle vidensområder baseret på sprogteknologi og store sprogmodeller (Large Language Models, LLM). Det gør de, fordi sprogmodellens anvendelighed og pålidelighed er en direkte funktion af kvaliteten af de data, som de er trænet på. Som leverandører af kulturarvsdata af høj kvalitet forudser vi, at den editionsfilologiske fagstand vil indtræde som hjørnesteinen i fremtidens AI-baserede vidensøkologi. I denne artikel vil vi imidlertid begrænse os til en illustration fra forskningsverdenen: vi vil give eksempler på, hvorledes den digitale editionsfaglige berigelse af ét

---

<sup>1</sup>Patrick Sahle redigerer et katalog over DSE's, hvorfra man fremsøge nordiske udgaver: [https://www.digitale-edition.de/exist/apps/editions-browser/\\$app/index.html](https://www.digitale-edition.de/exist/apps/editions-browser/$app/index.html). Kataloget viser 18 nordiske udgaver (fordelt på 6 danske, 5 finske, 3 norske, 3 svenske og 1 islandsk), men kataloget er ufuldstændigt. Eksempelvis mangler de danske udgaver *H.C. Andersens Eventyr og Historier. Den digitale manuskriptudgave* (2020 f.), <https://hcams.andersen.sdu.dk/exist/apps/andersen/index.html>, og *Diplomatarium Danicum* (afsluttet), <https://diplomatarium.dk/>.

væsentligt dansk kulturarvsdatasæt *Grundtvigs Værker* har muliggjort computationel udforskning, som for sin del har kastet afgørende, nye forskningsresultater af sig. Med afsæt i en skitsering af 1) filologiens fundamentale position i humanioras udviklingshistorie med særinteresse for 2) editionsfilologiens placering i den humanvidenskabelige fagfamilie og forskningshistorie vil vi derfor i det følgende 3) beskrive de digitale editionsfaglige projekters nøglerolle i produktionen af højkvalitetskulturarvsdata for 4) at illustrere, hvorledes dette arbejde konkret baner vej for kvalitativt nye former for forskningsresultater.

## **2. Filologien er hjørnестenen i humanvidenskaberne og den Computationelle Humaniora**

Få vil modsige, at filologien er udgangspunkt og centrum for humanvidenskaberne, som vi kender dem i dag. Filologien som begreb har været med til at forme udforskningen af mennesket med fokus på sprog og tekst siden Platon.[3]

Videnskabelig udgivelse er en helt særlig form for filologi, der har eksisteret stort set lige så længe. Formålet har altid været at gøre tekster tilgængelige på tværs af tid og sikre deres videre eksistens. Metoderne og de videnskabelige principper har varieret, ligesom man nu kan tale om forskellige filologiske grene afhængig af materialet, der gøres tilgængeligt. Den videnskabelige udgivelsesvirksomhed har i nyere tid desuden forandret sig fra at have været betragtet som en slags hjælpedisciplin for de øvrige videnskaber til at udgøre en selvstændig videnskabelig disciplin med et betydeligt teoretisk opsving i slutningen af det 20. århundrede.[4, s. 33] Videnskabelige udgaver har været en central del af nationsopbygningen, og de har inden for alle grene af humaniora tilvejebragt og formidlet vores tekstbårne historie.

Fokus i denne artikel er editionsfilologien, en gren af udgivelsesfilologien, der hovedsageligt beskæftiger sig med nysprogligt materiale fra efter bogtrykkets indførelse. Johnny Kondrup inddeler feltet i fem genstandsområder: «tekstkritik, variantapparater, dokumentbeskrivelse (bibliografi og manuskriptbeskrivelse), kommentering og elektronisk udgivelse».[5, s. 17] Inddelingen kan diskuteres, hvilket Kondrup også gør opmærksom på, men desuagtet er det passende, at tekstkritik nævnes som det første punkt, da det står som det helt centrale element i videnskabelig udgivelse, jf. også ovenfor. Overordnet set beskrives tekstkritikken «lämpligen som en samling verksamheter som utvecklats sedan antiken för att på metodisk grund välja bättre läsarter framför andra, när det föreligger textuell variation och/eller misstanke om korruption i det transponerade textmaterialet.»[6, s. 129-130] Bedre læsemåder skal ikke forstås æstetisk, men snarere hermeneutisk. Når tekster vandrer fra en medialisering til en anden, altså transponeres, kan der opstå fejl. Det kan både være egentlige fejl i teksten, f.eks. ord der ændrer form, men det kan også være materiel korrumpning. Afhængig af videnskabelig metode, tekstopfattelse og retningslinjer for transponering kan betydningen af disse korrumpninger vægtes forskelligt. Eksempelvis kan en transponering lægge stor vægt på den grafiske form og derfor søge at gengive de grafiske elementer så eksakt som muligt.

Centralt er opmærksomheden på, at tekster aldrig bare er tekster – og ikke mindst, at de aldrig kan gengives uden informationstab. Derudover kan tekster



også i sig selv indeholde fejl. «Det er et sundt filologisk (og hermeneutisk) princip at gå ud fra, at enhver konkret foreliggende tekst – en realtekst – indeholder fejl.»[5, s. 139] Realtekster er i denne terminologi teksten, som den fremtræder for os, hvorimod idealkteksten er en idealisering. Emendationer uden et foreliggende tekstvidne med en variant gøres med udgangspunkt i forestillingen om en idealktekst. En længere teoretisk diskussion af tekstbegrebet ligger uden for denne artikels ramme. Derimod er det vigtigt at understrege, at videnskabelig tekstgengivelse har indbyggede teoretiske og metodiske refleksioner, der både er historisk foranderlige og – hvad vigtigere er – videnskabeligt forankrede, men altid afhængige af det menneske, der udgiver teksten.[5, s. 140][7, s. 20] Ingen computer kan forventeligt foretage komplekse emendationer, som når vi i Grundtvigs Værker har rettet «Almagt» til «Afmagt» i Grundtvigs artikel «Om Religion og Liturgie».<sup>2</sup> Derimod kan man forvente, at man med AI kan træne en model til at rette mere trivielle tekstrettelser, for eksempel omvendte bogstaver. Mennesker må altså komplementere maskiner for retteligt at udøve den tekstkritik, der ligger til grund for kvalificeret reproduktion og -distribution af tekster.

En central af det videnskabelige udgivelsesarbejde, som på sin vis falder inden for tekstkritikken, men går forud for denne, er selve tilvejebringelsen af tekster ved hjælp af scanning og tekstgenkendelse. I Grundtvigs Værker kollationerer vi den digitale version af grundteksten tre gange, både for at identificere varians, men i særdeleshed også for at genoprette den automatiske tekstgenkendelse, så vores digitale tekster er i overensstemmelse med det trykte forlæg. Ingen automatisk tekstgenkendelse er endnu fejlfri, og selvom det kan virke som et trivielt skridt at kollationere, bliver denne del af tekstkritikken først rigtig kvalificeret i kraft af editionsfilologen, der også kan identificere eventuelle tekstfejl uden forlæg. Nøjagtigheden af genkendelse er dog ikke kun relevant for den kvalificerede tekstgengivelse, den er også afgørende for den potentielle brug af teksterne: «The accuracy of Optical Character Recognition (OCR) technologies considerably impacts the way digital documents are indexed, consulted and exploited.»[10] Det har været almindeligt at opfatte en vis mængde støj som irrelevant for udnyttelsen af tekster inden for den Computationselle Humaniora, eksempelvis til fjernlæsning ('distant reading'), men ved at acceptere støj risikerer man at lukke af for visse typer forskningsspørgsmål og -interesser. Selvom et korpora fuld af støj statistisk set kan være nok for indeværende, er det ikke nødvendigvis også sandt i fremtiden. En bedre OCR-proces kan ikke blot have betydning for den tekstnære forskning, men også påvirke læsere til en mere positiv, og dermed brugbar, opfattelse af materialet.[11] Således handler opretning af OCR-fejl ikke kun om materialet selv, men om hele den vidensøkonomi og -økologi teksterne indgår i.

Af editionsfilologiens fem genstandsområder, som blev omtalt tidligere, skal yderligere to behandles her, nemlig dokumentbeskrivelse og kommentering. Førstnævnte skal ideelt set gøre teksten nærværende for brugeren, men de færreste har nok den oplevelse, når de eksempelvis læser en bibliografisk formel eller en manuskriptbeskrivelse. Derudover vil mange, der lægger det fysiske dokumenter til grund for studier, behøve at se og arbejde med dokumentet selv.

---

<sup>2</sup>Eksemplet blev først brugt i en forelæsning af Johnny Kondrup på Københavns Universitet. [8, s. 146] Se også [9]

Dokumentbeskrivelsen kan med andre ord ikke erstatte det fysiske eksemplar. Alligevel indsamler dokumentbeskrivelsen værdifuld metadata, der kan bruges i større digitale udforskninger af materialesamlinger. For *Grundtvigs Værkers* vedkommende kunne dette være oplysninger om tidsskrifters størrelser (fysisk og oplag) og redaktører. Dokumentbeskrivelser er derfor ikke blot beskrivende, men også berigende i forhold til det digitale korpus en videnskabelig udgave også er.

Endnu mere berigende er kommenteringen, som vi vil demonstrere med eksempler fra forskningsstudier nedenfor. I *Grundtvigs Værker* har vi to typer kommentarer: overgribende kommentarer og punktkommentarer. Den overgribende er indledningerne og tekstredegørelserne, mens punktkommentarerne består i verbalkommentarer (der udarbejdes lokalt for hvert værk) og registeropmærkninger (der er globale i udgaven, dvs. identisk gentaget overalt for samme lemma). I registeropmærkningerne findes når personer, steder, titler, mytologiske entiteter og bibelreferencer nævnes i teksterne. Med et specifikt id-nummer henvises til en central fil med globale kommentarer, der vises lokalt i tekststedet. Dette gør redigering af posterne lettere og sparer tid ved størrelser, der allerede findes forklaringer på. Kommentering er, ved siden af tekstkritikken, editionsfilologiens kerneopgave, der i den videnskabelige udgave har til opgave at bygge bro over den historiske kløft mellem tekster og læsere, men som også i forskningen kan udnyttes til computationelle studier. Vi vil nu se nærmere på editionsfilologiens rolle mere bredt i etableringen og udforskningen af digital kulturarvsdata.

### 3. DSE som HITL i produktionen af høj kvalitetskulturarvsdata

2002 anslås ofte som begyndelsepunktet for den såkaldt digitale æra. I årene umiddelbart op til og umiddelbart efter 2010 udmøntede det sig i begrebet *Big Data*. Herom vidner søgninger på Google Trends, i forskningslitteraturlister og på platforme, der tilgængeliggør offentlige mediedata. Begrebet voksede sig til i datalogiske fagmiljøer og dækker over datasæt, hvis størrelse og kompleksitet omkalfatrede traditionelle metoder til indsamling, opbevaring, behandling og fortolkning af data. Et omsiggribende internet førte til en betydelig forandring af målestokken for tilgængelige data; i disse år gik man fra at tælle data i gigabytes (10<sup>9</sup>) til at opgøre dem i petabytes (10<sup>15</sup>). Udviklingen i størrelse ledsagedes af en udvidelse af datamaterialets variation samt en drastisk forøgelse af behandlingshastigheden. Det oplevedes som et væsentligt faseskifte.

Begejstring og videnskabelig optimisme omgærdede denne oplevelse og begrebet spredte sig nu på vingerne af ånden i den irske forfatter Oscar Wildes udbredte *bon mot*: «Nothing succeeds like excess». Skønt begrebet Big Data ikke har nogen formel definition voksede det med andre ord frem som et løfte. De nye rammer for datamængder, datavariation og databehandlingshastighed forventede man ville vise sig som løsningen på en broget mængde opgaver – af såvel videnskabelig som industriel, samfundsmæssig og kulturel natur – man ikke før var i stand til at tage sig af. Selvom man stadig kan støde på annoncer, der holder fast i den begejstrede tone og eksempelvis ønsker at redegøre for «How Big Data Can Help You Do Wonders In Your Business»,<sup>[12]</sup> er den generelle entusiasme aftaget og begrebet afmattet. Det skyldes først og fremmest, at forventningerne til de afgørende nybrud er blevet

omdirigeret til kunstig intelligensteknologien – ikke mindst til den generative AI.

Denne drejning er også et resultat af, at de massive datamængder, som Big Data-hypen tog afsæt i, var kendetegnet ved en anseelig grad af rod eller skidt. Som tiden gik, stod det altså klart, at Big Data i et vist omfang var synonymt med *Dirty Data*. Det vil sige, at mega-repositorierne som hovedregel indeholdt ufuldstændige, upræcise eller ligefrem ukorrekte data. I første omgang var den generelle holdning, at dataomfanget i sig selv udlignede skidtet eller at en automatiseret rensning kunne sikre dataintegriteten. Men gradvist stod det klart, at dette ikke var tilfældet. Og i en lang række faglige og industrielle sammenhænge er det efterhånden blevet uomtvisteligt, at pålidelige og bæredygtige data kræver menneskeligt tilsyn og håndtering.

Denne indsigt har fremelsket akronymet, der gemmer sig i denne artikels overskrift: HITL eller Human-In-The-Loop. Det anvendes i en righoldig mængde kontekster, hvor automatisering har vist sig at skabe praktiske såvel som etiske udfordringer. HITL har tilstrækkelig spændvidde til også at omfatte det editionsfaglige bidrag til produktionen og vedligeholdelsen af højkvalitetskulturarvsdata. Editionsfilologer indtager i kraft af deres møjsommelige arbejde med at tilvejebringe og kuratere digitale udgaver rollen som humans-in-the-loop – som den menneskelige valideringsinstans i den digitale kulturarvsbevarings kredsløb. I den forstand bør de opfattes som det første led i den kæde, som er den fremvoksende computationelle humaniora består af. Uden menneskelig fejlsikring, opmærkning eller anden berigelse af de relevante datasæt, er det ikke rimeligt at stille disse datasæt hermeneutisk komplekse spørgsmål. Med andre ord: uden denne sikring af en given datasamlings kvalitet vil udsagnskraften af en beregningsbaseret tekstanalyse, som bygger herpå, være begrænset.

Uden at fortabe os i spørgsmålet er det på sin plads at fremdrage en oplagt og uheldig konsekvens af disse omstændigheder: Menneskelig validering af et datasæt er en omkostningsrig affære – DSE-projekter kræver mange ressourcer i form af tid og penge. Offentlige og private fonde vil derfor naturligt nok være særdeles kritiske med, hvilke datasæt, de ønsker at understøtte berigelsen af. I praksis betyder det, at det gerne er i forvejen kendt og anerkendt materiale, som modtager støtte. Med andre ord flyder der en konservativ kanon- eller kulturreificerende tendens ud af kravet om HITL og højkvalitetsdata. For at skære situationen hård op: det er oftere projekter, som tager afsæt i kulturbærende figurer som N.F.S. Grundtvig, Zacharias Topelius, Henrik Ibsen, August Strindberg, der modtager støtte end projekter, der kredser om ukendte arbejderkvinders efterladte papirer. Logikken, der driver dette forhold frem, er forståelig; vi håber ikke desto mindre, at fremtidige DSE-ansøgere såvel som DSE-bevillingsgivere vil forsøge at modvirke de åbenlyse misforhold i kulturarvsbevaringen, som følger af tankerefleksion og bevillingsimpulsen. Derfor har vi på Center for Grundtvigforskning søgt at udvide vores dataetableringspligt fra Grundtvigs forfatterskab til hele den brogede og mindre prominente brugskultur, som forfatterskabet satte i gang. Ét håndgribeligt resultat af disse bestræbelser er et af *Grundtvigs Værkers* ledsageprojekter: en højkvalitetsdigitalisering af det grundtvigsk sindede tidsskrift *Højskolebladet* (1876-2020), som vi i 2021 modtog midler fra Carlsberg Fondet til at digitalisere. I nær fremtid vil korpuset være at finde på Det Kongelige Biblioteks side [tidsskrift.dk](http://tidsskrift.dk).

Begge samlinger er høj kvalitetsmateriale, der muliggør udforskning af Grundtvigs liv, værk og efterliv på måder, som hidtil ikke har kunnet lade sig gøre. Det er eksempler på denne udforskning og dens brug og afhængighed af en DSE, som vi nu skal give et indblik i. Vi har begrænset os til to eksempler. Det første angår et tvillingestudie af Grundtvigs emfasetilskrivninger ved hjælp af henholdsvis spatiering og udråbstegn; det andet angår en begrebsmæssig efterprøvning af Grundtvigs farvebrug.

#### 4. Betydningen af DSE for den computationelle humanvidenskab: Eller om udråbstegn og farvebegreber i *Grundtvigs Værker*

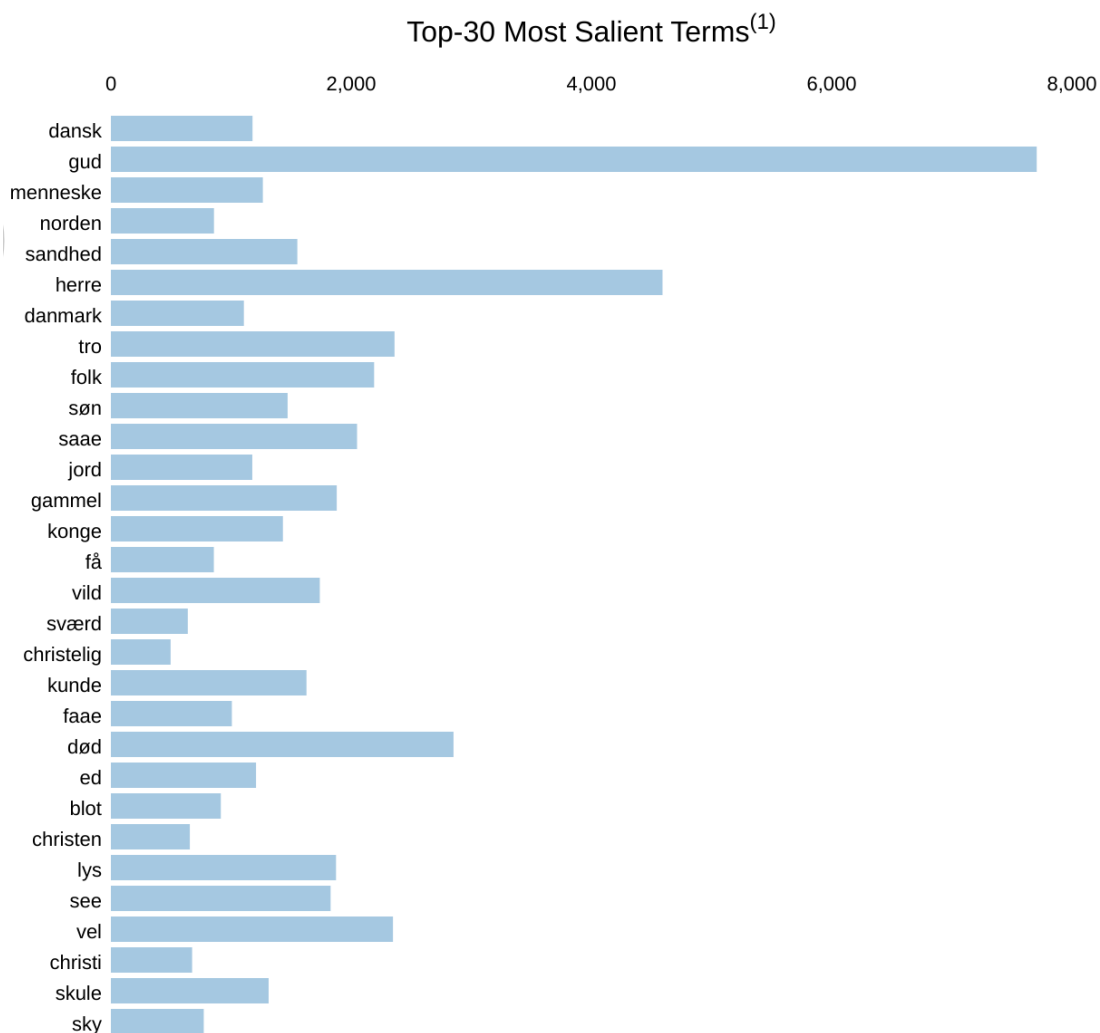
Det første demonstrationseksempel vi har valgt er metodisk simpelt og bygger først og fremmest på to specifikke træk ved *Grundtvigs Værkers* opmærkning (emfase-studiet). Det andet eksempel baserer sig på en matematisk mere kompleks fremgangsmåde – såkaldt simple neurale netværk eller *word embeddings* – der for sin del inddrager alle niveauer af opmærkningsarbejdet (farve-studiet). Der er med andre ord tale om to særdeles DSE-afhængige computationelle studier omend afhængigheden gestalter sig på meget forskelligartet vis. Begge studier har bragt Grundtvigforskningen til klarhed over hver sit felt, hver sit spørgsmål. Emfase-studiet har blotlagt, hvad der var de semantiske tyngdepunkter i forfatterskabet, mens farve-studiet har efterprøvet, om Grundtvig selv efterlevede sin berømte regel: «Med Farver kan man ingen Christus male».[13, s. 171] Begge undersøgelser har forlenet forskningen med ikke-trivielle indsigter og beriget vores forståelse af Grundtvigs tankeverden.

##### 4.1 De kristne eftertryk og de danske betoning

De to artikler «Emotional Imprints: Letter-Spacing in N.F.S. Grundtvig's Writings» og «Emotional Imprints: Exclamation Marks in N.F.S. Grundtvig's Writings»[14][15] tager afsæt i den indvending, man undertiden støder på i Grundtvig-forskningen og -formidlingen, at Grundtvigs kernesager ikke flugter med receptionen af ham som kirkefader og landsfader[16, s. 92-99] – at eftertiden ikke har begrebet ham og hans ærinde. Tvillingeundersøgelserne tillader sig at tage denne påstand på ordet ved at teste den. De tager desuden begge afsæt i to forhold: 1) det digitale, opmærkede Grundtvig-materiale og 2) den præmis, at Grundtvig var påvirket af en romantisk 'emotionaliseringsstrømning', der var kendetegnet ved en bestræbelse på at mætte de fleste udtryksformer – også tryksager – med følelser. Emfasetilskrivning som *s p a t i e r i n g*, der som fremhævelsesmiddel var frakturens parallel til antikvaens kursivering, og udråbstegn var de mest almindelige redskaber hertil – redskaber, som Grundtvig var en endog særdeles flittig bruger af.

Spatieringerne er konsistent opmærket i hele Grundtvigs Værker, hvilket reducerede den tekniske del af undersøgelsen til et simpelt udtræk. Vi undersøgte distributionen over tid og muligheden for påvirkning fra skiftende forlæggere. Begge dele syntes dog ikke at vise nogen påfaldende endsige signifikante tendenser. Fordelingen viser med andre ord ikke noget klart mønster. Det samme gør sig gældende for Grundtvigs brug af udråbstegn, som også meget simpelt lod sig

udtrække. Hvad der imidlertid tegner et klart mønster for både spatieringer og udråbstegn, er de termer, som emfaseredskaberne knytter sig til. Grundtvig understreger og udråber på den ene side ord, der knytter sig tæt til Danmark og til kristendommen – til 'dansk' samt til 'Jesus' og 'Kristus'. Denne påstand underbygges af nedenstående tabeller.



Figur 1. Grundtvigs 30 mest brugte ord i sætninger, der munder ud i udråbstegn.

I figur 1 ser vi de ord, som hyppigst indgår i sætninger, der munder ud i et udråbstegn. I figur 2 ser vi listen over de hyppigst spatierede termer. Det fremgår heraf, at der er to tyngdefelter i Grundtvigs tænkning. At Grundtvig sidenhen forbindes med tanker om danskhed og om dansk kristendom må på den baggrund siges at være i overensstemmelse med det indlejrede budskab i værkerne.

#### 4.2 Grundtvigs farvebrug

Grundtvig var billedskeptisk, men han var samtidig overmåde farveglad. Det er udgangspunktet for en undersøgelse, der er grundlaget for artiklen «Med Farver kan man ingen Christus male». Joakim Skovgaard og Grundtvigs farver».[17]

Artiklen forsøger at besvare, om den farveglade Grundtvig ligefrem holder Kristus fri af farverne i sine sprogbilleder, når han nu er så forhuppet på at oplyse sine kunstnerkolleger om, at billedkunst og kristendom ikke har noget med hinanden at skaffe. Til forskel fra de to forrige undersøgelser gør denne undersøgelse brug af stor regnekraft og er baseret på en algoritme trænet på det samlede digitaliserede korpus bestående af de over 8 millioner ord fordelt på over 1000 værker, som Grundtvig udgav i løbet af en periode på 68 år (1804-1872). På samme måde, som et landkort komprimerer og forenkler information fra et virkeligt landskab,[18] kan Grundtvig-algoritmen forstås som en sammenpresning og forenkling af det komplekse og multidimensionale rum, Grundtvigs forfatterskab og ordforråd repræsenterer. Både kort og algoritme kan betragtes som navigationsudstyr: begge muliggør orientering i et flerdimensionalt rum. I denne sammenhæng sætter algoritmen forskere i stand til at udarbejde såkaldte simple neurale netværk kaldet word embeddings, som svarer til kortlægningen af et ords associationsstruktur i et bestemt tekstkorpus.[19][20] Denne fremgangsmåde hjælper i denne sammenhæng til et overblik over de farver, hvormed Grundtvig i sine sprogbilleder fremmaler sin forståelse af verden, af naturen, af folkefærd. Den editionsfaglige møjsommelige oprensning og opmærkning af Grundtvig-materialet er nødvendig, så der ikke opstår fejl eller støj i kompressionen eller disproportioner på associationskortet. Undersøgelsen af Grundtvigs farvebegreber har hjulpet til nye indsigter i forskningen.

Denne visualisering af en *word embedding* viser, hvorledes 'Christus' forholder sig til serien af *blaa, grøn, gul, rød, brun, sort, hvid, guld* og *sølv*. Her ser vi tydeligt, at 'Christus' ikke danner betydningsforbindelser med de jordiske farver. I stedet lægger han sig tilrette på i en akavet treenighed, hvor han har fået føljeskab af med sin egen jordiske modpart Jesus samt Vorherre. Denne treenighed er 'akavet', fordi Helligånden mangler. Den står ellers centralt i Grundtvigs tænkning.[21] At den er helt fraværende fra hans farveassociationer er derfor bemærkelsesværdigt. Visualiseringen hjælper til at vise, at Kristus og særligt Helligånden er uden for farvernes rækkevidde. Det er de antageligvis, fordi de begge to først og fremmest er lydfænomener – Grundtvig fremstiller ligefrem ånden som en klangkuppel, hvor tonerne fra den himmelske oven og fra fortiden runger og resonnerer.[21] Når Grundtvig skal beskrive den guddommelige treenighed, styres han med andre ord af en akustisk logik – der fx er knyttet til Johannesevangeliet 1,14 og forståelsen af Jesus som «Ordet», der «blev kød». Og det er givetvis derfor han oftere slår ud i de larmende sprogbilleder fremfor de farverige, når han skal beskrive dem. Denne indsigt, som er faciliteret af Grundtvig-algoritmen, er ikke før blevet ført så håndfast igennem.

## 5. Afslutning

Vi er overbeviste om, at editionsfilologien har en afgørende og central rolle at spille i humaniora generelt og den computationelle humaniora specifikt. Editionsfilologiske kompetencer er værdifulde i den digitale transformation af tekster, der har stået på længe, men som nu ikke længere er en sær eksistens for tekster. Nu er alle tekster digitale.

Studierne vi henviser til i denne artikel, er mulige at foretage på tekst, der blot er

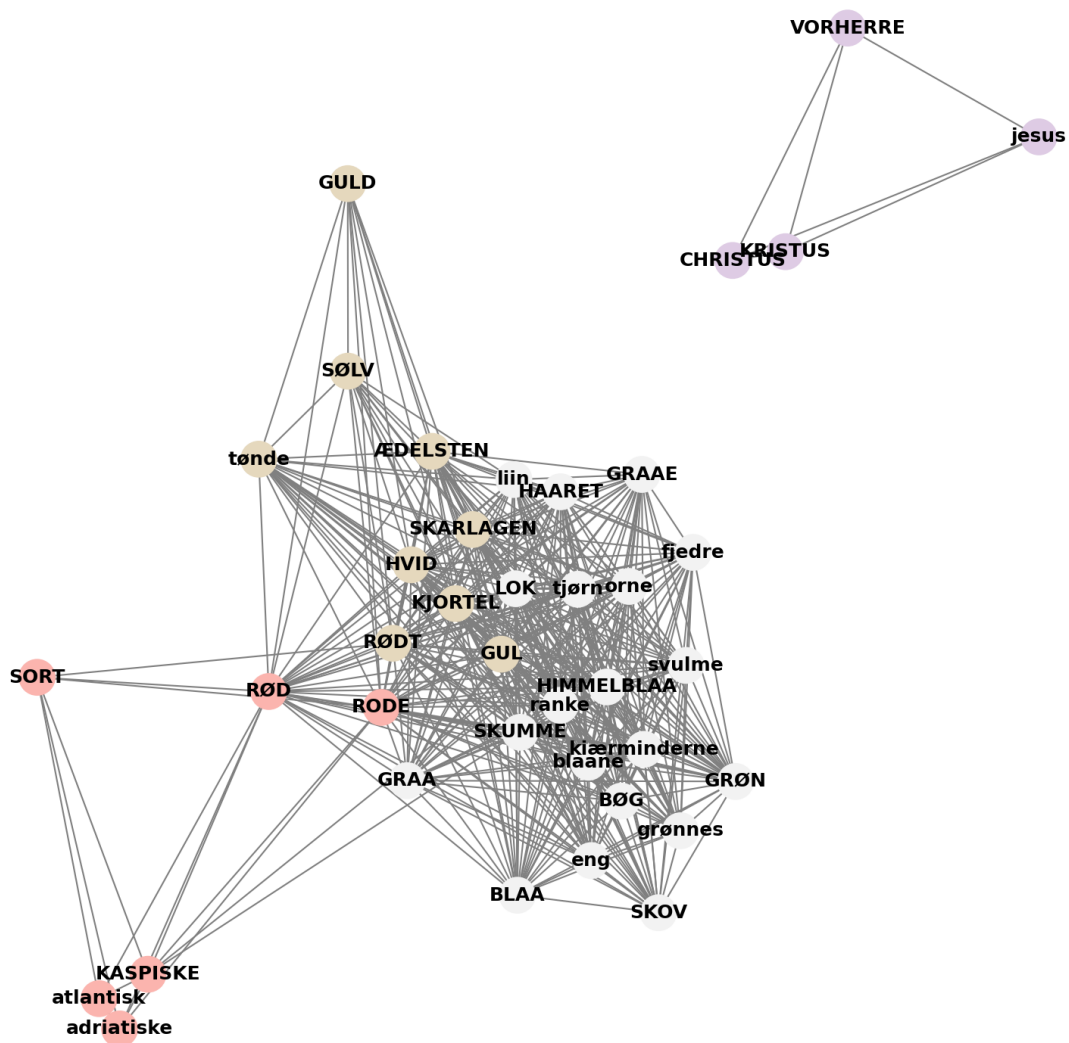
OCR-behandlet, men pålideligheden af undersøgelsernes resultater afhænger af, at teksterne er blevet rensede, dvs. har været igennem menneskelige hænder. Ligeledes kan de her nævnte studiers brug af *word embeddings* i princippet gennemføres uden de videnskabelige udgavers opmærkede tekster, men dette finder vi ikke attraktivt. Det er vigtigt, at materialet har høj standard. Eksempelvis blev emfasestudiet gennemført på et tidspunkt, hvor blot 40 % af GV-korpusset var rensat og opmærket, hvilket giver fejl som lemmaet «få» i figur 1, der rettelig burde være «faa». Jo længere vi er med udgaven, jo bedre bliver korpus, og jo bedre fungerer modellerne. Principielt kunne man til disse formål argumentere for, at et rensat OCR-korpus var rigeligt, men så ville studierne ikke kunne gå videre end dette. Aktuelt arbejder vi med at træne taggers, der blandt andet muliggør nye former for automatiseret opmærkning, men – og dette er en vigtig pointe – nye studier, hvis metoder vi endnu ikke kender. Muligheder vi opfatter som betinget af editionsfilologisk ekspertise og arbejde.

**Table 2.** 20 Most Frequent Letter-Spaced Words.

	L-S Words	L-S Words (English)	Total Amount of Occurrences	First Occurrence in Text
1	Danmark	Denmark	1106	78
2	Danske	Danish	1105	51
3	ikke	not	735	16
4	Jesus	Jesus	601	62
5	Daaben	The baptism	463	78
6	dansk	Danish	451	47
7	Jesus Christus	Jesus Christ	449	82
8	Damerne	The ladies	444	5
9	Ordet	The word	434	80
10	Odin	Odin	422	13
11	levende	living	415	49
12	Aanden	The spirit	415	57
13	det	it	414	15
14	Norden	The North	401	27
15	den	it	383	18
16	Danmarks	Denmark's	380	43
17	Frihed	Freedom	380	69
18	Thor	Thor	372	21
19	Zeus	Zeus	337	5
20	Jesu	Jesus'	328	28

Figur 2. Grundtvigs 20 hyppigst spatierede ord.





Figur 3. Det semantiske væv, der binder termerne 'Christus' og blaa, grøn, gul, rød, brun, sort, hvid, guld samt sølv sammen i Grundtvigs publicerede forfatterskab.

# Bibliografi

- [1] Vikram Singh Bisen. *What Is Human in the Loop Machine Learning: Why & How Used in AI?* VSINGHBISEN, May 20, 2020. URL: <https://medium.com/vsinghbisen/what-is-human-in-the-loop-machine-learning-why-how-used-in-ai-60c7b44eb2c0..>
- [2] Katrine Frøkjær Baunvig. “A Computational Future? Distant Reading in the Historical Study of Religion”. In: *Stepping Back and Looking Ahead: Twelve Years of Studying Religious Contact at the Käte Hamburger Kolleg Bochum*. Ed. by Maren Freudenberg et al. Leiden: Brill, 2023. DOI: 10.1163/9789004549319\_013.
- [3] James Turner. *Philology: The Forgotten Origins of the Modern Humanities*. Princeton: Princeton University Press, 2014. DOI: 10.1515/9781400850150..
- [4] Johnny Kondrup. “Indledning til Dansk Editionshistorie”. In: *Dansk Editionshistorie. Udgivelse af græsk og latinsk litteratur*. Ed. by Christian Troelsgård and David Bloch. Vol. 1. København: Museum Tusulanums Forlag, 2021, pp. 31–48.
- [5] Johnny Kondrup. *Editionsfilologi*. København: Museum Tusulanums Forlag, 2011.
- [6] Mats Dahlström. *Under utgivning. Den vetenskapliga utgivningens bibliografiska funktion*. Borås: Valfrid, 2006.
- [7] David Greetham. “A History of Textual Scholarship”. In: *The Cambridge Companion to Textual Scholarship*. Ed. by Neil Fraistat and Julia Flanders. Cambridge: Cambridge University Press, 2013, pp. 16–41. DOI: 10.1017/CC09781139044073.002..
- [8] Nikolai Frederik Severin Grundtvig. “Om Religion og Liturgie”. In: *Ny Minerva* (Feb. 1807), pp. 129–33. URL: <http://www.xn--grundtvigsvrker-71b.dk/tekstvisning/650/0#..>

- [9] Kirsten Vad and Krista Stinne Greve Rasmussen. “Kommentarer, registre og data”. In: *Kommentering og præsentation af nordiske tekster i en digital tidsalder*. Ed. by Simon Skovgaard Boeck, Karen Skovgaard-Petersen, and Jeppe Barnwell. København: Det Danske Sprog- og Litteraturselskab, upcoming.
- [10] Guillaume Chiron et al. “Impact of OCR Errors on the Use of Digital Libraries: Towards a Better Access to Information”. In: *2017 ACM/IEEE Joint Conference on Digital Libraries (JCDL)*. Joint Conference on Digital Libraries. 2017, pp. 1–4. DOI: 10.1109/JCDL.2017.7991582.. URL: <https://doi.org/10.1109/JCDL.2017.7991582..>
- [11] Kimmo Kettunen et al. “Optical Character Recognition Quality Affects Subjective User Perception of Historical Newspaper Clippings”. In: *Journal of Documentation* 79.7 (2023), pp. 137–56. DOI: 10.1108/JD-01-2023-0002.. URL: <https://doi.org/10.1108/JD-01-2023-0002..>
- [12] Simplilearn. *How Big Data Can Help You Do Wonders In Your Business*. URL: <https://www.simplilearn.com/how-big-data-can-help-do-wonders-in-business-rar398-article..>
- [13] Nikolai Frederik Severin Grundtvig. “Kiærminde-Bladet”. In: *Danne-Virke et Tids-Skrift* 3 (Oct. 1817), pp. 155–72. URL: <http://www.xn--grundtvigsvrker-71b.dk/tekstvisning/18413/0#..>
- [14] Katrine F. Baunvig, Oliver Jarvis, and Kristoffer L. Nielbo. “Emotional Imprints: Letter-Spacing in N.F.S. Grundtvig’s Writings”. In: *Proceedings of the Digital Humanities in the Nordic Countries 5th Conference, CEUR*. Digital Humanities in the Nordic Countries 5th Conference. Vol. 2612. 2020, pp. 192–202. URL: <https://ceur-ws.org/Vol-2612/short3.pdf..>
- [15] Katrine F. Baunvig, Oliver Jarvis, and Kristoffer L. Nielbo. “Emotional Imprints: Exclamation Marks in N.F.S. Grundtvig’s Writings”. In: *Post-Proceedings of the 5th Conference Digital Humanities in the Nordic Countries, CEUR*. Digital Humanities in the Nordic Countries. Vol. 2865. 2021, pp. 156–169. URL: <https://ceur-ws.org/Vol-2865/short7.pdf..>
- [16] Jes Fabricius Møller. *Grundtvigs død*. Aarhus: Aarhus Universitetsforlag, 2019.
- [17] Katrine Frøkjær Baunvig. “Med Farver kan man ingen Christus male’. Joakim Skovgaard og Grundtvigs farver”. In: *Kraft og skønhed. Joakim Skovgaards udsmykning i Viborg Domkirke*. Ed. by Iben Overgaard. Viborg/København: Skovgaard Museet: Strandberg Publishing, 2023, pp. 228–239.
- [18] Alfred Korzybski. *Science and Sanity: An Introduction to Non-Aristotelian Systems and General Semantics*. Lancaster: Penn, 1933.
- [19] Vincent D. Blondel et al. “Fast Unfolding of Communities in Large Networks”. In: *Journal of Statistical Mechanics: Theory and Experiment* 2008 (Oct. 2008). DOI: 10.1088/1742-5468/2008/10/P10008..
- [20] Xin Rong. *Word2vec Parameter Learning Explained*. arXiv. Nov. 1, 2014. URL: <https://doi.org/10.48550/arXiv.1411.2738..>

- [21] Katrine Frøkjær Baunvig. “Ånd hos Grundtvig: En digitalhumanistisk skitse af begrebsbrugen i Grundtvigs værker”. In: *Åndsfrihed: i det moderne samfund*. Ed. by Ove Korsgaard and Michael Agerbo Mørch. Aarhus: Aarhus Universitetsforlag, 2022, pp. 65–102.
-







# Biografier



## Redaktører

*Annika Rockenberger* (f. 1982). Ph.D., førstebibliotekar for digitale forskningsmetoder innen humaniora og samfunnsfag ved Universitetsbiblioteket i Oslo. Leder prosjektet BærUt! Bærekraftige digitale vitenskapelige utgaver (2024-2026) som sørger for effektiv utnyttelse av ressurser, kunnskapsutvikling og kompetanseoverføring innen digital edisjonsfilologi. Arbeider med en digital, to-språklig dansk-tysk utgave av *Ethica Complementoria & Trencherbok* (1678) og utarbeider sammen med Maria Kartveit og Øystein Kleiven med en digital utgave av Bernt Oscar Malling Mamens reisedagbøker (1901-1955) og det upubliserte bokmanuskriptet om Mamens siste reise til Mongolia *Atter et år i Østen* (1938).

*Aasta Marie Bjorvand Bjørkøy* (f. 1977). Ph.D., professor i nordisk litteratur og instituttleder ved Universitetet i Oslo. Har nylig utgitt bøkene *Samtidslitterære alderdommer* (2020) og *Norske noveller. En introduksjon* (2023), samt artikler om den norske biografien til Onkel Toms hytte i tidsskriftene *Nordica* (39/2022) og *Norwegian-American Studies* (2023). Medlem av NNEs planleggingsgruppe.

*Nina Marie Evensen* (f. 1969). Cand.philol., datakurator og edisjonsfilolog ved Senter for Ibsen-studier, Universitetet i Oslo. Redaktør for den digitale utgaven av Henrik Ibsens skrifter (HISe) og Det Virtuelle Ibsensenteret, en plattform som samler de digitale ressursene som senteret har bygget opp knyttet til Ibsens liv og forfatterskap. Har tidligere arbeidet med de tekstkritiske utgaveprosjektene *Henrik Ibsens skrifter* (1999–2010), *Ludvig Holbergs skrifter* (2011–2015), *M.B. Landstads skrifter* (2010–2015) og *A.O. Vinjes korrespondentbrev i Drammens Tidende 1851–59* (2015–2018), og har utgitt tekstkritiske utgaver av Cora Sandels *Alberte-trilogi* (2002) og Sigrid Undsets *Kransen* (2020) og *Husfrue* (2023). Medlem av NNEs planleggingsgruppe.

*Ellen Nessheim Wiger* (f. 1974). Cand.philol., forsker og edisjonsfilolog ved Nasjonalbiblioteket i Oslo. Redaktør for e-bokportalen Bokselskap, som driftes av Det norske språk- og litteraturselskap i samarbeid med Nasjonalbiblioteket. Har tidligere arbeidet med de tekstkritiske utgaveprosjektene *Henrik Ibsens skrifter* (1999–2010) og *Norske mellomalderballadar* (2010–2016), og har også utgitt tekstkritiske utgaver av Johan Nordahl Bruns *Einer Tamberskielver* (2002), Camilla Colletts *Amtmandens Døttre* og Sigrid Undsets *Kransen* (2020) og *Husfrue* (2023).

## Forfattere

*Katrine Frøkjær Baunvig* (f. 1983). Ph.d., professor og senterleder ved Center for Grundtvigforskning, Aarhus Universitet. Disputerte i 2013 med avhandlingen *Forsamlingen først. N.F.S. Grundtvigs og Émile Durkheims syn på fellesskap*. Har publisert artikler om Grundtvig, religionsteori og «text mining» av digitale tekstsamlinger. Deltar i prosjektene The Golden Array of Danish Cultural Heritage, Golden Imprints of Danish Cultural Heritage og ComSing: A singing society?.

*Jenny Bergenmar* (f. 1973). Professor i litteraturvitenskap ved Göteborgs universitet. Hun spesialiserer seg på kjønn og litteraturhistorie, med fokus på perioden fra slutten av 1800-tallet til midten av 1900-tallet, samt digitala humaniora og lesingens historie. Hennes siste prosjekt, Queerlit: Utveckling av metadata och sökbarhet för HBTQI-litteratur, ble gjennomført i samarbeid med Kungliga biblioteket og KvinnSam, Nationellt bibliotek för genusforskning. Hennes siste bok, *Lagerlöfs läsare* (2022), utforsker offentlige brev til Selma Lagerlöf fra 1890 til 1940. Hun er en av redaktørene for den kommende utgivelsen *Critical Neurodiversity Studies: Divergent Textualities in Literature and Culture* (Bloomsbury 2025).

*Mats Malm* (f. 1964). Professor i litteraturvitenskap ved Göteborgs universitet. Styremedlem i Svenska Vitterhetssamfundet. Har utgitt verk fra 1600- og 1700-tallet og er leder for Litteraturbanken, som samler inn utgaver av svensk litteratur og legger til rette for bruk av digitale forskningsmetoder på slikt materiale. Han har siden 2019 vært permanent sekretær ved Svenska Akademien.

*Ljubica Miočević* (f. 1977). Assisterende hovedredaktør ved Svenska Vitterhetssamfundet, samt redaktør og forskningsansvarlig ved Litteraturbanken. Styremedlem i Svenska Vitterhetssamfundet. Disputerte i 2017 ved Stockholms universitet med avhandlingen *Fantasiens morgonrodnad. En studie i Clas Livijns romaner*. Jobber med en tekstkritisk utgave av Livijns romaner. Har publisert artikler om tekstgenese, kollektiv tekstskrivning, materialitet, samt språk og autentisitet hos sørslaviske forfattere. Har tidligere vært redaktør ved Strindbergprosjektet.

*Krista Stine Greve Rasmussen* (f. 1979). Ph.d., førsteamanuensis og utgaveleder ved Grundtvigs Verker, Center for Grundtvigforskning, samt leder av Center for Digital Tekstarv (CDHtxt), begge ved Aarhus Universitet. Har skrevet artikler om utgivelseshistorie og tekstvitenskapelige og litterære emner, fører tilsyn med vitenskapelige utgaver og arbeider med digitale, vitenskapelige utgavers livssyklus. Medlem av NNEs planleggingsgruppe samt Det Danske Sprog- og Litteraturselskab.

*Elizabeth Svarstad* (f. 1977). Nordisk mastergrad (2005) og Ph.D. (2018) i dansevitenskap fra Norges teknisk-naturvitenskapelige universitet (NTNU). Utdannet dansekunstner fra Den norske balletthøyskole, underviser ved Norges

musikkhøgskole og Operahøgskolen og jobber som danser og koreograf. Redaksjonsmedlem i *The Nordic Journal of Dance* og fagansvarlig for «Ulike lands dansehistorie» i *Store norske leksikon*. Har mottatt Pareliusprisen for sitt kunstneriske virke.



# Personregister

- Adler  
  Melissa, 28
- Alenkjær  
  Christian Smidt, 9, 10
- Almqvist  
  C.J.L., 16
- Andersen  
  Hans Christian, 9, 50
- Aurora  
  Federico, 10
- Barabucci  
  Gioele, 5
- Baunvig  
  Katrine Frøkjær, 5, 50, 69
- Benedictsson  
  Victoria, 17
- Berg  
  Nils Holger Németh, 10
- Bergdahl  
  Liv Saga, 30
- Bergemar  
  Jenny, 4, 69
- Bergenmar  
  Jenny, 22
- Bjørkøy  
  Aasta Marie Bjorvand, 3, 68
- Blake  
  William, 26
- Bode  
  Katherine, 24, 25
- Bordalejo  
  Barbara, 25
- Boye  
  Karin, 17
- Bremer  
  Fredrika, 17
- Brown  
  Susan, 26, 28, 29
- Bruus  
  Johannes, 44
- Bue  
  Margrethe Støkken, 10, 39, 40, 44
- Bøe  
  Hilde, 10
- Christensen  
  Johanne Emilie, 10
- Clements  
  Patricia, 26
- Collett  
  Camilla, 7
- Darnton  
  Robert, 23
- Edelfeldt  
  Inger, 30
- Ekker  
  Mette Gismerøy, 8
- Evensen  
  Nina Marie, 3, 68
- Flanders  
  Julia, 26, 27
- Flygare-Carlén  
  Emilie, 17
- Foster  
  Susan, 43
- Freccero  
  Carla, 31
- Gailey  
  Amanda, 25, 27–29
- Gray  
  Thomas, 26
- Grundtvig  
  N.F.S., 52, 54–57
- Grundy  
  Isobel, 26
- Hauge  
  Hans Nielsen, 8
- Henrikson  
  Paula, 27
- Holland  
  Kathryn, 28, 29
- Hyttinen  
  Elsi, 24
- Ibsen  
  Henrik, 54

Jockers  
     Matthew, 25

Kaasa  
     Janicke Stensvaag, 10

Kjeldstadli  
     Knut, 43

Klein  
     Lauren F., 24

Knorring  
     Sophie von, 17

Kondrup  
     Johnny, 51, 52

Krusenstjerna  
     Agnes von, 17

Kvamen  
     Ingolf, 8

Kåsen  
     André, 7, 8

Köhler  
     Sebastian, 10

Lagerlöf  
     Selma, 16, 17

Lorin  
     André, 41

Lundegård  
     Axel, 17

Lyttken  
     Alice, 30

Malm  
     Mats, 4, 14, 69

Mandell  
     Laura, 26, 31

Masten  
     Jeffrey, 22, 28, 29, 31

McGann  
     Jerome J., 25–27

McKenzie  
     D.F., 26, 27

Melville  
     Herman, 26

Miočević  
     Ljubica, 4, 14, 18, 69

Molinari  
     Alessandra, 9

Moore  
     Michael Edward, 31

Moretti  
     Franco, 25

Nordenflycht  
     Hedvig Charlotta, 16

Ozment  
     Kate, 26, 27

Playford  
     John, 40

Rafto  
     Heidi Rohde, 10

Rasmussen  
     Krista Stine Greve, 5, 10, 50, 69

Rockenberger  
     Annika, 3, 6, 68

Rosetti  
     Dante Gabriel, 26

Sahle  
     Patrick, 50

Sappho, 30

Shakespeare  
     William, 26

Sjöberg  
     Birger, 17

Skovgaard  
     Joakim, 56

Smith  
     Martha Nell, 26, 27

Stagnelius  
     Erik Johan, 16

Strindberg  
     August, 16, 54

Svarstad  
     Elizabeth, 5, 39, 70

Tanselle  
     G. Thomas, 26, 27

Topelius  
     Zacharias, 54

Tuchman  
     Gaye, 27

Warme  
     Reidar, 44

Whitman

Walter, 26

Wiger

Ellen Nessheim, 3, 68

Wilde

Oscar, 53

Witting

Mette, 7

## Nordisk nettverk for edisjonsfilologer. Skrifter

ISSN 1601-1562

1. Vid texternas vägskäl. Textkritiska uppsatser, red. av Lars Burman & Barbro Ståhle Sjönell. Stockholm: Svenska Vitterhetssamfundet 1999. ISBN 91-7230-087-6
2. Megen viden i forskellige hoveder. Om kommentering af nordiske klassikere, red. af Johnny Kondrup & Karsten Kynde. København: C.A. Reitzels Forlag 2000. ISBN 87-7876-200-6
3. Bok og skjerm. Forholdet mellom bokbasert og digitalt basert tekstutgivelse, red. av Jon Gunnar Jørgensen, Espen S. Ore & Vigdis Ystad. Oslo: Fagbokforlaget 2001. ISBN 82-7322-171-7
4. Text och tradition. Om textedering och kanonbildning, red. av Lars Burman & Barbro Ståhle Sjönell. Stockholm: Svenska Vitterhetssamfundet 2002. ISBN 91-7230-107-4
5. Varianter och bibliografisk beskrivning, red. av Pia Forssell & Rainer Knapas. Helsingfors: Svenska litteratursällskapet i Finland 2003. ISBN 951-583-102-4
6. Læsemåder. Udgavetyper og målgrupper, red. af Per Dahl, Johnny Kondrup & Karsten Kynde. København: C.A. Reitzels Forlag 2005. ISBN 87-7876-447-5
7. Filologi og hermeneutikk, red. av Odd Einar Haugen, Christian Janss & Tone Modalsli. Oslo: Solum Forlag 2007. ISBN 978-82-560-1576-4
8. Bokens materialitet. Bokhistoria och bibliografi, red. av Mats Malm, Barbro Ståhle Sjönell & Petra Söderlund. Stockholm: Svenska Vitterhetssamfundet 2009. ISBN 978-91-7230-149-8
9. Digitala och tryckta utgåvor. Erfarenheter, planering och teknik i förändring, red. av Pia Forssell & Carola Herberts. Helsingfors: Svenska litteratursällskapet i Finland 2011. ISBN 978-951-583-228-3
10. Editionshistoire, red. af Johnny Kondrup & Klaus Nielsen. København: Museum Tusulanums Forlag 2014. ISBN 978-87-635-4243-2
11. Filologi og sensur, red. av Christian Janss, Stine Brenna Taugbøl & Hilde Bøe. Oslo: Novus forlag 2015. ISBN 978-82-7099-828-9
12. Textkritik som analysmetod, red. av Paula Henrikson, Mats Malm & Petra Söderlund. Stockholm: Svenska Vitterhetssamfundet 2017. ISBN 978-91-7230-187-0



13. Utgåvor i användning, red. av Pia Forssell, Carola Herberts, Maren Jonasson & Sebastian Köhler. Helsingfors: Svenska litteratursällskapet i Finland 2022. ISBN 978-951-583-574-1
14. Kildeudgivelse og tekstkritik i forskellige traditioner, red. af Annette Lassen, Krista Stinne Greve Rasmussen & Kirsten Vad. Odense: Syddansk Universitetsforlag 2022. ISBN 978-87-408-3458-1
15. Masedigitalisering og edisjonsfilologi, red. av Annika Rockenberger, Aasta Maria Bjorvand Bjørkøy, Nina Maria Evensen & Ellen Nessheim Wiger. Oslo: Universitetsbiblioteket i Oslo 2024. ISBN 978-82-8037-203-1