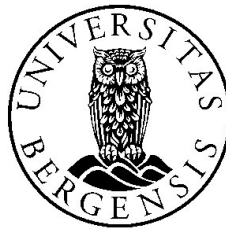


Skandalepolitikk og medienes rolle

En emnemodellering av mediedekningen av
metoo-kampanjen

Vilde Smedsvik



Masteroppgave

[Vår/2020]

Institutt for sammenliknende politikk

Universitetet i Bergen

Abstract

Throughout the past decades, political scandals have been increasingly widespread across the world. The key elements of political scandals are typically sex, money, and power. Today these scandals could not have existed without the media. Scandals have gained increasing status in the political public, and during scandals, the trust in politicians is rigorously tested. The scandals are often directed at individual politicians, where the characteristics of the politician are subject to considerable speculation. The study of political scandals is of importance to students of comparative politics because scandals affect the political system, and as a result of the widespread awareness of scandals, the opportunity is offered to examine what effect they have in society as a whole. The #metoo campaign became a global campaign for women fighting against sexual harassment, confronting several political actors, and the media played a central role. In December 2017, 2143 articles was written on #metoo by the Norwegian press. Whereas the media see it as their social duty to uncover power abuse and other fundamental norm violations, the scholars are claiming that this is contributing to a personification of politics, which is understood as a negative consequence. This thesis applies topic modelling of ten different Norwegian newspapers, examining which aspects of the political scandal the media has been focusing on in their coverage. Scandals affect voters, and readers' interpretations or perspectives may be modified based on the medias presentation. Thus, the role of media during political scandals is a significant matter. The result indicate that the media had an overwhelmingly large proportion of their focus directed at one political actor, and one may argue that this have contributed to personalization. Whether this is an absolute negative aspect, is up for discussion. Furthermore, the analyses in this thesis accumulate several questions that can not be answered by topic modelling but may be interesting and meaningful for further research. Although many of the results are significant, this thesis also display many of the pitfalls of automatic content analysis methods.

Forord

Jeg vil benytte anledningen til å takke veilederen min, Kristin Strømsnes, for å ha gitt meg mye god hjelp og rettleiding gjennom prosessen. Til tross for covid-19 har hun hele tiden gjort meg trygg på at hun var tilgjengelig, og har kommet med gode innspill til det endelige resultatet.

Covid-19 førte også til at mye av masteroppgaven har vært skrevet fra «hjemmekontor», som ble barnerommet hjemme hos mor. Jeg vil derfor takke henne for husly under det siste halvåret av oppgaveskrivingen, og også for masse støttende ord og kjærlighet.

Jeg vil takke kjæresten min som hele veien har støttet meg og overbevist meg om at det vil ende bra. Og en liten ekstra takk for å ha holdt ut med meg og humørsvingningene mine.

Til sist vil jeg takke min bror for å ha hjulpet meg med korrekturlesing og for å gi gode tips til hva som har vært viktig å få fram i diskusjonen.

Innholdsfortegnelse

Kapittel 1: Innledning	8
1.1 Tema og problemstilling	11
1.2 Tekst som data i statsvitenskap	12
Kapittel 2: Skandalepolitikk	14
2.1 Begrepsavklaring	14
2.2 Litteraturgjennomgang	15
2.2.1 «Personifisering og partifavorisering»	18
2.2.2 Media, skandaler og partipopularitet	19
Kapittel 3: Media	21
3.1 Media – presse og politikk	21
3.1.1 Formålet med mediene	21
3.1.2 Hvorfor og hvordan utgjør medieskandaler etiske problemer for pressen?	23
3.1.3 Hvordan håndteres disse diskuterbare mediepraksiser i et anerkjent og meget profesjonelt selvregulerende system?	23
3.2 Valg av media	25
3.2.1 Forskjell i mediedekning ut ifra politisk tilhørighet?	25
3.2.2 Fra partiaviser til partipolitisk uavhengig?	26
3.2.3 Norges største mediehus – presentasjon av valg av medier	27
Kapittel 4: Hvorfor metoo	30
4.1 Emneknaggsandalen metoo	30
4.2 metoo, norske mediehus og politiske partier	32
4.3 Oppfyller metoo kriteriene til en politisk skandale?	34
Kapittel 5: Metode	35
5.1 Fire prinsipper for automatisert tekstanalyse	35
5.2 Betydningen av «språk»	36
5.3 Innhenting av tekst	37

5.4	Tekstbryting	38
5.4.1	<i>Utfordringer ved tekstbryting</i>	38
5.4.2	<i>Grunnleggende tekstbehandling</i>	39
5.5	Følelsesanalyse	42
5.6	Emnemodellering	44
5.6.1	<i>Laten dirichlet allocation</i>	44
5.6.2	<i>Strukturell emnemodellering</i>	46
5.6.3	<i>Validere emner</i>	47
5.7	Klyngeanalyse	48
5.7.1	<i>K-gjennomsnittets klynging</i>	48
Kapittel 6: Resultater og funn i analysen		49
6.1	Tekstbryting	49
6.1.1	<i>Ordfrekvens</i>	50
6.1.2	<i>Tekstklynging og problemet med sparsitet</i>	51
6.1.3	<i>Korrelasjon</i>	52
6.2	Følelsesanalyse	54
6.3	Emnemodellering	57
6.3.1	<i>Latent Dirichlet Allocation</i>	57
6.3.2	<i>Strukturell emnemodellering</i>	61
6.4	Klyngeanalyse	65
6.5	Ordtelling av ord som kan ha betydning for diskusjonen	66
6.6	Eksempel på utfordringer med metoden som viste seg i resultatene	68
Kapittel 7: Diskusjon av resultatene		69
7.1	Seksuell trakassering som et generelt sensitivt tema	69
7.2	Medienes bruk av følelser	70
7.3	Individualisering eller privatisering?	72
7.4	Medienes makt – samfunnsoppdraget	74
7.5	Metoo-emner	76

7.6 Skandalens påvirkning på det politiske system	77
7.7 Innholdsmessige politiske nyhetsinnholdet	80
7.8 Konsekvensene av personfokuset mot Giske under metoo	81
Kapittel 8: Konklusjon og forslag til videre forskning	83
8.1 Forslag til videre forskning	84

Figurtabell

Figur 1: Opplag: Norges største aviser – over 25 000	28
Figur 2: Ordfrekvens	50
Figur 3: Følelsesanalyse ved bruk av ordbok: nrc	55
Figur 4: Latent dirichlet allocation, $k=2$	57
Figur 5: Latent dirichlet allocation, $k=10$	59
Figur 6: Grafisk framvisning av optimale antall k undersøkt ved hjelp av Griffiths2004, CaoJuan2009 og Arun2010	60
Figur 7: Grafisk visning av estimert hyppighet for hvert emne	62
Figur 8: Klyngeanalyse	65
Figur 9: Klyngeanalyse av objekter, $k=2$	66
Figur 11: Visualisering av ordfrekvens gjennom ordsky.	75
Figur 12: Gjennomsnitt av nasjonale meningsmålinger om stortingsvalg	79
Figur 10: Antall objekter, K	97

Tabell-liste

Tabell 5: Følelsesanalyse ved bruk av ordbok: bing	54
Tabell 6: Framstilling av det optimale antall k .	60
Tabell 9: Ordtelling for de ulike partiene	67
Tabell 10: Ordtelling for «metoo», «politiker» og «tillit»	67
Tabell 1: Ord som korrelerer over 98% med «metoo»	98
Tabell 2: Ord som korrelerer over 98% med «politiker»	99
Tabell 3: Ord som korrelerer over 98% med «tillit»	99
Tabell 4: Ord som korrelerer over 98% med «egenskaper»	100
Tabell 7: Vanlige ord for hvert emne	101
Tabell 8: Korrelasjonstest for de 20 emnene vist i figur 7	103
Tabell 11: Ordfrekvens	104

Kapittel 1: Innledning

I løpet av de siste tiårene har politiske skandaler blitt stadig vanligere over hele verden (Isotalus og Almonkari 2014, 3). De sentrale elementene i politiske skandaler er typisk sex, makt og penger, og i dagens samfunn kunne ikke disse skandalene eksistert uten media. Videre har også skandaler fått en økende status i politisk offentlighet, og under skandaler testes tilliten til politikere strengt. Thompson (2000) har sagt at «*skandaler er en kamp om symbolsk makt hvor omdømme og tillit står på spill*». Dette er fordi skandalene ofte er rettet mot enkeltpolitikere, hvor egenskaper ved politikeren blir gjenstand for betydelige spekulasjoner. Studiet av politiske skandaler er av betydning for studenter av sammenlignende politikk fordi at skandaler påvirker det politiske systemet (Kerby og Chari 2002, 409). Som et resultat av den utbredte oppmerksomheten rundt skandaler, tilbys muligheten til å undersøke hvilken effekt de har på samfunnet som helhet. Utfallet av skandalene kan omfatte slutten av etablerte politiske karrierer eller velt av en regjering – et tilfelle som kan utfordre det representative demokratiet.

I oktober 2017 ble det avslørt en rekke hendelser av seksuell trakassering i filmindustrien. Skandalen ble fronten med en emneknaggskampanje som i hovedsak forløp på sosiale medier. Innlegg med emneknaggen «metoo» skulle vise til tilfeller av mennesker som hadde blitt utsatt for seksuell trakassering og hadde som formål å øke bevissthet om denne seksuelle trakasseringen som et reelt problem. Emneknaggen spredte seg til aktører uavhengig av yrke, og etter kort tid kom det også fram varslinger om seksuell trakassering i politikken. Mediene svarte med å sette på det høyeste giret de hadde, og mengden artikler om metoo økte raskt. I desember 2017 ble det skrevet 2143 artikler om metoo i norsk presse (Atekst 2020). metoo-skandalen førte riktig nok ikke til velt av regjering – men flere politikere måtte trå ut av sin politiske stilling (Orgeret 2019).

Jenssen (2014) legger fram en hypotese om at mediene øker personifisering av politikerne (personifiseringshypotesen), og fremstiller dette som et problem med tanke på at det er de politiske partiene som er hovedaktørene i norsk politikk. Mediene på sin side ser på det som sin samfunnsplikt å avdekke maktmisbruk og andre fundamentale normbrudd. Disse holdningene gjør det derfor interessant å undersøke om personifisering under politiske

skandaler er et faktum, og om dette må bli sett på som rimelig eller urimelig. Videre er det interessant å se på tankene til Van Aelst, Sheaffer og Stanyer (2011) som hevder at det burde foreligge et skille mellom individualisering av politikk og privatisering av politikk.

Individualisering tar da for seg de aspektene ved representantene for de politiske partiene som er relevant informasjon for velgerne, mens privatisering handler om det som burde få bli holdt utenfor mediernes søkelys.

Ved å undersøke om mediene har hatt et personfokus må man undersøke hva mediene har skrevet av artikler om det aktuelle emnet. I dette tilfelle er mediene begrenset til aviser, som betyr at analyseenheten blir tekst. Men språk er komplekst, og å gjennomføre analyser av store mengder tekst kan være svært utfordrende. Automatiserte innholdsmetoder – som er tekstanalyser ved hjelp av avanserte dataprogrammer – kan også heller aldri erstattes med nøye lesning av tekst (Grimmer og Stewart 2013, 268). Videre blir det presisert at slike automatiserte metoder kan betegnes som «uriktige språkmodeller». Dette betyr at ytelsen til en hvilken som helst metode i et nytt datasett ikke kan garanteres, og at validering blir ekstremt viktig. På en annen side er det ved behandling av store data (stor N) en bred tilgjengelighet av rimelige programvarepakker og programmeringsspråk. Dette gjør det enklere å gjennomføre tekstanalysesamlinger av tekstdata som ganske enkelt er for store til å kunne leses og tolkes kostnadseffektivt av en gruppe med menneskelige kodere. Det gir altså en mulighet til å kunne gjennomføre analyser av tekstdata med moderne statistiske teknikker som drar nytte av den statistiske kraften som gis av datasett med et stort antall saker, og med tekstutvinningsverktøy utviklet av informatikere (Ignatow og Mihalcea 2018, 3).

En emnemodellering av emneknaggskampanjen metoo vil kunne finne fram til hvor mediene har lagt fokuset under mediedekningen av skandalen. Den vil vise hvilke aspekter som er trukket fram og hvilke ord som har vært sentrale. En slik skandale berører velgerne, og det vil være rimelig å anta at lesernes tolkning vil basere seg på mediens fremstilling – mediens rolle under politiske skandaler vil derfor være et sentralt tema. Det vil også bli gjort ulike analyser som vil prøve å finne fram til hvilke følelser som genereres hos leseren, og på bakgrunn av at tillit er et sentralt emne under politiske skandaler, vil det bli undersøkt for om det kan trekkes noen sammenheng mellom endringer i tillitsforholdet mellom politikere og velgere under slike skandaler.

Oppgaven er delt inn i ulike kapitler for å legge frem analysen på en strukturell måte. I dette første kapitlet introduseres temaet som skal analyseres, og det blir kortfattet lagt fram teoretisk og metodisk grunnlag for forskningsspørsmålet. Kapittel 2 gjør videre et mer grundig arbeid av teoretisk framstilling, hvor det blir gjort rede for definisjoner av sentrale begreper, samt en litteraturgjennomgang for å vise til relevant teori og empiri som underbygger hvorfor det vil være interessant å undersøke temaet, og som gir grunnlag for ulike diskusjonsemner senere i oppgaven. For å gi en forklaring på mediens funksjon og rollen de har i samfunnet har dette fått et eget kapittel, som er kapittel 3, før det i kapittel 4 legges frem en forklaring på hva metoo-kampanjen gikk ut på, og hvor det blir lagt fram en grundig forklaring av hvorfor metoo er et interessant forskningsområde. Valg av metode og forklaring av metoden blir lagt fram i kapittel 5, og kapittel 6 tar for seg resultatene av analysen. Kapittel 7 vil videre diskutere funnene fra analysen, og sette de i perspektiv sammen med den presenterte teorien fra de øvrige kapitlene, før kapittel 8 vil forsøke å legge fram en konklusjon av forskningsspørsmålet, samt vise til forskning som kan være relevant å undersøke videre.

1.1 Tema og problemstilling

Ved politiske skandaler er det en risiko for at man mister politisk innflytelse, og i verste fall kan man miste karrieren sin (Midtbø 2007, 12). Dette medfører at politikerne er nødt til å ta skandalen på alvor. Medie-omtaler er viktige i definisjonen av politiske skandaler, og det vil være interessant å undersøke hvordan media har opptrådt under denne stormen. Det finnes en rekke teori som forklarer at politikken har blitt mer personifisert, og om hvordan politiske skandaler og deres personfokus i stor grad har vært med på å øke en slik personifisering.

Tema for oppgaven er skandalepolitikk, og fokuset ligger som nevnt på mediedekningen av emneknaggkampanjen. Kampanjen er dekket av redaksjonelle medier, og flere politikere fra ulike partier er blitt nevnt. I oppgaven gjennomfører jeg en emnemodellering av en rekke avisartikler for å forsøke å kartlegge sakene. Det er ganske mange tusen artikler, og ved å kartlegge mediedekningen kan man få et bilde av hvilke emner som har vært sentrale for den politiske skandalen, og man får et innblikk i hvordan person versus politikk blir behandlet under politiske skandaler.

Oppgaven vil på bakgrunn av dette ta for seg hvordan mediene har gjennomført sin rolle som formidler under emneknagg-skandalen «metoo» og hvilke effekter det kan ha hatt med tanke på personifiseringshypotesen til Jenssen (2014). Forskningsspørsmålet som skal undersøkes lyder som følger:

«Har mediedekningen av metoo-kampanjen vært med på en personifisering av politikk, og hvordan har dette påvirket nyhetsbildet?»

Det er videre viktig å presisere at denne analysen er direkte rettet mot mediedekningen av metoo-skandalen, og vil på ingen måte stille spørsmål til de fundamentale problemene som seksuell trakassering bringer fram. Det vil ikke på noen tidspunkt bli stilt spørsmål til aktuelle

varslingssaker og deres troverdighet, men undersøkes for hvordan de ulike sakene er framstilt og mediens holdningsskapende funksjon.

1.2 Tekst som data i statsvitenskap

Tekst har alltid vært en viktig kilde i statsvitenskapen (Wilkerson og Casas 2017, 529). Det som har endret seg de siste årene er muligheten for å undersøke store mengder tekst med kvantitative metoder. Internett gir statsvitere mer data enn mentorene deres kunne forestille seg, og forskersamfunnet leverer tilgjengelige programvarepakker for tekstanalyse, sammen med opplæring og støtte. Som et resultat blir tekst som data en vanlig forskningsmetode i statsvitenskapen. Forskere benytter seg av nye datakilder, de bruker mer forskjellige metoder, og de blir kritiske forbrukere av funn basert på disse metodene.

Automatiserte innholdsanalysemetoder gir et bredt spekter av verktøy for å måle forskjellige emner av interesse (Grimmer og Stewart 2013, 294-295). Metodene varierer fra klassifisering av dokumenter – enten i eksisterende eller ennå ikke bestemte kategorier – eller skalering av politiske aktører i politisk rom. Tekstanalysemetoder er på sett og vis ukorrekte modeller av språk, og derfor kreves det alltid nøye validering av dataene. Utover metodologisk innovasjon er det enorme bunker med tekster som nå kan analyseres effektivt ved hjelp av automatisert tekstanalyse, og forskere kan lære mye av anvendelsen av tekstanalysemetoder til deres interesseområde.

Nytte teknologier har gjort store mengder digital tekst tilgjengelig, og registrerer en stadig større andel av menneskelig interaksjon, kommunikasjon og kultur (Gentzkow et al. 2019, 535). For samfunnsvitere er informasjonen kodet i tekst et rikt supplement til de mer strukturerte typer data som tradisjonelt brukes i forskning, og de siste årene har vist en eksplosjon av empirisk forskning som bruker tekst som data. I medier brukes blant annet tekst fra nyheter og sosiale medier for å studere effektene av politiske skrånlikk.

Når mennesker leser tekst ser de ikke en vektor av dummyvariabler, og heller ikke en sekvens av ikke-relaterte symboler. De tolker ord i lys av andre ord, og trekker ut mening fra teksten

som helhet. Det kan dermed virke åpenbart at ethvert forsøk på å destillere tekst til meningsfulle data må på lignende måte ta hensyn til komplekse grammatiske strukturer og rike samspill mellom ord. Algoritmer eksisterer for effektivt å analysere grammatiske strukturer, utvetydige forskjellige sanser av ord, skille viktige punkter fra sekundære sider og så videre. Likevel ignorerer praktisk talt all analyse av tekst i samfunnsvitenskapen, som mye av tekstanalysen i maskinlæring mer generelt brorparten av denne kompleksiteten. Råtekst består av en ordnet rekkefølge av språkelementer: ord tegnsetting og hvitt mellomrom. For å redusere dette til en enklere representasjon som er egnet for statistiske analyser gjør man vanligvis tre typer forenkling: dele teksten i enkeltdokumenter, redusere antall språkelementer vi vurderer og begrenser i hvilken grad vi koder avhengighet mellom elementer i dokumenter. Resultatet er kartlegging. Fra rå tekst, til en numerisk matrise.

Digital tekst gir et rikt arkiv av informasjon om både økonomiske og sosiale aktiviteter (Gentzkow et al. 2019, 569). Moderne statistiske verktøy gir forskere muligheten til å trekke ut denne informasjonen og kode den i en kvantitativ form som er mottagelig for beskrivende eller kausal analyse. Både tilgjengeligheten av tekstdata og grensen til metoder utvides raskt, og det forventes at viktigheten av tekst i empirisk forskning vil fortsette å vokse. Fremveksten av tekstanalyse er altså del av en bredere trend mot større bruk av maskinlæring og relaterte statistiske metoder.

Kapittel 2: Skandalepolitikk

2.1 Begrepsavklaring

En politisk skandale kan defineres som brudd på verdier, normer, og moralske koder (Allern og Pollack 2012, 11). Kerby og Chari (2002) hevder videre at politiske skandaler dreier seg om handlinger som blir begått av individer som befinner seg innenfor et politisk felt og som har påvirkning på holdninger innen det feltet hvor normbruddet blir begått (Kerby og Chari 2002, 410). Men denne oppgaven vil basere seg på egenskapene presentert av Midtbø (2012) som blir listet opp under.

Ifølge Midtbø (2012) har en nemlig en politisk skandale de følgende egenskapene; (1) skandaler er en overtredelse av normer, (2) skandaler er offentlige – en skandale blir ikke en skandale før alle vet om den, (3) det er alltid noen som blir fornærmet i sammenheng med en overtredelse av en norm, (4) avsløringen av overtredelse og den påfølgende fordømmelsen vil true en politikers omdømme og (5) skandaler er karakterisert med intens spenning og usikkerhet (Midtbø 2012, 152-153). Midtbø (2007) har også omtalt en skandale som en hendelse som skjer «*av amatørmessige trivialiteter og er bent frem komiske brudd på oppleste normer*» (Midtbø 2007, 11). Skandalene fungerer som en påminnelse om hvor grenser går, og de omhandler ofte personer som åpenbart burde visst bedre. Det kan også på sett og vis bli sett på som en overtramp av egne regler – da politikere selv som lovgiver har definert disse grensene. Videre er det ingen faste grenser mellom kontroversielle spørsmål og skandaler, og skandaler bør dermed ikke defineres som dikotome hendelser, men snarer kontinuerlige fenomener (Midtbø 2012, 153). Vi burde skille mellom små og større skandaler, og ikke lage et skille mellom enten/eller, og det er ulike tiltak for å finne ut av hvor på skalaen fra liten til stor skandalen ligger: (1) jo høyere status på politiker – jo høyere skade, (2) omfanget av mediedekningen, (3) noen overtredelser er regelrett verre enn andre, noe som fører videre til at (4) konsekvensen av overtredelsen, og (5) reaksjonen fra velgerne burde ha noe å si for størrelsen av skandalen (Midtbø 2012, 153). Hvordan metoo-skandalen oppfyller kravene for definisjonen av en politisk skandale vil bli forklart under kapittelet om hvorfor jeg valgte å skrive om metoo.

Som tidligere nevnt inngår normene ofte i de tre teoretiske kategoriene: sex, makt eller penger (Isotalus og Almonkari 2014, 3). Nøkkelen til å definere begrepet handler om *tillit og omdømme*, hvor skandalene ikke nødvendigvis ødelegger disse faktorene, men har kapasitet til det (Kerby og Chari 2002, 410). Det er denne kapasiteten, altså det potensialet for å skade omdømme og tillitsforhold, som gjør at skandaler har så stor betydning på det politiske feltet. Omdømme og tillit er hovedformene for symbolsk kapital som politiske aktører er avhengige av for å tilegne seg og utnytte symbolsk makt. Troikaen fra «sex, makt og penger»-skandaler er derfor spesielt farlige ettersom de nesten øyeblikkelig kan redusere den symbolske kapitalen som en politisk aktør kan ha bygget opp over tid. Likevel er det tilfeller der politiske aktører kan ha nok symbolsk kapital og makt til å ri av stormen av skandalen og dukke opp uskadd (Kerby og Chari 2002, 410).

2.2 Litteraturgjennomgang

Tor Midtbø (2007) beskriver i sin bok «Skandaler i norsk politikk» hvordan skandaler har ødelagt politiske karrierer, uten at oppslutning til partiet faller, og dette blir omtalt som å stride mot «sunn fornuft» (Midtbø 2007). Antall politiske skandaler har de siste tiårene økt betydelig både i Norge og resten av Norden (Jenssen 2014; Isotalus og Almonkari 2014), og de politiske skandalene er stort sett medierte skandaler hvor «*mediene inntar gjerne rollene både som anklager og dommer og spiller en avgjørende rolle som redskap for aktører som vil skandalisere politikere av ulike grunner*» (Jenssen 2014, 100). Skandalene har en utrolig kraft i seg, og det er store muligheter for å ende på forsiden av flere avviser i ukesvis. Utfallet ender ofte med at politikere trer av fra sin stilling, og forskning viser at tillit til politikere blir svekket i sammenheng med politiske skandaler (Bowler og Karp 2004; Maier 2014).

Men store skandaler snur ikke bare politikernes liv og karrierer opp-ned, de kan også ha effekter på institusjoner og den politiske kulturen (Midtbø 2012, 151). På slutten av 1980-tallet omtalte forskere Norge og andre skandinaviske land som en idyllisk himmel hvor politikere ikke var utsatt for å bli «nedverdiget». Norsk politikk er i dag neppe så fredelig som den som er beskrevet av utenlandske observatører, og både antall og typer av skandaler har blitt mer variert. Sex-skandaler, blant annet, har funnet sin vei inn i den norske politikken, og i en tid hvor velgerne har blitt mer opptatt av politiske personligheter og mindre lojale til

partiene, vil man kunne forvente at disse skandalene vil ha en viss innvirkning på partienes popularitet.

«Mediene kaster seg over de politiske skandalene med stadig større fryd og stadig færre hemninger» (Midtbø 2007, 12), og rollefordelingen i politiske skandaler er dermed klare: mediene jager og politikerne jages, noe som fører til at politikerne blir nødt til å ta skandalene på alvor. Tidligere er det lagt vekt på at området for politiske skandaler og deres påvirkning på den politiske virksomheten burde vært undersøkt nærmere. Skandaler er i minste forstand god sladder, men befolkningen viser seg uansett å være svært opptatt av dem.

Clinton/Lewinsky-skandalen i USA, hvor president Bill Clinton skal ha hatt et seksuelt forhold til praktikanten Monica Lewinsky, klargjorde, og på en måte omdefinerte, hvordan politisk kultur og underholdningskultur har blitt smeltet sammen til å angå parlamentariske ledere (Owen 2000, 161). Rutinemessig ble publikums stilling til saken formidlet av massemedier når det gjaldt borgernes motstridende syn på presidentens jobbprestasjon og hans karakter. Media lagde artikler og rapporterte om presidenten som opptrådte i sin profesjonelle rolle, samtidig som de la fram sexskandalen. Resultater fra en undersøkelse som ble publisert i pressen den gang, så dog ut til å indikere at publikum skilte mellom presidentens egnethet som leder og hans personlige oppførsel (Owen 2000, 166). Da de ble spurt, hevdet publikum at pressen ga for mye dekning til Clinton/Lewinsky-saken, men nyhetsorganisasjoner fant imidlertid ut at skandaledekningen økte seer- og lyttertallene. På en annen side viste et eksempel fra Finland hvor fatale konsekvenser en affære kan få, da den finske utenriksministeren Ilkka Kanerva i april 2008 ble tvunget til å trekke seg som følger av et enormt antall tekstmeldinger han hadde sent til en ung, kvinnelig erotisk danser (Isotalus og Almonkari 2014, 3). Under skandalen fulgte mediene denne episoden intenst, og mye av rapporteringen fokuserte på politikeren som enkeltpersoner, hvor hans egenskaper ble evaluert i større grad enn vanlig. Resultatene fra en analyse gjennomført av Isotalus og Almonkari (2014) viser at mediene fokuserer på fem kategorier: faglig kompetanse, personlige egenskaper og atferd, pålitelighet, opprettholdelse av relasjoner og kommunikasjonsevner (Isotalus og Almonkari 2014, 3).

Diskusjonen av politikernes egenskaper gjør metoo-skandalen interessant for forskning da skandaler seg ut til å ha blitt et fast innslag i den moderne politikk (Isotalus og Almonkari

2014, 4). Det viser seg at politiske skandaler etter hvert har blitt et mer regelmessig trekk i nyhetsmediers dekning av politikk i Norden, selv om nordiske skandaler som omhandler seksuell atferd imidlertid må regnes som sjeldne om man skal sammenligne dem med blant annet Storbritannia og USA (Allern et al. 2012, 29). Perspektivet som omhandler kvaliteten og egenskapene til politiske ledere blir interessante under skandalene fordi de typisk er basert på et normbrudd. Når normene brytes i en skandale blir det lagt merke til, og det dukker opp i offentligheten. Derfor utgjør en skandale en god mulighet til å vurdere normene. Normene for politikerens egenskaper er normalt underforstått; men vi kan anta at de under en skandale blir tydeligere i media (Isotalus og Almonkari 2014, 5). Studiene viser også at nyhetsdekning av skandaler er preget av å være sterkt orienterte mot personlige normbrudd og ansvar (Allern og Pollack 2012, 183). Det er også viktig å huske at det er mange kulturelle forskjeller i medieskandaler: personifisering og privatisering har utviklet seg ulikt, og skandaler er forskjellige mellom land. Å trekke seg på grunn av en skandale ser også ut til å være sjelden internasjonalt (Bynander og Hart 2007, 55), selv om det ser ut til å bli stadig mer vanlig (i alle fall i Norden). Hyppigheten av sex-skandaler i USA kan være med på å forklare deler av hvordan amerikanerne kunne skille mellom private forhold og arbeidsforhold ved Clinton-skandalen, mens dette ikke var tilfellet for Finland. I oppgaven vil denne problemstillingen bli belyst ved hjelp av de ulike emnene som kommer fram av analysen, og det vil bli tatt i bruk meningsmålinger for å se om mediedekningen kan ha hatt noen påvirkning på partipreferanse.

Det enorme individfokuset under mediedekning av skandaler kan sees på som en personifisering av politikken. Dette fordi artiklene er mer knyttet til individer enn de er til ideologier og partier (Isotalus og Almonkari 2014, 4). Personifisering er et konsept som beskriver prosessen der den politiske tyngde til den enkelte aktør i politikken øker over tid (Karvonen 2010, 4-5). Muligheten for å finne bevis for den historiske utviklingen av dette fenomenet er blitt diskutert, men Karvonen (2010) analyserte studier av personifisering og fant at utviklingen ikke er ensartet mellom land (Karvonen 2010). Han fant imidlertid ut at det er mange indikasjoner på at individet har blitt mer fremtredende innen valg- og utøvende politikk, og at betydningen av rollen som statsminister, partileder og individuell kandidat har blitt viktigere. Van Aelst, Sheafer og Stanyer (2011) har en mer avansert definisjon av begrepet om personifisering enn den som er nevnt ovenfor, hvor de deler begrepet inn i to former: individualisering og privatisering. Individualisering omhandler fokuset på individuelle politikere som sentrale aktører for den politiske arena, mens privatiseringen går

ut på at det skjer en forskyvning i mediernes fokus fra politikeren som okkupant av offentlig rolle, til politikeren som privatperson (Van Aelst et al. 2011, 205). De hevder at individualisering ikke nødvendigvis står i motsetning til det innholdsmessige politiske nyhetsinnholdet, men at det likevel innebærer et skifte i mediasynligheten fra partier til enkeltpolitikere, eller fra regjeringer til individuelle ministre. Privatisering er ellers den delen av personifisering som de hevder de ser en økning av, og hvor det oppstår et skille fra politikeren offentlige rolle til politikeren private liv. Videre deler de individualisering opp i to kategorier, hvor det ene handler om politikeren generelle synlighet, og det andre om politikeren konsentrerte synlighet – hvor sistnevnte involverer et fokus på et begrenset antall politiske ledere. De angir også at privatisering kan deles inn i oppmerksomheten rundt det personlige liv og oppmerksomheten på mer personlige egenskaper.

Medierapportering er et viktig område der personifisering av politikk kan observeres (Isotalus og Almonkari 2014, 4), og Adam og Maier (2010) konkluderer med at dekningen i media er det eneste området i politikken hvor forskning støtter uttalelsen om at politikk har blitt mer personlig (Adam og Maier 2010, 231). I følge Isotalus og Almonkari (2014) kan personifisering sees i politiske skandaler fordi de i dag ofte er relatert til personlige liv for enkeltpolitikere, og spesielt i tilfelle med sexskandaler som innebærer offentlige avsløringer av aspekter ved privatlivet som er i strid med normene og kodene for hva som er akseptabelt (Isotalus og Almonkari 2014, 5). I forbindelse med metoo-skandalen er det interessant å se på perspektivet om at en individualisering blir sett på som rett og rimelig, men spørsmålet blir om slike skandaler som en selvfølge blir rettet direkte mot politikeren offentlige rolle, eller om det blir forsøkt å fokusere på normbruddet som en privat handling.

2.2.1 «Personifisering og partifavorisering»

I Jenssens (2014) forskning viser han til en *personifiseringshypotese* hvor han hevder at mediene har et større fokus på politikere som private individer framfor representanter for partiene (Jenssen 2014, 102). Videre hevder han at det mediene i økt grad har begynt å lage et skille mellom politikere og parti ved eventuelle skandaler. Det er de politiske partiene som er hovedaktørene i den norske politikken, men likevel har mediene over tid lagt interessen for partiet som helhet til side, og blitt mer og mer interessert i den enkelte politiker og deres

handlinger. En slik personifisering har blitt lagt merke til i lang tid, og dette ser man gjennom hvordan politikerne har utviklet en evne til å tilpasse seg mediene for å sette sitt preg på hva som blir skrevet. Politikerne åpner opp dørene til sitt private liv, og partiene er blitt mer og mer opptatt av å fronte sine ledere som omgjengelige og folkelige, samtidig som mediene har fått et økt fokus på persondrama – særlig når de skal sette politikere til veggs (Johannessen 2013; Jenssen 2014).

Om man studerer effekten av sosiale medier i politikken kan man se en økt individualisert kampanje og et større kandidatfokus (Karlsen 2011, 22), og i moderne demokratier er konstruksjonen av den politiske ledelsen drevet av kommunikasjonsstrategier med større vekt på image over substans, og personlighet framfor ideologi (Campus 2010, 221). Politiske kandidater bruker altså markedsføringsteknikker for å forme valgbildene sine. At politikken er blitt mer personifisert er dog ikke noe nytt fenomen, og har vært under forskernes radar lenge. Blant annet var Kaase (1994) tidlig ute og hevdet at persondrama bidro til avpolitisering (Kaase 1994, 226), og det er på bakgrunn av dette og den informasjonen som er nevnt ovenfor at Jenssen (2014) kom med sin personifiseringshypotese. Som et supplement utviklet Jenssen (2014) også en hypotese om *partifavorisering*, som går ut på at publikummets respons på de politiske skandalene er avhengig av hvilket parti man foretrekker (Jenssen 2014, 103). Denne hypotesen bygger på informasjon om at politiske aktører blir vurdert annerledes ut i fra om man sympatiserer med dem eller ei. Videre forklarer han at skandaleinnholdet har betydning for velgerne, og man ser her at personer med ulik partitilhørighet har ulik toleranse for de ulike skandalene – og at dette munner ut i at det foreligger ulike verdiprioriteringer (Jenssen 2014, 104). Men uavhengig av dette viser det seg også at skandaler som handler om normbrudd gjort gjennom sin politiske rolle anses som mer utilgivelig enn et normbrudd som er gjennomført på «privaten» (Jenssen 2014, 107). Og uavhengig av hvilke normbrudd og skandale som skjer, så går ikke dette noe særlig ut over velgernes tanker om partiet de sympatiserer med, men det har heller større utfall for de partiene de står lengst fra politisk.

2.2.2 Media, skandaler og partipopularitet

Medieeffekten på valgatferd er kontroversiell (Midtbø 2012, 156). Noen hevder at effekten er øyeblikkelig, direkte og sterk, mens andre sier at media bare forsterker velgernes allerede

eksisterende holdninger. Nyere forskning har imidlertid kreditert medier en mye sterkere effekt på opinionen enn tidligere antatt. Mediene har blitt viktigere delvis fordi andre faktorer har blitt mindre viktige. Effekten av sosial bakgrunn og partiidentifisering har for eksempel blitt dårligere de siste tiårene. Ifølge nyere litteratur er det ikke lenger spørsmål «om», men «hvordan», media påvirker stemmeatferden (Midtbø 2012, 157). Ett perspektiv hevder at mediene fremdeles bestemmer hva folk – og faktisk hva politikere – skal tenke på. Media sies å utføre en agenda-innstillingsfunksjon, hvor nyere studier antyder at media, spesielt gjennom deres innramming av nyhetene, til og med påvirker velgernes meninger også. Nyheter som får bred dekning, bidrar til hvordan folk tenker på politiske spørsmål, og ubalansert medieomtale viser seg å tiltrekke spesiell oppmerksomhet.

Hvis media påvirker velgeroppførselen, bør det samme sannsynligvis også gjelde for politiske skandaler (Midtbø 2012, 157). Skandaler dominerer nyhetene og skandaler er også i sin natur ubalanserte. Til tross for forskjellige journalistiske rammer og perspektiver, er pressedekningen overveldende negativ. For å danne seg en mening om skandaler krever ikke så mye politisk kunnskap, og enkelte politikere er smertelig klar over at overtredelser ikke glemmes lett. Studier viser at folk har en tendens til å identifisere seg med majoriteten (Midtbø 2012, 157). Hvis media er samlet i sin fordømmelse av den eller de berørte politikere, kan det hende at de kan representere det populære flertallet. Dette kan i seg selv gi et tilsvarende flertall i valget. På den annen side kan velgerne se på skandaløse nyheter som god, men politisk irrelevant underholdning (som tilfellet Clinton/Lewinsky) (Midtbø 2012, 157). Om en skandale påvirker valgdeltagelse i det hele tatt, kan det og hende det fører til passivitet.

Kapittel 3: Media

3.1 Media – presse og politikk

Nyhetsmediene fungerer som et offentlig marked for skandaler på minst to måter (Allern og Pollack 2012, 181). For det første, som et sted der tabloidjournalistikk tilbyr skandaløse avsløringer til et publikum som er sulten på mer. Skandaler selger aviser, tiltrekker oppmerksomhet på internett og får folk til å samles foran fjernsynet. For det andre utgjør medier en markeds plass i en annen og mer overført forstand, hvor informasjon lekkes med vilje for å gi drivstoff til skandalen (Allern og Pollack 2012, 181). Informasjon utveksles for oppmerksomhet. Ryktene er plantet; dokumenter blir publisert – ofte skjult og med løfter om anonymitet. Mange agenter strømmer sammen i markedet for politiske skandaler.

3.1.1 Formålet med mediene

For å få en forståelse av mediens betydning for samfunnet og hvor stor påvirkningskraft de har, er det også verdt å nevne hva de selv tenker om egen gjennomføring og hva som er viktig når man driver med journalistikk. Medieskandaler reiser ekte profesjonsetiske dilemmaer (Bjerke 2012, 165), noe som mediene erkjenner. Mediene hevder selv at et av de viktigste temaene er etikk, og har laget en egen profesjons- og bransjeetikke som i bunn og grunn er styrt av bransjen selv gjennom paraplyorganisasjonen Norsk Presseforbund (NP) (Bjerke 2009, 176). NP er et organ for norske medier og skal sørge for å legge føring for etiske spørsmål, samt faglige spørsmål rettet mot det redaksjonelle arbeidet. Vær Varsom-plakaten er grunnmuren av reglene om etikk, og plakaten har fått høy prestisje blant journalistene i Norge.

Vær Varsom-plakaten tar for seg følgende:

- Pressen ivaretar viktige oppgaver som informasjon, debatt og samfunnskritikk.
- Pressen har et spesielt ansvar for at ulike syn kommer til uttrykk.
- Pressen skal verne om ytringsfriheten, trykkefriheten og offentlighetsprinsippet. Den kan ikke gi etter for press fra noen som vil hindre åpen debatt, fri informasjonsformidling og fri adgang til kildene.

- Det er pressens rett å informere om det som skjer i samfunnet og avdekke kritikkverdige forhold.
- Det er pressens oppgave å beskytte enkeltmennesker og grupper mot overgrep eller forsømmelser fra offentlige myndigheter og institusjoner, private foretak eller andre. (*Mediebedriftene*, s.v «Mediebransjens samfunnsansvar»)

Medienes Landsforbund (MBL) mener «*mediene bør ha rammebetingelser som gir dem et sunt og solid økonomisk fundament for sin virksomhet og gode utviklingsmuligheter*» (*Mediebedriftene*, s.v «Medienes rammebetingelser»). For å opprettholde et demokratisk samfunn krever dette at folk flest har tilgang til så mye og så god informasjon som mulig. Mediene er tenkt som en arena for å dele meninger, samt å reflektere, lære og å vise engasjement og deltagelse. Et mangfold av nyheter, analyser og offentlig debatt står sentralt, og det er medienes samfunnsoppdrag å opprettholde en plattform som legger til rette for dette. Vær Varsom-plakaten er en sentral del av hva som former medienes mål og mening, og her blir dette samfunnsoppdraget presisert.

De etiske normene er i all hovedsak et tiltak som skal sørge for å beskytte enkeltpersoner fra overgrep i mediene, og Vær varsom-plakaten kommer inn som er en oversikt som er laget for å legge føringer for hvordan man skal holde seg innenfor disse rammene. Føringerne Vær Varsom-plakaten legger for medienes samfunnsrolle viser hvilke rettigheter pressen har, så vel som plikter. Fra mediebedriftenes syn har medienes involvering i enkeltsaker flere ganger vist seg å være den utløsende faktoren for å beskytte og å rette opp i feilaktige hendelser i samfunnet (*Mediebedriftene*, s.v «Medienes rammebetingelser»). Å være bevisst på presseetiske vurderinger er altså med på å legitimere at man setter kritiske spørsmål ved handlinger gjennomført av offentlige institusjoner eller personer. Offentlige personer vil da være personer man kan regne som offentlige gjennom deres stilling eller tillitsverv (*Mediebedriftene*, s.v «Mediebransjens samfunnsansvar»), og politikere må helt klart kunne sies å falle under denne definisjonen. Det sies at «*presseetikken skal ikke beskytte det kritikkverdige, men de presseetiske normer skal hjelpe redaksjonen til å ikke påføre mennesker en unødig lidelse*» (*Mediebedriftene*, s.v «Mediebransjens samfunnsansvar»). Etiske konflikter skal vurderes nøye av redaktørene og journalistene gjennom en grundig drøfting, og på en sånn måte kommer man fram til hva som burde publiseres og ikke.

3.1.2 Hvorfor og hvordan utgjør medieskandaler etiske problemer for pressen?

Det finnes lite skandinavisk litteratur om hvordan ofre opplever medierte skandaler og mediejakten – det meste av litteraturen er selvbiografier (Bjerke 2012, 165). I tillegg har flere journalister skrevet bøker som inneholder intervjuer med medieofre. Ofrene og deres familier oppfatter at deres privatliv er blitt invadert og ødelagt, og de anser deknningen for å være urettferdig og utenfor proporsjoner (Bjerke 2012, 166). De oppfatter at «alle» stirrer på dem, og at de noen ganger er besatt av hva media kommer til å publisere. Noen medieofre forteller hvordan de måtte stå opp midt på natten for å lese avisen før noen andre kunne få tilgang til dem. Konklusjonen til Bjerke (2012) er at medieskandaler kan forårsake betydelig nød for ofrene og deres familier og føre til psykologisk skade, i verste fall selvmord (Bjerke 2012, 168). «Du skal ikke skade dine medmennesker» er et sentralt aspekt ved vanlig moral, og medierte skandaler reiser viktige og vanskelige etiske utfordringer i journalistikken ved å offentliggjøre personer.

3.1.3 Hvordan håndteres disse diskuterbare mediepraksiser i et anerkjent og meget profesjonelt selvregulerende system?

Rapportene fra ofrene og forskningsbidragene sier at det vanligvis ikke er identifikasjonen av skandalen, gjennomgangen av skandalen, eller kritikken i seg selv som skaper problemer (Bjerke 2012, 168). Snarere peker ofrene på størrelsen av skandalen og den lange varigheten – så vel som ensidigheten, gjentakelsen og den smale tolkningsrammen.

En del av definisjonen av mediedrevne politiske skandaler er at det påståtte bruddet på normer fører til langvarig og bred mediedekning i flere parallelle kanaler. En rekke casestudier har dokumentert det solide omfanget av denne deknningen, men disse forskningsfunnene gir fortsatt bare et ganske blekt inntrykk av dimensjonen (Bjerke 2012, 168-169). I løpet av de siste 20 årene har nyhetstjenestene i Norden økt betydelig og inkluderer nå 24/7 nyhetskanaler på TV, mainstream TV og radiokanaler, samt kontinuerlige oppdateringer på nett. En medieramme etablerer gjennom valg av kilde, hvilke fremtredende fakta som skal rapporteres, og hvilke som skal glemmes bort. De velger også de såkalte «sentrale organiseringstemaene»,

som er et kulturelt element som binder sammen de utvalgte kildene og faktaelementene i et stykke journalistikk, og som dermed er med på å formidle manifesterte og latente resonneringsretninger som kausale forklaringer og evalueringer av hendelsene som formidles (Bjerke 2012, 169).

Når medieskandaler utvikler seg, blir historier og informasjon som støtter og/eller underbygger det dominerende perspektivet til skandalene prioritert i nyhetsrommene (Bjerke 2012, 169), mens nyhetshistorier og/eller informasjon som ikke støtter, eller som svekker den dominerende innrammingen av skandaler ikke blir prioritert (Midtbø 2007; Allern og Pollack 2009). Med andre ord: historier som støtter ideen om at heksen er en heks, skrives ut eller sendes. Historier som presenterer forskjellige perspektiver blir droppet på morgenkonferansen, kastet i papirkurvene på pulten, eller skrevet om av kveldsredaktøren. Når medieskandalen utvikler seg, blir innrammingen ofte langt sterkere og smalere.

I de tøffeste medieskandaler kan altså rammen utvikle seg i retning av å demonisere dem som er utpekt som skurker (Bjerke 2012, 169). Demonisering er en handling som i seg selv bryter med en rekke etiske overbevisninger. I tillegg fratrar demonisering mennesker deres menneskelighet, og dermed svekker de barrierer folk vanligvis har som stopper dem fra å oppføre seg på en uetisk måte overfor hverandre. I medieskandaler kan dermed journalistenes etiske skjønn svekkes, og medieskandalene bringer i høysetet noen av de essensielle moralske og etiske utfordringene i journalistikken: proporsjonalitet, rettferdighet, respekt for personvern og korrekt sitering. Hva slags ensidig, negativt mediepress er det moralsk akseptabelt å utsette andre mennesker for? Hvem vil trekke grensen, og hvordan?

Det er bred enighet om at journalistikk skal avdekke uakseptable forhold, hvor medienes økonomiske og politiske privilegier er basert på ideen om at en fri presse er essensielt for et demokratisk samfunn (Bjerke 2012, 170). Slik er det blitt blant annet fordi det kan avdekke maktmisbruk, urettferdige privilegier og dårlig behandling av enkeltpersoner og grupper. Når journalistikken oppfyller sin «del» av denne sosiale kontrakten, vil den uunngåelig skade den/de ansvarlige. Og når norske journalister blir spurt om hvorfor de har en profesjonell rett til å skade andre i jobben sin, henviser de enstemmig til journalistikkens samfunnsoppdrag.

3.2 Valg av media

3.2.1 Forskjell i mediedekning ut ifra politisk tilhørighet?

Jenssen (2014) legger fram argumenter for to forskjellige hypoteser, og med særlig fokus på hypotesen om partifavorisering kan man undre seg i hvilken grad politisk tilhørighet til et parti har betydning for leserens valg av medie. Partifavoriseringen omhandler som tidligere nevnt at responsen til leserne er avhengig av hvilke parti man holder med (Jenssen 2014, 103). Det vil derfor være naturlig å anta at de ulike avisenes forhold til politiske partier vil forme måten partiene blir omtalt på, og på den måten kan også leserne velge aviser etter politisk ståsted. Vi skal likevel se at det gjennom tidene har skjedd store forandringer i forholdet mellom aviser og de politiske partiene (NOU 2000:15, 6). På slutten av sekstitallet hadde mesteparten av avisopplag klar tilknytning til bestemte politiske partier, mens aviser i dag prøver å være politisk uavhengig. Noen hevder at denne utviklingen har ført til for stor likhet i mediedekningen – at både avisenes dekning og vurdering av ulike saker blir for like. Det er viktig å ha ordninger som sørger for at slike tendenser ikke skjer, og NOU2000:15 finner i sin konklusjon at det er viktig med et «*bredt spekter av aviser og et ytre mangfold der det vil være aviser som inntar og kan innta avvikende synspunkter, vurderinger og meninger*» (NOU 2000:15, 6).

For å sikre at det foreligger en bredde i samfunnsdebatten (differensiert presse) har det vært et mål å sikre aviser med ulike politiske ståsted, men dagens aviser har ikke lenger som formål å sikre et mangfold innenfor politikken. Dette skyldes en omvelting av norsk presse hvor aviser ikke lenger skal være forbeholdt et politisk ståsted (NOU 2000:15, 113). Likevel kan man gjennom å undersøke avisers formålsparagrafer se at enkelte aviser ikke klarer å legge bort tradisjonen, og at avisene fremdeles har en politisk plattform. Aviser blir ofte omtalt som «den fjerde statsmakt», og dette begrepet kommer fram av mediernes rolle som kritiker og overvåker. Et nært forhold mellom presse og parti kan i den sammenheng tidvis vise seg å være problematisk. Hovedgrunnen for at norsk presse skulle stå som politisk uavhengig var nettopp for å sikre troverdighet under tittelen «den fjerde statsmakt», og gjennomføringen av frigjøringen fra de politiske partiene har i all hovedsak vært drevet av mediene selv (NOU 2000:15, 113).

3.2.2 Fra partiaviser til partipolitisk uavhengig?

Det er seks grunner for at partiene ønsket egne aviser: (1) øke velgerstøtte for partiet, (2) positiv omtale øker status, (3) støttepunkt, (4) partier kan ansette og utdanne personer gjennom avisene, (5) informasjonskilde ved informasjon som ikke gjøres kjent for allmenheten, og (6) holde oversikt på andre medier og opptre deretter (NOU 2000:15, 115-116). Til tross for at formålet med aviser har endret seg, er det fremdeles mange aviser som har tatt vare på sitt ideologiske ståsted (NOU 2000:15, 116). Dette ser man tydeligst i A-pressen (i dag Amedia (*Amedia s.v.* «Vår historie»)), hvor det har vokst fram noen mønstervedtekter som i stor grad er tilpasset lokale selskaper (NOU 2000:15, 117). Et godt eksempel her er avisen Nordlands Framtid: «*Avisen skal drive saklig og uavhengig journalistikk, forankret i arbeiderbevegelsens ideer om frihet, demokrati og likeverd*» (NOU 2000:15, 117). Likevel skal de forsøke å framstå «frie og uavhengige». De konservative avisene blir de tradisjonelle borgerlige avisene, som har samme grunnsteiner som A-pressen, men den politiske forankringen er på høyresiden (NOU 2000:15, 118). De liberale har gjerne ideologisk tilknytning til partiet Venstre, med sosialliberale verdier, men også disse anser seg som partipolitisk uavhengig (NOU 2000:118).

Likhetstrekkene for de ulike mediene er at de fleste avisene beholder sin ideologiske forankring, selv om de anser seg som partipolitisk uavhengig. Nøkkelordet her er altså parti. De liberale har gjennom tidene vært aviser knyttet til Orkla Dagspresse (i dag kjøpt opp av A-pressen/Amedia (Garvik 2019)). Disse avisene var delt på et liberalt og et konservativt tankesett, og som nevnt ovenfor, støttet de partiene Venstre og Høyre (NOU 2000:15, 119). Orkla-avisene lagde aldri mønstervedtekter slik som A-pressen gjorde, og da A-pressen kjøpte opp Orkla-avisene ble kutymen å ikke ta i bruk disse mønstervedtektene, men å fortsette på samme måte som de gjorde før. Som nummer tre av de store aviskonsernene finnes Schibsted. Også de stiller seg bak tanken om å være partipolitisk uavhengig. Verdens Gang (VG) er blant avisene som er eid av dette konsernet, og hevder selv å være en avis som «*bygger på humanistiske idealer og grunnleggende demokratiske verdier*» (NOU 2000:15, 119). I motsetning til enkelte andre aviser er dette en avis som skal sørge for å dekke hele landet. Men hva med de avisene som ikke tilhører noen «avisfamilie»? Fædrelandsvennen, for eksempel, er en regionavis i Kristiansand. Denne avisen hevder seg som en «*uavhengig*

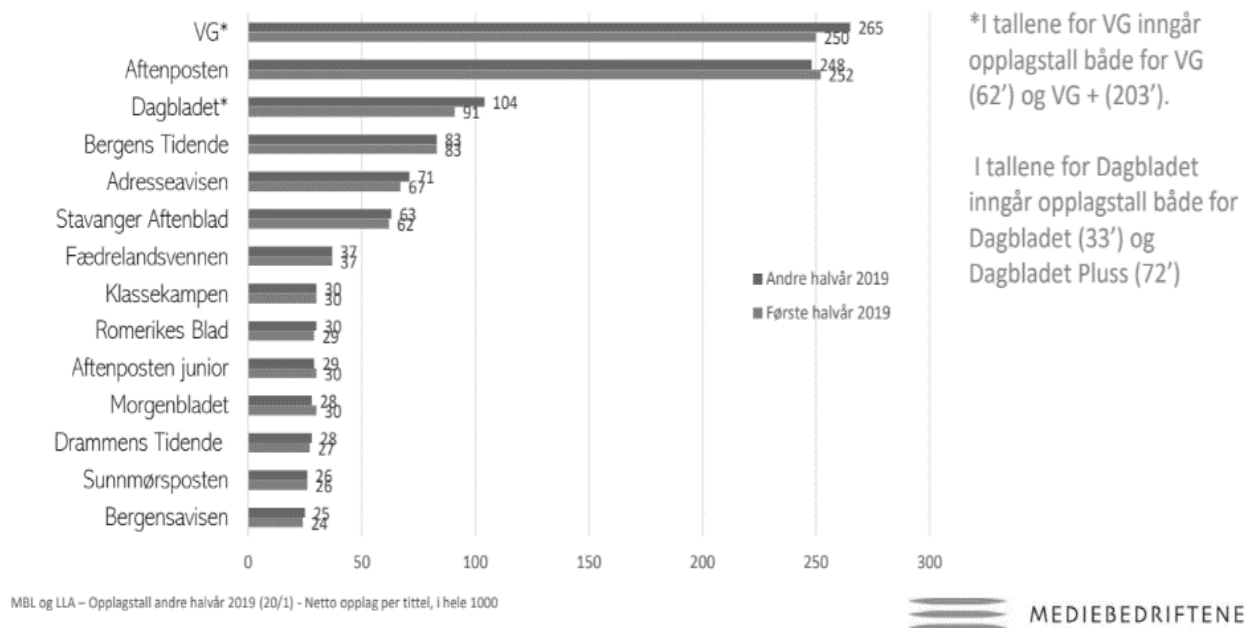
liberal abonnementsavis for Sørlandet» (NOU 2000:15, 120). At avisen anser seg som liberal betyr at de «*bekjenner seg til klassiske liberale og humanistiske idealer om demokrati, tanke-, møte- og ytringsfrihet, til åpenhet i styringssystemet, til solidaritet, likestilling og toleranse*» (NOU 2000:15, 120). Videre har den politiske avdelingen i hver avis som oppgave å prøve å følge disse retningslinjene, hvor demokratiske verdier står sterkt i fokus, samtidig som de skal være aktive i samfunnsdebatten.

3.2.3 Norges største mediehus – presentasjon av valg av medier

Mediebedriftenes Landsforbund (MBL) legger hvert år frem en statistikk om Norges største mediehus. Bransje- og utgiverorganisasjonen jobber ut ifra det de hevder er grunnsteiner i et demokratisk og åpent samfunn, og har som formål å både styrke og verne om medienes ytringsfrihet, pressefrihet og informasjonsfrihet. Videre er de opptatt av å opprettholde «*et mangfold av økonomisk sunne medier*» (*Mediebedriftene*, s.v «Om Mediebedriftenes landsforening»). Organisasjonen har 325 medlemmer, som inneholder avisbedrifter, trykkerier, distribusjonsselskaper, fjernsynsselskaper, annonsesamkjøringer, konserner og driftselskaper, ukepresse, multimedieselskaper, pressebyråer, reklamebyråer, radioer, m.fl. Totalt representerer Mediebedriftenes Landsforening om lag 97 prosent av totalt antall avisopplag. Grunnet det høye antallet av representasjon er det rimelig å bruke MBL som kilde for valg av medier for å gjennomføre emnemodelleringen. Og det er også derfor det er foreningens rammebetingelser som er lagt som grunnstein for hva som er formålet med mediedekningen.

Ved utregning av Norges største mediehus tar statistikken til MBL utgangspunkt i «digital total» og «papir total» som utgjør «netto opplagstall» (*Mediebedriftene*, s.v. «Opplagstall 19/2»). Dette betyr at det er kombinert alt av digitale aviser og papiraviser, samt abonnement og løssalg fra begge.

Figur 1: Opplag: Norges største aviser – over 25 000



(Mediebedriftene, s.v. «Opplagstall 20/1»).

Valget av medietypen avis er valgt på bakgrunn av at det er fundamentale likheter i måten de driver journalistikk på, samt at aviser er trykket skrift, som er nødvendig for den valgte metoden for analysen. Videre ville det absolutt enkleste være å velge ut de ti øverste avisene på listen som vises i figur 1, men for å sørge for et representativt utvalg ble det gjort enkelte endringer i utvelgelsen av hvilke ti aviser som ble en del av analysen. Utvelgelsen av mediene var ikke tilfeldig, og det ble tenkt nøye gjennom hvilke aviser som kunne gi et grunnlag for å representere hele landet, uten å i for stor grad bli påvirket av lokale faktorer. Listen viser 14 av de største avisene, og utvalget viser aviser både på nasjonalt og regionalt nivå. Målet er å analysere hvilke emner mediene har tatt for seg under mediedekningen av metoo-skandalen, og antagelsen er at ved å undersøke de avisene med et størst antall lesere, øker både representasjonen av de emnene som kommer fram, og det vil gjøre det lettere å generalisere ulike aspekter som kommer fram av analysen. Et høyt antall lesere blir i denne antagelsen altså assosiert med at informasjonen avisene kommer med blir spredt til et stort antall personer, og at man derfor kan anta at ulik informasjon som avisen deler har stor påvirkningskraft på mennesker over hele Norge.

Ved utvalget av avisene valgte jeg først og fremst bort Aftenposten Junior, da dette er en avis som er opprettet for å gi et nyhetsbilde som barn enkelt kan forstå (Aftenposten Junior 2020). I tillegg vil en tilnærming til temaet om politiske skandaler som er laget for barn, og da særlig skandaler som inneholder sex, ikke være representativ for å få et inntrykk av hvordan mediedekningen generelt har vært. Videre er det for å sørge for en landsdekkende regional dekning valgt å ta bort Drammens Tidene, som er en regionsavis med nær tilknytning til Buskerud og nordre Vestfold (Smith-Meyer og Pettersen 2016), til fordel for Nord-Norges største avis, Nordlys (Tjelmeland og Norsk presses historie 2019). Ved å bytte ut de to avisene dekkes et større område av Norge, noe som står i stil til målet som er nevnt ovenfor. Romerikes Blad er videre byttet ut til fordel for Morgenbladet da leserantallet blant de to avisene ikke er av betydelig forskjell, og fordi Morgenbladet blir ansett som en riksdekkende avis, i motsetning til Romerikes Blad som dekker regionen i den nordøstlige delen av fylket Viken utenfor Oslo (Pettersen 2017). Til sist er det tatt en vurdering av den partipolitiske tilhørigheten til avisene. Det er innledningsvis nevnt at de fleste aviser helst vil hevde at de står som «politisk uavhengig», men det blir også vist til at dette ikke er helt troverdig. Med VG, som den store nasjonale avisen, er det vanskelig å finne noen tilknytning til noen parti. Også Fædrelandsvennen kommer seg unna kilder om partipolitisk historie. Aftenposten, Adresseavisen og Morgenbladet er de avisene som kan knyttes til den konservative siden (Pettersen 2020), mens Dagbladet, Bergens Tidende og Stavanger Aftenblad har røtter som trekker seg tilbake til den liberale siden (Pettersen 2019; Pettersen 2017). Klassekampen er klar på at de er en uavhengig radikal avis (Pettersen 2016), mens Nordlys har bakgrunn fra arbeidergruppen (Tjelmeland og Norsk presses historie 2019).

Avisene som er med i analysen er dermed VG, Aftenposten, Dagbladet, Bergens Tidende, Adresseavisen, Stavanger Aftenblad, Fædrelandsvennen, Klassekampen, Morgenbladet og Nordlys.

Kapittel 4: Hvorfor metoo

4.1 Emneknaggskandalen metoo

Seksuell vold er et polariserende emne (Fileborn og Loney-Howes 2019, 1). På den ene siden fører seksuell vold til forargelse både for publikum og politikere, og på den andre siden møter ofrene for seksuell vold rutinemessig granskning av sine venner, familie, politi og publikum. Mange blir beskyldt for å lyve om opplevelsene sine, og andre for å ikke være «autentiske» ofre eller traumatisert nok. Noen får også skylden for at de blir overfalt; hvor de «ba om det». Feministiske aktivister har lenge søkt å utfordre disse synspunktene, samt tanker og antagelser om at voldtekt, seksuelle overgrep og seksuell trakassering er produkter fra tilfeldige handlinger gjennomført av enkeltmennesker som må bli sett på som «syke» eller «sosiale avvik». Det hevdes også at ofrene som oftest er ukjent for dem som gjennomfører disse overgrepene (Fileborn og Loney-Howes 2019, 1). Flere feminister hevder derimot at disse handlingene er en refleksjon av en voldskultur som refererer til de sosiale, kulturelle og politiske prosessene som kondolerer vold mot kvinner, men som også klandrer kvinner hvis og når vold blir utført mot dem. Flommen av deltakelse i metoo bekrefter offentlig hvor utbredt seksuelle overgrep og trakassering faktisk er, at de fleste overlevende kjenner den fornærmede, og at disse opplevelsene er rutinemessige og normaliserte som bekrefter mange feministiske argumenter om «voldskulturen» (Fileborn og Loney-Howes 2019, 2).

En metoo-sak handler om seksuell trakassering, som ifølge loven omhandler *«enhver form for uønsket seksuell oppmerksomhet som har som formål eller virkning å være krenkende, skremmende, fiendtlig, nedverdiggende, ydmykende eller plagsom»*. I «seksuell» ligger at det spiller på kropp, kjønn og seksualitet. «Oppmerksomhet» omfatter både verbal og fysisk atferd» (Orgeret 2019). De sentrale komponentene i metoo-kampanjen er kjønn og makt, og målet er å få folk til å forstå hvordan makt ofte blir misbrukt og hvordan en asymmetrisk relasjon kan øke muligheten for maktmisbruk. En metoo-handling kan variere i alvorlighetsgrad, men ethvert overtramp av seksuell karakter kan ansees som en metoo-handling.

I oktober 2017 eksploderte det for emneknaggen «metoo» på sosiale medier (Orgeret 2019). I kjølevannet av en streng seksuell trakassering og overgrepspåstand mot den høyprofilerte Hollywood-produsenten Harvey Weinstein, tok skuespiller Alyssa Milano til Twitter og oppfordret kvinner til å dele sine egne erfaringer om seksuell vold ved å legge ved «metoo» i innleggene sine (Fileborn og Loney-Howes 2019, 3). «*Hvis alle kvinner som har blitt seksuelt trakassert eller overfalt skriver «Metoo» som status, kan vi gi folk en følelse av størrelsen på problemet*» var beskjednen Milano skrev på sin twitter-konto. Og responsen var uten tvil fenomenal. Flere millioner av enkeltpersoner delte historier og de redaksjonelle mediene dekket en lang rekke av dem, hvor kampanjen også etter kort tid fikk preg av å være en global bevegelse. Hovedfokuset i kampanjen var seksuell trakassering fra overordnede personer i arbeidslivet, samt andre lignende maktposisjoner (Orgeret 2019). Som følge av de ulike delingene på sosiale medier og flere varslingssaker måtte en rekke profilerte mannlige skuespillere, produsenter, næringslivsledere og politikere trekke seg fra sitt arbeid og sine verv. Metoo-kampanjen anses å ha bidratt til lavere terskel for å si ifra om seksuell trakassering, og har vært en sterk bidragsyter til at det slås hardere ned på seksuell trakassering i arbeidslivet. Mer generelt kan kampanjen ha ført til en kulturell normendring i hvordan man ser på uønsket seksuell oppmerksomhet, ikke minst i relasjoner med skjev maktfordeling. Når emneknaggen ble brukt over 12 millioner ganger i løpet av det første døgnet ble størrelsen på problemet med seksuell vold i kvinner (og andres) liv svært tydelig.

Kampanjen på sosiale medier utviklet seg raskt på tvers av lokale sammenhenger, hvor forskjellige land utviklet emneknaggkampanjen med forskjellige slagord. I Norge var det til sammen 29 yrkesgrupper som hang seg på metoo-oppropet (Orgeret 2019). Blant annet brukte skuespillere #stilleførøpoptak, fagforeningene #ikketilforhandling, og eiendomsbransjen #nårgrunnmurensprekker. Kampanjen ble også brukt til å styrke andre saker, hvor de eksempelvis i Argentina utnyttet reaksjonene på metoo for å videreføre målet om å legalisere abort (Fileborn og Loney-Howes 2019, 3-4). At aktivistbevegelser tok særlig tak ved å vise støtte til de som opplevde seksuell trakassering og vold på arbeidsplassen, førte til at kampanjen videreutviklet seg fra «bare snakk» til «handling». metoo-bevegelsen kan videre anees som meget vellykket; en handling av solidaritet med og mellom overlevende som uten tvil er i en skala som vi ikke har vært vitne til før (Fileborn og Loney-Howes 2019, 5). Bevegelsen har generert konkrete handlinger og konsekvenser for noen gjerningsmenn, og drevet frem substansiell offentlig debatt om seksuell vold. I tillegg var bruken av sosiale

medier medvirkende til spisspolitisk diskusjon på moderne og nyanserte måter. Likevel ville det å stoppe her føre til en altfor optimistisk og forenklet gjengivelse av kampanjen. Det krever en mer kritisk utpakking av bevegelsen i dialog med bredere diskusjon knyttet til kjønn, seksualitet, rase, geografi, jus, politikk og historie for å forstå det forstyrrende potensialet for metoo. Og mange undrer på om man faktisk kan forvente at noe vil endre seg eller være annerledes framover.

Konsekvensen av å snakke ut om seksuell trakassering genererer flere utfordringer i en globalisert verden som er veldig avhengig av sosiale medier for deling av erfaring og generere støtte til kampanjer som metoo. Å skaffe seg et digitalt munnstykke betyr ikke nødvendigvis at enkeltpersoner og grupper vil bli sett eller hørt – men fordelene er at noen vil sitte i bedre posisjon til å bli lagt merke til enn andre. Denne gruppen må bli trodd og hørt, som fører til validering av andre (Fileborn og Loney-Howes 2019, 9). Utover spørsmålet om hvem som har tilgang til digitale plattformer ligger spørsmålet om hvem som i det hele tatt er i stand til å snakke og under hvilke forhold. Enkelte privilegier og tilgang til ressurser gjør det mulig for noen overlevende å snakke om sine opplevelser og i noen tilfeller også navngi gjerningsmennene. Bevegelsens tilsynelatende verdi for nyhets-mediene genererte en substansiell og pågående medierapportering i både tradisjonelle og online nyhetsmedier (Fileborn og Loney-Howes 2019, 14-15), hvor blant annet norske medier har vært svært viktige for metoo-kampanjen i Norge (Orgeret 2019).

4.2 metoo, norske mediehus og politiske partier

I løpet av 3 måneder ble metoo-kampanjen omtalt over 4 000 ganger i norske medier (Atekst 2020), og innen 2018 var saken omtalt over 15 000 ganger (Orgeret 2019). De norske mediehusene har vært en viktig brikke i utviklingen av bevegelsen, men har også vært nødt til å møte seg selv i døren et par ganger da det kom til å handle om trakassering innenfor egne vegger i ulike mediehus. Om man sammenligner med Sverige kan man si at navngivning av både varslere og de som ble anklaget har vært holdt på et relativt lavt nivå, men likevel har mediedekningen vært preget av en rekke enkelthistorier (Orgeret 2019). I hvilke grad enkeltpersoner blir dratt fram er noe av det undersøkelsen vil kunne finne svar på, og det blir på forhånd antatt at de norske navnene som er omtalt i stor grad vil være forbeholdt navn i

politikken. 13. desember 2017 gikk metoo-skandalen over i en politisk fase i Norge, hvor Dagens Næringsliv meldte om at det hadde kommet en intern varsling om at Arbeiderpartiets Trond Giske hadde oppført seg upassende mot en yngre kvinne i partiet. Det hele endte med at Giske trakk seg fra sin stilling som nestleder i partiet 7. januar 2018. Men det stoppet ikke bare her, og i løpet av januar 2018 trakk ytterligere to politikere seg fra sine politiske verv, og flere har også måtte gå ved senere anledninger. Dette var leder for Unge Høyre, Kristian Tønning Riise, og stortingsrepresentant Ulf Leirstein (Orgeret 2019).

Dagsavisen fulgte opp anklagene ved å gjennomføre en spørreundersøkelse hvor de spurte alle partiene på Stortinget om antall varslinger som hvert parti hadde mottatt om seksuell trakassering eller upassende oppførsel og holdninger de siste fem årene (Prestgård 2018). Undersøkelsen viste at Høyre og Unge Høyre hadde mottatt 21 varsler siden bare siden 11. januar 2018, Arbeiderpartiet hadde mottatt et par og tjue varsler siden midten av desember, og FrP hevdet å bare ha mottatt åtte varsler i samme tidsrom, noe som kom som en overraskelse på reporterens siden de hadde hevdet å bare ha mottatt tre varsler de siste fem årene ved en anledning like før jul (Prestgård 2018). Videre hadde Senterpartiet fem saker de siste fem årene, Kristelig Folkeparti hadde mottatt tre saker, Venstre hadde tre varslet siden desember 2017, og før den tid hadde de seks varsler med varierende alvorlighetsgrad. Til sist hadde Sosialistisk Venstreparti to varsler siden november, mens ungdomsorganisasjonen deres hadde tre varsler i løpet av en toårs-periode. Miljøpartiet De Grønne hadde mottatt fem varslet siden 2014, og Rødt hadde fire varsler spredt utover en femårsperiode.

Videre viste Medieundersøkelsen 2018 at så mange som 82 prosent av leserne og 93 prosent av journalistene i Norge hevdet at metoo hadde vært viktig i kampen for å begrense seksuell trakassering, og forsvarsminister Frank Bakke-Jensen (per 2018) uttalte at «*Metoo er en kulturrevolusjon. Det som stilltiende ble godtatt i går er uhørt i morgen*». En slik uttalelse gir tydelig indikasjon på at også politikerne tok metoo-kampanjen på alvor, og det vil være rimelig å anta at en slik uttalelse gir et bilde av politikeres holdning generelt, framfor hans holdning som enkeltmann.

4.3 Oppfyller metoo kriteriene til en politisk skandale?

De fem kriteriene for å være en politisk skandale ble oppgitt av Midtbø (2012, 152-153), og man kan vel trygt si at metoo-kampanjen førte til skandaler både i politikken og næringslivet. Metoo-kampanjen viste klare tegn til (1) overtredelse av normer, og (2) skandalen ble i all sin forstand offentlig. Det var ellers (3) en rekke fornærmede i sammenheng med normbruddet, og (4) omdømmet til enkelte politikere ble betydelig svertet under kampanjen. Med tanke på politikere som trakk seg, og en stadig forventning om at det kunne komme fram flere anklager mot enten samme politiker eller andre politikere var det også en høy grad av (5) spenning og usikkerhet innblandet i mediedekningen av metoo.

Kapittel 5: Metode

En tekstanalyse kan sees på som en oppdagelse av ny, tidligere ukjent informasjon fra ustrukturerte tekstdata (Anandarajan et al. 2019, 1). Som et resultat av de mange historiene som kom fram om metoo-skandalen, eksploderte medieoppslagene. Fra oktober 2017 og fram til juni 2020 er det skrevet over 30 000 artikler om metoo (Atekst 2020). En emnemodellering av et utvalg av disse artiklene vil som tidligere nevnt kunne gi svar på hvilke emner som har stått i fokus ved å strukturere et utvalg av mediedekningen. Analysen vil basere seg på tekstbryting (text mining) og emnemodellering (topic modeling).

Tekstbryting handler om prosessen som gjennomføres for å utvinne data fra datasamlinger med tekst (Cai og Sun 2009, 3061), og emnemodellering er en type statistisk modellering for å oppdage de abstrakte "emnene" som oppstår i en samling av dokumenter. Dette kan enten være enkeltmedlemskapsmodeller, der dokumenter hører til ett enkelt emne, eller modeller med blandet medlemskap, der dokumenter er en blanding av flere emner (Anandarajan et al. 2019, 117). Språk er sammensatt, men ikke all språkets kompleksitet er nødvendig for å kunne analysere tekster effektivt (Grimmer og Stewart 2013, 272). Alle trinnene er designet for å beholde informasjon som vil bli brukt i analysen, samtidig som den kaster informasjon som sannsynligvis vil være lite nyttig eller for komplisert til å bli brukt i en statistisk modell. Dokumentene blir antatt å ligge i en uvilkårlig rekkefølge, og dette må man derfor ta hensyn til i analysen.

5.1 Fire prinsipper for automatisert tekstanalyse

Først vil jeg presentere fire prinsipper som ligger til grunn for enhver kvantitativ tekstanalyse: (1) *Alle kvantitative språkmodeller er feil – men noen er nyttige*: Datagenereringsprosessen for enhver tekst er et mysterium. Hvis en setning har en komplisert avhengighetsstruktur, kan betydningen endres drastisk om man inkluderer nye ord, og det samme kan skje med setningens kontekst. Kompleksiteten i språket innebærer at ikke alle metoder nødvendigvis klarer å gi en nøyaktig redegjørelse for den datagenererende prosessen som brukes til å produsere tekster (Grimmer og Stewart 2013, 270). Automatiserte innholdsanalysemetoder bruker innsiktsfulle, men feile modeller av politisk tekst for å hjelpe forskere til å finne

slutninger fra dataene sine. Det vanlige rådet for modellbygging av kausal inferens (årsakssammenheng) er at det er viktig å inkludere alle relevante funksjoner i den datagenererende prosessen, men dette rådet gjelder ikke for kvantitative tekstanalysemetoder.

(2) *Kvantitative metoder forsterker mennesker, det erstatter dem ikke*: framfor å erstatte mennesker, forsterker datamaskiner menneskelige evner (Grimmer og Stewart 2013, 270).

Den mest produktive undersøkelseslinjen er å identifisere den beste måten å bruke både mennesker og automatiserte metoder for å analysere tekster. (3) *Det er ingen verdens beste metode for automatisert tekstanalyse*: Det er forskjellige forskningsspørsmål og design som innebærer at forskjellige typer modeller vil være nyttige. Grimmer og Stewart (2013) hevder at en av de viktigste undersøkelseslinjene er å identifisere effektive måter å kombinere tidligere distinkte metoder for kvantitativ tekstanalyse og å videreutvikle disse (Grimmer og Stewart 2013, 270). (4) *Validering*: Det påhviler forskeren å validere bruken av tekstanalyser (Grimmer og Stewart 2013, 271). For å validere dataene fra en metode uten tilsyn, slik som kvantitative metoder er, må man kombinere eksperimentelle, materielle og statistiske bevis for å påvise at tiltakene er like konseptuelt gyldige som tiltak fra en tilsvarende overvåket metode, som vil være en eventuell kvalitativ metode.

5.2 Betydningen av «språk»

Språk er mediet for politikk og politisk konflikt (Grimmer og Stewart 2013, 267). Gjennom språk signaliseres motivasjoner, og på politisk plan formulerer individuelle kandidater og politiske partier sine synspunkter gjennom partiplattformen og manifesten. Disse eksemplene viser at for å forstå hva politikk handler om, trenger vi å vite hva politiske aktører sier og skriver. Å erkjenne at språk er sentralt i studiet av politikk er ikke nytt, tvert imot. Politikere har lenge erkjent at mye av politikken kommer til uttrykk i ord, men forskere har hatt det vanskelig når de har tatt i bruk tekster for å gjøre slutninger om politikk: Det primære problemet er volum; det er rett og slett for mange tekster.

Til tross for et stort antall tekster og utallige uttrykk, skal jeg gjennomføre en systematisk analyse av store tekstsamlinger uten massiv støtte, for å finne svar på min problemstilling. Avisartiklene har vært en av plattformene for både ofrene og politikerne for å nå ut med sine

budskap. Mens ofrene har fått fortelle sine opplevelser, har det også vært rom for uttalelser fra de aktuelle politikerne som hendelsene har handlet om.

5.3 Innhenting av tekst

Ved gjennomføring av tekstanalyse kreves det filer som er lagret i rent tekstformat (Grimmer og Stewart 2013, 271). Den enkleste måten å skaffe tekst på er fra elektroniske databaser av artikler. I denne oppgaven er dataene hentet fra mediearkivet Atekst hos retriever.no. Abonnementet som Universitetet i Bergen har hos dem gir tilgang til å søke i mediearkivet etter gamle artikler og å laste dem ned. I utgangspunktet ble alt av avisoppslag som inneholdt metoo, samt alle de ni partiene som holder til på Stortinget og i regjering hentet fram (Søkeordene inneholdt; Arbeiderpartiet, Ap, Fremskrittspartiet, Frp, Høyre, H, Kristelig Folkeparti, KrF, Miljøpartiet De Grønne MDG, Rødt, R, Senterpartiet, Sp, Sosialistisk Venstreparti, SV, Venstre, V). Tanken ved å hente fram artikler som hovedsakelig inneholdt et av de politiske partiene var for å sikre at utvalget inneholdt metoo-handlinger fra norsk politikk. Alle artiklene måtte ikke inneholde alle partiene, men artiklene måtte inneholde minst et av de nevnte partiene. Etter at alle artikler var hentet inn, fikk jeg dessverre tekniske problemer. Først viste det seg at både Dagbladet pluss og VG pluss ikke viste fullstendige artikler, men bare overskrift, samt et lite avsnitt på et par setninger. Etter vurdering fant jeg fram til at dette kunne bli misvisende, da det i dag er en trend å bruke overskrifter og lokkende små avsnitt for å få folk til å 1) trykke seg inn på artikkelen og 2) velge å kjøpe abonnement til avisen. For å få svar som kunne representere mediene på en troverdig måte framfor å få med en rekke «clickbaits» til å dominere utfallet, ble det derfor vurdert at disse artiklene ikke var representative i en slik undersøkelse som jeg ville gjennomføre, og de ble dermed fjernet. Det andre problemet skulle vise seg å være at datamaskinen som ble brukt i gjennomføringen av undersøkelsen ikke hadde kapasitet til å laste opp alle artiklene i R. Siden jeg ikke hadde mulighet til å få tilgang på noen datamaskin med større kapasitet, ble det derfor nødvendig å kutte deler av artiklene, og konklusjonen ble å hente ut oppslag fra de plattformene i avisene som hadde størst lesertall. Dette ble nettavisene til Adresseavisen, Bergens Tidene, Dagbladet, Fedrelandsvennen, Nordlys, Stavanger Aftenblad og VG, samt Aftenpostens betalte nettavis, og papiravisene til Klassekampen og Morgenbladet (*Mediebedriftene*, s.v. «Lesertall»).

5.4 Tekstbryting

Målet med tekstbryting er å oppdage kunnskap, informasjon eller mønstre fra tekstdata, som i sin natur er ustrukturert eller semistrukturert. Tekstbryting bygger på informasjonsinnhenting, datamining, maskinlæring, statistikk og beregningsspråklighet (Ignatow og Mihaleca 2018, 3). Teknologien blir stort sett brukt av myndigheter, i akademisk forskning og i næringslivet, og på grunn av at de fleste av verdens digitale opplysninger lagres som ustrukturerte data, anerkjennes disse teknologiene og har en høy potensiell kommersiell verdi.

I databehandlingen er oppgaven å filtrere ut støyende informasjon for å sørge for at de følgende trinnene blir mer nøyaktig (Cai og Sun 2009, 3062), og en slik forbehandling av tekst er en strek bidragsyter for at utfallet av tekstbrytingen skal bli en suksess. Å skape et uttrykk eller å dele innspillene inn i ord er et viktig førsteskritt som i utgangspunktet kan virke enkelt, men som er full av små avgjørelser: hvordan man skal takle apostrofer og bindestreker, store bokstaver, tegnsetting, tall, alfanumeriske strenger, om mengden hvit plass er betydelig, om du skal pålegge symboler en maksimal lengde, hva du skal gjøre med tegn som ikke skriver ut, og så videre. Uttrykk kan standardiseres ved å bruke en ordbok for å kartlegge forskjellige, men likeverdige varianter av et begrep til en enkelt form. Når dataene har fått et uttrykk er det vanligvis nødvendig å gjennomføre en syntaktisk behandling (Cai og Sun 2009, 3062). Det vil si en operasjon for å fjerne stoppeord. En annen er å identifisere vanlige setninger og kartlegge dem i enkeltfunksjoner.

5.4.1 Utfordringer ved tekstbryting

Tekstanalyser og tekstbryting har gitt nye måter å forske på, men selv om det er en overbevisning om at dette er en effektiv metode, er det likevel viktig å anerkjenne noen av risikoene og ulempene ved å bruke slike metoder i sin forskning (Ignatow og Mihaleca 2017, 3). Samlet internett, store data (stor N), og billige, tilgjengelig og kraftige programvarer som blant annet R (som blir benyttet i denne forskningen) har potensialet til å revolusjonere tekstanalyse i samfunnsvitenskapen (Ignatow og Mihaleca 2017, 4). Men mens forskere som er interessert i å skaffe og analysere store mengder tekstdata produsert av individer, grupper,

samfunn og organisasjoner har lavere inngangsbarrierer enn noen gang før, står de også overfor nye utfordringer med å maksimere forskningsverdien til disse nye datakildene og metodene. En første stor utfordring innebærer å finne måter å realisere potensialet i store samlinger av tekstdata uten å skape tvil om den store størrelsen. Selv om store samlinger av tekstdata er veldig bra for å oppdage subtile korrelasjoner som analyser av mindre datasett kan gå glipp av, kan data i seg selv aldri fortelle forskere hvilke korrelasjoner som er meningsfulle. For det andre foreligger det en risiko for falsk korrelasjon. Man kan med all sannsynlighet finne korrelasjon mellom variabler, men det er viktig å være kritisk til tross for at de tilsynelatende er statistisk signifikante. For validering av korrelasjonen burde det undersøkes for empirisk forbindelse mellom variablene, da størrelsen på dataene kan forsterke slike feil (Ignatow og Mihalcea 2017, 5). Det tredje punktet å være obs på er utvalgsskjevhet og «selvsøkende» prøver. Det er en overdreven tillit til det engelske språket på bekostning av andre språk, noe som kan få stor betydning for analysen i denne oppgaven, da all tekst er på norsk. Videre er det også viktig å være forsiktig slik at det ikke forveksles mellom skriftlige uttrykk og faktisk oppførsel. Når algoritmene blir utviklet må man være klar over de tekniske utfordringene som vanlig språklig praksis gir. Dette kan for eksempel være bruk av forkortelser, ironi, humor og sarkasme.

5.4.2 Grunnleggende tekstbehandling

Grunnleggende tekstbehandling er ofte det første trinnet i forskningsstudier som involverer språklige innspill (Ignatow og Mihalcea 2017, 2). Tekstbryting er prosessen med å identifisere ordene i inngangssekvensen til ord, hovedsakelig ved å skille ord, men også ved å identifisere sammentrekninger, forkortelser osv. (Ignatow og Mihalcea 2017, 2-3). Det kan virke som en triviell prosess med det første, men det er noen tilfeller som krever spesiell oppmerksomhet. For eksempel må man på et tidspunkt skille mellom slutten på en setning og markeringer av forkortelser. Selv om tekstbryting stort sett er spørkuavhengig, kan flere av de spesielle sakene som må håndteres for korrekt «nedbrytning» være språkspesifikke. For eksempel er ofte forkortelser og sammentrekninger avhengig av språk, og man må derfor sette sammen en liste over slike ord for å sikre at disse ordene blir håndtert riktig. Det samme gjelder apostrofer og bindestreker. Stoppeord, også referert til som funksjonsord eller ord i lukket klasse, består av høyfrekvente ord som inkluderer pronomen (f.eks. jeg, vi, oss), determinativ (som det, en eller et), preposisjoner (som i eller på), osv. (Ignatow og Mihalcea 2017, 3). For

noen oppgaver kan stoppeord være nyttige, som for eksempel kjønnspronomen (Grimmer og Stewart 2013, 273), samt har det blitt funnet at de kan gi betydelig innsikt i menneskers personligheter og atferd (Pennebaker og King 1999, 1300). Uansett er det viktig å ha midler til å identifisere stoppeordene i en tekst. Generelt gjøres det ved å bruke en allerede eksisterende liste over stoppeord sammen med en effektiv oppslagsalgoritme (Ignatow og Mihalcea 2017, 3-4). Målet er å fjerne ord som i bunn og grunn bare er til stede for den grammatiske funksjonen, samt veldig uvanlige ord. Man fjerner vanligvis ord som vises i mindre enn 1% og mer enn 99% av dokumentene i korpuset, selv om disse valgene må gjøres betinget både av mangfold i ordforrådet, gjennomsnittslengden på dokumentet, og størrelsen på korpuset (Quinn et al. 2010).

Mange ord på det naturlige språket er relatert til hverandre, men de kan også ha mange forskjellige skriveformer som gjør at deres betydning kan være tvetydig om de sees alene. Mens noen av disse forholdene er av semantisk karakter og krever ordbokkunnskap, er det også mange relasjoner som lettere kan fanges opp gjennom enklere former for strenganalyse (Ignatow og Mihalcea 2017, 4). Den enkleste måten å identifisere den vanlige stammen for flere ord er gjennom prosessen som kalles «stemming». Enkelt sagt så bruker stemming et sett med regler på et inntastingsord for å fjerne suffikser og prefikser (for- og etterstavelser på ord) og oppnå stammen av ordet, som videre vil bli delt med andre relaterte ord. En enkel liste med ord, unigrams, er ofte tilstrekkelig for å formidle den generelle betydningen av tekst, men det er også muligheter for å inkludere bigrams (ord i par) eller trigrams (lengre sammensetninger av ord) i analysen. Tanken med å bruke en slik metode er for å unngå misforståelser i analysen og reduserer kompleksiteten ved å kartlegge ord som refererer til det samme grunnleggende konseptet til en enkel rot (Manning et al. 2008, 32). Til tross for disse fordelene viser det seg at det ikke alltid er like lurt å bruke stemming. Basert på observasjoner gjort av Farrar og Hayes (2019) ser det ut til at det ved enkelte tilfeller vil lønne seg å la være å bruke teknikken, til tross for at det i mange tilfeller har vist seg å være til stor nytte (Farrar og Hayes 2019, 42). Gevinsten fra stemming avtar ved bruk av større datasett, og ettersom modellen med mindre sannsynlighet klarer tilordne seg varianter av en felles stamme til de forskjellige temaene. I analysen ble begge metodene gjennomført for å se hva som kunne gi best resultat, og det viste seg til slutt at det klart beste resultatet kom fram ved å la være å bruke «stemming». Da stemmingen ble anvendt forsvant endinger på ord som hadde vesentlig betydning for oppgaven og det dukket opp en rekke stavelser uten mål og mening.

Hovedårsaken for at valget endte på å la være å stemme ordene var fordi funksjonen mistet sin verdi da ordboken som ble anvendt var engelsk, og det viste seg videre at det ikke fantes noen ordbok som bare baserte seg på det norske språket. Videre ble også bredden på emnene ansett å være av stor betydning for problemstillingen, og svarene som kom fram ved å gjennomføre analysen uten bruk av stemming var langt mer intuitive enn den med bruk av stemming.

En av de enkleste analysene man kan gjøre på en hvilken som helst tekstsamling er å telle ord og finne ordene med høyest frekvens (Ignatow og Mihalcea 2017, 5). Til tross for dets kompleksitet er det naturlige språket veldig forutsigbart: man kan for eksempel komme med en god gjetning om hva som vil være de hyppigste ordene i en hvilken som helst tekstsamling eller komme med forutsigelser om hva som blir størrelsen på ordforrådet (dvs. antall unike ord) i en ny samling. Tekstklynging er videre læring uten tilsyn der det ikke er forhåndsdefinerte kategorier eller klasse, men det søkes grupper av dokumenter som hører sammen (Cai og Sun 2009, 3063). Klyngepaneler krever ikke at opplæringsdata blir forhåndsklassifisert, og algoritmene i seg selv er generelt langt mer beregningsintensive enn metoder av tilsyn. Vanligvis anses kvaliteten på klynger som bedre hvis innholdet i dokumentene i en klynge er mer like og mellom klyngene mer forskjellige. Klynging er en grunnleggende oppgave innen tekstbryting, hvor effektiv organisering, navigasjon og oppsummering av tekst kan oppnås (Al Sumait og Domeniconi 2008, 87). Klynge av tekster byr på vanskelige utfordringer på grunn av sparsiteten og den høye dimensjonen av tekstdata, og den komplekse semantikken i det naturlige språket. Ordskyer gir en visuell visning av ordtelling fra en eller flere tekster (Ignatow og Mihalcea 2017, 11). Jo hyppigere et ord vises, jo større vises ordet i et «word cloud visual». Til tross for at ordskyene viser dramatiske bilder, har bruken blitt noe kritisert. En bekymring er nemlig at ordskyene stoler helt og holdent på ordfrekvensen, og gir ikke kontekst for leseren til å forstå hvordan ordene er brukt i teksten.

Ordskyene er ikke i stand til å skille mellom ord med positiv eller negativ konnotasjoner, og de kan være visuelt misvisende fordi lengre ord tar mer plass i skyen (Ignatow og Mihalcea 2017, 11). Til tross for disse bekymringene, kan ordskyer være et praktisk verktøy når de blir brukt sparsomt og med begrensning. Selv om de ikke er veldig nyttige for komplekse

analyser, kan de brukes i prosjektets tidlige faser for å hjelpe forskere å identifisere nøkkelord i tekster eller sammenligne flere dokumenter. Sammen med skriftlig analyse og forklaring kan ordskyer brukes til å illustrere ideer eller temaer lekent for publikum (Ignatow og Mihalcea 2017, 11). Matriser er også verktøy for visualisering, og er ofte brukt til å fremme temaer (Ignatow og Mihalcea 2017, 12). Siden prosessen med å identifisere temaer krever minst en innledende analyse av teksten, er visualisering av temaer mer verdifulle i analyse- og rapporteringsfasen i et prosjekt enn i den utforskende fasen. Fordi forskere kan rangere temaer eller plassere dem i ikke-ordinære kategorier, gir visualisering av et korpus på tematisk nivå flere alternativer og dimensjoner for visuelle fremstillinger enn ord- eller setningsnivå. Matriser er et sett med tall som er ordnet i rader og/eller kolonner (Ignatow og Mihalcea 2017, 12). De fleste kvalitative programvarepakker har muligheten til å lage en matrise basert på temafrekvens med lenker til den tilsvarende teksten i programmet, men fordi matriser ikke gir historiene eller konteksten bak temaene de organiserer, ideelt når de lager matriser eller andre visualiseringer, bør forskere knytte de enkelte boksene til sitater som støtter leserens forståelse av temaet.

5.5 Følelsesanalyse

En viktig type informasjon som formidles i mange typer skriftlig og muntlig diskurs, er den mentale eller emosjonelle tilstanden til forfatteren (Ignatow og Mihalcea 2017, 3).

Nyhetsartikler rapporterer for eksempel ofte emosjonelle svar på en historie i tillegg til fakta. Å forstå følelser er et av de viktigste aspektene ved personlig utvikling, og man kan se på det som en nøkkel til utvikling av menneskelig intelligens (Cambria et al. 2017, 1).

Emosjonsbehandling er også viktig for den nær beslektede oppgaven polaritetsdeteksjon, som handler om å finne fram til positive og negative ord. Muligheten til å automatisk fange følelser fra allmenheten om sosiale begivenheter, politiske bevegelser, markedsføringskampanjer og produktpreferanser har skapt økende interesse både i det vitenskapelige samfunnet, næringslivet, for markedsføring, og for prediksjon i det finansielle markedet. Dette har ført til de nye feltene med affektiv databehandling og følelsesanalyse, som utnytter menneskelig datamaskininteraksjon, informasjonsinnhenting og sammensatt signalbehandling for å filtrere folks stemninger fra den stadig økende mengden sosiale data online.

De grunnleggende oppgavene til affektiv databehandling og følelsesanalyse er følelsesgjenkjenning og polaritetsdeteksjon (Cambria et al. 2017, 4). Mens førstnevnte fokuserer på å trekke ut et sett med følelsesetiketter, er sistnevnte vanligvis en binær klassifiseringsoppgave med utganger som «positive» versus «negative». Men de to oppgavene er sterkt innbyrdes relatert og avhengige av hverandre i den grad noen følelsesklassifiseringsmodeller behandler den som en unik oppgave ved å utlede polariteten knyttet til en setning direkte fra følelsene dette formidler. I mange tilfeller regnes faktisk følelsesgjenkjenning som en underoppgave av polaritetsdeteksjon. Hvis vi avstår fra antakelsen om at innspillene som blir undersøkt er meningsdyktige og det handler om ett enkelt tema eller element, oppstår nye utfordrende oppgaver, for eksempel subjektivitetsdeteksjon, identifisering av meningsmål og mer.

Eksisterende tilnærminger til affektiv databehandling og følelsesanalyse kan grupperes i tre hovedkategorier: kunnskapsbaserte teknikker, statistiske metoder og hybridtilnærminger (Cambria et al. 2017, 5). Statistisk metode, som er utgangspunktet for denne analysen, har vært populære for å påvirke klassifisering av tekster og har tidligere blitt brukt av en rekke forskere (Cambria et al. 2017, 6). Ved å mate en maskinlæringsalgoritme et stort treningskorpus av affektive kommenterte tekster, er det mulig for systemet å både lære den affektive verdien til påvirkende nøkkelord og å samtidig ta hensyn til andre vilkårlige nøkkelord, samt undersøke hyppighetsfrekvensen av ordene. Det første steget er å utvikle et subjektivitetsleksikon på ordnivå, en liste over nøkkelord som er samlet sammen fordi de har subjektive bruksområder, og så blir polaritetsinformasjon ofte lagt til i slike leksikon. Ord- og sansnivå-subjektivitetsleksikoner er viktige fordi de er nyttige ressurser for kontekstuell subjektivitetsanalyse, å kunne gjenkjenne og trekke ut private uttrykk i en faktisk tekst eller dialog. Vi kan bedømme teksters subjektivitet og polaritet på flere forskjellige nivåer. Disse leksikonene ble i analysen hentet fra pakken i R som heter tidytext. Denne pakken gir flere leksikoner, blant annet «bing» fra Bing Liu og samarbeidspartnere, og «nrc» fra Saif Mohammad og Peter Turney (Silge og Robinson 2020). Begge leksikonene er basert på unigrams, dvs. enkeltord. Leksikonene inneholder mange engelske ord, og grunnen til at det dessverre ikke er mulig å få et leksikon med norske ord er at det rett og slett ikke finnes ennå (Velldal 2017). Dette setter noen begrensinger for validering av bruken av følelsesanalyse, men blir likevel tatt med som en del av analysen da det gir tilstrekkelig informasjon som kan støtte oppom andre deler av analysen. Ordene i «bing»-leksikonet er tildelt en score for

positiv/negativ følelse, mens i «nrc»-leksikonet kategoriseres ord på binær måte (ja/nei) i kategorier av positive, negative, sinne, forventning, avsky, frykt, glede, tristhet, overraskelse og tillit. På dokumentnivå kan vi spørre om en tekst er meningsfull og i så fall om den hovedsakelig er positiv eller negativ (Ignatow og Mihalcea 2017, 3).

5.6 Emnemodellering

Emnemodellering er en nyttig teknikk når man skal identifisere hvilke temaer som diskuteres i tekstene, og hvordan temaene som diskuteres skifter over tid etter hvert som hendelser utspiller seg (Ignatow og Mihalcea 2017, 2). Det innebærer automatiserte prosedyrer for koding av samlinger av tekster i form av meningsfulle kategorier som representerer hovedtemaene som diskuteres i tekstene. Emnemodeller antar at betydninger er relasjonelle, og at betydningene knyttet til et samtaleemne kan forstås som et sett med ordklynger. De behandler tekster som det lingvister kaller «en pose med ord», og fanger ordforekomster uavhengig av syntaks, fortelling eller plassering i en tekst. Et emne kan sees på som en klynge av ord som har en tendens til å komme opp i en diskusjon, og som derfor forekommer hyppigere enn de eller ville gjort når et tema diskuteres.

I analysen blir det brukt to former for emnemodellering; latent dirichlet allocation (LDA) og strukturell emneanalyse (STM). LDA er den vanligste formen for emnemodellering og er en metode som antar at dokumenter og ordene i dem er avledet fra en «generativ sannsynlighetsmodell» (Kuhn 2018, 109). STM er videre en sannsynlig måte å beskrive dokumenter på i forhold til emner. Enkelte dokument er ofte knyttet til mange emner, og samtidig er emner til stede i mange dokumenter.

5.6.1 Latent dirichlet allocation

Latent dirichlet allocation (LDA) er en statistisk modell for språk introdusert av Blei, Ng og Jordan (Grimmer og Stewart 2013; Blei, Ng og Jordan 2003). Denne modellen antar at hvert dokument er en blanding av emner og er basert på ideen om at hver tekst i en tekstsamling tilsvarer en pose med ord produsert i henhold til en blanding av temaer som forfatteren hadde tenkt å diskutere. Hvert emne er en fordeling på alle observerte ord i tekstene slik at ord som

er sterkt assosiert med tekstens dominerende emner har større sjans for å bli inkludert i tekstens pose med ord. Målet med emnemodellering er å finne parametere til LDA-prosessen som har generert den endelige teksten eller tekstsamlingen, en prosess referert til som «inferens» i LDA-litteraturen. Blant resultatene fra inferensen er et sett emnefordelinger per ord som knytter en sannsynlighet til hvert emneordpar og et lignende sett med tekstdistribusjoner per emne som beskriver sannsynligheten for å velge et bestemt emne for hver spesifikke tekst. LDA er med andre ord en sannsynlighetsfunksjon over et sett med ord, og inferensen forteller oss hvor ofte et gitt ord brukes. Modellen undersøker en samling dokumenter for å lære hvilke ord som pleier å bli brukt i dokumentene. Dette er et klassisk klyngeproblem, hvor disse klyngene er temaene – et særegent element i LDA er at disse klyngene ikke er forskjellige. Det står ingen steder hva selve emnet er, estimerer av parametere til modellen gir oss data om temarepresentasjon i hvert dokument og i korpuset, og man må bruke intuisjon og finne logikken bak ordene som blir oppgitt i figurene (Kuhn 2018, 109). Emnemodellering er videre generelt mer induktiv enn de fleste andre tilnærminger til tekstanalyse (Ignatow og Mihalcea 2017, 2). I stedet for å starte med forhåndsdefinerte koder eller kategorier avledet fra teori, starter man med å spesifisere k , som er antall emner man ønsker at algoritmen skal finne. Valg av for få emner kan gi resultater som er for brede, mens hvis man øker mengden veldig vil dette føre til for mange små, overfløydige emner. Men når k er satt, identifiserer programvaren det angitte antallet emner, returnerer sannsynligheten for ord som brukes i et emne, og gir et regnskap for fordeling av disse emnene på tvers av tekstene.

I pakkene «`ldatuning`» og «`topicmodels`» finnes det noen hjelpemidler for å finne et rimelig antall k ; Griffiths2004, CaoJuan2009 og Arun2010 (Nidhi 2017). Dette er tre ulike funksjoner for å finne fram til antall latente emner, med noe ulik tilnærming. Enkelt forklart vil Griffiths2004 og CaoJuan2009 være modeller med algoritmer som søker null-punktet i en matrise, mens Arun2010 er en «maksimeringsmodell» som søker det antall modeller som ligger nærmest 1. Griffiths2004 bruker en metode for å undersøke konsekvensen av å endre antall emner, mens CaoJuan2009 velger den beste LDA-modellen basert på tetthet (Nidhi 2017). Arun2010 derimot observerer avvikerverdiene, og baserer seg på at denne verdien er høyere for et ikke-optimalt antall emner. Den laveste verdien vil derfor vise til det mest optimale antall emner.

5.6.2 Strukturell emnemodellering

En strukturell emnemodellering (STM) utvider LDA-rammeverket ved å gi mulighet for å finne korrelasjoner, altså sammenhenger, mellom emner (Kuhn 2018, 110). De kovariate dataene inkluderer metadata fordi dokumenter påvirker emnets utbredelse i dokumenter. Metadata er «*data om data*» (Gjersdal 2019), som betyr at det er informasjon som kan utdype de dataene vi allerede har. Emneord er et eksempel på metadata, og STM bruker da disse dokumentspesifikke dataene for å definere fordelinger av ordbruk i et emne. I denne typen av emnemodellering brukes en multinominal logit-modell for ordfordelinger hvor ordets forekomst er basert på emne, samt dokumentoverføringsdata og emne-samvariære interaksjoner (Kuhn 2018, 109). Indeksen i metadata er individuelle ord i relevante ordforråd for mulige ord med grunnlinje og frekvens for ord, og metadataene er i denne undersøkelsen satt til å være listen av alle de ulike ordene som forekommer i teksten. Metadataene er derfor med på at STM kan finne emner og sammenheng mellom emner, samtidig som det også blir tatt høyde for ordtelling av de ulike ord og emnene.

Et naturlig første problem når du bruker STM innebærer å identifisere antall emner. Det er ingen eneste riktig måte å ta opp dette problemet på, men en mulighet innebærer å studere handlingen mellom semantisk sammenheng og eksklusivitet (Kuhn 2018, 111). Semantisk sammenheng er basert på målinger av hvor ofte individuelle ord forekommer, samt hvor ofte par av enkelte ord forekommer. Slike tiltak kan hjelpe med å unngå å definere temaer som er problematiske av en av flere spesifikke grunner. Når antallet emner i en modell øker, vil den semantiske sammenheng generelt synke (Kuhn 2018, 111). Et emne anses å være eksklusivt hvis ordene som har stor sannsynlighet for å virke betinget av det emnet, har lave sannsynligheter betinget av andre emner. Og når antallet emner i en modell øker, vil vanligvis eksklusiviteten til modellen som helhet øke.

Etter bruk av STM kobles spesifikke ord til spesifikke emner. Eksplisitte definisjoner, etiketter eller intuitive betydninger av emner genereres ikke automatisk. Det er opp til analytikeren eller forbrukeren av analysen å bestemme hva temaene representerer basert på hvor ofte forskjellige ord brukes i de forskjellige emnene. Dette krever å vurdere hvert emne etter tur og undersøke, for hver en liten samling av ord som er assosiert med emnet (Kuhn

2018, 112). Den mest rett frem tilnærmingen ville være å velge ordene som har størst sannsynlighet for å forekomme betinget av emnet i en STM-modell. Problemet her er at visse ord som «trakassering» og «kvinner» vil med stor sannsynlighet vise seg som ord med stor sannsynlighet for mange emner. Dette er sentrale ord, og ord som har en relativt dominerende betydning. Det vil altså være vanskelig å kunne sette klare skillelinjer mellom emnene med mindre de resterende ordene har en like klar og intuitiv mening.

Som et forsøk på å sette skillelinjer mellom emnene brukes kommandoen «labelTopics» fra R-pakken «stm» for å finne fram til enkelte deskriptive funksjoner. Denne kommandoen finner fram til de ordene med høyest sannsynlighet for å forekomme i et emne forklart av kolonnen med «highets prob», samt hvilke ord som både forekommer hyppig, men som også kan være mest interessante for emnet, som vises i kolonnen som heter «FREX» (Roberts et al. 2019, 13). «Lift» er også med i tabellen som kommer fram, og er basert på en funksjon som vektet ord ved å dele dem med frekvensen de har i andre emner, og gir derfor høyere vekt på ord som vises sjeldnere i andre emner. Siste kolonne er «score». Dette er en tilsvarende funksjon som «lift», men som bruker loggefrequensen for ordet i et emne og sammenligner med loggefrequensen i andre emner for å finne fram til ord som er sentrale for et spesielt emnet.

5.6.3 Validere emner

Å validere emner handler om å bestemme hva hvert emne måler (Grimmer og Stewart 2013, 286). En metode for merking er lesing: prøvetaking av ti til femten dokumenter som er tilordnet et emne. Det er også viktig å identifisere ord som er veldig prediktive for dokumenter som tilhører bestemte emner (Grimmer og Stewart 2013, 287). Hvis et emne er gyldig, skal eksterne hendelser forklare plutselige økninger i oppmerksomheten til emne. I dette tilfelle ble det slik at metoo-dekning fra opprøret i USA rettet mot Weinstein økte fokuset på seksuell trakassering i resten av verden. Og etter hvert som mediedekningen begynte å omhandle politikere i Norge økte anklagene også her.

Det er videre viktig å erkjenne at emnemodellering er en form for uovervåket læring (Kuhn 2018, 111). Det finnes ingen grunnleggende «sannhetsdata», noe som gjør det utfordrende å

validere dataene fra en emnemodell. En vanlig strategi er å la leserne inspirere resultatene og finne gyldighet i resultatene for seg selv. Mer nylig har Chang et al. (2009) bemerket at bruken av emnemodeller avhenger av hvor sammenhengende og relevante (de menneskelige) analytikere finner resultatene sine.

5.7 Klyngeanalyse

Klyngeanalyse gir innsikt i dataene ved å dele objektene (ordene) inn i grupper (klynger) av objekter (Wu 2012, 2). Siden den ikke bruker ekstern informasjon som klasseetiketter, kalles også klyngebasert analyse uten tilsyn på noen tradisjonelle felt som maskinlæring og mønstergjenkjenning. Generelt er det to formål å bruke klyngeanalyse til; forståelse og nytteverdi. «Clustering» (klynging) er å benytte klynge-analyse for automatisk å finne konseptuelt meningsfulle grupper av objekter som har felles egenskaper. Det spiller en viktig rolle i å hjelpe til å analysere, beskrive og bruke verdifull informasjon som er skjult i gruppene. Klynging for å finne nytteverdi av ordene er å abstrahere prototypene eller de representative objektene fra enkeltobjekter i de samme klyngene. Disse «prototypene» tjener da som grunnlag for en rekke databehandlingsteknikker som oppsummering, komprimering og funn av «nærmeste nabo» av ord (Wu 2012, 2). Klyngeanalyse har lenge spilt en viktig rolle i en lang rekke anvendelsesdomener som forretningsintelligens, psykologi, samfunnsvitenskap, informasjonsinnhenting, mønsterklassifisering og bioinformatikk.

5.7.1 K-gjennomsnitts klynging

K-gjennomsnitt er en grupperings-algoritme som prøver å finne K (klynger) som ikke overlapper hverandre (Wu 2012, 7). Disse klyngene er representert av sine «centroids» (en klynge centroid er vanligvis gjennomsnittet av punktene i klyngen). Klyngeprosessen til K-gjennomsnittet er som følger; Først må antall K (klynger) velges, og deretter blir hvert punkt i dataene tildelt den nærmeste «centroiden» og hver samling av poeng som er tilordnet en «centroid» danner en klynge. «Centroiden» i hver klynge blir deretter oppdatert basert på poengene som er tildelt den klyngen, og denne prosessen gjentas til intet punkt endrer klynger (Wu 2012, 7).

Kapittel 6: Resultater og funn i analysen

6.1 Tekstbryting

I analysen er det ti forskjellige dokumenter, hvor hvert dokument representerer en avis fra mediene som er valgt som utgangspunkt for analysen. Lengden varierer fra avis til avis da det er ulikt hvor mange artikler hver av avisene har skrevet om metoo-skandalen, men i utgangspunktet er det snakk om over 15 millioner ord, tegn og tall som skal gjøres rede for. Og ved hjelp av tekstbryting vil jeg finne fram til et representativt utvalg med ord for å gjennomføre analysen.

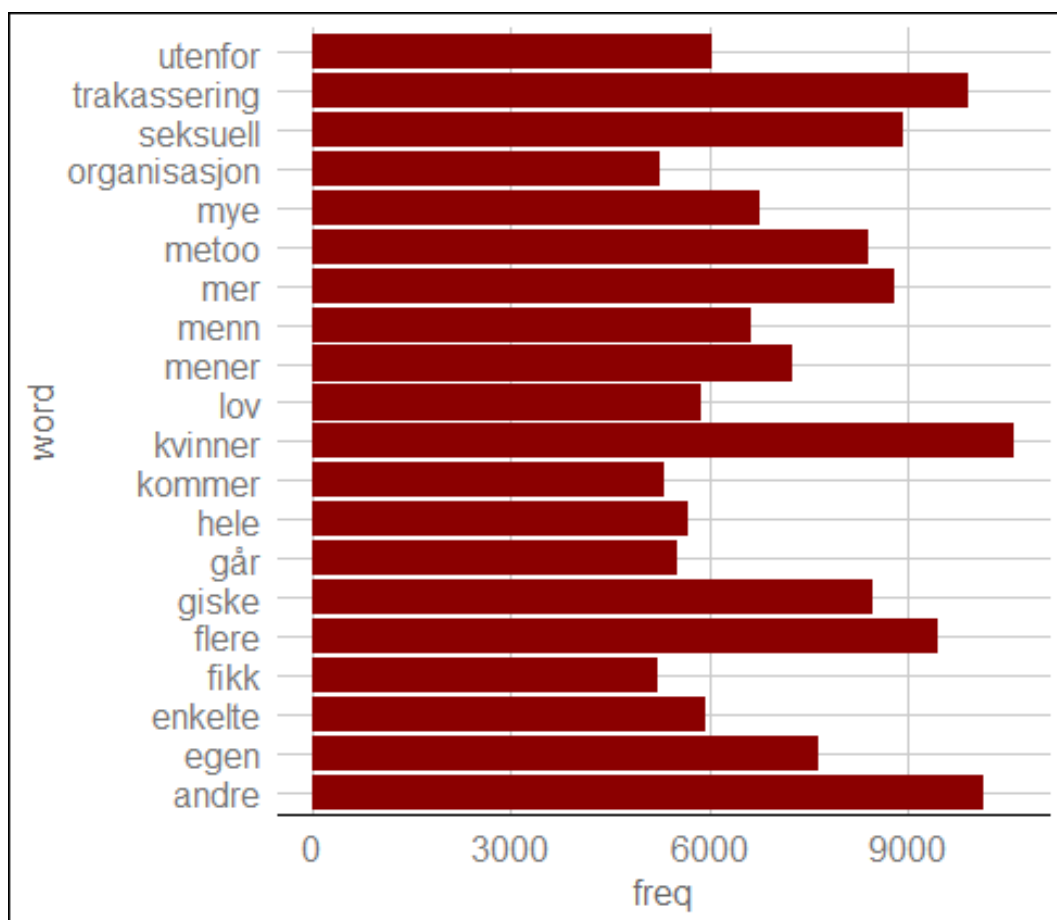
Ved gjennomføringen av tekstbrytingen ble det gjennomført standardprosedyrer som å fjerne tall, mellomrom, tegnsetting og stoppeord for å bryte ned tekstene til data som kunne behandles under emnemodelleringen. Videre var det også enkelte ord som ble tydelig at måtte fjernes helt fra datamatriksen, da disse ordene ble tillagt for stor vekt i analysen uten å ha noe relevant betydning for emnet. Når man gjennomfører ordtelling tas det nemlig ikke hensyn til hvor i dokumentet man finner ordene. Analysen tok derfor for seg både sidetall, overskrifter og alt annet med liten skrift i dokumentene.

Det første som ble fjernet var derfor alle avisenes navn. Dette ble gjort fordi alle avisene er delt inn i hvert sitt dokument, og man trenger derfor ikke navnet på avisen for å vise til hvilke aviser informasjonen er hentet ut fra. I bunn og grunn er det heller ikke slik det gjøres i analysen. I datamatriksen ble avisene delt inn med hver sine kolonner for ordtelling, og man kunne enkelt finne fram der, og i eventuelle andre metoder hvor man trengte å vite hvor informasjonen kom fra ble det referert til pdf-filen (dokumentet) som informasjonen var hentet fra. Utenom avisenes navn måtte ordene «Universitetet» og «Bergen» også fjernes, da disse ordene viste seg å være sentral i ordfrekvensen. Etter noe om og men forsto jeg at dette var fordi datamaterialet var hentet fra atekst gjennom tilgangen Universitetet i Bergen har hos dem, og både topp teksten og bunnteksten viste seg å inneholde disse ordene. Det var altså ingen sammenheng mellom ordene og temaet analysen skulle belyse. Ved oppdagelsen av dette kom også andre ord som ikke var relevant for analysen, men som ble tillagt stor vekt oppdaget. Ord som «foto», «sier», «saken», «nyhet», «avis», «loginn», «nett», «oppdatert» og

«artikkel» ble alle fjernet. Dette fordi det var rent formelle ord som ble brukt som formaliteter ved publisering av nyhetsartikler. Også ordet «sier» som endte som en av de mest brukte ordene i analysen ble fjernet fordi ordet ikke ga noen tegn til forklaring av emnene, men heller tegn til at journalistene har snakket med noen som forteller noe. Ordet hadde med andre ord ingen form for forklaringskraft i analysen, og ble derfor fjernet.

6.1.1 Ordfrekvens

Figur 2: Ordfrekvens



Figur 1 viser fordeling av de ordene som forekommer mest i løpet av artiklene i analysen. Ordet som blir nevnt flest ganger er ordet «kvinner», mens «trakassering», «flere» og «andre» følger tett etter. Selve tabellen gir dog ikke noen tegn på hvilken retning man skal tolke disse ordene. Uten et teoretisk grunnlag er det lett å gå seg vill i analysen, og man kunne for all del tenke seg at det er kvinner som driver med trakassering, men et godt teorigrunnlag gjør at man

kan forstå sammenhengen av ordene, og disse ordene gir indikasjonen på at trakassering mot kvinner øker, og i hvert fall er det et tegn på at tilfeller av trakassering blir stadig mer belyst. Ordet «flere» bidrar til tolkningen om at situasjonene enten øker, eller at det er et mer utbredt faktum enn man tror. Akkurat hvem som trakasserer kan ikke figuren gi noe godt svar på, men annen informasjon om metoo-skandalen gir klare indikasjoner på at det er mye snakk om hendelsen hvor Giske har blitt anklaget for å trakassere en rekke kvinner.

Det forekommer videre en rekke ord som uten å bli satt i kontekst ikke kan gi noen direkte mening; «mer», «går», «kommer» og «hele». Samtidig er det et par ord som kan være en smule tvetydig; «utenfor», «mye», «mener», «fikk» og «enkelte». Uten et teoretisk grunnlag er dette ord som i stor grad vil føre til synsing og en rekke antagelser. Mens «trakassering», «seksuell», «organisasjon», «metoo», «menn», «lov», «kvinner», «Giske» og «egen» er ord med en noe mer klar betydning, og som i større grad kan være med på å danne et bilde av emnet.

6.1.2 Tekstklynging og problemet med sparsitet

Resultatet av forbehandlingstrinnene er at hvert dokument er representert som en vektor som teller antall ganger hvert av de unike ordene forekommer (Grimmer og Stewart 2013, 273). Samlingen av tellevektorer i en matrise kalles ofte den dokumentbegrensede matrisen. For et moderat volum av dokumenter uten et spesialisert ordforråd, vil denne matrisen ha mellom tre tusen og fem tusen funksjoner eller vilkår og vil for det meste inneholde nuller – en tilstand kalt sparsitet. Det vil med andre ord si at ord som forekommer i artiklene gjerne bare er nevnt i et av dokumentene, og gjerne bare én gang. Dette fører til at det kommer en rekke nuller i ordtellingen hos de dokumentene som ikke inneholder ordet. Forskere har likevel vist at det er mulig å komme fram til vesentlige interessante egenskaper ved tekster til tross for høy grad av sparsitet i datamatriksen (Grimmer og Stewart 2013, 273).

Under undersøkelsen av sparsitet i analysen viste det seg at om man tok med alle ordene i alle dokumentene, ville det være en sparsitet på 77%. Det betyr med andre ord at 77% av analysen inneholdt ord som bare var talt én gang. Til tross for teori som forteller oss at dette ikke trenger være noe problem, viste det seg at det da var snakk om en ordtelling på over 138 000

ord, hvor mange av ordene var stavelser og ord satt sammen med ulike tall som ikke ga noen som helst mening. Det er også vanskelig å si noe om hvor disse kommer fra, men det kan for eksempel ha vært en del av en URL-link eller lignende som er referert til i en artikkel. På bakgrunn av dette ble det rimelig å minske datamatriksen, og resultatet endte på å beholde 79% av ordene, og sparsiteten sank til 44%. Ved å gjøre datamatriksen mindre endte ordtellingen på 37 959 ord. På den måten ble ord som nesten ikke ble brukt i analysen fjernet. Dette vil være ord som eventuelt bare var nevnt ved én enkelt anledning, og som ikke kan være med på å fortelle oss noe om mediedekningens fokus.

6.1.3 Korrelasjon

Korrelasjon handler om samvariasjon, altså sammenhengen mellom ulike variabler (Frøslie 2019), som i dette tilfelle er gitt som ord. I analysen ble det kjørt fire forskjellige tester for korrelasjon, hvor undersøkelsen gikk ut på å søke opp de ordene som korrelerte 98% og mer med «metoo», «politiker», «tillit» og «egenskaper». Korrelasjonskoeffisienten blir da oppgitt som et tall mellom -1 og 1 (Frøslie 2019), og når man søker etter de på over 98%, vises verdiene som 0.98, 0.99 og 1. Et positivt tall er det samme som positiv korrelasjon, som betyr at det er en positiv sammenheng mellom de to ordene.

En slik korrelasjonstest blir brukt for å undersøke eventuelle diskusjonsemner fra emnemodelleringen som kan trenge støtte. Testen er altså gjort som et forsøk på å forsterke valideringen av emnemodelleringen. Begrepet «metoo» er hovedsakelig valgt fordi det er dette som er hovedtemaet, og for å se om det er noen faktorer som kan diskuteres utenom hva emnemodellene kan vise. Siden emnemodelleringen symboliseres ved å referere til ord som blir nevnt i analysen er det også sjekket for ord som korrelerer med «politiker». Dette i tilfelle det skulle være noen tvil om emnene i det hele tatt kan knyttes til politikk eller ikke, og for å understreke at det politiske aspektet er viktig for analysen. «tillit» er valgt som et ord som undersøkes for sammenhenger ved fordi at ordet er sentralt i omtalen av skandalepolitikk. Ved en rekke anledninger nevnes det at medieskandaler kan lage sprekker i tillitsforholdet mellom velgerne og politikerne, samt at mediene er holdningsskapende. Tillit er også et begrep som er enormt vanskelig å måle, så funnen tillegges ikke svært stor vekt, men det er likevel ekstremt viktig å få med. Til sist er det kjørt en korrelasjonstest med «egenskaper».

Dette ordet er også valgt på bakgrunn av hva som kommer fram i teorien. Denne bygger nemlig på at politikernes egenskaper blir omtalt under politiske skandaler, og det vil derfor være interessant å se på om noen av ordene i emnene kan ha sammenheng med omtale om nettopp dette. Funnene i undersøkelsen til Isotalus og Almonkari (2014) understreker dette.

Når det ble kjørt korrelasjonstest for «metoo» var der flere ord med en koeffisient på 1. Disse ordene var; bevis, ledende, ofre, sender, sikre og urovekkende. Videre var det 60 ord som korrelerte på 99% (0.99). De mest interessante ordene her var; hensyn, kontroll, kritiske, kvinner, menn, næringsliv, penger, prinsipielle, regjeringen, sosiale, stortinget, støttet, svake, og toppledelse. På 98% kom det også fram ord som lære, oppvask, reduserte og varsel, som kan tyde på tema for å forebygge flere metoo-hendelser. Noen av ordene som korrelerte med «politiker» var; kvinnekamp, nettverk, oppføre, tillit, berørte, dårlige, jentene, opprørt, sure, bråk, drama, frister, maktposisjon, mediestormen, samarbeide, tilliten, useriøst og utvikles. Og enkelte ord som korrelerte med «tillit» var; politiker, dårlige, forsvant, kvinnekamp, jentene, ungdommen, voksen, voldsomt, alder, berørte, drama, konkretisere, mediestormen, partikollega, personalet, skjønne, svekkede og useriøst. Til sist korrelerer «egenskaper» med en lang rekke ord. Noen av ordene som kan være relevante for oppgaven er; nærmest, ordinært, aktualitet, avgjørende, bevissthet, bruk, demokratiet, dører, faglige, fall, folkeparti, kompetanse, kvaliteter, manns, mer, ambisiøse, fordømmene, forestillinger, mangler, politisk, tyder, utvikling og verdighet.

«metoo» viser ikke noe overraskende ved å referere til kvinner, storting og toppledelse. Dette kommer nok som følge av at de som er anklaget, og som det har vært størst storm rundt, er enten politikere eller på toppen av næringslivet. Ved korrelasjoner med ordene til «politiker» må det bli sett på som en selvfølge at ord som «partikollega» nevnes, men «oppføre» og «maktposisjon» kan være med på å gi et hint om at det følger visse forventinger til politikerne. Ordene ved «tillit» og «egenskaper» er dog de mest interessante, da det er her de mest uttrykksfulle ordene kommer fram. Om man setter alle ordene i en sammenheng kan man se et mønster av misnøye og alvor, og også her en del ord som kan kobles til forventinger.

Alle korrelasjonstestene kan sees i vedlegg 2, som tabell 1, 2, 3 og 4.

6.2 Følelsesanalyse

Tabell 5: Følelsesanalyse ved bruk av ordbok: bing

Dokument	Negativ	Positiv	Sentiment
<i>Nordlys</i>	89	86	-3
<i>Morgenbladet</i>	1462	1487	25
<i>Klassekampen</i>	2728	2763	35
<i>Bergens Tidende</i>	426	474	48
<i>Fædrelandsvennen</i>	258	328	70
<i>Aftenposten</i>	725	830	105
<i>Stavanger Aftenblad</i>	552	740	188
<i>VG</i>	806	1093	287
<i>Adresseavisen</i>	1905	2236	331
<i>Dagbladet</i>	1223	1657	434

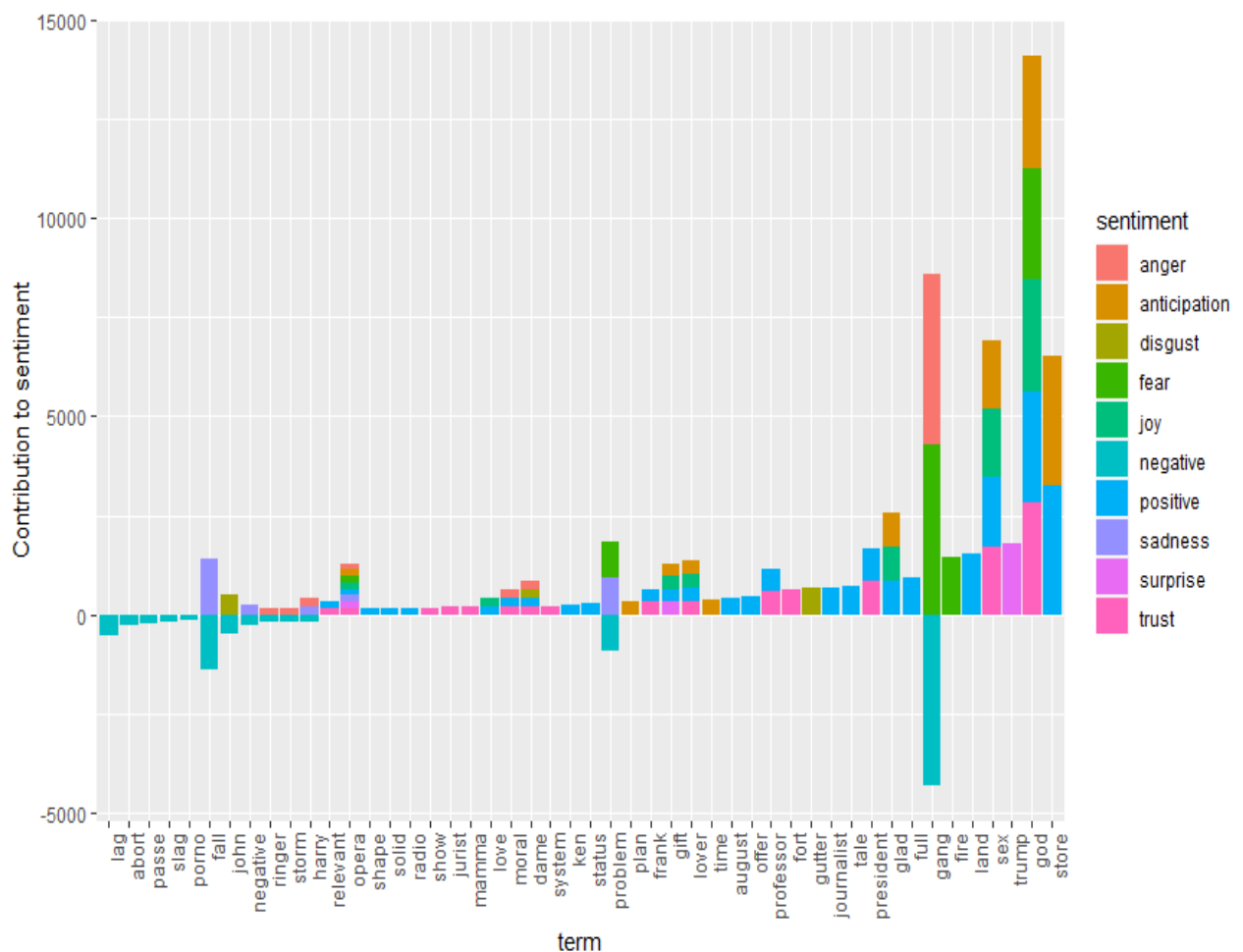
Tabell 5 viser en følelsesanalyse av ordene som er brukt under mediedekningen av metoo-skandalen. Tallene som står under kategoriene «negativ» og «positiv» viser antall ord i hvert av dokumentene som er hevdet å ha enten en negativ eller positiv betydning, og under «sentiment» ser vi differansen mellom de to polene, og om dokumentet i sin helhet framstår som negativt eller positivt. Hvert av dokumentene inneholder alle artikler fra den aktuelle avisen som inneholdt de kriteriene som ble satt for å dekke temaet som analysen ville belyse.

Man kan ut ifra tabellen se at Nordlys blir eneste avis som i all sin forstand blir regnet som «negativ». Denne avisen har også det laveste antallet av både negative og positive ord, og differansen mellom positiv og negativ som skyver avisen over på et gjennomsnittlig negativt perspektiv er heller ikke så stor. Ved Aftenbladet begynner trenden å ta seg opp til verdier som kan ha en viss betydning, men med tanke på forskjellen av antall artikler som er skrevet av de representative avisene er det jevnt over en del negativitet ute og går blant dette temaet.

Den store overraskelsen er at det ikke er et større utbredt negativt preg blant artiklene enn hva som kommer fram av denne tabellen.

Til tross for at resultatene ser rimelig grei ut i tabellen, er det verdt å nevne at man skal være forsiktig med å tillegge følelsesanalysen for mye vekt. Dette er fordi ordbøkene som er blitt brukt til å gjennomføre analysen er engelske ordbøker, som betyr at man ikke kan stole på at alle ordene er forstått ut ifra sin riktige betydning. Dette blir nærmere forklart i teksten under figur 3, hvor også konsekvensen av den engelske ordboken blir langt mer tydelig når man tillegger ordene mer detaljerte følelser enn de som er nevnt i ordboken: *bing*.

Figur 3: Følelsesanalyse ved bruk av ordbok: nrc



Figur 3 er ment å gi et bilde på forskjellige følelser som blir satt i sving under mediedekningen av metoo-skandalen, men som dessverre ikke er med på annet enn å vise stor svakhet ved tekstanalyse for andre språk enn engelsk.

Den store svakheten skyldes at det ikke ennå er produsert noen ordbok som inneholder norske ord, og for å gjennomføre analysen ble jeg derfor nødt til å bruke to ordbøker som begge var basert på engelske ord. Alle ordene i analysen er derfor tolket med engelsk betydning, og dette kommer fram av åpenbare ord hvor for eksempel «offer» er blitt kategorisert som et positivt ord. Det er derfor rimelig å anta at dette er blitt tolket med engelsk betydning, og i analysen blir antatt å bety «tilbud». Ordene «gang» og «god» to ord som åpenbart frembringer sterke følelser. «gang» blir koblet til negativitet, frykt og sinne, mens «god» gir utslag på tillit, positivitet, glede, frykt og forventning. Det er også her rimelig å anta at ordene er knyttet til de engelske betydningene, som vil være gjeng og Gud. Og videre er eksempler og til tross for utslaget i figur 2 bør ikke de tillegges noen vekt.

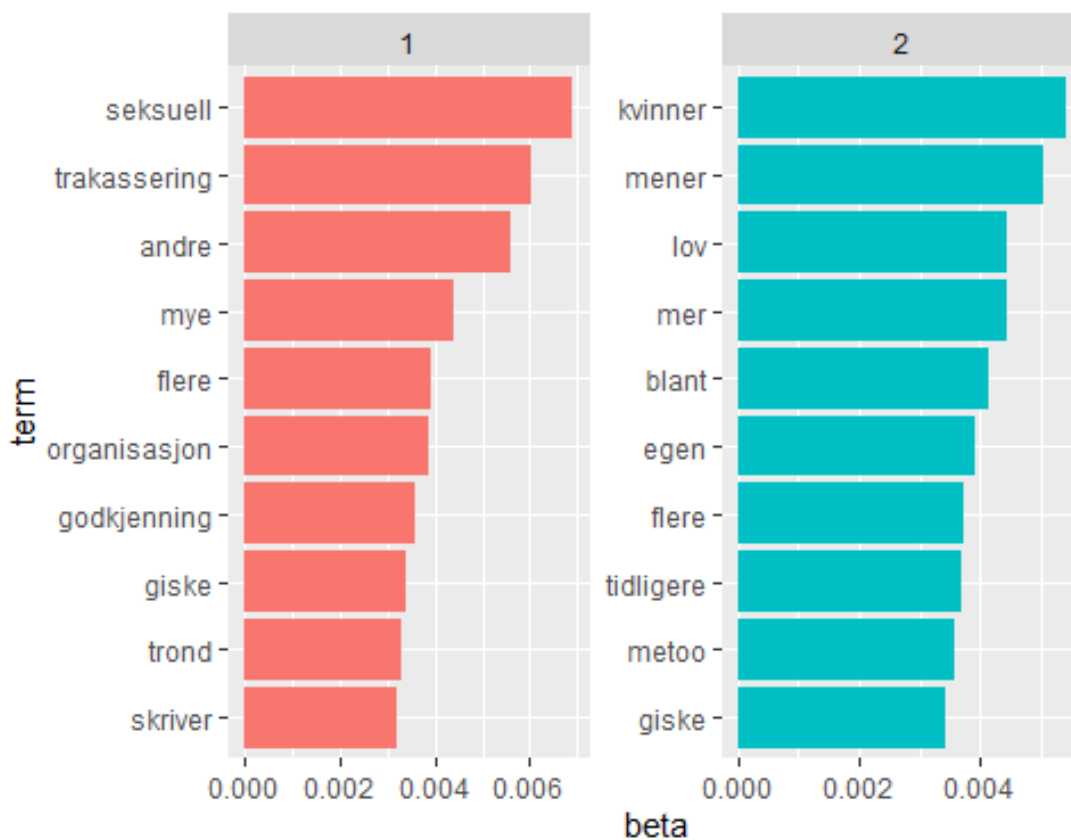
Det er signifikante bevis på at rasjonell læring hos mennesker er avhengig av følelser (Cambria et al. 2017, 2), og det er noen ord som gir interessante utslag i figuren, og som kan være sentrale i en diskusjon av temaet. For det første kan man se at både «sex» og «lover» i utgangspunktet er ord som skaper tillit, positivitet, glede og forventinger. Til tross for at ordet «lover» med all sannsynlighet her er tolket som «elsker» gir dette oss et bilde på at intime handlinger i bunn og grunn er handlinger som blir sett på som positive. Videre forekommer det flere ord som med sin engelske betydning likevel kan bli trukket over til det norske språket. Eksempelvis er positive ord «journalist» og «status», mens ord som «system» og «moral» blir koblet til tillit (selv om ordet «moral» også får et lite utslag på følelsen «sinne»). «Problem» er derimot negativt, trist og skaper frykt, noe som med sikkerhet kan assosieres med problemet med seksuell trakassering.

6.3 Emnemodellering

6.3.1 Latent Dirichlet Allocation

Figur 4 viser en emnemodellering med utgangspunkt i to emner. Ordene på y-aksen forteller oss hvilke ord som blir brukt i temaet, og tallene på x-aksen forteller oss hvor ofte det er sannsynlig at hvert ord blir brukt under omtale om emnet. Begge temaene omhandler seksuelt overgrep, men de har forskjellig vinkling. Emne en tar for seg trakassering, hvor Trond Giske blir tatt opp som en del av emnet. Giske ble et av de store politiske navnene under metoo-kampanjen, så at saker som omhandler han kommer fram som eget tema er ikke særlig overraskende. Det andre emnet kan tilsynelatende se ut til å omhandle at noe øker, eller forskjell fra nå og da. «Kvinner» og «mener» er de ordene som er nevnt mest, og med ord som «mer», «flere» og «tidligere» tyder dette på et tema som omhandler metoo som økende fenomen.

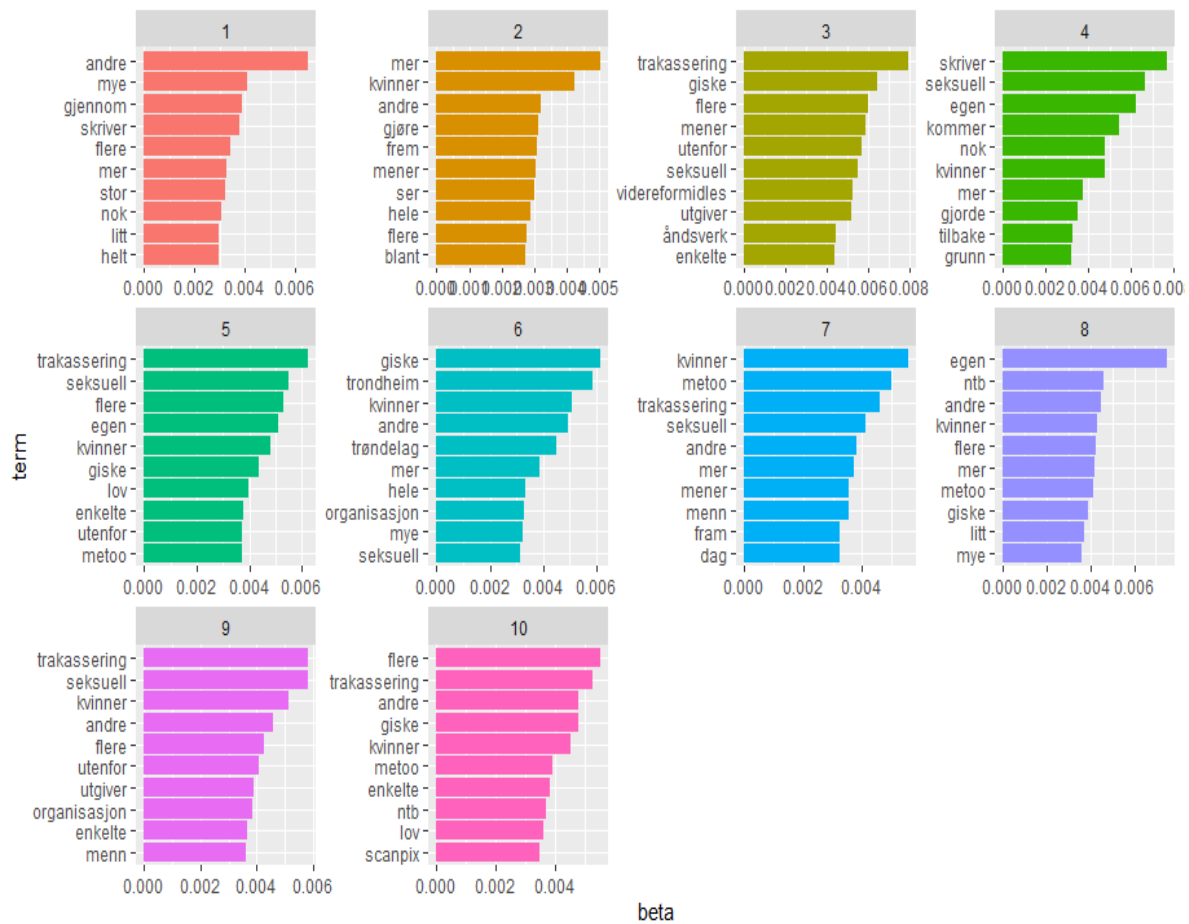
Figur 4: Latent dirichlet allocation, $k=2$



I figur 5 ser vi en oversikt over ti forskjellige emnemodeller. Her er k satt til ti for å undersøke om det kan forekomme ulike emner i de ulike avisene. Det er ikke overraskende at emnene inneholder en rekke av de samme ordene da artiklene allerede er valgt ut i fra et tema om seksuell trakassering i den politiske sfæren, og som forventet er «kvinner», «trakassering» og «seksuell» eksempel på ord som nevnes i over halvparten av emnene som kommer fram. Også «Giske» blir nevnt i fem av ti av emnene (emne tre, fem, seks, åtte og ti). Emne tre og fem er noe uklart, men emne tre snakker om videreformidling og åndsverk, mens emne fem tar for seg kvinner, lov og metoo. Emne seks er dog noe lettere å tolke, og handler helt klart om hvordan anklagene mot Giske påvirket Arbeiderpartiet i trøndelag. Emne åtte er dominert av ordet «egen». Akkurat hva det ordet skal gi et tegn på er vanskelig å si, men «andre og «flere» er igjen sentrale ord, som kan vise til økning. «mer» og «mye» er også med på å støtte dette. Emne ti ellers ut til å ligne på emne åtte, men med et større fokus på handlingen «trakassering». I emne åtte stor dog ikke Giske like sentralt i emnet, som i emne 3, 6 og 10.

«Egen» er et ord som blir nevnt en rekke ganger, som kan være et uttrykk for personlige historier som kommer fram. Utover det er det vanskelig å finne noe intuitivt om emnene, og det kan tyde på at et antall $k = 10$ kan ha sørget for mange og diffuse emner.

Figur 5: Latent dirichlet allocation, k=10

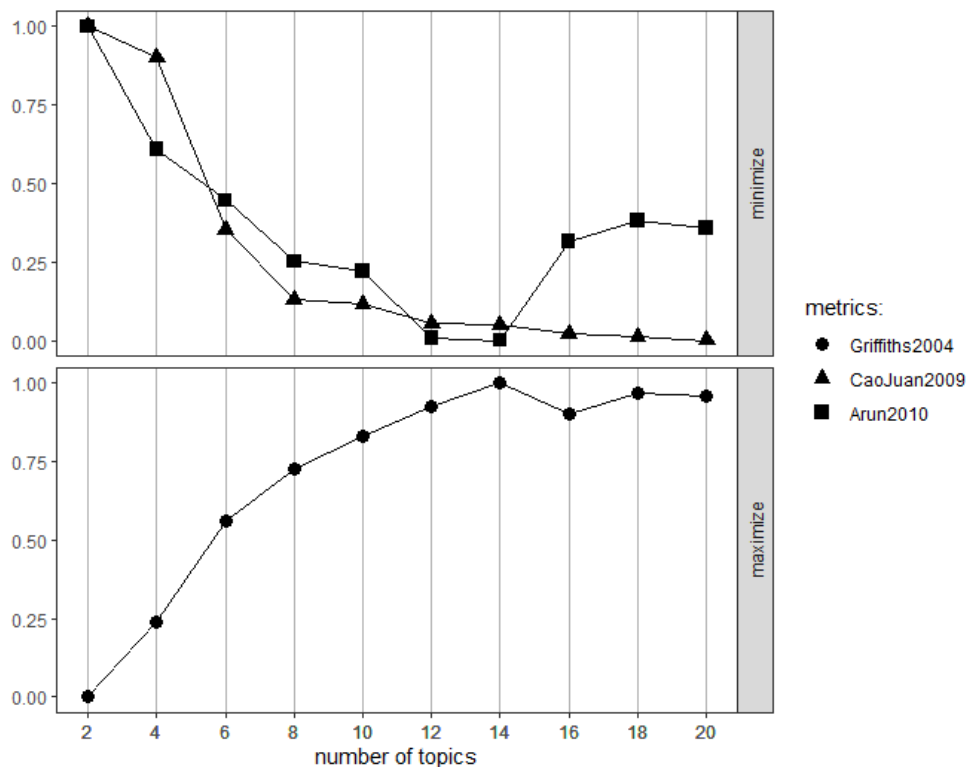


Ved hjelp av en metode fra pakken «topicmodels» er det noe uklare valget av antall emner undersøkt ved hjelp av Griffiths2004, CaoJuan2009 og Arun2010. Antagelsen om at $k = 10$ var for mange emner blir her motbesvist. Dette kommer fram av tabell 6 hvor Griffiths2004 og Arun2010 på hver sin side av +/- søker etter verdien som er lavest, og CaoJuan2009 søker nullpunktet. Til tross for at CaoJuan2009 er nærmere null på 20 emner, viser de to andre modellene at det optimale antall emner vil ligge på rundt 14. Figur 6 viser en grafisk framvisning av tabell 6, hvor det optimale antallet emner er vist ved det punktet hvor de tre ulike estimeringene treffer hverandre (punkt 14). X-aksen viser antall emner, og y-aksen viser verdiene av de tre forskjellige modellene.

Tabell 6: Framstilling av det optimale antall k.

Antall emner	Griffiths2004	CaoJuan2009	Arun2010
20	-19879929	0.09623331	15.11030
18	-19864290	0.10460325	15.30782
16	-19974677	0.11125638	14.76446
14	-19802920	0.12965947	12.16494
12	-19938834	0.13252634	12.23675
10	-20099234	0.17555872	13.98624
8	-20286310	0.18454384	14.23227
6	-20576714	0.33724461	15.83498
4	-21146860	0.71367042	17.16015
2	-21569959	0.78191870	20.37951

Figur 6: Grafisk framvisning av optimale antall k undersøkt ved hjelp av Griffiths2004, CaoJuan2009 og Arun2010

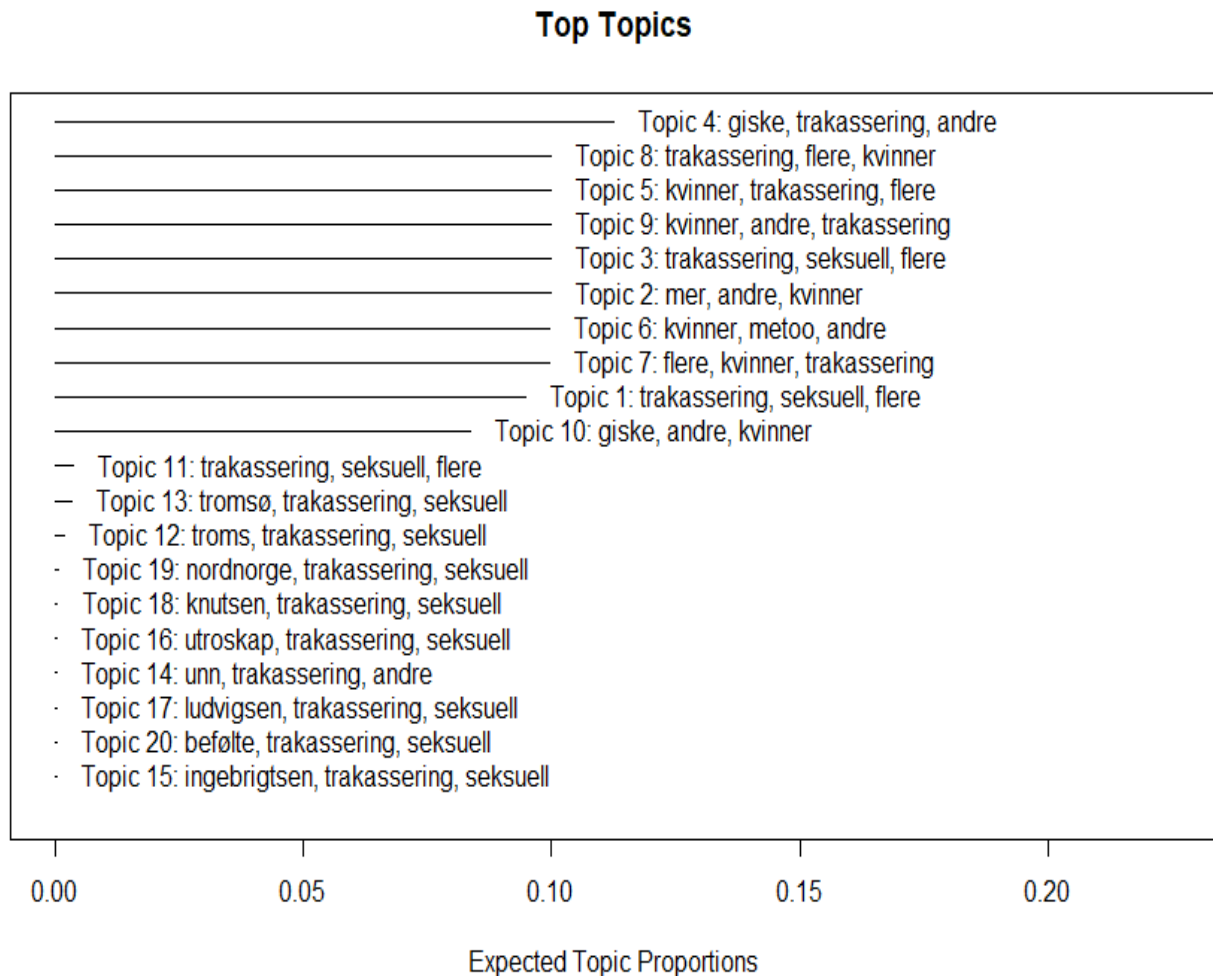


6.3.2 Strukturell emnemodellering

Som en videre hjelp til å finne antall og for å få en bedre forståelse emnene ble det også gjennomført en strukturell emnemodellering (STM). Å estimere forholdet mellom metadata og emner er en kjernefunksjon i stm-pakken (Roberts et al. 2019, 16). Disse forholdene kan også spille en nøkkelrolle i valideringen av emnemodellens nytte (Grimmer og Stewart 2013; Roberts et al. 2019). Figur 7 viser tjue emner fordelt etter hyppigheten av emnene fra dataene i analysen. De ni mest brukte emnene er estimert å bli nevnt med relativt lik hyppighet, videre er det et emner på et slags mellomnivå, før et brått hopp til de siste ti emnene. Syv av de ni mest hyppige emnene inneholder ordet «trakassering», og videre blir andre hyppige ord som «kvinner» og «flere» ofte nevnt. Emnene kan tyde på en økende oppmerksomhet mot seksuell trakassering mot kvinner. Giske viser seg også her i en del av det mest hyppige emnet, og blant de mindre hyppige emnene kommer navn som Knutsen, Ludvigsen og Ingebrigtsen fram. Her får man et møte med problemene med store data innenfor tekst, hvor disse navnene referer til ulike reportere som har omtalt sakene, samt andre skandaler både i og utenfor politikken.

Tromsø dukker opp som et nytt geografisk punkt når man utvider emnemodeller til større k, og er nevnt i to emner, i tillegg blir Nord-Norge nevnt i et annet. Det vil være rimelig å anta at det dreier seg om noenlunde det samme geografiske punktet.

Figur 7: Grafisk visning av estimert hyppighet for hvert emne



Tabell 7 (finnes i vedlegg 2) kan bekrefte denne antagelsen. Til forskjell fra figur 7, som bare gir oss de tre mest brukte ordene i hvert emne, gir tabell 7 oss en dypere forståelse av dem. Vi får innsikt i hvilke ord som forekommer mest, men tabellen viser også fram hva som skiller de ulike emnene fra hverandre. Dette er forklart gjennom highest prob, FREX, lift og score. For eksempel blir det her tydelig at emne 12, 13 og 19 som snakker om Tromsø og Nord-Norge i stor grad omhandler det samme, og det kommer også fram at Tromsø er nevnt i flere emner enn bare disse tre. Uavhengig av det er fellestrekkene for de tre emnene er at det blir åpenbart at Tromsø og Finnmark er de sentrale geografiske punktene, og det kommer blant annet fram en rekke navn i de tre emnene, hvor Silje og Sunniva er de mest sentrale. «Silje» er nylig kommet fram blant metoo-sakene og mediene generelt, og handler om en som har utgitt seg for å være en den åpenbart ikke er, og som har brukt en falsk profil på instagram for å

oppsøke jenter og for å få de til å sende bilder. Dette kommer fram i en reportasje fra NRK hvor det er en journalist fra NKR som har opplevd et møte med denne «Silje» (Kirsebom et al. 2020). Sunniva Andreassen er videre en av varslerne mot Giske som har valgt å stå fram i mediene med fullt navn, og som har vært sentral i kvinnekampen mot seksuell trakassering (Andreassen 2019).

Både «sjefen» og «næringslivet» blir sagt å være en del av emnene, og i tabellen kommer det fram at «sjefen» er et ord som har betydning for emne 13 (se FREX topic 13). Skillelinjene mellom 12 og 13 er få, men i emne 13 kan man se at det blir snakk om utroskap, noe som emne 12 ikke nevner. I emne 19 blir også «Fpu» og «næringslivet» dratt fram som faktorer som er særegent for emnet.

For å understreke at det foreligger en sammenheng mellom ulike emner ble det også gjennomført en korrelasjonstest på modellen med alle tjue emnene (se vedlegg 2, tabell 8). Tabellen er for eksempel med på å understreke at det med sikkerhet kan sies å være en sammenheng mellom emne 12, 13 og 19, da disse emnene korrelert med 0.99.

Andre funn i korrelasjonstesten er at emne 1 korrelerer med en verdi på over 0.90 med emne 12, 13, 16, 17, 19 og 20. Hvis man ser på figur 5 er emne 1 det emnet med høyest frekvens her, og hoppet ned til de andre er en del. Likevel er likhetstrekkene i emnene like, hvor det jevnt over er snakk om Tromsø, stillinger av høy rang og utroskap. «Sunniva» er også et sentralt element, da dette navnet kommer fram i alle emnene bortsett fra emne 1. Videre er emner som korrelerer på et høyt nivå emne 7 og 18, og emne 10 og 11. Emne 7 og 18 er svært like ord av de som har høyest sannsynlighet i og blir nevnt, men utenom det er det få likheter. Mens emne 7 tar for seg ulike personer i underholdningsbransjen, tar emne 18 for seg ulike politikere og metoo-ofre. Fellestrekket er at det er en økende grad av trakassering i begge miljøene. I emne 10 og 11 er det igjen et overveldende antall ord som er like i «hightest prob», men utenom det svært få fellestrekk. Emne 10 ser ut til å handle om situasjonen i Trondheim og nevner også aviser som et tegn på at det omhandler mediedekningen, mens emne 11 er fokusert på revyen og russelåt, hvor førstnevnte viser til at flere russerevyer tok for seg sketsjer med ulike aspekter av metoo-skandalen og trakassering som tema, mens

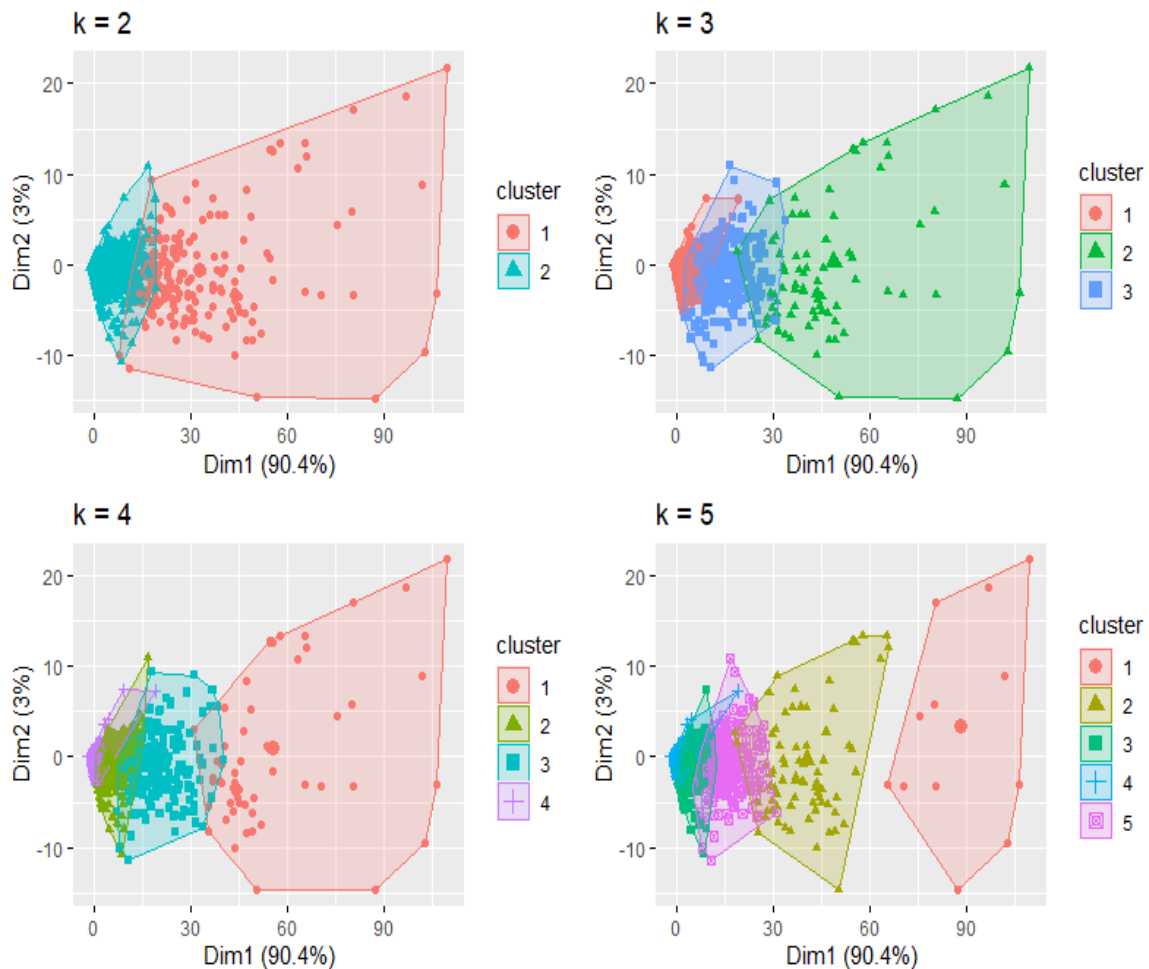
russelåt referer til et emne som fikk mye omtale i media. Dette handlet om at de aller fleste russelåters innhold handlet om at man skulle ligge med flest mulig, og hvor det kom fram flere referanser til at gutter hadde et eierskap over jenter (Pedersen 2019). Det var da tre jenter fra Brannfjell skole i Oslo som skrev et innlegg i Aftenpostens SiD-side, hvor de satte et nytt perspektiv på metoo-skandalen ved å problematisere at seksuell trakassering starter i ung alder. Innlegget deres fikk så stor oppmerksomhet at det ble tatt opp på Stortinget, og ble på den måten gjort til en politisk problemstilling om «hvordan sette en stopper for dette» (Baklund, Aure og Fyhn 2019).

Korrelasjonstesten finner også en grad av sammenheng mellom emnet nummer 15 og emnene 1, 4, 10, 11, 12, 13, 16, 17, 19 og 20. Forskjellen mellom de emnene som korrelerer med 15, fra de emnene som er nevnt tidligere, som f.eks. emnene som er nevnt ved å korrelere med emne nummer 1, er at det finnes ingen sammenheng mellom emnene som korrelerer med emne 15 (sett bort i fra de emnene som er nevnt at korreler lengre oppe på siden). Ved emnene som korrelerer med emne nummer en er det vist en gjensidig positiv korrelasjon, noe som betyr at sammenhengen mellom emnene går begge veier. Emnene er altså like og de er knyttet til hverandre. Ved emnene som viser seg å korrelere med emne 15 er korrelasjonen ensidig. Det vil si at man kan knytte emne nummer 15 til de nevnte emnene, men disse emnene har ikke verdier som kan sees i gjensidig sammenheng. Dette tyder på at emne 15 er et relativt bredt og åpent emne, som knytter seg til mange mindre emner, som ikke åpenbart kan kjenne seg igjen med emne 15.

Sett bort i fra korrelasjon viser figur 7 at de emnene som forekommer oftest er i synkende rekkefølge emne 4, 8, 5, 9, 3, 2, 6, 7, 1 og 10. Emne 6 er et av de mer interessante emne som omhandler politikk (hvis man følger intuisjon). Temaet tar for seg virkninger på høyre og venstresiden og nevner ord som «retningslinjer» og «ettertrykk», som indikerer at metoo-skandalen har hatt en åpenbar påvirkning på politiske parti.

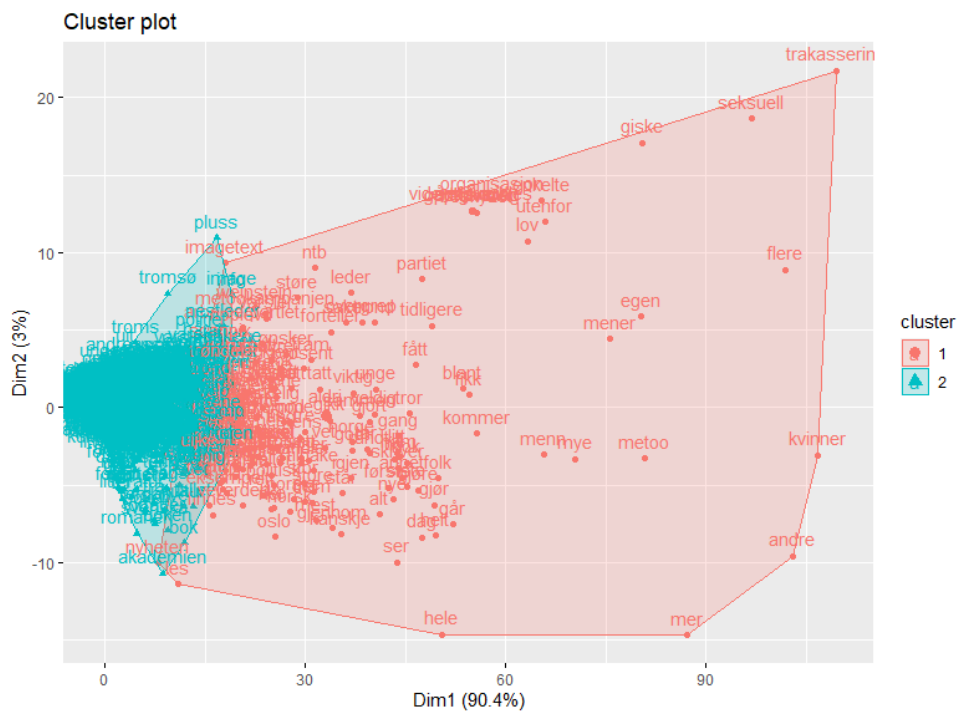
6.4 Klyngeanalyse

Figur 8: Klyngeanalyse



Figur 8 viser ulik fordeling av ordene som er brukt i analysen. De aller fleste ordene viser seg å ligge innenfor dimensjon 50, mens det finnes også ord som strekker seg langt utenfor. På figur 9 kan du se at dette er de ordene som stort sett beskriver emnene og som er oftest brukt i artiklene, det vil si «trakssering», «kvinner» og «seksuell». Ettersom k øker ser man også at emnene fordeles inn i mindre og mindre grupperinger, og figur 9 viser at k gjevnes ut fra 6-14 ved å øke k til 6. Dette styrker antagelsen om en optimal emnemodellering på rundt 14 emner. Figuren forklarer over 90% av objektene dimensjon, og klyngene fordeler seg inn i grupperinger av ord med lik forklaringskraft og betydning etter hvert som k øker.

Figur 9: Klyngeanalyse av objekter, k=2



6.5 Ordtelling av ord som kan ha betydning for diskusjonen

De to tabellene over viser ordtelling for ulike ord i analysen. Ordtellingen av de ulike partiene i tabell 9 er tatt med for å vise til de ulike avisenes bruk av politikk i mediedekningen for å understreke sammenhengen mellom skandalen og det politiske aspektet mediene legger fram. Tabell 10 viser ordtelling på de ulike ordene «metoo», «politiker» og «tillit». Dette for å undersøke hyppigheten av bruk av metoo i sammenheng med politikere, samtidig som man også kan undersøke denne sammenhengen med ordtellingen i tabell 9. Tillit er videre tatt med fordi det er et begrep som blir mye omtalt rundt teorien om skandalepolitikk, og for å kunne enten styrke eller svekke korrelasjonstesten som ble tatt i forbindelse med tillit som er vist til tidligere i resultatene. Ordtellingen blir dog presentert til slutt av resultatene da dette er som et supplement for å forsøke å validere analysen som er gjennomført.

Tabell 9: Ordtelling for de ulike partiene

	Ap	FrP	Høyre	Krf	MDG	Rødt	Sp	SV	Venstre
Adresseavisen	556	424	620	193	86	150	127	31	294
Aftenposten	250	222	278	97	49	59	70	16	143
Bergens Tidende	53	143	93	65	6	14	9	2	53
Dagbladet	379	286	552	118	13	64	57	18	152
Fedrelandsvennen	70	80	60	62	6	29	7	3	47
Klassekampen	273	546	446	290	132	343	105	66	398
Morgenbladet	109	45	170	33	13	13	15	0	91
Nordlys	48	6	5	5	1	1	3	0	4
Stavanger Aftenblad	77	137	128	87	14	22	9	1	68
VG	203	211	415	42	14	34	80	11	122

Tabell 10: Ordtelling for «metoo», «politiker» og «tillit»

	metoo	Politiker	Tillit
Adresseavisen	1536	182	504
Aftenposten	653	70	177
Bergens Tidende	400	21	66
Dagbladet	1159	61	149
Fedrelandsvennen	201	16	23
Klassekampen	2496	91	233
Morgenbladet	694	42	112
Nordlys	110	7	16
Stavanger Aftenblad	406	28	58
VG	746	39	104

6.6 Eksempel på utfordringer med metoden som viste seg i resultatene

Enkelte emner har noen utydelige ord, som kan gi hint om at noe har gått feil i analysen, og også ord som «imagetext» som helst skulle vært fjernet, men som dessverre ikke ble oppdaget er med på å gjøre tabellen noe uklar. En tekstanalyse som tar for seg så store mengder data har stor fare for å få med enkelte ord som er uten betydning, men har mer en formell funksjon i tekstene og derfor forekommer ofte. Som nevnt tidligere er det forsøkt å ta høyde for dette, men noen ord har falt igjennom. Dette er en av ulempene med automatiserte metoder. På en annen side er fordelen at å behandle store mengder data kan gjøres på betydelig kortere tid. Så lenge man er klar over at man kan støte på slike hinder, vil en automatisert metode være fordelaktig. Tabell 3 sammen med de andre visuelle figurene og deskriptive dataene har likevel et overveldende antall ord av stor betydning, som sørger for at man kan stole på analysen. For eksempel kan man se at emne 2 og 14 viser seg å korrelere med en verdi på 0.95 til tross for et noe utydelig språk på mange av ordene i emne 2. Den åpenbare sammenligningen er her svært vanskelig å se, utenom ordene som blir nevnt i «highest prob» som er ord som kvinner, trakassering, mer og andre. Dette uklare resultatet kan være med på å spre noen tvil om metoden som helhet, og det er viktig å ikke legge for stor vekt på enkeltresultater som kommer frem av analysen, men heller sammenligne resultater og bygge godt opp under antagelser som blir gjort.

Kapittel 7: Diskusjon av resultatene

7.1 Seksuell trakassering som et generelt sensitivt tema

Følelsesanalysen viser at ord som «sex» og «lover» (elsker) er ord som viser til tillit, positivitet, glede og forventinger, men seksuelle overgrep kan få disse følelsene hos folk til å ta dramatiske vendinger, og skape mer avsky enn glede. Temaet «seksuell trakassering» som metoo åpner opp for er et tema som berører de aller fleste, og etikken til journalistene blir derfor satt ekstra på prøver under disse forholdene. Dette vil være et faktum uavhengig av om du er enig eller uenig i ulike utsagn. Målet med mediene er nemlig ikke å uttrykke personlige meninger, men å dele informasjon. Under medieskandaler må journalistene derfor holde tungen bent i munnen for å unngå å krysse de etiske prinsippene som de selv har satt opp med tanke på proporsjonalitet, rettferdighet, respekt for personvern og korrekt sitering. Samtidig skal de forsøke å opprettholde både samfunnsplikten sin, og retningslinjene som Vær Varsom-plakaten legger fram.

Mediene skal spre informasjon, åpne for debatt, og samtidig kritisere de forhold ved samfunnet og politikken som er av vesentlig interesse for leserne, men proporsjonaliteten og rettferdigheten med tanke på mengde av informasjon og kritikk som ble utlevert om Giske under mediedekningen av metoo-skandalen kan bli satt spørsmål til ved å undersøke funnene som kom fram i emnemodelleringen. Giske blir nevnt i 5 av emnene i LDA-modellen og 5 av emnene under den strukturelle emnemodelleringen (tabell 7) – hvorav 4 av emnene er de 10 mest aktuelle emnene – mens de andre navnene som blir nevnt referer til varslere og politikere utenfor innblanding i selve handlingen som omhandler seksuell trakassering, og kommer heller ikke fram som et like stort omtalt emne. Med tanke på at flere andre ble navn kommer fram under teori om metoo-kampanjen, som blant annet de to politikerne fra Høyre og FrP som er nevnt i kapittel 4, kan det virke som en noe ensidig vinkling at Giske har fått et så stort fokus at det ikke har vært rom for noen andre politiske navn under mediedekningen.

Å øke antall emner under emnemodelleringen av for eksempel LDA-modellen hvor man setter antall k (antall emner som modellen skal finne) før man gjennomfører analysen kunne vært et alternativ, men analysen som viser at det optimale antall emner ligger rundt 14 emner,

og STM-modellen som viser 20 emner totalt, avskriver altså det å øke antall emner som en mulighet for å muligens kunne finne belegg for en bredde i personbruken av mediedekningen. På en annen side åpner LDA-modellen og STM-modellen sitt prinsipp om at det er forskerens oppgave å bruke intuisjon når man skal forstå hva de ulike emnene i de ulike tabellene handler om opp for at det kan forekomme misforståelser. Modellene er ikke helt åpenbart informative, og det krever at man har en god oversikt over temaet før man starter analysen. Hvis ikke kan dette føre til store feiltolkninger. Dette prinsippet svekker også noe av funksjonen ved å bruke automatiserte innholdsanalysemetoder, da det likevel viser seg at en del kvalitativt arbeid må legges ned for å få best mulig resultat. Et eksempel som kan belyse dette er det faktumet om at seksuell trakassering er et sensitivt tema, og hvor trakassering og kvinner er ord som er åpenbart sterkest representert i mange av emnene, blir det vanskelig å skille mellom emnenes holdningsretning. Artikkene tar for seg både kronikker og andre innlegg som også blir sendt inn av publikum, og ikke bare er skrevet av journalister. Dette gjør at det vil være grunnlag for å tro at enkelte av innleggene om Giske kan ha vært et innlegg som viser støtte framfor avsky, men det blir svært vanskelig å finne signifikante bevis for uten et kvalitativt arbeid. Til tross for dette skal det likevel ikke bli lagt noe tvil om at mesteparten av innleggene har blitt vinklet som en skandale, og derfor notorisk negativ. Følelser kan dog være med på å hjelpe oss i å tolke en retning av hvordan de ulike emnene er lagt fram.

7.2 Medienes bruk av følelser

Vær Varsom-plakaten viser til medienes rett til fri informasjonsdeling, og pressens rett i å dele fakta som avslører kritikkverdige forhold. metoo-skandalen er utvilsomt et slikt tilfelle, og en følelsesanalyse kan være med på å fortelle oss at å holde seg til ren faktaforklaring kan være vanskelig. Bruk av følelser kan være med på å skape holdninger, og bruken av følelsesladde ord er svært forskjellig fra avis til avis. I tabell 5 ser vi at Nordlys er den eneste avisen som har brukt mest negativt ladde ord, i forhold til positive. Men med tanke på det lave antallet av følelsesladde ord som er brukt i teksten generelt, i tillegg til antall artikler som omtaler temaet, og i sammenligning med andre aviser, blir det misvisende å bare se på differansen på polariseringen uten å ta høyde for omfanget av størrelsen på ordtellingen.

Dagbladet hadde desidert flest artikler, men ble fulgt tett etter av Adresseavisen, Klassekampen, VG og Aftenposten. Nordlys, som angivelig var mest negativ, er den avisen med desidert færrest antall artikler angående temaet. Klassekampen, Adresseavisen, Morgenbladet og Dagbladet er de avisene som avdekker mest bruk av følelser i deres mediedekning, hvor Klassekampen uten tvil stiller sterkest. Klassekampen blir sett på som en radikal sosialistisk avis, og det å være radikal betyr at man ønsker store endringer, og det raskt (Thorsen 2019). Med en slik tankegang er det naturlig at journalister ikke sparer på kruttet når det skjer regelmessige brudd på normer i samfunnet, og som undertrykker andre mennesker. Den ideologiske tilhørigheten kan også være med på å forklare hvordan Klassekampen kan uttrykke seg i et større spenn enn hva andre aviser som hevder at de er politisk uavhengige kan få seg til å gjøre. Dette fordi å vise bastante meninger ofte kan virke skremmende, og det kan også ofte vise tegn til at man tar parti med den ene eller andre siden, noe mediene dog hevder at de ikke gjør. Likevel viser de høye verdiene på avisene Adresseavisen, Morgenbladet og Dagbladet at det foreligger en del følelsesladde termer også i deres artikler. Med tanke på antallet brukte følelsesladde ord, kommer det ikke fram som noe eksepsjonelt positivt tema ut ifra noen av sentimentene, til tross for at Dagbladet har nesten en tredjedel av ordene som er brukt på den positive siden.

Med tanke på størrelsen av avisene og mediedekningen ser vi et åpenbart tegn på nøytralt språk hos VG. Avisen blir regnet som en tabloidavis, noe som kan være med på å forklare at selve innholdet ikke tar for seg de største aspektene ved emnet. Tradisjonelt har tabloidaviser vært forklaring på aviser som holder seg til lettlest stoff og kortfattethet (Smith-Meyer 2018). I bunn og grunn kan grunnlaget om tabloidaviser være med på å forklare at det er en generelt liten polariserende effekt i alle av avisene, da trenden er blitt at de aller fleste aviser blir skrevet i tabloidformat.

Ved å appellere til følelser setter gjerne leseren fornuft til side, og kan på den måten lettere la seg påvirke. Dette støttes av faktumet av at mediens innramming av nyhetene er med på å påvirke velgernes meninger. Størrelsen på mediedekningen er med på å tiltrekke oppmerksomhet, men mediene er også både holdningsskaper og holdningsendrende, som vil si at de er med på å skape endringer i holdninger hos folk. Deres måte å legge fram skandalen på er med på å bidra til hvordan leserne oppfatter temaet og er med på å forme hvordan man

ser på ulike politiske spørsmål. Midtbø (2012) nevner at velgerne har en tendens til å identifisere seg med majoriteten, og denne majoriteten kan mediene være med på å både styrke eller svekke. En ensidig politisk framstilling vil være med på å generere like holdning hos leserne, og dette kan få utfall for hvordan politiske endringer former seg hos leserne.

7.3 Individualisering eller privatisering?

Basert på defineringen av individualisering og privatisering, og skillelinjene denne definisjonen åpner opp for, blir det naturlig å stille seg spørsmålet om det rett og slett bare er rimelig at sex-skandaler også skal omhandle den aktuelle evner til å utføre sin rolle som politiker. Alvorlighetsgraden på normbruddet og den enorme mediedekningen er med på å klassifisere metoo som en definitivt politisk skandale. Seksuell trakassering er et begrep som går under de handlinger som vurderes som normalt uaksepterte handlinger, samtidig som det også bryter med loven (*Arbeidstilsynet s.v «Seksuell trakassering»*). Ved sitt politiske verv kommer videre byrden av at man har en mer offentlig posisjon i samfunnet. Jo høyere stilling i ulike parti, jo høyere blir også ofte graden av offentlighet, og en økt grad av offentlighet kan være med på å øke fallhøyden ved en eventuell skandale.

Undersøkelsen som Dagsavisen gjennomførte er et tegn på åpenhet og offentlighet inn mot de politiske stillingene i samfunnet. Til tross for at undersøkelsen ikke hadde noe personfokus, fikk man innsyn i de ulike politiske partiene og deres situasjon med håndtering av tilfeller av seksuell trakassering. Dette er for eksempel også grunnen til at Dagsavisen hadde mulighet til å gjøre undersøkelsen om antallet seksuelle varslinger, og å faktisk kunne forvente at alle partiene svarte. Her var ikke målet å ta enkeltpersoner, men denne muligheten ved å kunne gjøre noe slik er også med på å offentliggjøre stillingen man får når man deltar i politikken. Åpenhet er videre viktig for mange av de politiske partiene, hvor blant annet Rødt hevder at åpenhet stryker lojaliteten hos velgerne (*Rødt s.v «Åpenhet»*). Men til tross for denne åpenheten til politikken generelt, er det ikke gitt at den skal sees i sammenheng med personlige normbrudd. Med tanke på at normbrudd kan skje i forskjellige størrelser og at størrelsen ofte blir assosiert med alvorlighetsgrad, hviler vekten på om normbruddet kan trekkes til aktørens politiske verv på nettopp dette. Det politiske aspektet av metoo-skandalen tok for det meste for seg saker som omhandlet seksuell trakassering innad i partiet, noe som

gjør dette til et brudd på arbeidsmiljøloven §4-3 (*Arbeidstilsynet s.v «Seksuell trakassering»*). Et brudd på lov blir også ansett som normbrudd, da disse lovene kan defineres som nedskrevne normer.

Politikk blir i all hovedsak representert ved partier, men disse partiene blir representert av individer. Det lojalitetsbåndet som rødt referer til, avhenger av at disse representantene for partiene i sin minste forstand klarer å følge skikker og regler i samfunnet. Partiene er altså fundamentalt avhengig av å i det minste ha ledere i partiene sine som framstår som lovlidige og kunnskapsrike mennesker. Det å skulle klare å legge et skille mellom menneskers troverdighet som politiske ledere og normbrudd av en slik størrelse som metoo-skandalen har bevist er rett og slett urimelig å forvente. Uavhengig av hvilket yrke man har, om man er en privat eller en offentlig person, så vil et slikt normbrudd ha en påvirkning på menneskers oppfatning av deg som person, og jo større normbrudd, jo større konsekvens.

Som nestleder av et parti og som parlamentarisk nestleder, som Trond Giske og Ulf Leirstein, eller som leder, slik som Kristian Tonning Riise (leder for Unge Høyre) er man i det som må regnes som relativt offentlige posisjoner. Denne offentlige posisjonen legger også store forventninger til oppførsel, ikke bare av velgerne, men også av partiene, samtidig som den åpner opp for sårbarhet for politikerene. Lojalitetsbåndet strekker seg altså både mellom velger og politiker, og parti og politiker. Med tanke på medienes makt som holdningsskaper er de alle avhengig av å unngå skandaler. Når noen av disse offentlige personene da velger å bryte med normer av en slik alvorlighetsgrad vil det mange være naturlig at det blir dekket av mediene – og alvorlighetsgraden er uten tvil med på å svekke deres troverdighet som et redelig menneske.

På sett og vis kan normbrudd bli begått av flere personer samtidig, og man kan for eksempel begå normbrudd i sin offentlige stilling, men handlingen må likevel bli sett på som privat. Problemstillingen med privat handling blir nevnt i den sammenheng med at det blir sett på som rimelig å trekke det til rollen som politiker. Det er gjerne normer man kan bryte uten at det går noe særlig utover din stilling som politiker, men også her handler det om størrelsen på normbruddet. Seksuell trakassering har fått en enorm plass i kvinnekampen – som for øvrig er

et ord som kommer fram under korrelasjonstesten med både «politiker» og «tillit» (se vedlegg 2, tabell 2 og 3) – om å kunne bli likestilt med menn. Seksuell trakassering har nesten utelukkende handlet om menn som trakasserer kvinner, og tilknytningen av temaet til debatten om likestilling styrker grunnlaget for at et slik normbrudd blir satt i sammenheng med den aktuelle troverdighet til å være med på å lede landet. Avhengigheten av lojalitetsbåndet er altså med på å gjøre normbrudd av en slik alvorlighetsgrad som metoo-skandaler omfatter til å gå under definisjonen av individualisering, og det blir sett på som en naturlig prosess at mediene knytter normbruddet til vedrørens politiske verv.

7.4 Medienes makt – samfunnsoppdraget

Med tanke på individualisering, har mediene et stort samfunnsoppdrag med å legge fram informasjon og å dekke alle aspekter av temaet. Mediene har et stort ansvar i å dele all relevant informasjon og bør være nøye på å få fram alle sider av en sak, uavhengig av egne holdninger. Som tidligere nevnt inntar mediene rollen som både anklager og dommer, og spiller en svært avgjørende rolle når det kommer til å dele informasjon som kan sverte politikerne. Analysen indikerer at det optimale antallet av emner blant artiklene som er analysert ligger på rundt fjorten emner. Den umiddelbare reaksjonen kan være at det kan virke noe finurlig at et så konsentrert tema kan omhandle så mange ulike emner, men det er også med på å bevise bredden av seksuell trakassering som et problem, og omfanget av handlingene. Når medieskandaler utarter seg – i hvert fall i form av seksuell trakassering – er det historiene som underbygger det dominerende perspektivet av skandalen som blir prioritert i nyhetsrommet.

Ordfrekvensen som man kan se i figur 2, kapittel 6, og som blir visualisert gjennom figur 11 nedenfor, er første indikator på makten man kan se gjennom medienes bruk av ord. Samtidig som figur 2 avslører en rekke ord som blir brukt til å forklare ulike situasjoner, får man også et innblikk i essensielle ord som er med på å forklare det overfladiske ved metoo-skandalen. Trakassering viser seg å i perioder bli brukt uten bruk av ordet seksuell, noe som tyder på at en generell trakassering kan ha blitt trukket utover sin seksuelle forstand. Ordet «kvinner» er nevnt 10 622 ganger, «trakassering» er nevnt 9923 ganger, og «seksuell» er bruk 8941 ganger (se vedlegg 2, tabell 11). Med tanke på refereringen til kvinnekamp har mediene et stort

7.5 Metoo-emner

Emnemodelleringen viser til en rekke emner som kommer fram under mediedekningen av metoo-skandalen, og til tross for store likhetstrekk mellom emnene viser det store spennet av emner at fenomenet seksuell trakassering er stort. Klyngeanalysen viser at ordene som forekommer i tekstene er lagt til et snevert område, og at disse ordene i og for seg ikke har så stor forklaringskraft, men når de settes sammen med emnene i tabell 7 får vi forklart bredden av fenomenet på en litt annen måte.

De små variasjonene er nemlig også preget av vinklingen som de ulike mediene har valgt å ta i bruk. Til tross for at mediene skal være redelig i deres framvisning av nyheter, er det likevel rift om leserne, og enkle små vinklinger kan gjøre mye med tekst. Dette kan også være med på å forårsake at det kan foreligge så like, men så mange ulike emner. Blant de tjue emnene som kommer fram av den strukturelle emnemodelleringen er ti av de angitt å bli trukket fram relativt hyppig. Og i det absolutt mest omtalte emnet er Giske et sentralt tema. Giske er nevnt i fem av de tjue temaene som blir lagt fram av analysen i STM-modellen, hvorav fire av disse emnene er blant de ti emnene som er sagt å bli nevnt oftest. Dette faktumet tyder på et stort personfokus under mediedekningen av metoo-skandalen.

Enigheten om at journalistikken skal informere om uakseptable forhold og maktmisbruk kan på sett og vis være med på å styrke en holdning om at personifisering av politikken er nødvendig. I figur 3 kan man se at også ulike navn kan være med på å skape følelser hos leserne. Dette blir synlig da «Harry» viser å være assosiert med negativitet, tristhet og sinne, og «Trump» er i dette tilfelle blitt merket som overraskende. Med Giske som et sentralt tema i mediedekningen av metoo ville det vært svært interessant å undersøke hvilke følelser navnet hans blir assosiert med blant leserne i dag.

For å kunne avdekke maktmisbruk er det også ofte svært nødvendig å presisere person. Det kan for så vidt være grupper som utøver et kollektivt maktmisbruk, men uavhengig av hvor mange aktører, så er det særdeles sjeldent at man kan trekke maktmisbruk til å være en standard for et politisk parti. Igjen blir dette lojalitetsbåndet mellom parti og velger en viktig

faktor for å kunne overleve i politikken. I tillegg er det under metoo-dekningen hovedsakelig enkeltaktører som blir avslørt å ha begått seksuell trakassering i form av maktmisbruk, og for å avsløre dette er det ofte essensielt å nevne navn. I dette tilfelle blir det plutselig viktig å skille mellom person og parti for å forsøke å skåne partiets omdømme mest mulig. Ved å ikke nevne navn ville man i tilfelle brukt partiet den enkelte aktør tilhører for å avdekke fakta som omhandler saken, og det er flere grunner til at dette ville vært mer negativt enn positivt. Ved å begrense skandalen til å omhandle den enkelte aktør kan politisk tilknytning bli forsøkt å holdes utenfor, men ved å omtale en hendelse som har skjedd i et parti, er det lett å dra hele partiet under en kam. På en annen side er en seksuell handling noe en person styrer selv, og hvor det ikke kan tildeles noe skyld i partiet som sådan. Men politikeres rolle som offentlige personer, og som ambassadører for et politisk parti er kanskje grunn nok til å sette spørsmål til de generelle holdningene i de ulike partiene likevel.

7.6 Skandalens påvirkning på det politiske system

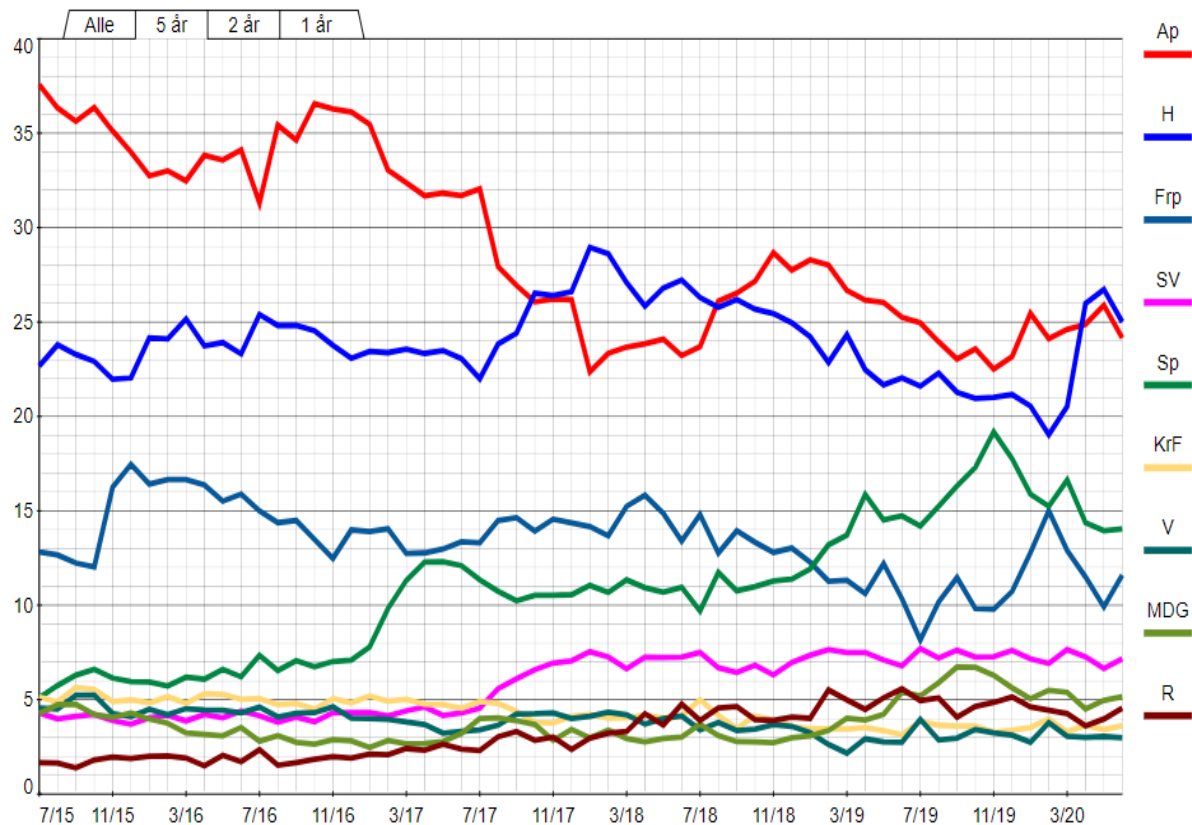
At politikere, som ledere for lovgivende og utøvende makt ikke er i stand til å følge lover og regler må som tidligere nevnt bli sett på som tydelige personlighetstrekk som er med på å spre et negativt inntrykk hos velgere. Når politikere bryter med norske lover bryter de altså med et system som de selv har vært med på å forme. I denne situasjonen vil jeg i bredeste forstand anta at partiene unødig vil assosieres med normbruddene gjennomført av en av deres representanter, og det blir igjen et tema om hvem som egentlig tjener på en individualisering av politiske skandaler? Gjerne ville de politiske partiene ønsket at slike situasjoner skulle bli trukket hele veien til privatisering, uten tilknytning til partiet.

Metoo-skandalen blir videre viktig for det politiske systemet med tanke på at kampanjen har fått flere politikere til å tre av fra sine politiske verv. Anklagen om Trond Giske kom fram 13. desember 2017 og allerede 7. januar 2018 trakk han seg fra sin stilling. Det tok altså ikke lange tiden fra anklagene kom fram i avisen før nestlederen for Ap valgte å trekke seg, noe som gir klare indikatorer om alvorlighetsgraden av skandalen.

Ifølge Midtbø (2012) er det ikke lenger et spørsmål om mediene påvirker stemmeatferden, men hvordan de påvirker den. Politiske skandaler kan eksempelvis være et slikt tilfelle. Via tabell 10 ser man en relativt svak referering av politikk i forhold til metoo, men tabell 9 viser at de aktuelle politiske partiene ble nevnt vesentlig flere ganger enn hva tabell 10 tilsier. Politiske parti er heller ikke særlig fremtredende under emnemodellene. Men når man sammenligner de avisene som har skrevet flest artikler om temaet, som i synkende rekkefølge er Dagbladet, Adresseavisen, Klassekampen og VG, er det også disse avisene som har de høyeste verdiene på bruken av de forskjellige ordene.

For å undersøke om mediepresset har påvirket velgerne er det blitt sett på meningsmåling for partipreferanse hentet fra «Poll of polls». Figur 12 viser gjennomsnittet av nasjonale meningsmålinger ved stortingsvalg, og den viser at under perioden hvor skandalene kom fram, som politisk var i desember 2017, er det et byks i både Høyres og Arbeiderpartiets linje. Mens Høyre så ut til å gå oppover rundt januar 2018, går Ap brått nedover. Dette kan tyde på en umiddelbar reaksjon på nyheten om Giske, og hvor vi ser at Høyre også går nedover like etter kan ha sammenheng med nyheten om Unge Høyre-leder Tonning Riise. Fra januar 2018 til april 2018 går Høyre ned med 3,2 prosentpoeng, og Ap opp med 1,4 (*Poll of Polls s.v* «Gjennomsnitt av nasjonale meningsmålinger om stortingsvalg»). Meningsmålingen rundt de to nevnte partiene ser dog ut til å jevnes noe ut etter juli 2018, og det er vanskelig å trekke noen konklusjon om den videre trenden kan ha påvirket forholdene noe særlig. For de andre partiene er det rimelig stabilt i det nevnte månedene, og det blir også vanskelig å trekke slutninger blant andre partier i de senere månedene, til tross for enkelte endringer. De interessante partiene å følge gjennom meningsmålingene er Ap og Høyre, og i oktober 2018 sammen med at mediedekningen av metoo øker igjen, går også Ap over Høyre på meningsmålingene. Dette øker til så neste måned og Ap holder seg stabilt over høyre på meningsmålingene. Til tross for at det ikke var særlig utslagsgivende er det uansett et aspekt som er verdt å ta med, da teorien antydte at målingene kunne fått store utslag i sammenheng med skandalen.

Figur 12: Gjennomsnitt av nasjonale meningsmålinger om stortingsvalg



(Poll of polls s.v «Stortingsvalg: Hele landet»).

Jenssen (2014) viser til at normbrudd som er begått i form av sitt politiske verv og offentlige stilling blir ansett som verre og mer utilgivelig enn normbrudd som er gjennomført ved private anledninger. Og hevder at uavhengig av hvilke normbrudd eller andre former for skandale, så er ikke dette noe særlig utslagsgivende for de personer som sympatiserer med det aktuelle partiet som politikeren som representerer partiet hører til. Det har videre større utfall for de partiene som står lengst fra dette partiet politisk. Dette går heller ikke an å forklares ved å se på figur 12, da det ikke påvirker noen endring i hvilket parti man stemmer på, men det kan ha noe å si for engasjementet rundt skandalen.

7.7 Innholdsmessige politiske nyhetsinnholdet

Emnemodelleringen viser ikke veldig klare tegn til et innholdsmessig politisk nyhetsinnhold. LDA-modellen viser i figur 5 at emne 6 tar for seg hvordan metoo-skandalen med Giske påvirket Arbeiderpartiet i Trondheim, men utover det er det vanskelig å finne noen emner her som er direkte rettet mot politikken. Og til sammenligning får man heller ikke et inntrykk av politisk fokus i figur 7 som viser resultater fra den strukturelle emnemodelleringen.

Tabell 7 i vedlegg 2 viser derimot til noen flere politiske aspekter, hvor man for eksempel kan se at ordene samfunnstopp og ledertype er nevnt i tabell 1, Giske saken tas opp igjen i emne 4 og emne 10 hvor de viser til Rita Ottervik som er ordfører i Trondheim, samt får varsleren Sunniva Andreassen mye oppmerksomhet i flere emner; 11, 12, 13, 14, 15, 16, 17, 18, 19 og 20. Samfunnstopp er ikke utelukkende en del av det politiske aspekt, men man må kunne anse det som en del av politikken fordi en samfunnstopp blir assosiert med viktige personer eller «*ledende skikkelser i et samfunn*» (NAOB s.v «samfunnstopp»), noe som går godt innenfor hva det innebærer å være politiker. Videre blir elleve av de nevnte emnene knyttet til Trond Giske hvor også noen av emnene tar for seg hvordan varslene og hans avskjed fra sin stilling blir diskutert i henhold til hvordan det påvirket partiet i Trondheim. I de resterende emnene er Sunniva Andreassen som er den varsleren som har stått fremst i stormen som er knyttet til at hun varslet mot Giske. Utover at navnet hennes er nevnt, er emnene dog ikke noe åpenbart annerledes enn de andre emnene. Hovedfaktorene er de samme; kvinner og trakassering, og man får altså ikke noe spesielt innsyn i akkurat hva som gjør disse emnene veldig mye mer annerledes enn de andre, men man får en forståelse for hvor omfattende Giske og Andreassen har vært for å belyse seksuell trakassering generelt.

Emne 6 tar videre for seg meninger angående temaet fra både høyre og venstresiden, og russelåter får oppmerksomheten over på skolegården i emne 11. Dette emnet er for øvrig spesielt fordi det tydelig viser hvordan aktualisering av et samfunnsproblem kan øke oppmerksomheten på de ulike aspektene i samfunnet. Før de tre jentene skrev et innlegg i Aftenposten var hovedfokus på seksuell trakassering i arbeidslivet, men ved å få innlegget

på trykk i mediene får man belyst omfanget av seksuell trakassering som et økende samfunnsproblem.

Den absolutt store bredden av politisk innhold viser altså emnemodelleringen ikke, men tabell 9 er likevel med på å styrke inntrykket om at politikk har vært en stor del av mediedekningen til de ulike avisene. Det skal også sies at når problemer blir store nok, og de sees på som et samfunnsproblem, så blir det uansett knyttet til politikk. Dette fordi det da umiddelbart vil komme et krav fra folket om at her må politikerne inn for å få bukt med problemet.

7.8 Konsekvensene av personfokuset mot Giske under metoo

Mediene hevder også selv at involvering av enkeltsaker kan være med på å rette opp i feilaktige hendelser i samfunnet. Dette prinsippet kan på sett og vis bekreftes, da Giske som står som et sentralt tema under denne skandalen, mistet sinn politiske stilling som følge av skandalen. Det viser seg dog i teorien at Høyre har vel så mange saker innad i partiet som hva Arbeiderpartiet har, likevel er Giske omtalt mye mer. Ute å legge noen tvil om at seksuell trakassering er et alvorlig faktum, kan det virke som at Giske ble offer for å være et slags symbol for metoo-skandalene i politikken. Som nevnt er seksuell trakassering et sårt tema, og de som blir utsatt for seksuell trakassering kan ofte være personer av «lavere rang» enn de som gjennomfører trakasseringen, og med begrepet lavere rang viser jeg da til personer med lavere stilling på arbeidsplasser, eller i samfunnets struktur generelt (hvor også begrepet om å begå maktmisbruk er med på å beskrive den som gjennomfører seksuell trakassering som høyrere stilt, eller i hvert fall i besittelse av mer makt enn den som eventuelt blir offer for seksuell trakassering). Terskelen for å tørre å si ifra og å komme med anklager mot en som er bedre stilt i partiet, på arbeidsplassen, eller i samfunnet enn dem selv kan derfor gjerne være høy. I tillegg kan den offentlige statusen også være med på å fungere skremmende å gå imot. Frykten for å kanskje ikke bli trodd kan stå sterkt i en slik situasjon, og den offentlige stillingen assosiert med individualisering, som betyr at dette kan bli en mediesak, kan i dette tilfelle bli sett på som noe positivt for politikerens trygghet. Så hvem drar egentlig mest nytte av individualiseringen?

Et annet aspekt ved politiske skandaler og personifisering er at man får innblikk i personlige liv til enkeltpolitikere. Dette er heller ikke et aspekt som dekkes ved bruk av emnemodellering, men som gjerne krever en mer kvalitativ forskningsmetode. Til tross for personfokus, så viser ikke analysen noen stor grad av fokus på privatliv. Videre er det også riktig nok et stort fokus rettet mot Giske under mediedekningen, men sammenlignet med andre land, som for eksempel Sverige, så har navngivning blitt holdt på et relativt lavt nivå i Norge (Orgeret 2019). Sverige har også gått i fellen av å bryte presseetiske regler, som har endt med 17 fellelser under metoo-dekningen. Dette i motsetning til Norge, hvor PFU bare har tatt imot én klage, og hvor ingen av mediehusene er felt.

Kapittel 8: Konklusjon og forslag til videre forskning

Problemstillingen som her skal besvares er:

«Har mediedekningen av metoo-kampanjen vært med på en personifisering av politikk, og hvordan har dette påvirket nyhetsbildet?»

Ved hjelp av emnemodelleringen viser det seg at den absolutt mest nevnte politikeren har vært Trond Giske. Hans navn blir nevnt i mange av emnene, hvor han i flere emner er et av de ordene med høyest frekvens, og ut ifra emnemodellene får man inntrykk av at Giske har blitt brukt som et symbol på konsekvensene av seksuell trakassering. Det vil altså være rimelig å svare ja på at mediedekningen har vært med på en personifisering av politikken.

Hva dette har å si for politikken generelt er vanskelig å si, men diskusjonen tar for seg flere aspekter som viser at det kan være en fordel for de politiske partiene at skandalen har et personfokus. At fokuset tas bort fra parti under politiske skandaler er ikke nødvendigvis negativt, og at egenskaper ved politikere blir omdiskutert ved skandaler blir en selvfølge om man skal strebe etter å opprettholde et lojalitetsbånd mellom velger og politiker.

Emnemodelleringen kan det tyde på at nyhetsbildet har vært påvirket av en ensidig politisk framstilling, noe som kan ha formet holdninger til leserne. Normbrudd av en slik alvorlighetsgrad er også med på å svekke troverdighet for leserne til at politikerne kan behandle andre viktige politiske tema, som for eksempel kampen om likestilling. I dette tilfelle vil det derfor være rimelig å anta et nyhetsbilde som fremmer Giske som sentral aktør av seksuell trakassering vil svekke hans politiske troverdighet sterkt. Videre er det store antallet emner med på å vise omfanget av seksuell trakassering som et samfunnsproblem, og det viser seg at økt aktualisering av temaet kan føre til videre avsløringer i samfunnet. For å kunne avdekke maktmisbruk er det også nødvendig å presisere person, og et mediebilde som fremmer personer som gjennomfører normbrudd kan være med på å gjøre emnet mindre farlig for andre som ikke tørr å si ifra. For partiet sin del viser diskusjonen også til at det kan være fordelaktig for dem at normbrudd av slik karakter blir forbeholdt til å omhandle den aktuelle personen, da det kan være fare for at økt fokus på politisk parti framfor person kan føre til at

de generelle holdningene som blir stilt til veggs ved gjennomføring av alvorlige normbrudd er utbredte holdninger for hele partiet, og ikke bare et tilfelle av en enkelt aktør. Skandaler er også viktig for det politiske system fordi det kan bidra til at politikere må tre bort fra sin stilling, men det er ikke mulig å se hvordan et nyhetsbilde med fokus på denne politiske skandalen kan påvirke holdninger hos velgerne. Men til tross for at man ikke kan se noe tilknytning til partifavorisering som resultat av skandalen i nyhetsbilde, åpner funnene opp til en rekke interessante emner som kan være relevant for videre forskning.

8.1 Forslag til videre forskning

Politiske skandaler og politisk tilhørighet

Partifavorisering og hvordan folk stiller seg til skandaler basert på hvilke parti de tilhører, og i hvilke parti den aktuelle skandalen skjer, er et interessant tema som gjerne burde undersøkes. Dette finner man dog ikke noe belegg for å si noe om i denne analysen. Meningsmålingen i figur 12 ga også et overraskende svakt bilde på om man kunne trekke noen sammenligninger fra mediene til partipreferanse, og at det burde undersøkes på en annen måte er det liten tvil om.

Det blir også vanskelig å si noe om hvordan avisene har vinklet emnene fordi det for eksempel var utelukkende vanskelig å skrive om metoo uten å nevne Arbeiderpartiet, som vi også ser i tabell 9 er blitt nevnt ved en rekke anledninger. Eneste avisen vi kan si noe om partitilhørighet er Klassekampen. Som uten tvil har hatt størst fokus på FrP under omtalen av politiske partier. Rødt er også nevnt en rekke ganger, men det vil være rimelig å anta at de har fått en noe større grad av positiv omtale enn hva FrP har fått. Til tross for at andre aviser også har hatt stort fokus på FrP, er Klassekampen eneste avisen man kan omtale på grunn av sitt klare politiske ståsted. At Høyre er nevnt ved mange anledninger er sikkert et reelt faktum, med tanke på at det med all sin sikkerhet blant annet ble skrevet et par saker om trakassering i partiet, om ikke ved andre anledninger, så i hvert fall ved avskjeden til Unge Høyre-lederen. Men ordtellingen kan også være et resultat av formelle forklaringer i artikler, som for eksempel at det ved bilder gjerne vil stå noe om personer til høyre, eller ved navngivning hvor

de skriver «fra høyre» og så ramser opp navnene. På grunn av denne risikoen vil det altså ikke være rimelig å tillegge ordtellingen hos Høyre noe vekt.

Avpolitisering konsekvens av personifisering?

Kaase (1994) legger fram forskning som viser at persondrama bidrar til avpolitisering, som førte til personifiseringshypotesen til Jenssen (2014). Videre ser vi at Bill Clinton i USA ikke mistet spesielt politisk kraft av skandalen hvor han hadde en affære, mens utenriksministeren Kanerva i Finland måtte si opp sin stilling. Funnene om avpolitisering er fra en studie i Tyskland, noe som kan tyde på at forholdene i USA og Tyskland kan være annerledes enn for Norge og Finland. Dette vil likevel være interessant å undersøke i forbindelse med Giske. Hvordan er egentlig hans politiske påvirkning i dag, i forhold til før skandalen?

«Hvordan har metoo-skandalen påvirket tillitsforholdet mellom politiker og velger?»

Dette er en problemstilling som jeg anser som svært interessant, og som et resultat av den utbredte oppmerksomheten rundt skandaler, tilbys muligheten til å undersøke hvilken effekt de har på samfunnet som helhet. Forskning viser at tillit til politikerne blir svekket i sammenheng med politiske skandaler (Bowler og Karp 2004; Marier 2014). Jeg er under den oppfatning at en skandale som omhandler seksuell trakassering har påvirket tillitsforholdet til, om ikke Arbeiderpartiet generelt, så i hvert fall til Trond Giske. Giske er ikke ute av politikken, men er medlem av Familie- og kulturkomiteen. At man kan fortsette i politikken kan tyde på at folk lett glemmer, og er også et bevis på at noen har en egenskap på å ri av stormen fra skandaler.

Tillit er også nevnt under tabell 10, og man ser at dette begrepet blir brukt relativt hyppig av enkelte aviser. Den avisen med størst fokus på tillit viser seg å være Adresseavisen, mens Klassekampen ikke har et like hyppig bruk av dette ordet som man gjerne kunne antatt. Tillit er et sentralt begrep under definisjonen av skandalepolitikk, men det er svært vanskelig å finne signifikante bevis som sier noe i den ene eller andre retningen. Dette fordi det er en

tilstand som er svært vanskelig å måle. Tilliten kan sees i sammenheng med andre faktorer som blir funnet i analysen, men det er viktig å ikke tillegge dette for stor vekt.

Litteraturliste

- Adam, Silke og Michaela Maier. 2010. «Personalization of Politics A Critical Review and Agenda for Research». *Annals of the International Communication Association*. 34(1) 213-257. <https://www.tandfonline.com/doi/abs/10.1080/23808985.2010.11679101>.
- Aftenposten Junior. 2020. Lest 16.april.2020
<https://www.aftenposten.no/kampanje/aftenpostenjunior/>.
- Allern, Sigurd og Ester Pollack. 2009. *Skandalenes markeds plass: politikk, moral og mediedrev*. Bergen: Fagbokforlaget.
- Allern, Sigurd og Ester Pollack. 2012. «Mediated Scandals» i *Scandalous! The Mediated Construction of Political Scandals in Four Nordic Countries* redigert av Sigurd Allern og Ester Pollack. 9-28. Gøteborg: Nordicom.
- Allern, Sigurd, Anu Kantola, Ester Pollack og Mark Blach-Ørsten. 2012. «Increased Scandalization. Nordic Political Scandals 1980-2010» i *Scandalous! The Mediated Construction of Political Scandals in Four Nordic Countries* redigert av Sigurd Allern og Ester Pollack. 29-50. Gøteborg: Nordicom.
- Allern, Sigurd og Ester Pollack. 2012. «The Marketplace of Scandal» i *Scandalous! The Mediated Construction of Political Scandals in Four Nordic Countries* redigert av Sigurd Allern og Ester Pollack. 181-190. Gøteborg: Nordicom.
- Al Sumait Loulwah, og Carlotta Domeniconi. 2008. «Text Clustering with Local Semantic Kernels» i *Survey of Text Mining II: Clustering, Classification, and Retrieval*, redigert av Berry, Michael W. og Malu Castellanos 2end ed. 87-105. London: Springer.
<https://link.springer.com/content/pdf/10.1007%2F978-1-84800-046-9.pdf>.
- Amedia s,v. «Vår historie» lest 15.april.2020.
<https://www.amedia.no/virksomheten/vare-mediehus/>
- Anandarajan, Murugan, Chelsey Hill og Thomas Nolan. 2019. *Practical Text Analytics. Maximizing the Value of Text Data*. Østerriket: Springer.

- Andreassen, Sunniva. 2019. «Kysset. Prøvde å kysse. Stryke på ryggen. Er det den alvorlige saken?». *Aftenposten*. 4.oktober 2019.
<https://www.aftenposten.no/meninger/kronikk/i/70J37B/kysset-proevde-aa-kysse-stryke-paa-ryggen-er-det-den-alvorlige-saken?>.
- Arbeidstilsynet s.v «Seksuell trakassering». Lest 29.juni.2020,
<https://www.arbeidstilsynet.no/tema/trakassering/seksuell-trakassering/>.
- Atekst. 2020. <https://web.retriever-info.com/services/archive.html?redir=true>.
- Baklund, Ea, Ellisiv Aure, og Ella Fyhn. 2019. ««Billig» og «pulbar» - #metoo kom aldri til skolegården. *Aftenposten*. 4.mars.2019.
<https://www.aftenposten.no/meninger/sid/i/e1VOPQ/billig-og-pulbar-metoo-kom-aldri-til-skolegaarden?>.
- Bíró, István og Jácint Szabó 2018. *Large scale link based Dirichlet allocation for web document classification*. Masteroppgave, Cornell University.
- Bjerke, Paul. 2009. «Skandaler og presse-moral» i *Skandalenes markeds plass: politikk, moral og mediedrev* redigert av Sigurd Allern og Ester Pollack 165-179. Bergen: Fagbokforlaget.
- Bjerke, Paul. 2012. «Media Victims and Media Morals» i *Scandalous! The Mediated Construction of Political Scandals in Four Nordic Countries* redigert av Sigurd Allern og Ester Pollack. 151-163. Gøteborg: Nordicom.
- Blei, David, M., Andrew Y. Ng, og Michael I. Jordan. 2003. «Latent Dirichlet Allocation». *Journal of Machine Learning Research*. 3 993-1022.
<http://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>.
- Bowler, Shaun og Jeffrey A. Karp. 2004. «Politicians, Scandals and Trust in Government». *Political Behavior* 26(3) 271-287.
<https://link.springer.com/content/pdf/10.1023/B:POBE.0000043456.87303.3a.pdf>.
- Bynander, Fredrik og Paul 't Hart. 2007. «The politics of party leader survival and succession:

- Australia in comparative perspective», *Australian Journal of Political Science* 42(1) 47-72.
<https://www.tandfonline.com/doi/pdf/10.1080/10361140601158542?needAccess=true>.
- Cai, Yanli, og Jian-Tao Sun. 2009. «Text Mining» i Liu, Ling, og M. Tamer Özsu (eds) *Encyclopedia of Database Systems*. Boston:Springer. 3062-3065.
https://link.springer.com/content/pdf/10.1007%2F978-0-387-39940-9_418.pdf.
- Cambria, Erik, Dipankar Das, Sivaji Bandyopadhyay og Antonio Feraco. 2017. «Affective Computing and Sentiment Analysis» i *A Practical Guide to Sentiment Analysis* redigert av Erik Cambria, Dipankar Das, Sivaji Bandyopadhyay og Antonio Feraco 1-10. Switzerland: Springer. <https://link.springer.com/content/pdf/10.1007%2F978-3-319-55394-8.pdf>.
- Campus, Donatella. 2010. «Mediatization and Personalization of Politics in Italy and in France: The Case of Berlusconi and Sarkozy». *The International Journal of Press/Politics*, 15(2) 219-235.
<https://journals.sagepub.com/doi/pdf/10.1177/1940161209358762>.
- Elden, John Christian. 2016. «Offentlig handling». *Store Norske Leksikon*. Oppdatert 1.januar. 2016. https://snl.no/offentlig_handling.
- Farrar, David og Jane Huffman Hayes. 2019. «A Comparison of Stemming Techniques in Tracing». *10th International Symposium on Software and Systems Traceability (SST)*. 37-44. <https://dl.acm.org/doi/pdf/10.1109/SST.2019.00017>.
- Fileborn, Bianca og Rachel Loney-Howes. 2019. *#MeToo and the Politics of Social Change*. Switzerland: Springer. <https://link.springer.com/content/pdf/10.1007%2F978-3-030-15213-0.pdf>.
- Frøslie, Kathrine Frey. 2019. «Korrelasjon». *Store Norske Leksikon*. Oppdatert 20.september.2019.<https://snl.no/korrelasjon>.
- Garvik, Olav. 2019. «Edda Media AS». *Store Norske Leksikon*. Oppdatert 30.desember.2019. https://snl.no/Edda_Media_AS.

- Gentzkow, Matthew, Bryan Kelly, og Matt Taddy. 2019. «Text as Data». *Journal of Economic Literature*. 57(3) 535-574.
<https://web.stanford.edu/~gentzkow/research/text-as-data.pdf>.
- Gjersdal, Aud. 2019. «metadata». *Store Norske Leksikon*. Oppdatert 11.november.2019.
<https://snl.no/metadata>.
- Grimmer, Justin, og Brandon M. Stewart. 2013. «The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts». *Political Analysis* 21(3) 267-297.
https://www.cambridge.org/core/services/aop-cambridge-core/content/view/F7AAC8B2909441603FEB25C156448F20/S1047198700013401a.pdf/text_as_data_the_promise_and_pitfalls_of_automatic_content_analysis_methods_for_political_texts.pdf.
- Ignatow, Gabe og Rada Mihalcea. 2017. «Social Science and the Digital Text Revolution» i *Text Mining: a Guidebook for the Social Sciences*, 1-13
<https://methods.sagepub.com/base/download/BookChapter/text-mining/i440.xml>.
- Ignatow, Gabe og Rada Mihalcea. 2017. «Basic Text Processing» i *Text Mining: a Guidebook for the Social Sciences*, 1-10.
<https://methods.sagepub.com/base/download/BookChapter/text-mining/i659.xml>.
- Ignatow, Gabe og Rada Mihalcea. 2017. «Thematic Analysis, Qualitative Data Analysis Software, and Visualization» i *Text Mining: a Guidebook for the Social Sciences*, 1-13. <https://methods.sagepub.com/base/download/BookChapter/text-mining/i761.xml>.
- Ignatow, Gabe og Rada Mihalcea. 2017. «Sentiment Analysis» i *Text Mining: a Guidebook for the Social Sciences* 1-7.
<https://methods.sagepub.com/base/download/BookChapter/text-mining/i1084.xml>.
- Ignatow, Gabe og Rada Mihalcea. 2017. «Topic Models» i *Text Mining: a Guidebook for the Social Sciences* 1-6.
<https://methods.sagepub.com/base/download/BookChapter/text-mining/i1110.xml>.
- Isotalus, Pekka og Merja Almonkari. (2014). «Political Scandal Test Trust in Politicians. The

- Case of the Finnish Minister Who Resigned Because of His Text Messages». *Nordicom Review* 35(2). 3-16.
file:///C:/Users/vilde/Downloads/Political_Scandal_Tests_Trust_in_Politicians.pdf.
- Jenssen, T. Anders. 2014. «Medierte politiske skandaler: Sårbare politikere – usårbare partier?». *Norsk Medietidsskrift* 21(2) 100-118.
https://www.idunn.no/file/pdf/66705607/medierte_politiske_skandaler_saarbare_politikere_-_usaarbare.pdf.
- Johannessen, Christine Rian. 2013. *Partiledersdebatt i norske medier. En innholdsanalyse av VG og Aftenposten*. Master, Universitetet i Oslo, Oslo.
- Kaase, Max. 1994. «Is There Personalization in Politics?» Candidates and Voting Behavior in Germany. *International Political Science Review*. 15(3) 211-230.
<https://journals.sagepub.com/doi/pdf/10.1177/019251219401500301>.
- Karlsen, Rune. 2011. «A Platform for Individualized Campaigning? Social Media and Parliamentary Candidates in the 2009 Norwegian Election Campaign». *Policy & Internet* 3(4) 1-27. <https://onlinelibrary.wiley.com/doi/epdf/10.2202/1944-2866.1137>.
- Karvonen, L. 2010. *The personalisation of politics. A study of parliamentary democracies*. Colchester: ECPR Press.
https://books.google.no/books?hl=no&lr=&id=zcdMAQAAQBAJ&oi=fnd&pg=PP7&dq=Karvonen,+L.+2010.+The+personalisation+of+politics.+A+study+of+parliamentary+democracies.+ECPR+Press:+Colchester.&ots=dfSTk_EdOr&sig=fa6CZ2EzXWkazkaiDURFAhPMitw&redir_esc=y#v=onepage&q&f=false.
- Kerby, Matthew og Raj S. Chari. 2002. «‘Policy Scandals’: A Spanish Case». *Government and Opposition* 37(3) 409-426. <https://onlinelibrary.wiley.com/doi/epdf/10.1111/1477-7053.00107>.
- Kirsebom, Kaja, Ruben Solvang, Magnus Helle, Arne Fredrik Næss, Vidar Kvien og Marco Vaglieri. 2020. «Hvem er Silje». *NRK dokumentar*. Lest 29.juni. 2020,
https://www.nrk.no/dokumentar/xl/hvem-er-silje_-1.15005229#authors--expand.
- Kuhn, Kenneth D. 2018. «Using structural topic modeling to identify latent topics and trends

- in aviation incident reports». *Transportation Research Part C*. 87 105-122.
<https://reader.elsevier.com/reader/sd/pii/S0968090X17303881?token=09B3F9384EC6DBAF9DE0A2F01AC5D91F72F79CF984A0E4BF3B3BBF733FD62CE1CC708FBA67D574D21A007DA4484E2987>.
- Lovdata. 2018. «Lov om likestilling og forbud mot diskriminering (likestillings- og diskrimineringsloven)». Oppdatert 01.januar, 2018
https://lovdata.no/dokument/NL/lov/2017-06-16-51#KAPITTEL_2
- Maier, Jürgen. 2014. «The impact of political scandals on political support: An experimental test of two theories». *International Political Science Review* 32(3) 283-302.
<https://journals.sagepub.com/doi/pdf/10.1177/0192512110378056>.
- Manikonda, Lydia, Ghazaleh Beigi, Subbarao Kambhampati, og Huan Liu. 2018. «#metoo Through the Lens of Social Media» i *Social, Cultural, and Behavioral Modeling*, redigert av ed. Thomson Robert, Christopher Dancy, Ayaz Hyder og Halil Bisgin. Switzerland: Springer. 104-110. https://quanttext.com/wp-content/uploads/2018/09/2018_Book_SocialCulturalAndBehavioralMod.pdf.
- Manning, Christopher D., Prabhakar Raghavan, og Hinrich Schütze. 2008. «Introduction to Information Retrieval». *Cambridge University Press*.
<https://www.cambridge.org/core/books/introduction-to-information-retrieval/669D108D20F556C5C30957D63B5AB65C#>.
- Mediebedriftene, s.v. «Lesertall» lest 27.april.2020
<https://www.mediebedriftene.no/tall-og-fakta/lesertall/>.
- Mediebedriftene, s.v. «Medienes rammebetingelser» lest 18.februar.2020,
<https://www.mediebedriftene.no/mediepolitikk/medienes-rammebetingelser/>.
- Mediebedriftene, s.v «Mediebransjens samfunnsansvar» lest 26.februar.2020,
<https://www.mediebedriftene.no/mediepolitikk/mediebransjens-samfunnsansvar/>.
- Mediebedriftene, s.v. «Om Mediebedriftenes landsforening» lest 18.januar.2020,
<https://www.mediebedriftene.no/om-mbl/om-mediebedriftenes-landsforening-mbl/>.

Mediebedriftene, s.v. «Opplagstall 20/1» lest 07.april.2020,

https://www.mediebedriftene.no/siteassets/dokumenter/tall-og-fakta/medietall/20_1/medietall_20_1_presentasjon.pdf.

Midtbø, Tor. 2007. *Skandaler i Norsk Politikk*. Oslo: Universitetsforlaget AS.

Midtbø, Tor. 2012. «Do Mediated Political Scandals Affect Party Popularity in Norway?» i *Scandalous! The Mediated Construction of Political Scandals in Four Nordic Countries* redigert av Sigurd Allern og Ester Pollack. 151-163. Gøteborg: Nordicom.

NABO s.v «samfunnstopp». Det Norske Akademis Ordbok. Lest 30.juni.2020

<https://naob.no/ordbok/samfunnstopp>.

Nidhi. 2017. «Number of Topics for LDA on poems from Elliston Poetry Archive»

RPubs by RStudio lest 23.juni.2020 <https://rpubs.com/MNidhi/NumberoftopicsLDA>.

Norges Offentlige Utredning (NOU) 2000:15 (Kulturdepartementet). 2000. «Pressepolitikk ved tusenårsskifte». *Statens Forvaltningstjeneste Informasjonsforvaltning*. Oslo. lest 27.01.2020 <https://www.regjeringen.no/no/dokumenter/nou-2000-15/id376239/?ch=1>

Orgeret, Kristin Skare. 2019. «metoo». *Store Norske Leksikon*. Oppdatert 9.juli.2019

<https://snl.no/metoo>.

Owen, Diana. 2000. «Popular Politics and the Clinton/Lewinsky Affair: The Implications of Leadership». *Political Psychology* 21(1) 161-177.

<https://onlinelibrary.wiley.com/doi/epdf/10.1111/0162-895X.00182>.

Pedersen, Erik, Bernt. 2019. «Musikk er en del av problemet». *Dagsavisen*. 4.april.2019.

<https://www.dagsavisen.no/kultur/musikk-er-en-del-av-problemet-1.1460039>.

Pennebaker, James. W, og Laura A. King. 1999. «Linguistic Styles: Language Use as an Individual Difference». *Journal of Personality and Social Psychology* 77(6) 1296-1312. <https://psycnet.apa.org/fulltext/1999-15054-015.pdf>.

Pettersen, Øyvind, Breivik. 2016. «Aftenposten». *Store Norske Leksikon*. Oppdatert

10.juni.2016. <https://snl.no/Aftenposten>.

- Pettersen, Øyvind, Breivik. 2016. «Klassekampen». *Store Norske Leksikon*. Oppdatert 19.august.2016. <https://snl.no/Klassekampen>.
- Pettersen, Øyvind, Breivik. 2017. «Romerikes Blad». *Store Norske Leksikon*. Oppdatert 20.januar.2017. https://snl.no/Romerikes_Blad.
- Pettersen, Øyving, Breivik. 2017. «Stavanger Aftenblad». *Store Norske Leksikon*. Oppdatert 6.mars.2017. https://snl.no/Stavanger_Aftenblad.
- Pettersen, Øyvind, Breivik. 2019. «Dagbladet». *Store Norske Leksikon*. Oppdatert 14.mars.2019 <https://snl.no/Dagbladet>.
- Pettersen, Øyvind, Breivik. 2019. «Bergens Tidende». *Store Norske Leksikon*. Oppdatert 31.juli.2019 https://snl.no/Bergens_Tidende.
- Pettersen, Øyvind, Breivik. 2019. «Adresseavisen». *Store Norske Leksikon*. Oppdatert 15.august.2019 <https://snl.no/Adresseavisen>.
- Pettersen, Øyvind, Breivik. 2020. «Morgenbladet». *Store Norske Leksikon*. Oppdatert 2.juni.2020 <https://snl.no/Morgenbladet>.
- Poll of Polls* s.v «Gjennomsnitt av nasjonale meningsmålinger om stortingsvalg» lest 29.juni.2020 <http://www.pollofpolls.no/?cmd=Stortinget&do=visallesnitt>.
- Poll of polls* s.v «Stortingsvalg: Hele landet» lest 27.juni.2020, <http://www.pollofpolls.no/?cmd=Stortinget>.
- Prestgård, Sofie. 2018. «Metoo-effekt i alle partiene.» *Dagsavisen* 12.februar.2020. <https://www.dagsavisen.no/nyheter/innenriks/metoo-effekt-i-alle-partiene-1.1089066>.
- Quinn, Kevin. M, Burt L. Monroe, Micheal Colaresi, Michael H. Crespin og Dragomir R. Radev. 2010. «How to Analyze Political Attention with Minimal Assumptions and Costs». *American Journal of Political Science* 54(1) 209-228. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-5907.2009.00427.x>.
- Retriever. 2020. Hentet 23.mars 2020 <https://web.retriever-info.com/services/archive/search>.

Roberts, Margaret E., Brandon M. Stewart og Dustin Tingley. 2019. «stm: An R Package for Structural Topic Models». *Journal of Statistical Software* 91(2) 1-40.

DOI: 10.18637/jss.v091.i02.

Rødt s.v «Åpenhet». Lest 1.juli.2020. <https://roedt.no/apenhet>.

Smith-Meyer, Trond og Øyvind Breivik Pettersen. 2016. «Drammens Tidene». *Store Norske Leksikon*. Oppdatert 27.mai.2016. https://snl.no/Drammens_Tidende.

Smith-Meyer, Trond. 2018. «tabloid». *Store Norske Leksikon*. Oppdatert 20.februar.2018. <https://snl.no/tabloid>.

Silge, Julia og David Robinson. 2020. «Text Mining With R». Oppdatert 7.mars.2020.

<https://www.tidytextmining.com/index.html>.

Thompson, John. B. (2000). «Political Scandal: Power and Visibility in the Media Age».

Cambridge: Wily 2000.

[https://books.google.no/books?hl=no&lr=&id=DBAYaHewWiUC&oi=fnd&pg=PP5&dq=Thompson,+J.B.+\(2000\)+Political+Scandal.+Cambridge:+Polity+Press.&ots=t4qgEuXLJB&sig=JJOWEcWnmJ3C8xXrfQZHFdrefEw&redir_esc=y#v=onepage&q&f=false](https://books.google.no/books?hl=no&lr=&id=DBAYaHewWiUC&oi=fnd&pg=PP5&dq=Thompson,+J.B.+(2000)+Political+Scandal.+Cambridge:+Polity+Press.&ots=t4qgEuXLJB&sig=JJOWEcWnmJ3C8xXrfQZHFdrefEw&redir_esc=y#v=onepage&q&f=false)

Van Aelst, Peter, Tamir Sheafer og James Stanyer. 2011. «The personalization of mediated political communication: A review on concepts, operationalizations and key findings». *Journalism* 13(2) 203-220.

<https://journals.sagepub.com/doi/pdf/10.1177/1464884911427802>.

Thorbjørnsen, Kjersti. 2010. «Med media som 7.sans. Norsk politikk som ekstremt case».

Sosiologi i dag, 40(4) 77-91.

<file:///C:/Users/vilde/Documents/SAMPOL2020/Masteroppgave/Teori/Thorbjørnsen,%20K%202010.pdf>.

Thorsen, Dag Einar. 2019. «radikal». *Store Norske Leksikon*. Oppdatert 26.juni.2019.

<https://snl.no/radikal>.

Tjelmeland, Hallvard og Norsk presses historie (Universitetsforlaget). 2019. «Nordlys». *Store*

Norske Leksikon. Oppdatert 23.august.2019 <https://snl.no/Nordlys>.

Velldal Erik. 2017. «Sentiment Analysis for Norwegian Text». Universitetet i Oslo:

Department of Informatics.

Wilkerson, John, og Andreu Casas. 2017. «Large-Scale Computerized Text Analysis in

Political Science: Opportunities and Challenges». *Annual Review of Political Science*.
20 529-544. <https://www.annualreviews.org/doi/pdf/10.1146/annurev-polisci-052615-025542>.

Wu, Junjie. 2012. «Cluster Analysis and K-means Clustering: An Introduction» i *Advances in*

K-means Clustering. Berlin: Springer 1-16.

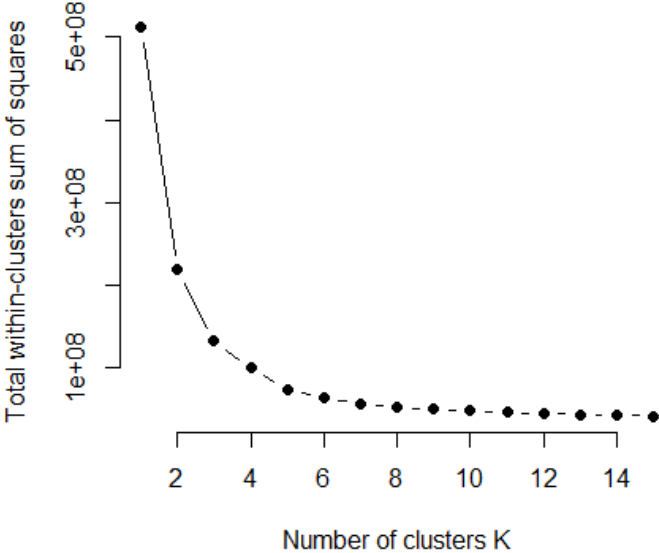
<https://link.springer.com/content/pdf/10.1007%2F978-3-642-29807-3.pdf>

2020 Presse.no «Vær Varsom-plakaten». *Pressens Faglige Utvalg*. Lest 25.mars.2020

<https://presse.no/pfu/etiske-regler/vaer-varsom-plakaten/>.

Vedlegg 1: Visualisering ved hjelp av figurer

Figur 10: Antall objekter, K



Vedlegg 2: Deskriptiv statistikk

Tabell 1: Ord som korrelerer over 98% med «metoo»

bevise	ledende	ofre	sender	sikre	urovekkende	aftenpostens	altfor
1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.99
andelen	ber	dag	dags	dekning	derfor	dersom	eiendom
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
fred	fylle	glem	grunner	går	helse	hensyn	hjemmet
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
høyeste	innført	island	kontroll	kritiske	kvinner	landets	lanserer
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
lengre	ligge	menn	ment	muslim	norge	norske	næringsliv
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
omtalen	opphold	penger	perioden	preges	prinsipielle	pålagt	reaksjonen
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
regjeringen	resultat	retorikk	rette	slått	sosiale	stortinget	støtten
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
støttet	svake	tid	tide	toppleidelse	tunge	tusen	vold
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
vårt	øvrig	begrense	betale	brukt	bør	diskusjon	dratt
0.99	0.99	0.98	0.98	0.98	0.98	0.98	0.98
eksempler	eksisterende	faktiske	feil	foreslo	front	godta	hijab
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
ille	januar	kampen	klokt	konkurransen	kontroversene	krever	likevel
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
lære	miste	målet	navngitt	negative	offentlighetens	oppvask	plage
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
politi	prestisjefylte	påstå	redusert	relativt	ressurser	samfunnet	sitert
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
skade	skiller	slo	stille	stilling	største	suppe	svein
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
sårbare	tilbakeslag	tillegg	tom	tvers	tåle	utføre	varsel
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
vedta	viser	årsakene					
0.98	0.98	0.98					

Tabell 2: Ord som korrelerer over 98% med «politikker»

bra	kvinnekamp	nettverk	oppføre	tillit	berørte	datter	dårlige	femten	forsvant
1.00	1.00	1.00	1.00	1.00	0.99	0.99	0.99	0.99	0.99
geir	glad	jentene	jurist	mase	nittitallet	opprørt	personalet	rakk	rulles
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
snudde	sure	sve	voldsomt	works	aufs	bråk	drama	dro	evnet
0.99	0.99	0.99	0.99	0.99	0.98	0.98	0.98	0.98	0.98
frister	høyest	høylytt	inntil	kjøre	konkretisere	lykkes	maktposisjon	mediestormen	oppe
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
oppfører	rart	samarbeide	spilles	stian	sønnene	tidsskillet	tilliten	trond	ungt
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
useriøst	utvikles	voksen	øvelser						
0.98	0.98	0.98	0.98						

Tabell 3: Ord som korrelerer over 98% med «tillit»

politikker	bra	dårlige	forsvant	inntil	jentene	jurist	kvinnekamp	landskap
1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
nettverk	nittitallet	operasjoner	oppføre	oppfører	opprørt	rulles	snudde	spilles
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
sve	tilliten	ungdommen	ungt	voksen	voldsomt	works	alder	assosiasjoner
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.98
berørte	bokprosjektet	borti	drama	femten	fingeren	frister	geir	ingrid
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
kandidat	kjøre	konkretisere	likene	mangle	mase	mediestormen	nettbrett	partikollega
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
personalet	rakk	selvfølge	sjekker	skjønne	stian	sure	svekkede	synde
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
torunn	tung	useriøst	øvelser					
0.98	0.98	0.98	0.98					

Tabell 4: Ord som korrelerer over 98% med «egenskaper»

nærmest	ordinært	aktualitet	albumet	avgjørende	bakt	bevissthet	bruk	bruke	bukk
1.00	1.00	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
bygge	demokratiet	drar	dører	ene	faglige	fall	fantasi	fantastisk	fint
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
folkeparti	forsøkes	fort	ganske	gjerner	grepene	henge	henger	henvist	indre
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
innflytter	johan	kompetanse	kulturmann	kvaliteter	lands	lengst	lett	ligger	manisk
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
manns	mer	merke	mil	mild	mørkere	mørkets	nei	nikolaj	overtok
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
papiraviser	reiser	rettes	rosa	samle	snakke	synlig	taper	tenkt	tette
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99
typisk	uformelt	ung	usikker	vanlige	visjoner	volden	work	aktiv	alltid
0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.98	0.98
ambisiøse	anstrøk	besitter	bidrar	bosatt	diskutert	dragning	dras	drift	ekko
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
ende	enkelt	eventuelt	finne	fordommene	forestillinger	gir	greier	helst	hold
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
hovedsakelig	høyst	illustrerer	kalles	kanskje	kjennes	kjøper	kroppen	liten	mangler
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
maskin	nytt	ofte	ord	overflaten	plan	plass	politisk	prosjektet	sitter
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
skape	skoler	skritt	skyldes	små	sprang	stadig	sterkere	stor	tider
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98
tilfeldige	troll	tror	tyder	utvikling	vanskeligere	veldig	verdighet		
0.98	0.98	0.98	0.98	0.98	0.98	0.98	0.98		

Tabell 7: Vanlige ord for hvert emne

Topic 1 Top Words:
Highest Prob: trakassering, seksuell, flere, mener, andre, mer, mye
FREX: uit, myrseth, samfunnstoppen, beldo, normann, itromsø, nustad
Lift: uit, samfunnstoppen, stykkene, beldo, ledertype, normann, nso
Score: uit, factbox, pluss, samfunnstoppen, myrseth, image, beldo
Topic 2 Top Words:
Highest Prob: mer, andre, kvinner, flere, hele, mye, helt
FREX: lig, ker, romanen, bergman, ver, nen, ger
Lift: marokko, wittbrattstrøm, morgenbladets, bachewiig, valeria, hassan, wild
Score: lig, nen, ger, ter, ner, kvin, ker
Topic 3 Top Words:
Highest Prob: trakassering, seksuell, flere, kvinner, egen, giske, utenfor
FREX: nominasjonen, water, spacey, shape, vgs, schau, navarsete
Lift: saugestad, mosveen, cara, broer, skrivet, downsizing, simmons
Score: imagetext, holmes, saugestad, ingvill, trump, mosveen, spacey
Topic 4 Top Words:
Highest Prob: giske, trakassering, andre, kvinner, flere, egen, seksuell
FREX: aglen, nyhetsbrevet, trondheim, ottervik, ntnu, meninger, innboksen
Lift: nyhetsbrevet, innboksen, abonner, reformasjonen, espira, ntbap, nettkampanjen
Score: imagetext, nyhetsbrevet, aglen, pluss, trondheim, tiller, abonner
Topic 5 Top Words:
Highest Prob: kvinner, trakassering, flere, andre, seksuell, mer, egen
FREX: kristiansand, agder, sørlandet, øgrey, niva, arendal, kilden
Lift: lillian, lillesand, avdelingslederen, øgrey, åsland, jannike, skagestad
Score: image, factbox, kristiansand, niva, røen, sørlandet, avdelingslederen
Topic 6 Top Words:
Highest Prob: kvinner, metoo, andre, seksuell, trakassering, mer, flere
FREX: frå, meiner, høgge, venstresida, gen, ring, braanen
Lift: bendik, mímir, opnar, bjørgulv, håland, retningsliner, ettertrykk
Score: frå, høgge, ger, gen, nen, bjørgulv, bendik
Topic 7 Top Words:
Highest Prob: flere, kvinner, trakassering, seksuell, giske, andre, egen
FREX: video, helena, dagbladets, imagetext, eurovision, spacey, løperen
Lift: annonsen, mogen, lea, bratland, markle, dolores, contest
Score: imagetext, eurovision, annonsen, helena, burøe, scanpix, ntb
Topic 8 Top Words:
Highest Prob: trakassering, flere, kvinner, andre, seksuell, egen, metoo
FREX: nhh, espedal, svendsen, eikefjord, jenten, eirin, hope
Lift: nemnden, studentforeningen, nhh, torgallmenningen, damsgaard, tillitsmannen, kråkevik
Score: nhh, factbox, image, jenten, info, arbeidspartiet, røen
Topic 9 Top Words:
Highest Prob: kvinner, andre, trakassering, flere, mer, egen, seksuell
FREX: aftenbladet, nybø, uis, rogaland, persbrandt, johnson, statoil
Lift: uis, marriage, wendy, heine, sauda, almaas, merkesaker
Score: image, factbox, uis, aftenbladet, jenten, persbrandt, info
Topic 10 Top Words:
Highest Prob: giske, andre, kvinner, mer, trondheim, flere, trakassering
FREX: trondheim, ottervik, aglen, ntnu, trøndelag, adressa, adresseavisens
Lift: papiravis, abonnent, sertifisert, belønner, newspaper, kalbakk, hjertesak
Score: trondheim, aglen, imagetext, trøndelag, husby, tiller, tlf

Topic 11 Top Words:
Highest Prob: trakassering, seksuell, flere, andre, mener, pluss, giske
FREX: holstad, pluss, sørensen, ofstad, revyen, dagfinn, russelåter
Lift: revyen, teit, dagfinn, russelåter, ofstad, andreassen, sissel
Score: pluss, andreassen, holstad, ofstad, sunniva, revyen, lundberg
Topic 12 Top Words:
Highest Prob: troms, trakassering, seksuell, flere, andre, mener, kvinner
FREX: troms, silje, ungdom, andreassen, kvinne, finnmark, larsen
Lift: troms, nettavisen, silje, finnmark, eilertsen, itromsø, cecilie
Score: troms, pluss, andreassen, eilertsen, sunniva, virtanen, seksuell
Topic 13 Top Words:
Highest Prob: tromsø, trakassering, seksuell, flere, andre, mener, mer
FREX: tromsø, silje, kvinne, ungdom, andreassen, sjefen, finnmark
Lift: tromsø, nettavisen, silje, utro, finnmark, arktis, eilertsen
Score: tromsø, pluss, andreassen, eilertsen, sunniva, virtanen, nettavisen
Topic 14 Top Words:
Highest Prob: unn, trakassering, andre, flere, seksuell, mer, kvinner
FREX: unn, bibliotek, utro, intelligens, nettavisen, akt, solstad
Lift: unn, utro, akt, kommunisere, nettavisen, aktørene, nordover
Score: unn, andreassen, utro, nettavisen, bibliotek, virtanen, nielsen
Topic 15 Top Words:
Highest Prob: ingebrigtsen, trakassering, seksuell, pluss, flere, andre, mener
FREX: ingebrigtsen, pluss, andreassen, varsle, undersøkelsen, uønsket, sunniva
Lift: ingebrigtsen, andreassen, silje, teit, pluss, ofstad, sunniva
Score: ingebrigtsen, pluss, andreassen, lundberg, holstad, ofstad, sunniva
Topic 16 Top Words:
Highest Prob: utroskap, trakassering, seksuell, flere, andre, kvinner, mener
FREX: utroskap, kvinne, veldig, tror, årene, jobb, litt
Lift: utroskap, nettavisen, kommunisere, akt, aktørene, silje, næringslivet
Score: utroskap, pluss, andreassen, sunniva, nettavisen, virtanen, seksuell
Topic 17 Top Words:
Highest Prob: ludvigsen, trakassering, seksuell, flere, andre, kvinner, mener
FREX: ludvigsen, andreassen, varsler, politiet, uønsket, varslene, undersøkelsen
Lift: ludvigsen, kommunisere, utro, eilertsen, pressedekningen, brystene, jørn
Score: ludvigsen, andreassen, pluss, lundberg, eilertsen, sunniva, martha
Topic 18 Top Words:
Highest Prob: knutsen, trakassering, seksuell, flere, andre, kvinner, mener
FREX: knutsen, eilertsen, mathilde, andreassen, natta, framover, amundsen
Lift: knutsen, eilertsen, mathilde, utro, cecilie, kommunisere, nettavisen
Score: knutsen, andreassen, eilertsen, sunniva, imagetext, pluss, lundberg
Topic 19 Top Words:
Highest Prob: nordnorge, trakassering, seksuell, andre, flere, kvinner, mener
FREX: nordnorge, pluss, seksuell, kvinne, trakassering, saker, veldig
Lift: nordnorge, finnmark, nettavisen, silje, karoline, teit, næringslivet
Score: nordnorge, pluss, andreassen, sunniva, ofstad, seksuell, fpu
Topic 20 Top Words:
Highest Prob: befølte, trakassering, seksuell, flere, andre, kvinner, mener
FREX: befølte, varsler, kvinne, seksuell, politiet, trakassering, ungdom
Lift: befølte, utro, kommunisere, jørn, nettavisen, silje, eilertsen
Score: befølte, pluss, andreassen, lundberg, eilertsen, virtanen, sunniva

Tabell 8: Korrelasjonstest for de 20 emnene vist i figur 7

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]	[,9]	[,10]	[,11]	[,12]	[,13]	[,14]	[,15]	[,16]	[,17]	[,18]	[,19]	[,20]
[1,]	1.00	-0.11	-0.11	-0.13	-0.11	-0.11	-0.11	-0.11	-0.11	-0.11	-0.11	1.00	1.00	-0.07	0.64	1.00	0.90	-0.08	0.99	0.95
[2,]	-0.11	1.00	-0.11	-0.13	-0.11	-0.11	-0.11	-0.11	-0.11	-0.11	-0.11	-0.12	-0.11	0.95	-0.33	-0.09	-0.15	-0.13	-0.16	-0.11
[3,]	-0.11	-0.11	1.00	-0.13	-0.11	-0.11	-0.11	-0.11	-0.11	-0.11	-0.11	-0.12	-0.11	-0.15	-0.08	-0.14	0.21	-0.12	-0.16	0.17
[4,]	-0.13	-0.13	-0.13	1.00	-0.13	-0.13	-0.13	-0.13	-0.13	0.02	0.02	-0.12	-0.12	-0.17	0.45	-0.12	-0.20	-0.13	-0.06	-0.19
[5,]	-0.11	-0.11	-0.11	-0.13	1.00	-0.11	-0.11	-0.11	-0.11	-0.11	-0.11	-0.11	-0.12	-0.16	-0.33	-0.16	-0.20	-0.13	-0.13	-0.17
[6,]	-0.11	-0.11	-0.11	-0.13	-0.11	1.00	-0.11	-0.11	-0.11	-0.11	-0.11	-0.09	-0.09	0.18	0.00	-0.09	-0.18	-0.07	-0.08	-0.14
[7,]	-0.11	-0.11	-0.11	-0.13	-0.11	-0.11	1.00	-0.11	-0.11	-0.11	-0.11	-0.10	-0.11	-0.16	0.06	-0.12	0.14	1.00	-0.14	-0.02
[8,]	-0.11	-0.11	-0.11	-0.13	-0.11	-0.11	-0.11	1.00	-0.11	-0.11	-0.11	-0.12	-0.12	-0.16	-0.34	-0.12	-0.20	-0.13	-0.14	-0.17
[9,]	-0.11	-0.11	-0.11	-0.13	-0.11	-0.11	-0.11	-0.11	1.00	-0.11	-0.11	-0.11	-0.12	-0.15	-0.33	-0.07	-0.20	-0.13	-0.04	-0.17
[10,]	-0.11	-0.11	-0.11	0.02	-0.11	-0.11	-0.11	-0.11	-0.11	1.00	1.00	-0.12	-0.10	-0.14	0.30	-0.09	-0.16	-0.11	-0.10	-0.17
[11,]	-0.11	-0.11	-0.11	0.02	-0.11	-0.11	-0.11	-0.11	-0.11	1.00	1.00	-0.12	-0.10	-0.14	0.30	-0.09	-0.16	-0.11	-0.10	-0.17
[12,]	1.00	-0.12	-0.12	-0.12	-0.11	-0.09	-0.10	-0.12	-0.11	-0.12	-0.12	1.00	1.00	-0.07	0.65	1.00	0.90	-0.07	0.99	0.95
[13,]	1.00	-0.11	-0.11	-0.12	-0.12	-0.09	-0.11	-0.12	-0.12	-0.10	-0.10	1.00	1.00	-0.06	0.65	1.00	0.90	-0.08	0.99	0.95
[14,]	-0.07	0.95	-0.15	-0.17	-0.16	0.18	-0.16	-0.16	-0.15	-0.14	-0.14	-0.07	-0.06	1.00	-0.27	-0.04	-0.13	-0.16	-0.10	-0.08
[15,]	0.64	-0.33	-0.08	0.45	-0.33	0.00	0.06	-0.34	-0.33	0.30	0.30	0.65	0.65	-0.27	1.00	0.64	0.62	0.10	0.67	0.61
[16,]	1.00	-0.09	-0.14	-0.12	-0.16	-0.09	-0.12	-0.12	-0.07	-0.09	-0.09	1.00	1.00	-0.04	0.64	1.00	0.89	-0.09	1.00	0.94
[17,]	0.90	-0.15	0.21	-0.20	-0.20	-0.18	0.14	-0.20	-0.20	-0.16	-0.16	0.90	0.90	-0.13	0.62	0.89	1.00	0.17	0.87	0.98
[18,]	-0.08	-0.13	-0.12	-0.13	-0.13	-0.07	1.00	-0.13	-0.13	-0.11	-0.11	-0.07	-0.08	-0.16	0.10	-0.09	0.17	1.00	-0.11	0.00
[19,]	0.99	-0.16	-0.16	-0.06	-0.13	-0.08	-0.14	-0.14	-0.04	-0.10	-0.10	0.99	0.99	-0.10	0.67	1.00	0.87	-0.11	1.00	0.92
[20,]	0.95	-0.11	0.17	-0.19	-0.17	-0.14	-0.02	-0.17	-0.17	-0.17	-0.17	0.95	0.95	-0.08	0.61	0.94	0.98	0.00	0.92	1.00

Tabell 1: Ordfrekvens

Ord	Frekvens
Kvinner	10622
Andre	10140
Trakassering	9923
Flere	9463
Seksuell	8941
Mer	8794
Giske	8476
Metoo	8401
Egen	7652
Mener	7247
Mye	6772
Menn	6627
Utenfor	6048
Enkelte	5956
Lov	5884
Hele	5673
Går	5536
Kommer	5326
Organisasjon	5270
Fikk	5236