

# Angreb i den offentlige debat på Facebook

Analyse & Tal F.M.B.A har med støtte fra TrygFonden udviklet den hidtil mest avancerede dansksprogede algoritme til detektion af angreb i den offentlige debat på Facebook. Baseret på assisteret maskinlæring og deep learning-teknologi er det nu muligt at kortlægge andelen af hhv. angreb, anstødeligt og hadefuldt sprog i kommentarsporene på offentlige danske Facebooksider.

&#



# Om rapporten

Facebook er blevet dét digitale forum, hvorigennem flest danskere tilgår og udveksler holdninger og information. Gennem årene har der løbende været stillet spørgsmålstejn ved tonen i debatten på Facebook. Er tonen for hård? Får bestemte samfundsgrupper en særligt hård medfart i kommentarsporene? Med brug af deep learning-teknologi er det nu for første gang muligt at kortlægge hadefuldt og anstødeligt sprog i den offentlige debat på Facebook.

Angrebsalgoritmen er den første af to algoritmer, der lanceres i forbindelse med projektet "De digitale medborgerhuse". Nummer to i rækken vil blive det første danske bud på en algoritme til detektion af anerkendelse i den offentlige debat på Facebook. Men vi starter med angrebene. Det er ikke altid en munter omgang at dykke ned i de vrede kommentarer på de sociale medier, men forhåbentlig kan kortlægningen bidrage til, at vi kan tale om tonen i debatten på et mere velfunderet grundlag, ligesom de nye sprogalgoritmer kan bidrage til at højne den danske forskning på området og med tiden videreudvikles til at understøtte moderationsarbejdet.

## Rapporten vil:

1. Redegøre for metoden bag og arbejdet med udviklingen af en dansksproget algoritme til detektion af sproglige angreb, herunder anstødeligt og hadefuldt sprogbrug. Algoritmerne gøres tilgængelige for forskere som open-source på [www.huggingface.co](http://www.huggingface.co).
2. Vise algoritmerne i aktion gennem en analyse af angreb, herunder hadefuldt og anstødeligt sprog i **63 mio.** kommentarer fra **199** politikeres og **477** mediers Facebooksider. Datasættet repræsenterer alle tilgængelige kommentarer på siderne for perioden januar 2019 til og med februar 2021.

God læselyst!



Analyse & Tal

Med støtte fra  
**TrygFonden**

# Danmarks Digitale Medborgerhuse



## Om projektet:

Analysen af anstødeligt og hadefuldt sprog på Facebook er den første udgivelse i en større kortlægning af "Danmarks Digitale Medborgerhuse". Projektet har som sit hovedformål at kortlægge og analysere medborgerskabet i det digitale demokratis vigtigste offentlige fora. Disse fora tæller mediernes, politikernes og borgernes offentlige Facebooksider og grupper. Til formålet udvikles nogle redskaber (algoritmer) til at analysere angreb og anerkendelse i den offentlige debat. Projektet vil resultere i open source sprogalgoritmer og fire rapporter:

- 1. Angreb i den offentlige debat på Facebook**
- 2. Anerkendende sprog i den offentlige debat på Facebook**
- 3. Casestudie: Coronapandemien på Facebook**
- 4. Danmarks Digitale Medborgerhuse. En kortlægning og analyse af 5.000 offentlige danske Facebookgrupper**



# Indhold

		SIDE
INTRODUKTION	Den offentlige debat på Facebook	5-6
	#. Rapportens vigtigste tal	7-10
METODE	0. Baggrund og datasæt	11-14
	1: Hvad detekterer algoritmerne?	15-20
	2: Sådan virker en algoritme, og sådan har vi bygget vores	21-29
RESULTATER	3: Angreb i den offentlige debat på Facebook	30-35
	4: Angreb på politikernes Facebooksider	36-43
	5 Angreb på mediernes Facebooksider	44-50
	6 Sprængfarlige emner i den offentlige debat	51-57
	7: Hadet retter sig mod bestemte grupper	58-59
	8: Afsluttende bemærkninger og tak	60-63
BILAG	Referencer og bilag	64-67





# INTRODUKTION

# Introduktion: Hvorfor interessere sig for den offentlige debat på Facebook?

## Facebook er dén digitale arena for offentlig debat

Facebook udgør Danmarks største arena for digital debat. Mellem 3,5 og 4 mio. danskere har en Facebookprofil.<sup>1</sup> Facebook er dermed dét digitale forum, hvorigennem flest danskere konsumerer og udveksler nyheder, information og holdninger. Den gennemsnitlige aktivitet for danske Facebookprofiler over 30 dage er 7 kommentarer, 12 interaktioner og 1 deling af indhold.<sup>2</sup> 81% af de danskere, der har en Facebookprofil svarer, at de bruger den til at tilgå nyheder<sup>3</sup>. Facebook er altså et helt centralt sted for den offentlige, digitale debat i Danmark.

I denne analyse zoomer vi specifikt ind på den offentlige debat, der foregår på mediernes og politikernes offentlige Facebooksider. Danske medier lægger dagligt nyheder ud på deres Facebookside, og her står det frit for danske Facebookbrugere at kommentere på historien og deltage i debatten. På politikernes Facebooksider deler politikerne typisk holdninger og standpunkter, hvilket Facebookbrugere ligeledes kan kommentere på.

Det er disse opslag og kommentarer på politikeres og mediers offentlige Facebooksider, vi undersøger i rapporten.

## Den hårde tone afholder mange fra at deltage i debatten

Det er ikke alle Facebookbrugere, der griber til tasterne og kommenterer på mediers og politikeres opslag. En megafonmåling for TV 2 viste i år, at 68 procent af danskerne afholder sig fra at deltage i debatter og samtaler på sociale medier, fordi de synes, at tonen er for hård.<sup>4</sup> Tilbage i 2016 var det ifølge en undersøgelse fra Institut for Menneskerettigheder 50 procent af danskerne, der afholdte sig fra at udtrykke deres mening og deltage i debatten på sociale medier pga. tonen i debatten.<sup>5</sup> Dét kunne tyde på, at flere afholder sig fra at deltage i den offentlige debat gennem sociale medier i dag end tidligere.

*Men hvor hård er debatten egentlig? Er der nogen medier eller politikere, der har en særligt hård debat i kommentarsporene? Hvem retter hadet sig imod? Og hvilke emner antænder især den hadefulde tone?*

Disse spørgsmål vil vi kaste lys over i rapporten.



# Rapportens vigtigste tal



# Hvor hård er debatten?



## 5,2% Sproglige angreb\*

Lidt mere end hvert **20.** indlæg, der står tilbage på mediers og politikeres Facebooksider er et sprogligt angreb, dvs. en stigmatiserende, nedsættende, krænkende, stereotypiserende, ekskluderende, chikanerende eller truende ytring. Det svarer til **3,3 mio.** ud af 63 mio. kommentarer siden 2019.



## 3,8% anstødelige 1,4 % hadefulde

Ud af de 3,3 mio. sproglige angreb er **902.029** hadefulde, svarende til 1,4 pct. af den samlede debat. De resterende **2.386.535** angrebskommentarer klassificerer algoritmen som anstødelige. Det svarer til 3,8 pct. af den samlede debat.



## 7,8% angreb hos politikere

Tonen i debatten er hårdere på politikersiderne end på mediesiderne. Topscoreren hos politikere har **17%** angreb i kommentarsporet på sin side. Der er forskel på, om politikere fortrinsvist er værter for angreb mod andre eller selv bliver udsat for angreb på deres side.

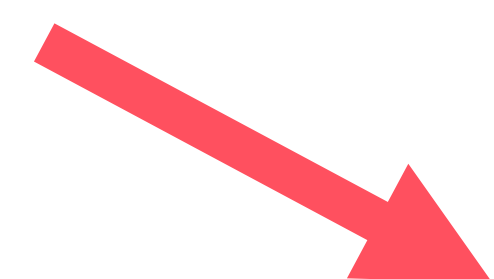


## 4,8% angreb hos medierne

Tonen i debatten på mediesiderne er mindre hård end på politikersiderne, men enkelte medier har en særligt høj andel angreb i deres kommentarspor. Topscoreren hos medierne har således **24%** angreb i kommentarsporet på siden.



# Hvordan er udviklingen over tid?



## Fra 2015 til 2021

Fra 2015 til 2021 ser vi et lille fald i andelen af angreb, herunder had og anstødelighed, der står tilbage i debatten på mediernes sider. Udviklingen skyldes formentlig et øget fokus på moderation både fra Facebook og medierne selv.



## Mellem 2019 og 2021

Siden 2019 har andelen af angreb, der står tilbage i debatten på mediernes og politikernes sider, været relativt stabilt.



## Større begivenheder kan ses online

Antallet og andelen af angreb påvirkes af fysiske begivenheder såsom folketingsvalg, terrorangreb, coronanedlukning og valg i USA.



## Agurketid ligeså

Når Folketinget er på sommerferie, falder andelen af angreb i politikernes kommentarspor. Politikernes følgere lader sig i høj grad inspirere af aktiviteten på siderne.

# Får nogen en særligt hård medfart?

Baseret på **2.600** emneord og **250** ord, der repræsenterer betegnelser for grupper og individer med beskyttede karakteristika, har vi kortlagt, hvilke emner og grupper der oftest omtales i de hadefulde og anstødelige kommentarer.



## Etniske minoriteter, især muslimer, er omdrejningspunkt for angreb

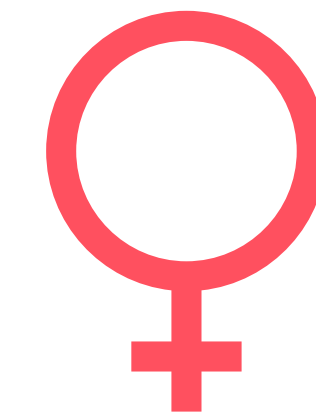
Debatten om ferier foregår ret civilt, men så snart snakken falder på islam, muslimer, indvandrere og integration ligger andelen af angreb i kommentarsporene på op mod **13%**.

Knap **50%** af de angreb, hvori vi kan identificere betegnelser for beskyttede karakteristika, er rettet mod nationalitet, etnicitet og religion.



## Kriminalitet, retspolitik og terror er svært at debattere med hovedet koldt

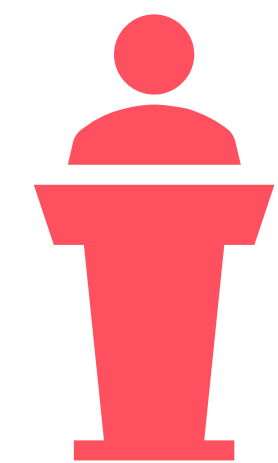
Der er ingen nåde i omtalen af kriminelle i den offentlige debat på Facebook. Her udgør angreb omkring **11%** af debatten. Vi ser et tydeligt overlap mellem disse angreb og angreb rettet mod etniske minoriteter, når politikere og medier deler nyheder og opslag om kriminelle med anden etnisk baggrund.



## Kvinder, køn og ligestilling er også effektivt benzin på bålet

Køn, feminisme, ligestilling og #metoo er også blandt de emner, der kan antænde debatten. **9%** af kommentarerne om emnet er et angreb.

Knap **14%** af de angreb, hvori vi kan identificere betegnelser for beskyttede karakteristika, er rettet mod kvinder.



## Danske og internationale politikere og politiske "yderfløje"

Kommentarer, der nævner danske og internationale politikere eller har karakter af ideologisk fløjkrig, indeholder relativt ofte angreb. Politisk orientering er målet for **14%** af de angreb, hvori vi kan identificere betegnelser relateret til beskyttede karakteristika. Angreb rettet mod politisk orientering er f.eks., når venstrefløjen omtales som "væmmeligboere" eller højrefløjen som "racister".



**METODE**



# Nye indsigter og nye metoder: Med algoritmer og big data kan vi undersøge *hele* debatten

## Meningsmåling med hundredetusindvis af respondenter

De seneste syv år har vi set en bølge af forskning inden for samfundsanalyse, demokrati og offentligheder, der har fået øje på, at vi gennem big data analyse af danskernes millioner af opslag, kommentarer, reaktioner og delinger kan få et unikt indtryk af, hvad vi danskere er optagede af, hvordan den offentlige mening udvikler sig, og hvordan vi taler til hinanden på sociale medier.

“Public opinion polling” baseret på big data giver os et helt andet indblik i den offentlige debat, end et survey eller en interviewrække kan. Ved at undersøge Facebook-debatten kan vi se på langt større datamængder, end man nogensinde vil kunne samle ind med f.eks. spørgeskemaer. Til forskel fra et spørgeskema har vi også mulighed for at undersøge data, som er genereret over tid. Frem for en stikprøve kan vi nu undersøge *hele* debatten. Endelig kigger vi ikke på folks opfattelse af deres egen adfærd, som man gør med spørgeskema. Vi kigger på folks faktiske digitale adfærd, selvfølgelig i anonymiseret form, jf. beskrivelsen af datagrundlaget.

## Algoritmer tager dansk social medieanalyse et skridt videre:

Med dette projekt kan vi for første gang anvende skræddersyet kunstig intelligens, der bedre end nogensinde før tillader os at undersøge ikke bare temaerne – men *tonen* i debatten på Facebook i stor målestok. Gennem brug af machine learning har vi udviklet en algoritme, der kan detektere hhv. om en kommentar er et angreb eller ej, og yderligere om angrebskommentarerne er hhv. hadefulde eller anstødelige.





# Baggrund: Det ved vi om hadet i debatten fra andre undersøgelser

I 2017 lavede Institut for Menneskerettigheder en kvalitativ stikprøvebaseret undersøgelse af debatten på TV 2's og DR's Facebooksider, som viste, at hver syvende kommentar (ca. 15 pct.) var hadefuld.<sup>6</sup>

Samme år fulgte vi op med en undersøgelse af hadet på politikeres og mediers Facebooksider ved hjælp af en simplere sprogalgoritme, som blandt andet viste, at tonen på politikernes sider generelt er hårdere end på mediernes. Analysen viste desuden, at emnerne flygtninge/integration, religion, ligestilling/køn og justits-/sikkerhedspolitik var særligt præget af en hadefuld tone.

I 2020 forsøgte vi at sætte tal på antallet og typen af slettede kommentarer sammen Common Consultancy og Tænk tanken Justitia.

Facebook offentliggør desværre ikke, hvor mange kommentarer og opslag der vurderes til at være hadefulde. Og hverken Facebook eller lokale sideadministratorer offentliggør mængden af modereret (dvs. slettet) indhold. Det er derfor svært at vide, præcis hvor mange hårde og hadefulde opslag og kommentarer, der bliver *afsendt* på Facebook.

Vores analyse viste, at der på fem nyhedsmediers sider slettes i omegnen af en halv million kommentarer om året, svarende til ca. seks procent af alle kommentarer på de fem mediesider.<sup>7</sup> Omkring en tredjedel af de slettede kommentarer blev vurderet til at være nedladende, grove, racistiske, truende og enkelte af dem strafbare.

Facebook har de senere år skruet op for moderationen af hadefuld tale. I 2020 fjernede Facebook 12 millioner opslag, der brød retningslinjerne for hadsk tale på verdensplan.<sup>8</sup> Men selvom Facebook i dag fjerner langt flere kommentarer end tidligere, er Facebookdebatten så enorm, at selv sletning af 12 millioner kommentarer er meget få i det store billede. Dertil er der et indbygget dilemma i, at det indhold, som grænser til forbudt (fx fordi det er hadefuld tale), skaber særligt meget interaktion og aktivitet på platformen. Som CEO Mark Zuckerberg har formuleret det: *"Our research suggests that no matter where we draw the lines for what is allowed, as a piece of content gets close to that line, people will engage with it more on average – even when they tell us afterwards they don't like the content."*<sup>9</sup>



# Datagrundlaget for rapporten

Vi har ved hjælp af algoritmen analyseret alle opslag og kommentarer fra mediernes og politikernes Facebooksider i perioden 1. januar 2019 til og med den 28. februar 2021. Til højre ses antallet af kommentarer, hvor mange opslag de er afgivet under, samt antallet af Facebooksider vi har indsamlet data fra. Der er dog kun tale om opslag og kommentarer tilgængelige via Facebooks API pr. april 2021. Data indsamlet gennem API'et indeholder ikke navne eller ID'er på brugerne. Vi har i analysen yderligere anonymiseret kommentarerne. I de citater, vi fremhæver i rapporten, har vi sikret, at der ikke nævnes navne på andre end offentlige personer.

Det er vigtigt at holde sig for øje, at modereret (dvs. slettet) indhold ikke indgår som grundlag for undersøgelsen. Vores opgørelser af angreb, anstødelighed og had i kommentarsporene på forskellige Facebooksider afspejler de kommentarer, der får lov at stå tilbage efter moderation fra såvel Facebook som lokale sideadministratorer.

## Mediesider

**54.492.711**

Kommentarer under

**994.466**

opslag fra

**477**

Pressenævnstilmeldte mediers  
Facebooksider\*

## Politikersider

**8.583.465**

kommentarer under

**102.646**

Opslag fra

**199**

Folkevalgte politikere og partiers  
Facebooksider



# Hvad detekterer algoritmerne?

Angreb, anstødeligt  
sprog og hadtale



# Definitionerne:

## Hvad prøver vi at finde?

Hvor går grænsen mellem ophedet debat, kontroversielle men legitime synspunkter, angreb og decideret hadtale? Vi har i første omgang skulle beslutte, hvad vi gerne vil lære algoritmen at finde. Sammen med medlemmerne af vores advisory board og TrygFonden nåede vi frem til, at målet var at udvikle en algoritme, der kan detektere **sproglige angreb** i den offentlige debat på Facebook.

At fokusere på sproglige angreb tillader os at kvantificere hårdheden i den digitale offentlige debat på et mere overordnet niveau, der også dækker angreb mod ikke-beskyttede individer og grupper som f.eks. offentlige og ikke offentlige personer, debattører, medborgere, politikere, journalister og andre samfunds- og faggrupper.

Fordi vi er en flok ambitiøse fagfolk, besluttede vi at træne endnu en algoritme, der yderligere inddeler **angrebene** i hhv. **anstødeligt sprog** og **hadtale**. Forskellen beskrives til højre, hvor rapportens fire centrale begreber står kort defineret.

De fire begreber til højre har været fundamentet for den kodemanual med regler, eksempler og undtagelser, der har været bibelen i den menneskeassisterede træning af algoritmen.

Sådan har vi defineret fænomenerne:

### Angreb

*Stigmatiserende, nedsættende, krænkende, stereotypiserende, ekskluderende, chikanerende eller truende ytringer.*

### Anstødeligt sprog – en underkategori af angreb

*Angreb mod en gruppe eller et individ, der ikke er baseret på personens eller gruppens beskyttede karakteristika.*

### Hadtale – en underkategori af angreb

*Angreb mod en gruppe eller individ baseret på dennes beskyttede karakteristika.*

### Beskyttede karakteristika

*Race/etnicitet, hudfarve, nationalitet og oprindelse, religion og tro, seksuel orientering, køn og kønsidentitet, socialklasse og social status, politisk orientering, alder eller handicap og seriøse sygdomme (både fysiske og psykiske).*

### Ikke beskyttede karakteristika

Dækker f.eks. over beskæftigelse, uddannelse, lokalt geografisk tilhørsforhold og lignende. Angreb baseret på disse karakteristika vil blive kategoriseret som anstødeligt sprog.



# Eksisterende definitioner på hadtale

*“Hate speech [...] entails [...] the advocacy, promotion or incitement of the denigration, hatred or vilification of a person or group of persons, as well any harassment, insult, negative stereotyping, stigmatization or threat of such person or persons and any justification of all these forms of expression – that is based on a non-exhaustive list of personal characteristics or status that includes “race”, colour, language, religion or belief, nationality or national or ethnic origin, as well as descent, age, disability, sex, gender, gender identity and sexual orientation.”*

– ECRI (European Commission against Racism and Intolerance)

*“We define hate speech as a direct attack on people based on what we call protected characteristics – race, ethnicity, national origin, religious affiliation, sexual orientation, caste, sex, gender, gender identity, and serious disease or disability. [...] We define attack as violent or dehumanizing speech, harmful stereotypes, statements of inferiority, or calls for exclusion or segregation.”*

– Facebook Community guidelines

Der findes ingen leksikalsk definition af hadtale i den danske ordbog. Det tætteste vi kommer på en definition af hadtale er formuleringen af straffelovens paragraf 266b (“racismeparagraffen”):

“Den, der offentligt eller med forsæt til udbredelse i en videre kreds fremsætter udtalelse eller anden meddelelse, ved hvilken en gruppe af personer trues, forhånes eller nedværdiges på grund af race, hudfarve, national eller etnisk oprindelse, tro eller seksuel orientering, straffes med bøde, hæfte eller fængsel indtil 2 år.”

Udgangspunktet for denne analyse er dog ikke blot at detektere ytringer, der i juridisk forstand er strafbare, men at pege på ytringer, der kan medvirke til at afskrække borgere fra at deltage i den demokratiske debat. Til dette formål fandt vi ECRIs (the European Commission against Racism and Intolerance) og Facebooks definitioner af hadtale brugbare. Vi har derfor valgt at læne os op ad disse to eksisterende definitioner på hadtale i udarbejdelsen af vores egen.

Begge definitioner bygger (ligesom vores) på formlen:

**Had = et sprogligt angreb baseret på beskyttede karakteristika**

ALGORITME 1:  
Angreb vs. ikke angreb



ANGREB



ANSTØDELIGHED



HADTALE

ALGORITME 2:  
Anstødeligt sprog vs. hadtale

# Algoritmerne: Angreb, anstødelighed og had

Egentlig er det en tilsnigelse at tale om "algoritmen" i ental. Teknologien, der ligger til grund for undersøgelsen, bygger på to algoritmer med hver deres funktion. Den første detekterer angreb, og den anden klassificerer disse som enten anstødelige eller hadefulde. De fungerer på følgende måde:

1. Den første algoritme detekterer om en kommentar er et angreb eller ikke-angreb, dvs. om der er tale om *stigmatiserende, nedsættende, krænkende, stereotypiserende, ekskluderende, chikanerende eller truende ytringer*.
2. Det giver os en pulje af angreb og ikke-angreb.
3. Den anden algoritme detekterer, om puljen af angreb er rettet mod beskyttede karakteristika eller ej. Hvis angrebet er rettet mod en gruppe eller et individs beskyttede karakteristika, defineres angrebet som hadtale. Hvis angrebet ikke er baseret på personens eller gruppens beskyttede karakteristika, defineres angrebet som anstødeligt sprog.

Begge algoritmer ligger frit tilgængelige på platformen Hugging Face.



# Eksempler: Angreb eller ikke angreb?

Hvor går grænsen for, hvornår et udsagn er nedsættende eller stigmatiserende? Som supplement til de leksikalske definitioner har vi udarbejdet en kodemanual, der operationaliserer definitionerne gennem regler, eksempler og undtagelser. Den første del af processen med at udvikle algoritmerne er normativ og kvalitativ. Herunder er vist et par eksempler på Facebookkommentarer, som vi har defineret som hhv. angreb og ikke-angreb med reference til udvalgte regler fra vores kodemanual.

## Angreb

Velbekomme dit klamme usle løgnersvin!

*Udtryk for afsky ved andre mennesker anses som et angreb ligesom brugen af ordet "svin" er dehumaniserende*

De burde hjælpe ham, ved slet ikke at give ham noget mad? Lede psykopat. Lad idioten dø altså... Ingen vil savne ham alligevel....

*Trusler om, forherligelse af, eller udtryk for ønske om skade på mennesker betragtes som et angreb.*

*At bruge kliniske diagnoser (psykopat) som fornærmelse betragtes også som et angreb*

De stjæler alt, der ikke er boltet fast

*Stigmatisering, generalisering, fordomme og anklager betragtes som angreb*

sådan en møg ko ! Og så bare krybe udenom, når hendes sande jeg kommer frem... Føj for en kost...

*Der anvendes nedsættende ord om beskyttede karakteristika (køn), og udtrykkes afsky ved et andet menneske*

Nogle bør ikke få lov til at yngle – de to f.eks.

*Udtryk for, at nogle bør ekskluderes eller fratages grundlæggende rettigheder betragtes som et angreb*

Dumme perker svin smid din uniform og få de tæsk der tilfalder der dit sorte møg svin

*Der anvendes nedsættende ord om beskyttede karakteristika (etnicitet), udtrykkes ønske om skade og anvendes dehumaniserende sprog (møgsvin)*

## Ikke-angreb

Uden KORRUPTTE who & Sundhedst. som er finansieret ca 80% af Big Pharma !

*Anklager om kriminelle forhold regnes for angreb, men angreb skal være rettet mod mennesker eller grupper af mennesker, ikke organisationer*

Det er det største lorteforslag. Farligt for Danmark!

*Det er forslaget men ikke mennesker, der bliver omtalt i latrinære vendinger.*

Hun lyder som en robot, der er gået i baglås. Det er den samme febrilske udenomssnak

*Kritik er ikke et angreb*

"Kunne vi ikke snart blive fri for at høre på ham!"

*Kommentaren udtrykker et personligt håb/ønske men ikke et generelt ønske om at modtageren får frataget sin ret til at ytre sig*

Men hendes sprogbrug er utvetydigt røvhulsagtig og et forsøg på at tækkes et segment af landets mest forhærdede racister...

*Kommentaren er ikke defineret som et angreb, fordi det er sprogbruget og ikke personen der kaldes "røvhulsagtig". Vi har vurderet, at det er kritisk, men ikke et angreb at kalde nogen for hykler.*

Danmark burde ikke modtage en eneste flygtning

*Udtalelsen er kontroversiel, men repræsenterer en legitim holdningsytring*

"Han er da en kæmpe klaphat... Og du Bernhard lukker det samme sludder ud gang på gang."

*Klaphat ligger ifølge kodemanualen under "bagatelgrænsen" for nedsættende sprog sammen med bl.a. klovn og dumrian*



# Eksempler: Anstødeligt eller hadefuldt sprog?

Forskellen på om et angreb er anstødeligt eller hadefuldt afgøres af, om angrebet er rettet mod beskyttede karakteristika; altså race/ethnicitet, hudfarve, nationalitet og oprindelse, religion og tro, seksuel orientering, køn og kønsidentitet, socialklasse og social status, politisk orientering, alder eller handicap og seriøse sygdomme (fysiske og psykiske). Forskellen mellem det hadefulde og anstødelige er ikke "hårdhedsgrad". Der kan være meget voldsomme anstødelige kommentarer, der f.eks. er truende, ligesom nogle kommentarer er hadefulde "bare" fordi de indeholder nedsættende eller stigmatiserende ord, der dog er nedsættende på baggrund af et beskyttet karakteristika (møgtøs, retarderet, svans).

## Anstødeligt sprog

Naser, du er uværdig og skadelig med din oeffen og leflen. Du er en god foelger af Trump. Og dermed er du skadelig.

*Nedsættende (uværdig, skadelig) og dehumaniserende (oeffen), men Khader bliver ikke angrebet på et beskyttet karakteristika.*

Knep din mor!

*Upassende opfordringer er et angreb, men angrebet er ikke baseret på beskyttede karakteristika*

Hun burde få samme behandling selv!

*Angrebet er ikke baseret på beskyttede karakteristika*

Sådan en flok bonderøve

*Beskæftigelse og lokalt geografisk tilhørsforhold er ikke beskyttede karakteristika*

du ligner ikke en der mangler mad .....smut hjem til dig selv du er utåelig t læse

*Kropstype er ikke et beskyttet karakteristika, men kommentaren er nedsættende*

Danske journalister er lort de har mistet al deres integritet. Boycott drs røde lejesvende!

*Journalist er ikke et beskyttet karakteristika*

## Hadtale

"Hva' fanden ka' der gøres v. de åndsformørkede overtroiske fjolser, når de er født her? ? Det' fucking uhyggeligt, at væmligboere m. iman Østergård i spidsen ikke ka' se skriften på væggen,"  
*Angreb baseret på politisk orientering og religion + brug af ord med religiøs tilknytning som nedsættende*

Hun laver revolution det er det er hun gør hun er psykisk syg kælling  
*Angreb baseret på køn + angreb mod psykisk syge ved at bruge betegnelser for begge som noget nedsættende*

"Den olding kan vidst ikke længere tænke klart"

*Angreb på baggrund af alder*

"Han sku ha fået et år... Sådan en svans sammen med sin bøsse søn! Svans!!"

*Angreb pba. seksuel orientering; anvendelse af ordet "bøsse" som nedsættende.*

"Det giver fuldstændig mening, for en kommunist. For normalt begavede mennesker, ikke så meget."

*En politisk orientering bliver sat i modsætning til at være normalt begavet*

Udlever ALDRIG et våben til en troende muslim. Det ville være selvmord...?

*Angreb baseret på religion*



2

2

2

2

**Sådan virker en  
algoritme,  
og sådan har vi  
bygget vores**

# Kort om algoritmer: Hvad kan de? Og hvordan virker de?

## Opskrift, funktion eller brugsanvisning: Hypet barn, mange navne

En algoritme kan defineres som *en opskrift på at løse et problem gennem en regelbaseret proces*. Algoritmer gives et input og returnerer et output afhængigt af "reglerne" i processen. I sin simpleste forståelse adskiller algoritmer sig ikke meget fra en opskrift eller en brugsanvisning, hvor man ved at følge en trinvis proces får returneret et ønsket output. De simpleste computeralgoritmer sorterer f.eks. tal på en liste efter deres størrelse, hvor de mest komplekse, som den nyligt lancerede sprogalgoritme GPT-3, har lært at producere tekst i alverdens genrer om alverdens emner (næsten) som et menneske, fordi den er trænet på alle tilgængelige tekster fra internettet.

## Input og output

Inputet til vores algoritme er en kommentar fra et Facebook-kommentarspor. Outputtet er en vurdering af, om kommentaren er et angreb, og dernæst om angrebet er hadtale eller anstødeligt sprogbrug.

## Neurale netværksalgoritmer

Moderne avancerede algoritmer kaldes neurale netværksalgoritmer, fordi de efterligner den struktur, som vi kender fra opbygningen af den menneskelige hjerne. Hjernen indeholder lag af neuroner (små processorer), der er forbundet i et netværk. Når et input sendes igennem netværket, aktiveres de små processorer efter tur. Hver neuron afkoder en lillebitte del af inputtet og sender signalet videre.

## Deep learning: Mange lag gør algoritmen "intelligent"

De enkelte "neuroner"/processorer i algoritmen repræsenterer relativt simple regler. Så det er dybden i det samlede netværk, altså antallet af lag, der gør algoritmen "intelligent". Intelligent i den forstand, at den i første omgang kan komme med et kvalificeret bud på, om en kommentar er et sprogligt angreb, og dernæst om det sproglige angreb er hhv. anstødeligt og eller hadefuldt. I de enkelte lag i algoritmen vil der være opstillet **tusindvis** af regler, der f.eks. omfatter kommentarens længde, forekomsten af ord og emojis, kapitaler og tegnsætning, rækkefølge og ordstilling, intern sætningskontekst, der alle har en vægt og påvirker algoritmens samlede dom over kommentaren.



# Superviseret maskinlæring: Rugbrødsarbejdet i algoritmetræning

Vores algoritmer er først trænet ved hjælp af en sprogmodel og dernæst beriget med inputdata, som vi har "fodret" den med. Algoritmen bruger en præ-trænet dansk sprogmodel, som hedder Ælectra\* (en dansk version af den engelske Electra model). Ælectra giver algoritmen en god forståelse for det danske sprog. Algoritmen lærer gennem sprogmodellen at "læse" kommentarer og genkende mønstre på tværs af kommentarer.

Når algoritmen dernæst skal lære at detektere sproglige angreb, bliver den trænet på kommentarer, der er kategoriseret af mennesker (annotører). Vores træningsdatasæt består af ~70.000 kommentarer fra den offentlige debat på Facebook, som mennesker har kategoriseret (annoteret) ud fra en vurdering af, om de indeholder angreb, anstødeligt sprog og hadtale. Denne form for træning kaldes for *superviseret maskinlæring*. Algoritmen bliver sat til at genkende mønstre i de kommentarer, som mennesker har kategoriseret som sproglige angreb.

Vores menneskelige vurderinger af, hvorvidt en kommentar indeholder et angreb, og hvorvidt dette angreb skal klassificeres som hhv. anstødeligt sprog eller hadtale, er baseret på de tidligere nævnte definitioner og den omfattende kodemanual med regler og eksempler.

At algoritmerne er trænet af mennesker giver mange styrker, men også forskellige begrænsninger og bias. Dels er mennesker ikke éns, og de er heller ikke altid indbyrdes enige. Der vil forekomme fejlannoteringer, og ingen kodemanual kan sikre stringens, især når det gælder evalueringen af naturligt sprog. Dette er både godt og skidt for algoritmen, der på grund af menneskers forskellighed får mindre bias, men også bliver begrænset i sin præcision. Forskellen på menneskers og algoritmers forståelse af sprog og medfølgende fejl vender vi tilbage til. Men overordnet er vi begejstrede for, at mennesker og maskiner i samarbejde har bygget en teknologi, der med ret stor succes kan kortlægge angreb i mere end 63 mio. indlæg i den danske offentlige debat siden 2019.

# Algoritmetræning i 10 trin: den forsimplede guide

1. Mennesker udarbejder en definition af sproglige angreb, anstødeligt sprog og hadtale.
2. Mennesker udarbejder en kodemanual til annotører (menneskelige "trænere") med regler for og eksempler på, hvordan man genkender sproglige angreb.
3. Mennesker kategoriserer ~70.000 kommentarer, efter om de indeholder sproglige angreb, og hvilke angreb der er hhv. anstødeligt sprog og hadtale. 4,5% af kommentarerne annoteres af mere end én annotør. De kategoriserede kommentarer udgør algoritmens trænings- og testdatasæt.
4. Mennesker designer en simpel algoritme bestående af en præ-trænede dansk sprogmodel og et klassificeringsmodul. Algoritmen har en god forståelse for dansk, men kan endnu ikke identificere anstødeligt eller hadefuldt sprog.
5. Algoritmen bliver givet størstedelen af de annoterede kommentarer (nogle gemmes som testdata) og begynder at genkende regler og mønstre for, hvornår en kommentar er klassificeret som et angreb, og hvornår den ikke er.
6. Den foreløbige algoritme testes på et input af kommentarer. Algoritmen returnerer kommentarerne sin vurdering af, om kommentaren indeholder et angreb.
7. Annotørerne "efterannoterer" både de kommentarer, som algoritmen er sikker på og i tvivl om. De fortæller algoritmen, hvornår den har "ret" og tager fejl i sine foreløbige antagelser.
8. Algoritmen får kommentarerne retur med de menneskelige input og sættes igen til at genkende mønstre og regler for, hvorvidt en kommentar indeholder et angreb. Algoritmen korrigeres dermed i sin forståelse. Denne form for iterativ træning kaldes for "aktiv læring".
9. Vi eksperimenterer parallelt med det bedst mulige inputdata til algoritmen. Algoritmen performede f.eks. ikke bedre af at kende konteksten for kommentarerne i form af navnet på Facebooksiden og opslaget i relation til kommentaren. Til gengæld har algoritmen lært at genkende de mest almindelige emojis, og disse indgår i algoritmens vurdering af, om kommentaren er et hadefuldt eller anstødeligt.
10. Algoritmen trænes og justeres, indtil den opnår de bedst mulige resultater på et testdatasæt. Gennem træningen forsøger vi at forbedre to kvaliteter ved algoritmen, nemlig algoritmens præcision og genkendelse.



# Resultater: Hvor god er algoritmen?

**Præcision (precision):** Hvis algoritmen har en høj præcision, betyder det, at en stor del af det output, algoritmen returnerer, er identificeret korrekt. I vores tilfælde vil en algoritme med en høj præcision returnere en stak kommentarer, der er korrekt identificeret som hhv. hadefulde eller anstødelige. Men et ensidigt fokus på høj præcision er også ensbetydende med, at algoritmen kun fanger en mindre del af dét samlede fænomen, som vi leder efter; altså *alle* sproglige angreb på danske mediers og politikeres Facebooksider.

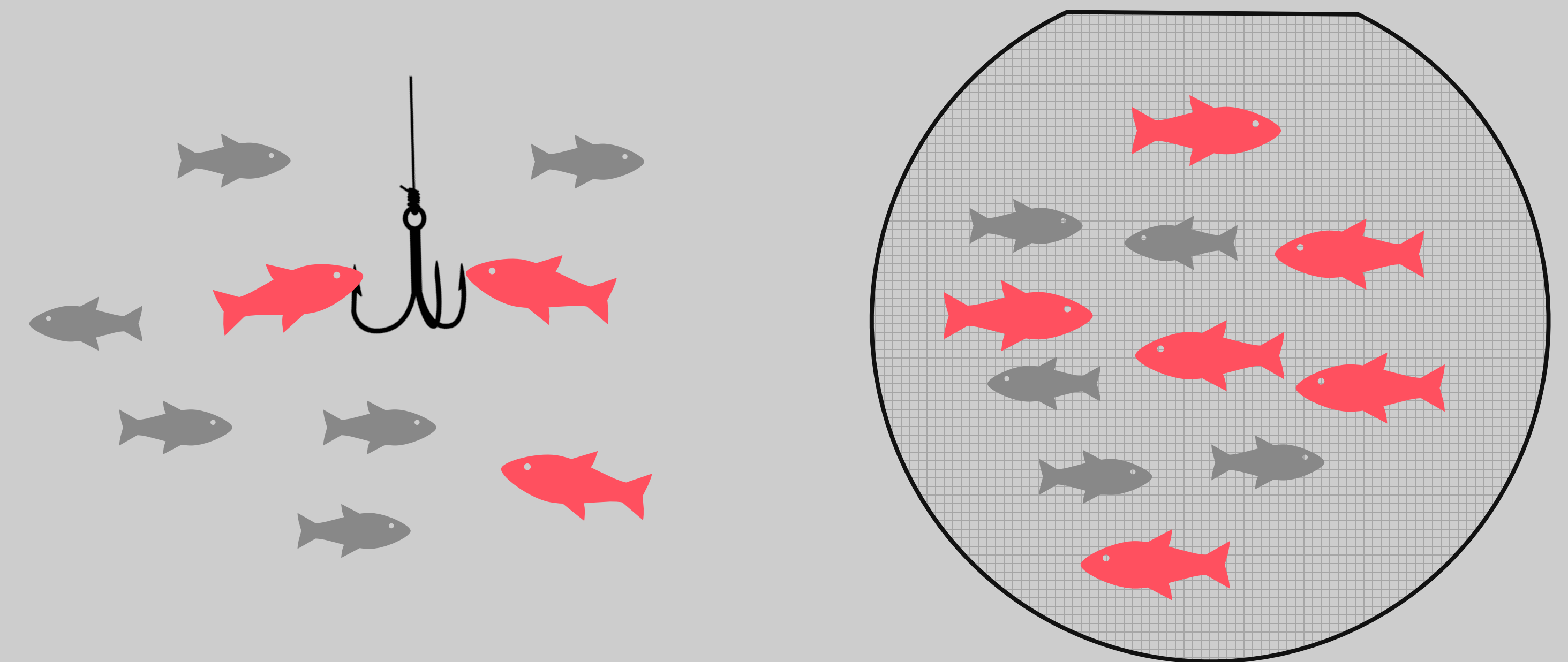
**Genkaldelse (recall):** Hvis vi skal gå efter at finde så mange sproglige angreb som muligt, skal vi forbedre algoritmens genkendelsesgrad. Med en høj *genkendelsesgrad* fanger algoritmen en større portion af dét, vi prøver at finde. Til gengæld vil outputtet også inkludere flere falske positive (støj).

**Tilsammen "macro averaged f1-score":** Den samlede vurdering af algoritmens præcision og genkaldelse opgøres i en "macro averaged f1-score". Scoren vægter algoritmens præcision og genkaldelse på alle dens parametre.

## På Fisketur

Man kan sammenligne balanceringen mellem algoritmens præcision og genkendelse med forskellige fiskestrategier.

Hvis vi fisker efter torsk, kan vi vælge at kaste en line ud, som især torsk bider på. På den måde undgår vi skidtfisk, men fanger knap så mange torsk. Hvis vi kaster et stort net ud, fanger vi rigtig mange torsk, men også en del skidtfisk. Kunsten er at justere vores net, så det fanger flest mulige torsk og færrest mulige skidtfisk.



### Algoritmens macro averaged F1-score:

Angreb/ikke angreb (At&tack): **0.8341**

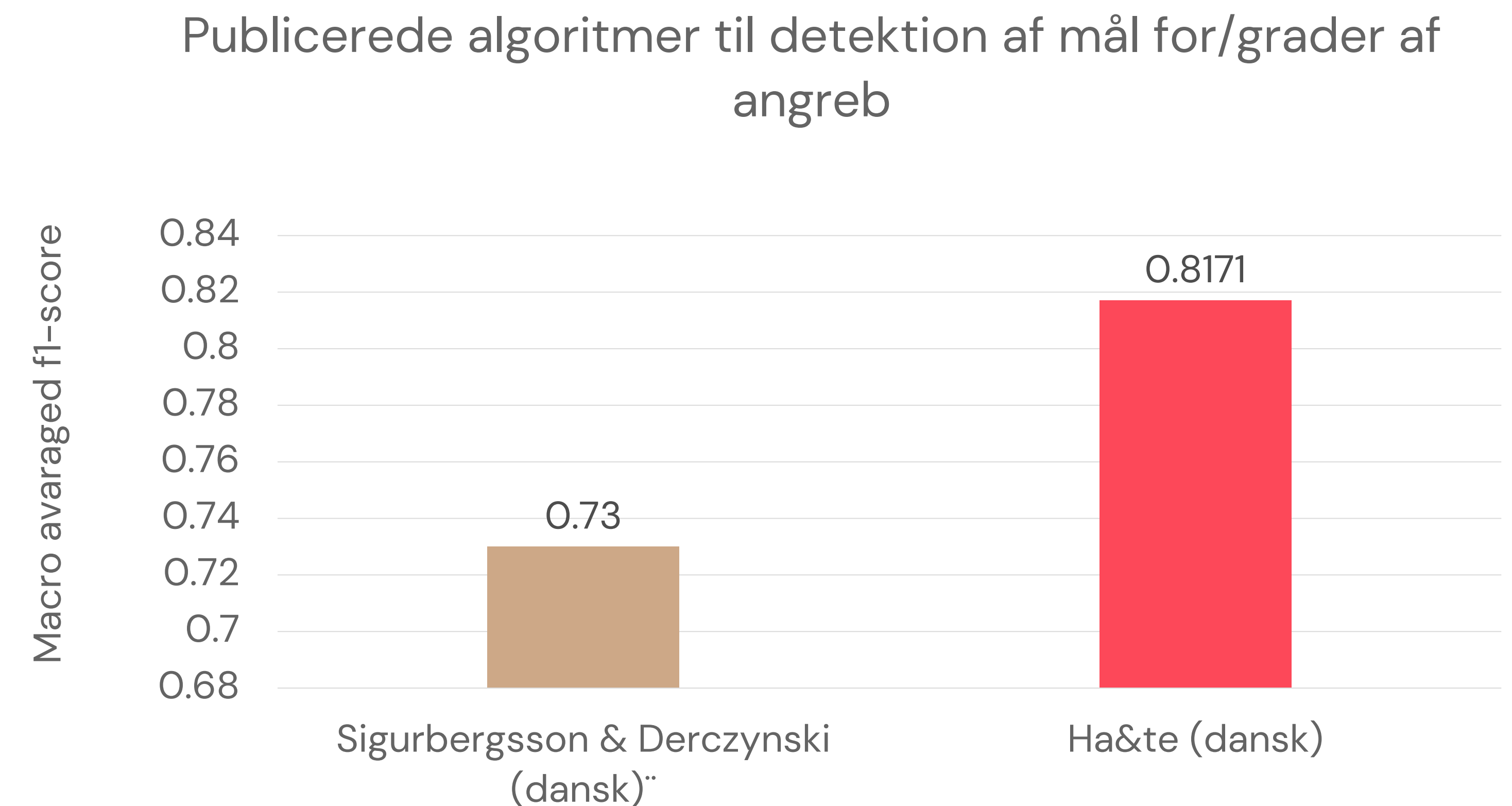
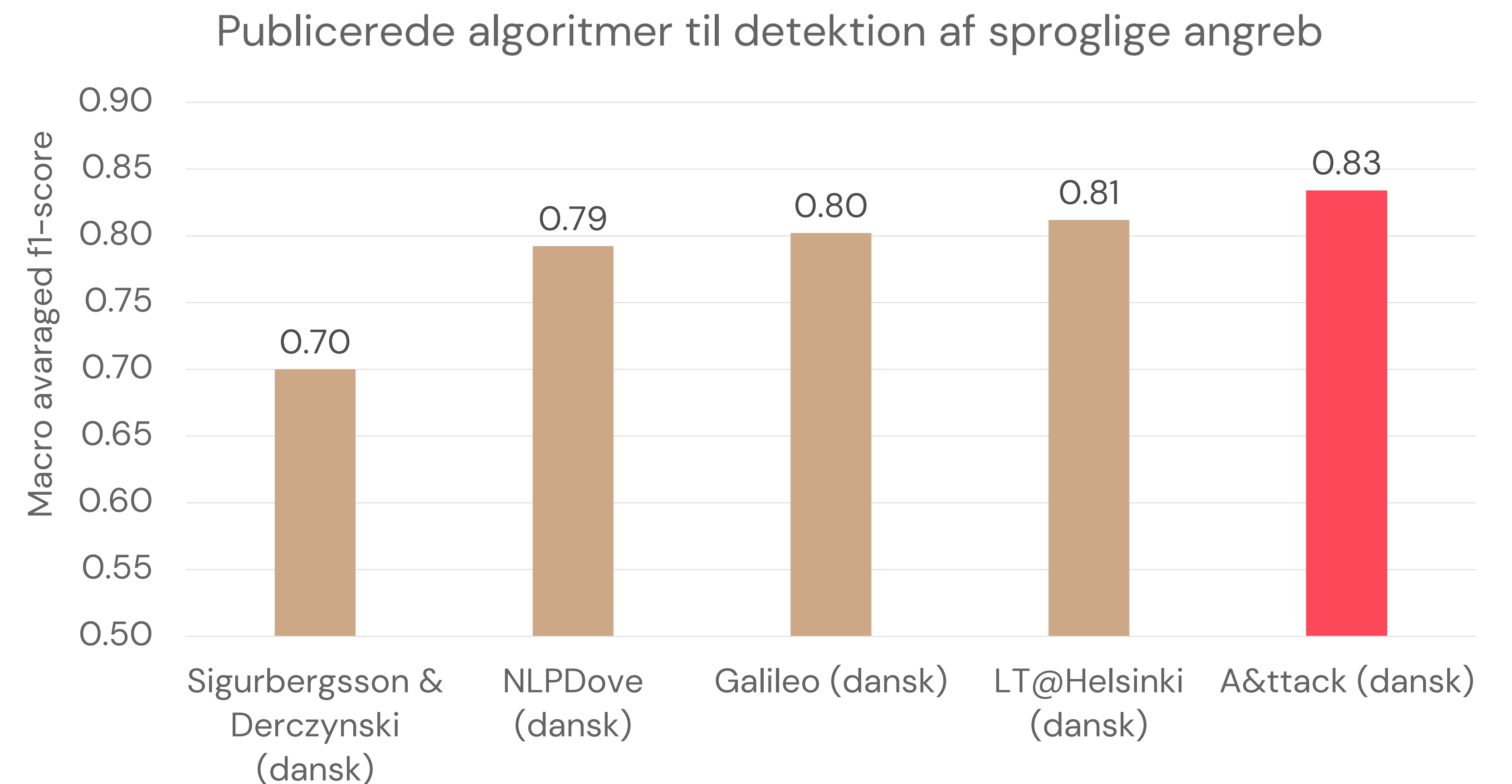
Had vs. anstødelighed (Ha&te): **0.8171**

Medmindre man forsker i NLP og kunstig intelligens betyder en sådan score ikke så meget. Så lad os kigge på resultaterne for andre algoritmer, der er udviklet til detektion af angreb og hadtale.

# Resultater: Vores algoritmer er klassens topscorerer

Teknisk set består vores setup af to algoritmer, som vi for nemheds skyld omtaler som én samlet algoritme. Den første (A&ttack) detekterer sproglige angreb vs. ikke-angreb. Nummer to (Ha&te) klassificerer angreb som hhv. anstødelige og hadefulde.

Sammenlignet med lignende algoritmer klarer vores sig godt. Vi har kigget i forskningslitteraturen efter andre dansksprogede algoritmer, som er udviklet med henblik på detektion og klassifikation af sproglige angreb.<sup>10,11,12,13</sup> I figurerne til højre har vi sammenlignet *macro averaged f1-score* for de danske algoritmer, vi fandt i vores gennemgang. Vi er også stødt på angrebsalgoritmer på andre sprog, og det er vores indtryk, at vores algoritmer også internationalt er med i førerfeltet, om end der helt sikkert findes bedre angrebsalgoritmer på engelsk og andre store sprog, hvor mulighederne for træningsdata er langt større. Algoritmers performance er især afhængig af mængden og kvaliteten af træningsdata. Jo mere og bedre træningsdata algoritmer trænes på, des bedre bliver de.





# Resultater: Rigtig god, men ikke perfekt

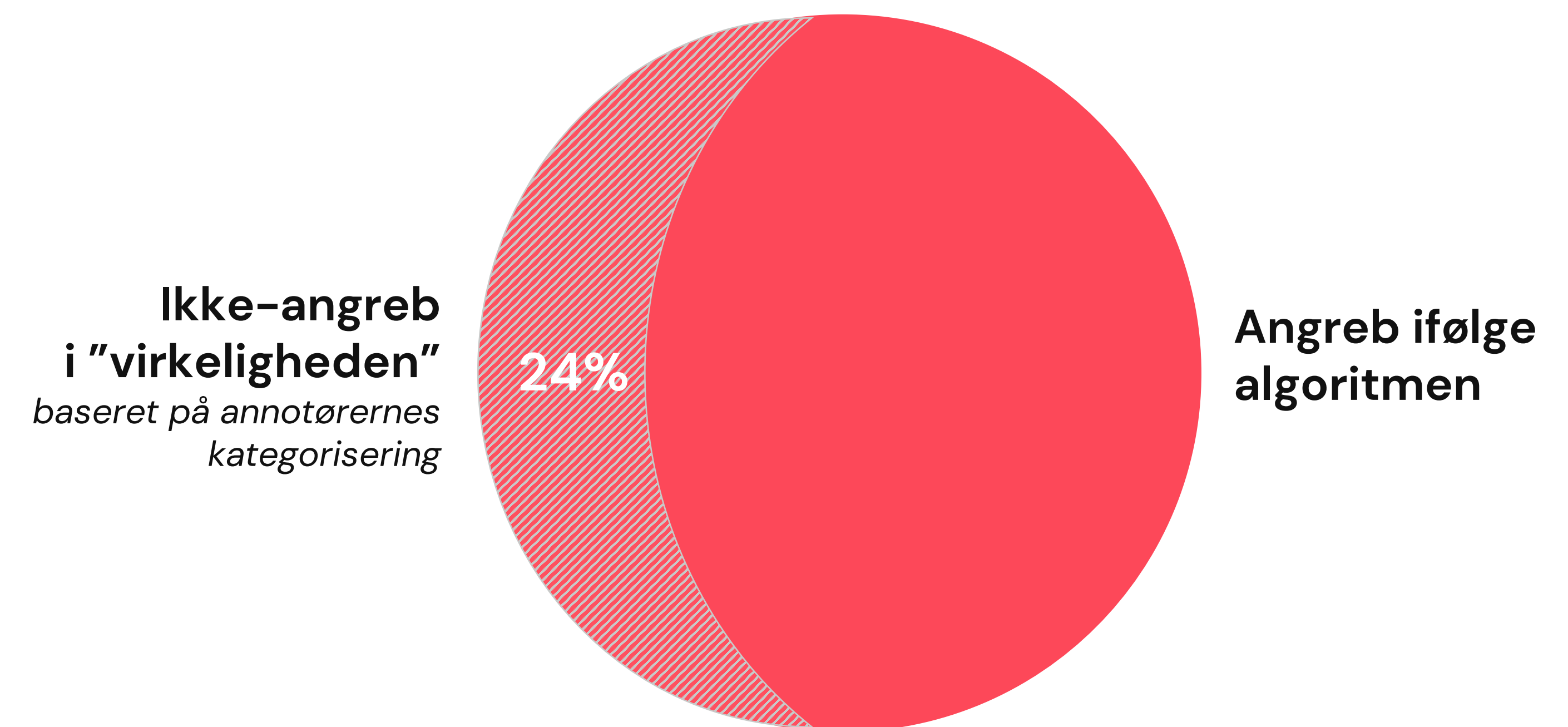
## Topscore betyder desværre ikke fejlfri

Selv med branchens flotteste macro avaraged F1-score vil der stadig være fejl i algoritmens output. Hvor stor denne fejlmargen er, kan vi kun forsøge at estimere ved at sammenligne algoritmens kategoriseringer med de menneskelige trænere (annotørernes) kategoriseringer på et evalueringssæt. Disse tests viser lidt forsimplet sagt, at 24 pct. af de kommentarer, som algoritmen estimerer er et angreb, i virkeligheden *ikke* er et angreb ifølge annotørerne. Derudover finder vores algoritme "kun" cirka 73 pct. af de angreb, som annotørerne identificerer. Den misser således i omegnen af 27 pct. af de angreb, der er derude.

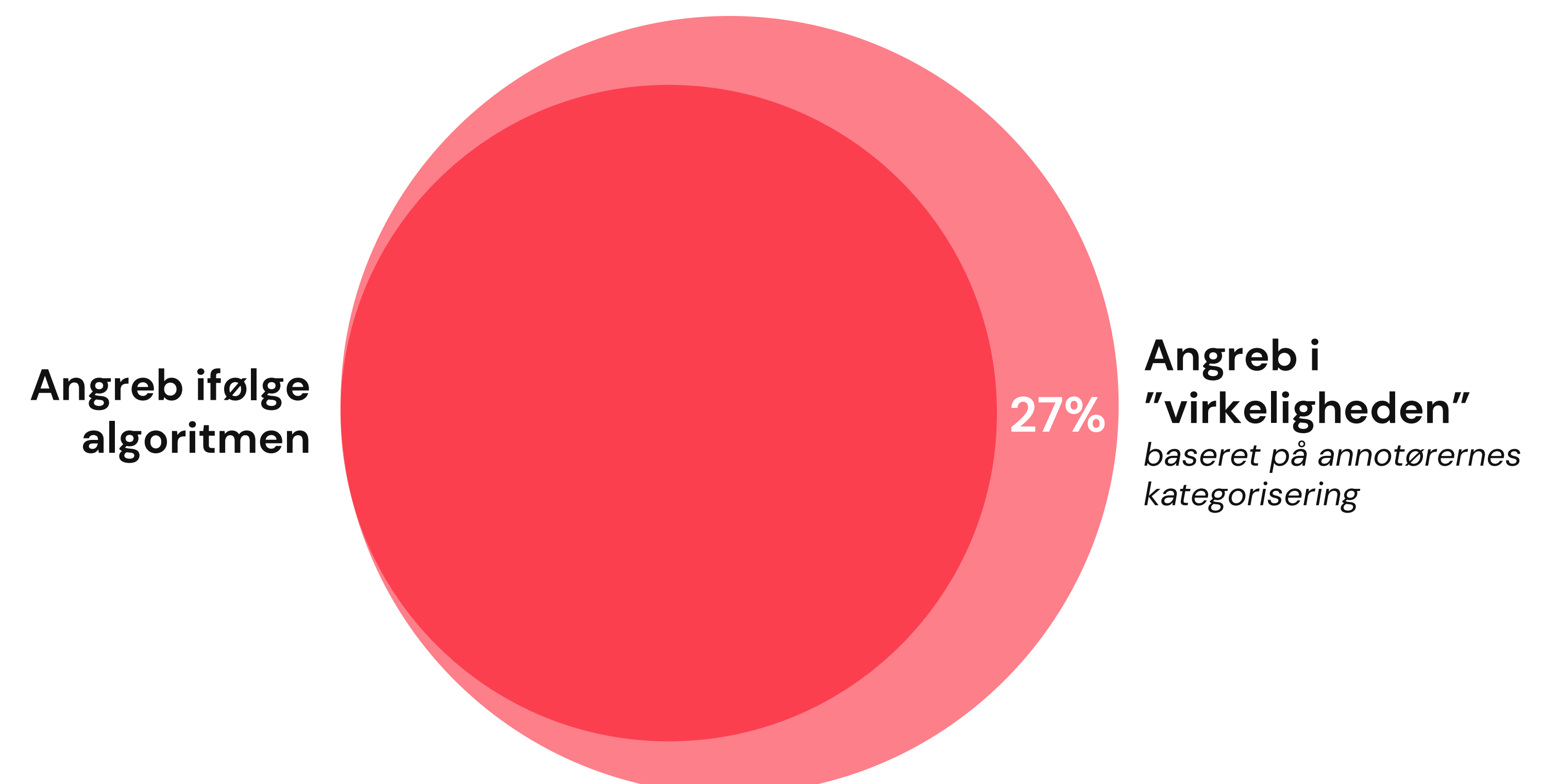
Det er svært at vide helt præcist, hvor store vores fejlmargener er. For der forekommer også tilfælde, hvor annotørerne tager fejl, og algoritmen har ret. Men hvis vi antager, at vores menneskelige trænere har anoteret korrekt hele vejen igennem, betyder vores fejlmargen, at vi ender med 24 pct. fejkategoriseringer i datasættet, og at algoritmen misser 27 pct. af angrebene i den offentlige debat.

## Algoritmens formodede usikkerhed

Estimat af andelen af ikke-angreb i algoritmens output af angreb



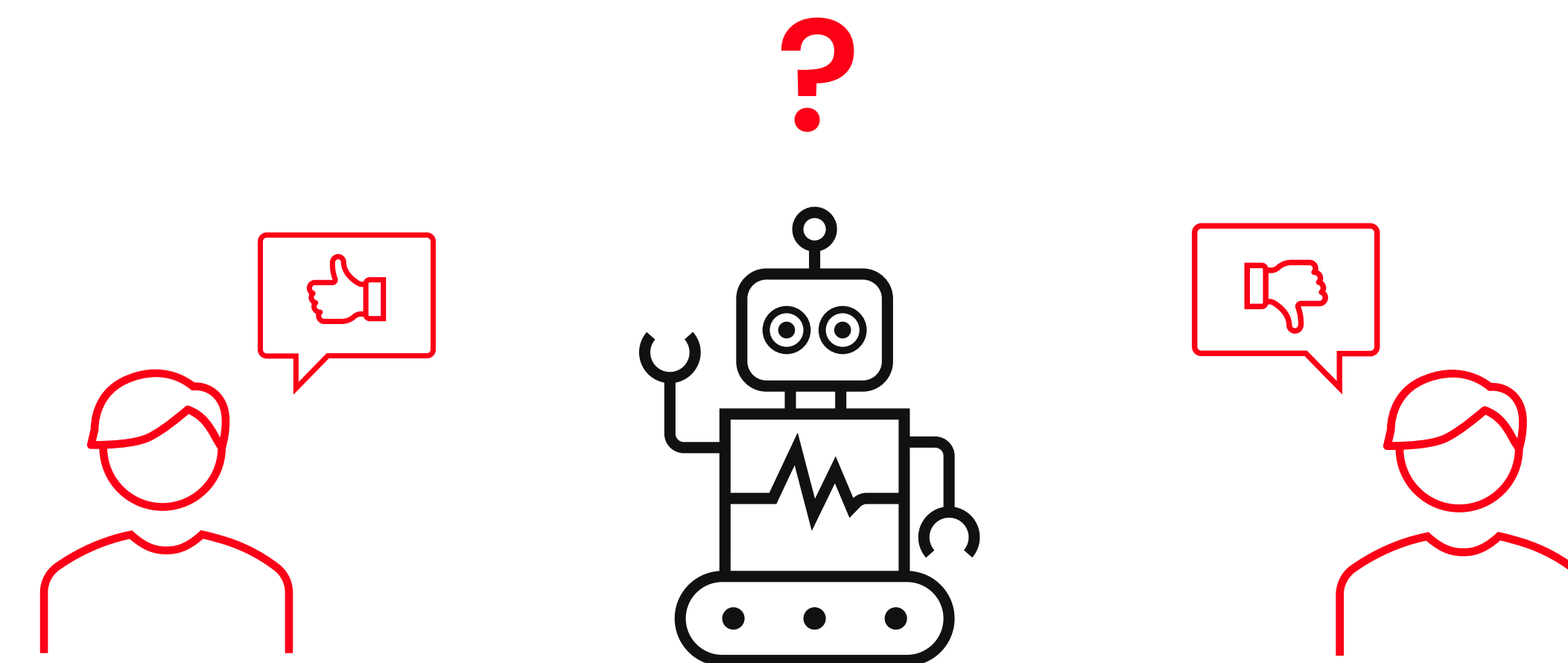
Estimat af manglende angreb i algoritmens output af angreb



# Menneskers nøjagtighed er sjældent bedre end algoritmens

Når vi peger fingre af teknologiens formåen, er det kun fair, hvis vi også tager et nærmere kig på de menneskelige "fejlrater", når det gælder kategorisering af indhold. Til projektet her har vi fået hjælp af fire annotører, der har kategoriseret i alt ~70.000 kommentarer, efter om de indeholder angreb, og om disse angreb er anstødelige eller hadefulde. De berigede kommentarer udgør algoritmens trænings- og testdata. 4,5 pct. af kommentarerne er blevet kategoriseret af mere end én annotør. Disse kommentarer bruges til at udregne en score for annotørernes indbyrdes enighed.

Opgørelsen af annotørernes enighed kaldes for vores inter-coder reliability. Jo mere kompleks kategoriseringsopgaven er, des mindre enighed vil der være mellem annotørerne. Vores annotører er enige i 77% af deres kategoriseringer. Og det på trods af at, de forsøger at være enige i afkodningen baseret på en kodemanual med eksempler og regler. Tallet lyder måske ikke overbevisende, men i forhold til opgavens kompleksitet og sensitivitet ift. subjektive bias er resultaterne overraskende gode sammenlignet med lignende undersøgelser<sup>14, 15</sup>.



## Algoritmen lærer af de mennesker, der træner den

Helt i starten af processen vil vi gerne teste, hvor "god" vores kodemanual er til at udstikke anvisninger for kategoriseringen af angreb, anstødelighed og had. Her hjælper opgørelsen af inter-coder reliability med at få udbedret de største huller i vores vejledninger til annotørerne. Men faktisk er både den indbyrdes enighed og uenighed også noget, vi udnytter til at bygge en bedre og "mere objektiv" algoritme. Hvor annotørerne er mest indbyrdes enige, vil algoritmen være mere sikker i sin vurdering.

Et tredje formål med at teste annotørernes indbyrdes "enighed" er at afstemme, hvor god vi med retfærdighed kan forvente, at algoritmen bliver. Manglende enighed betyder, at algoritmen i en del tilfælde har fået modstridende input fra annotørerne, og det sætter nogle naturlige begrænsninger for algoritmens formåen.



## Kommentarer, algoritmen fejlagtigt kategoriserer som angreb

”Det har altid været indianerne det går mest udover. Høvdingene skal nok klare frisag”

Her bruges betegnelser for en etnisk gruppe i en pointe om magthierarkier

”Du må ikke slå ihjel, du må ikke stjæle. Det er da et godt religiøst synspunkt.”

Religion, drab og tyveri er temaer,, der ofte er forbundet med angreb.

”Bubber er jo en sindssyg wingman?”

Engelske ord som algoritmen ikke forstår forvirrer den generelt. Og sindssyg er i mange andre tilfælde forbundet med et angreb.

”Luskebuks! Der tog du røven på mig”

Algoritmen gennemskuer at nogen bliver kaldt et kæle/øgenavn og at nogen ”gør noget” mod nogen andre.

”Der behøver kun at komme en eneste lille besmittet gris ind -- så efter min mening er hegnet blot en meget dyr absurditet :((”

Dehumanisering i form af sammenligninger med dyr vil ofte være angreb. Metaforer omkring sygdom og smitte optræder ofte i hadefulde kommentarer

”Gamle mennesker i dette land, bliver betragtet som affald.”

Alder er et beskyttet karakteristika og sammenligningen med affald kan være et angreb. Algoritmen forstår ikke, at der er tale om en metabetragtning.

”Lækkert svin. Det skal bare en tur i ilden i en times tid”

Algoritmen har svært ved at vide, at der er tale om mad og ikke et ønske om skade på mennesker, der omtales nedsættende

# Mennesker og algoritmer laver forskellige fejl

Når mennesker er uenige om kategoriseringer indenfor sprog, er det på en anden måde end algoritmer. Mennesker diskuterer nuancer: Hvilke fornærmelser er grove nok til at tælle som sproglige angreb; klaphat? Tosse? Luskebuks? Og hvad med budskaber der er underforståede eller sagt i spøg? Vores sprogforståelse bygger på livslang erfaring, sproghandlinger, kontekst og kultur.

Algoritmen derimod har lært sprog af en sprogmodel og derefter lært sig at genkende mønstre på tværs af kommentarer, der er kategoriseret af mennesker. Algoritmen genkender helt andre mønstre end os, og det resulterer i en anden type fejl. Hvis man som menneske gennemlæser algoritmens output, finder man derfor jævnligt fejl, der for mennesker er helt åbenlyse. Til venstre kan du se eksempler på kommentarer, som algoritmen fejlagtigt kategoriserer som angreb. Vi kan godt gætte på, hvad der medvirker til fejlestimeringerne. Dog er der ikke sådan, at vi systematisk kan se, at algoritmen tager fejl i de pågældende situationer. De mest systematiske fejl har vi korrigeret for gennem træningen af algoritmen.

# RESULTATER



3

3

3

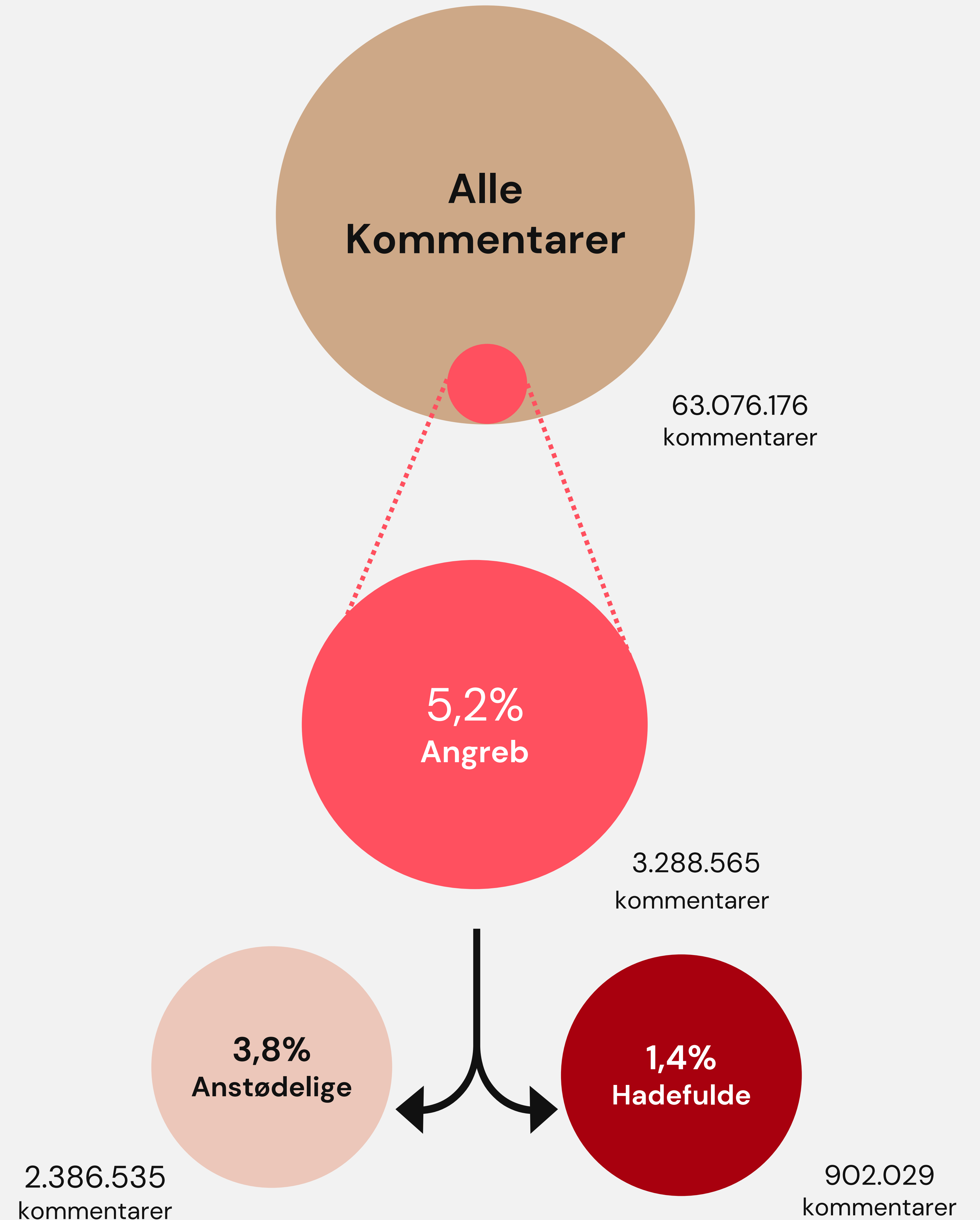
3

**Angreb i den  
offentlige debat  
på Facebook**

# Angreb i den samlede offentlige debat på Facebook

Vores kortlægning af den offentlige debat på mediernes og politikernes Facebooksider viser, at 5,2 pct. – dvs. lidt mere end hvert 20. kommentar – der får lov at stå tilbage i den offentlige debat, er et sprogligt angreb. Det indeholder altså en stereotyperende, stigmatiserende, nedsættende, krænkende, truende, ekskluderende eller chikanerende ytring.

Af 5,2 pct. sproglige angreb bliver en overvægt (3,8 pct. af totalen, 72,6 pct. af angrebene) klassificeret som anstødeligt sprog. Det efterlader en andel af hadefulde ytringer i den offentlige debat på 1,4 pct. (27,4 pct. af angrebene). Knap halvanden procent lyder måske ikke af meget, men de dækker over næsten en million kommentarer på danske mediers og politikeres Facebooksider siden 2019, der angriber grupper eller individer på baggrund af beskyttede karakteristika. Dertil kommer, at vi kun har at gøre med de kommentarer, som både Facebook, sideindehaverne og debattørerne selv har ladet blive stående.

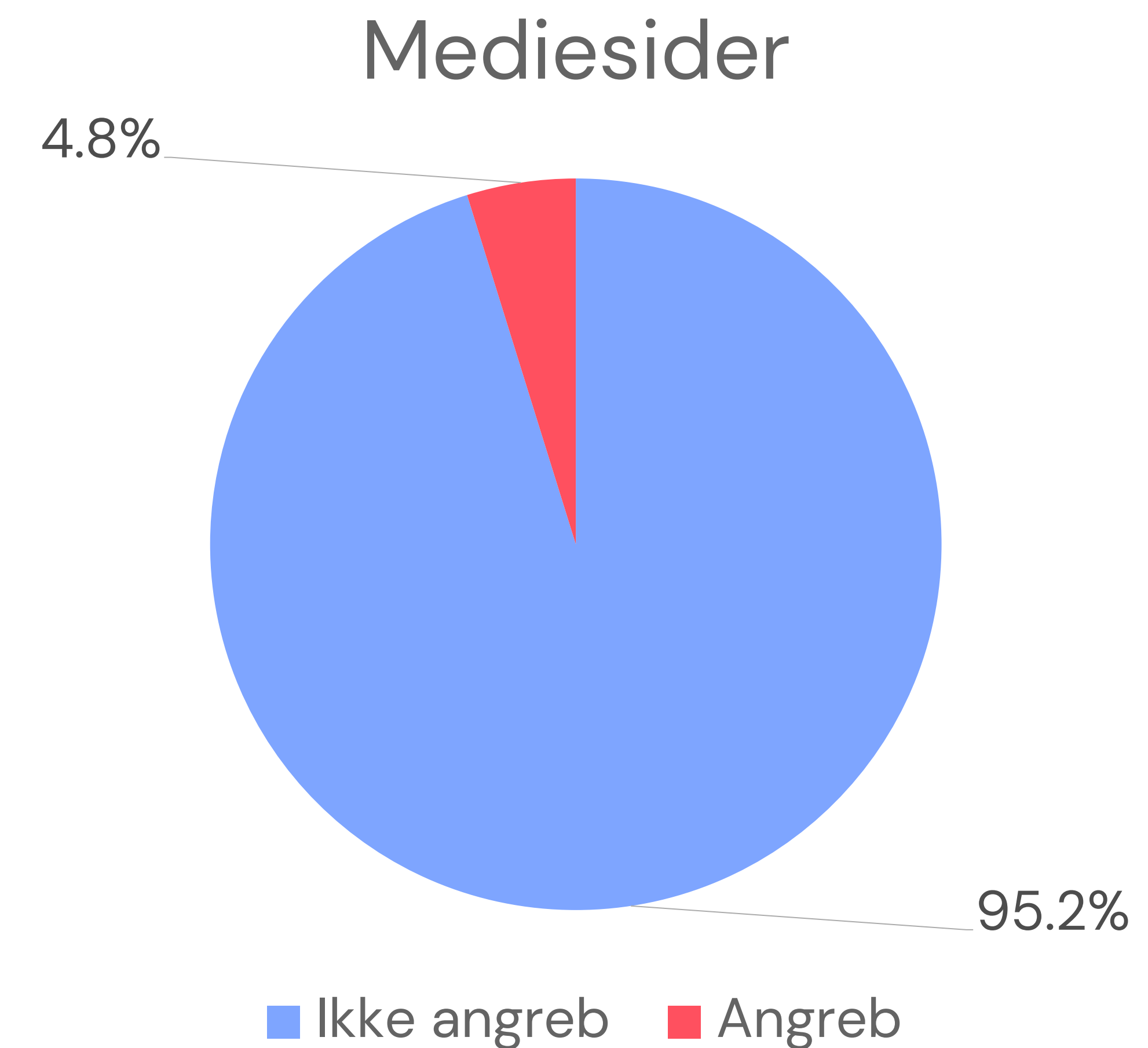




# Angreb på mediernes og politikernes sider

Sproglige angreb udgør knap 5 pct. af alle tilbageblivende kommentarer på mediernes Facebooksider siden 2019. Andelen er en del højere i kommentarsporene på politikernes Facebooksider, hvor angreb udgør 8 pct. af debatten. Angrebene på mediesiderne fordeler sig på ca. 75 pct. anstødelige og 25 pct. hadefulde. På politikersiderne fordeler angrebene sig på 67 pct. anstødelige og 33 pct. hadefulde. Angrebene er altså både hyppigere og hårdere på politikernes Facebooksider. Hvordan angrebene fordeler sig på enkelte sider og emner, ser vi på i de næste afsnit.

I vores undersøgelse fra 2020 så vi, at cirka 6 pct. af alle indlæg på fem nyhedsmediers sider modereres, og at en tredjedel af disse kvalitativt blev vurderet til at være nedladende, grove, racistiske, truende og enkelte af dem strafbare. De knap 5 pct. angrebskommentarer, som vi finder på mediesiderne, og de knap 8 pct. angrebskommentarer, som vi finder på politikernes Facebooksider, repræsenterer altså et minimum af angreb i debatten. Andre er blevet slettet af Facebook, sideadministratorer eller af debattørerne selv.



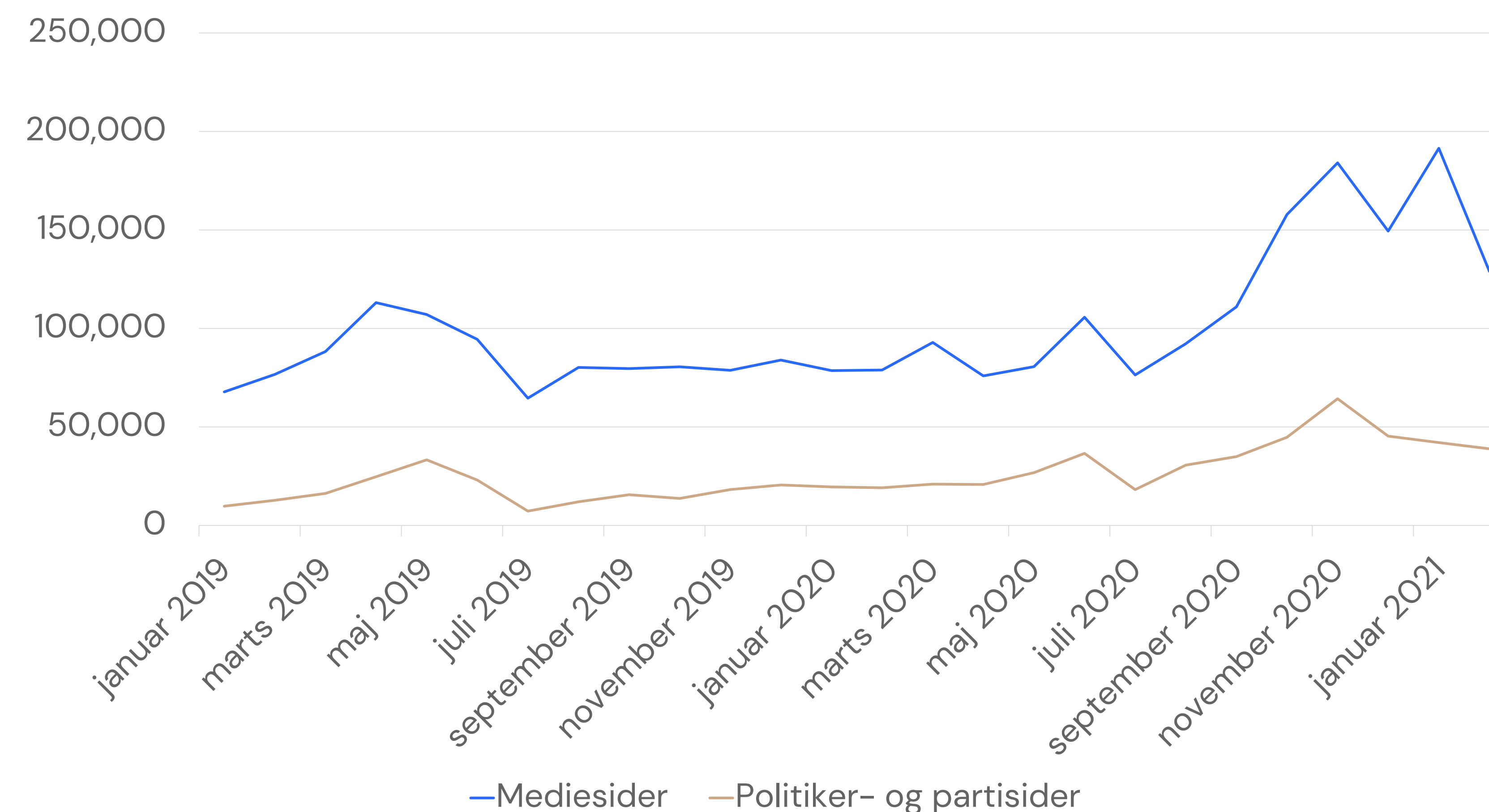
# Angreb i den offentlige debat på Facebook over tid

Angrebene i den offentlige debat påvirkes både af den generelle aktivitet på Facebook og af begivenheder ude i verden.

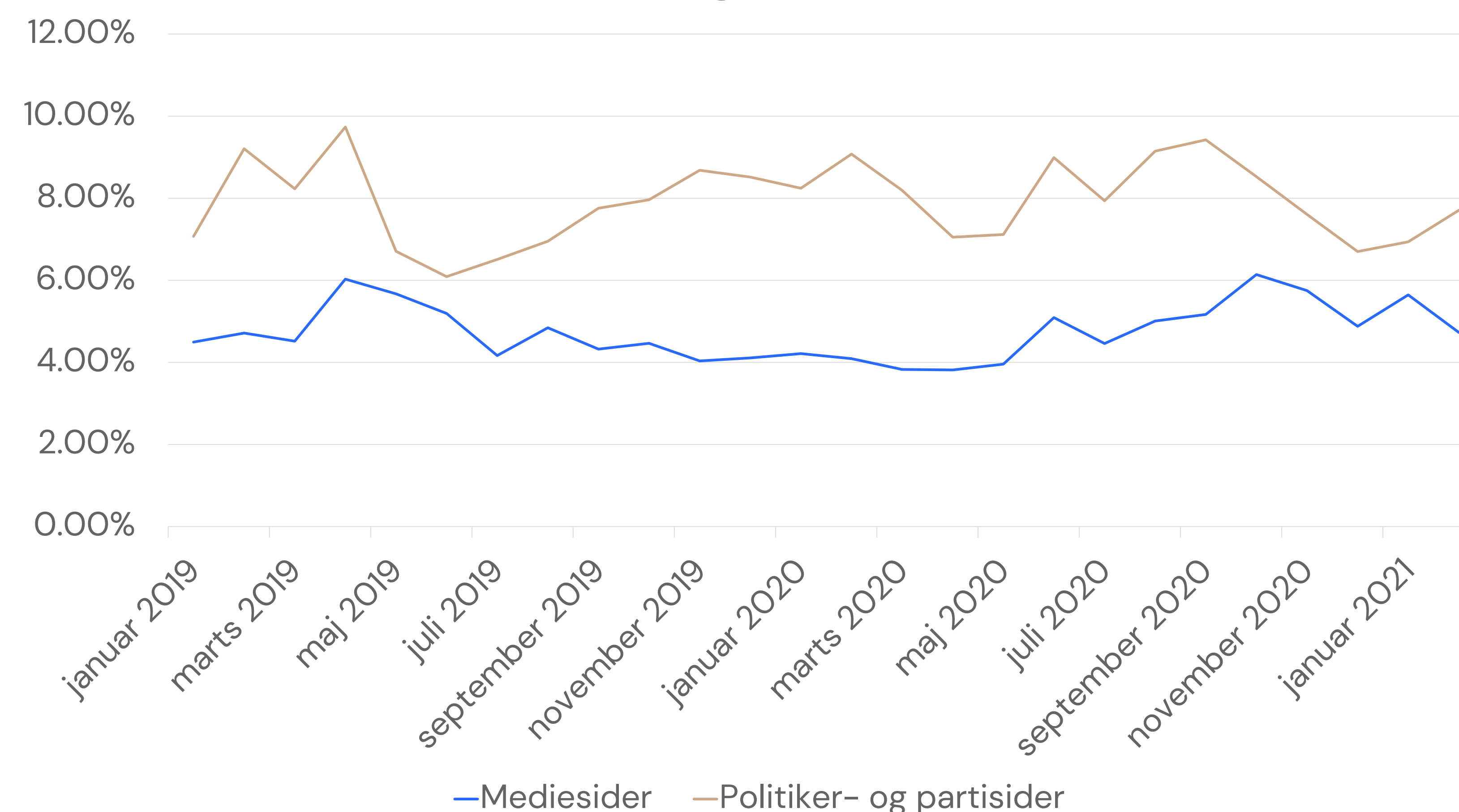
For eksempel kan vi se, at antallet af angreb i debatten stiger omkring Folketingsvalget (juni 2019), terrorangrebene på Sri Lanka (april 2019), Black Lives Matter-demonstrationerne (juni 2020) og markant i efteråret 2020, hvor Danmark både går ind i anden coronanedlukning, og det Amerikanske præsidentvalg afholdes. I nedlukningsperioden og valgperioden stiger antallet af angreb markant, men fordi den generelle aktivitet på de sociale medier også stiger, udgør de ikke en større andel.

Angreb på politikernes sider falder, når deres sommerferie starter, og sideindehaverne selv er mindre aktive. Denne tendens bekræfter vores tidligere arbejde med social medie analyse, der viser, at sideindehavernes aktivitetsniveau og opslag har indvirkning på danskernes interaktion.

## Antal angreb over tid



## Andel angreb over tid



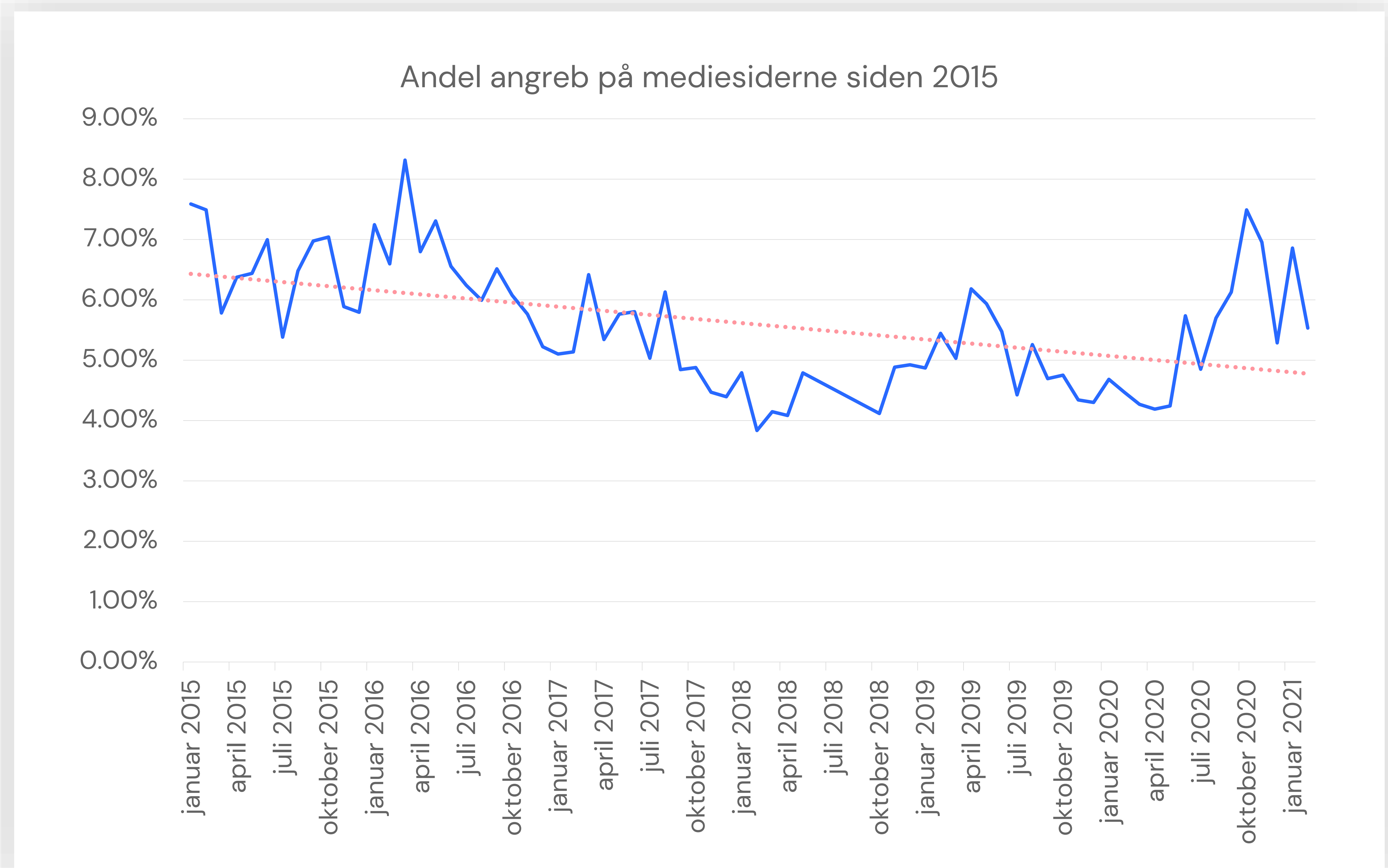


# Et kig på udviklingen endnu længere tilbage\*

I figuren til højre ser vi, at andelen af angreb i kommentarsporene på mediesiderne er svagt faldende siden 2015.

I 2016 er andelen af angreb størst, især på grund af terrorangrebet i Bruxelles i marts. Året 2018 er det mest fredelige år på mediernes Facebooksider.

Overordnet er andelen af angreb, der får lov at blive stående, faldet siden vores egen og IMRs kortlægninger i 2017. Resultaterne bekræfter Facebooks egen udmelding om øget moderation, ligesom vi nok ser resultatet af, at sideejerne selv er gået mere aktivt ind i at håndhæve nogle regler for debatten på deres sider. I hvert fald har den hårde tone været oppe at vende i medierne flere gange de seneste år.



\*Tal fra før 2019 er baseret på en stikprøve på 10% af den offentlige debat på mediesiderne helt tilbage til 2015.

4

4

4

4

**Angreb på  
politikernes  
Facebooksider**

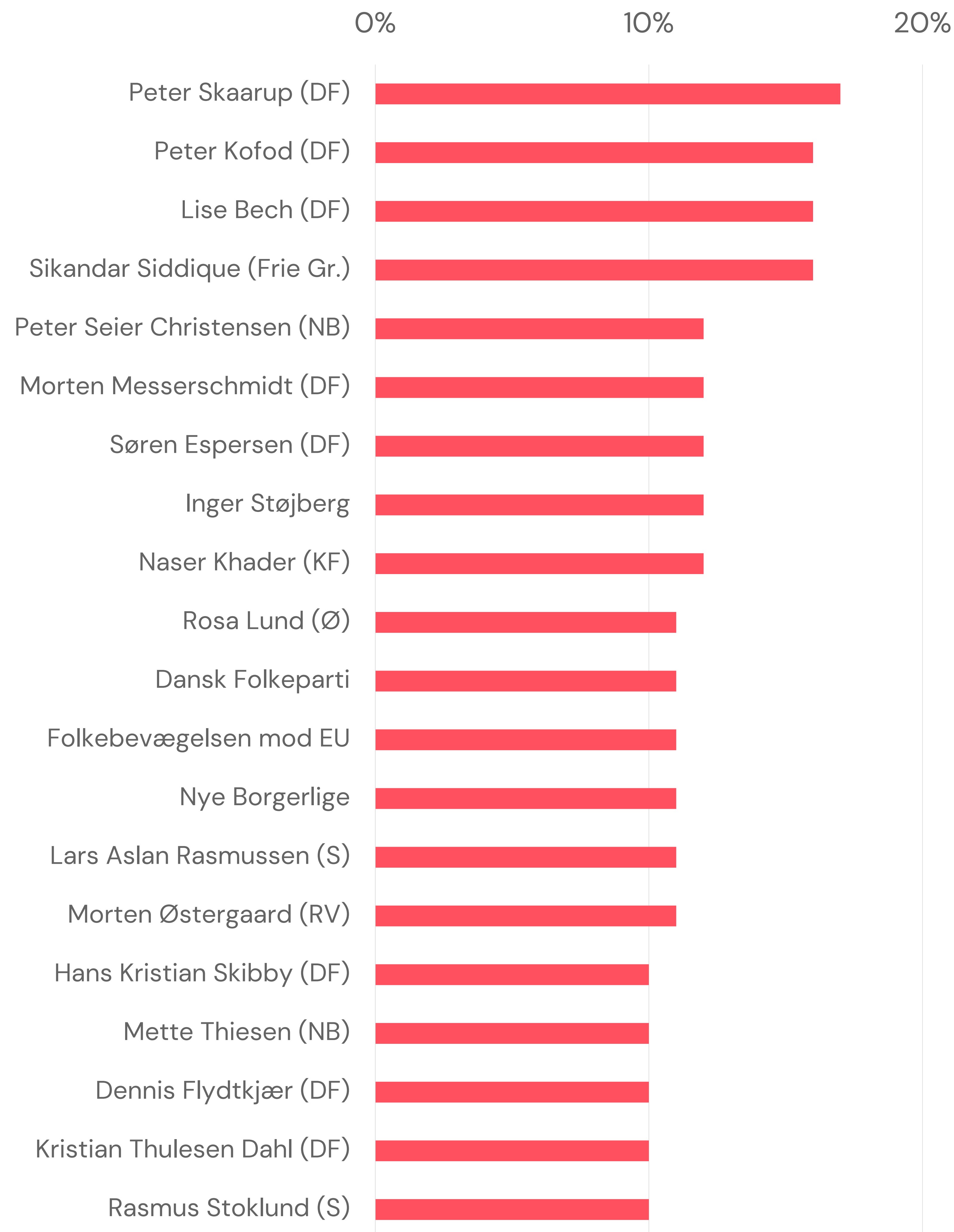


# Politikere og partier med størst andel angreb

På listen over politiker- og partisider med den største andel af angreb i kommentarsporene finder vi en del politikere fra Dansk Folkeparti. Der er også enkelte politikere fra andre partier samt partisiderne for Dansk Folkeparti, Nye Borgerlige og Folkebevægelsens mod EU. Fælles for mange af de politikere og partier, der optræder på listen, er, at de ofte skriver om (især muslimsk) indvandring, integration og udlændinge samt kriminalitet- og retsspørgsmål. På listen finder vi flere politikere, som netop er udlænding-, integrations-, indfødsrets- og retsordførere.

Peter Skaarup fra Dansk Folkeparti er den politiker, som har den største andel angreb i kommentarfelterne. Algoritmen finder, at 17 pct. af kommentarerne på Skaarups side er angreb. Dernæst følger hans partikolleger Peter Kofod og Lise Bech samt Sikandar Siddique fra Frie Grønne, som alle har ca. 16 pct. angreb i kommentarsporene. De 20 nederste politikere på listen har til sammenligning mellem 0-2% angreb stående tilbage i deres kommentarspor.

Politiker- og partisider med størst andel angreb

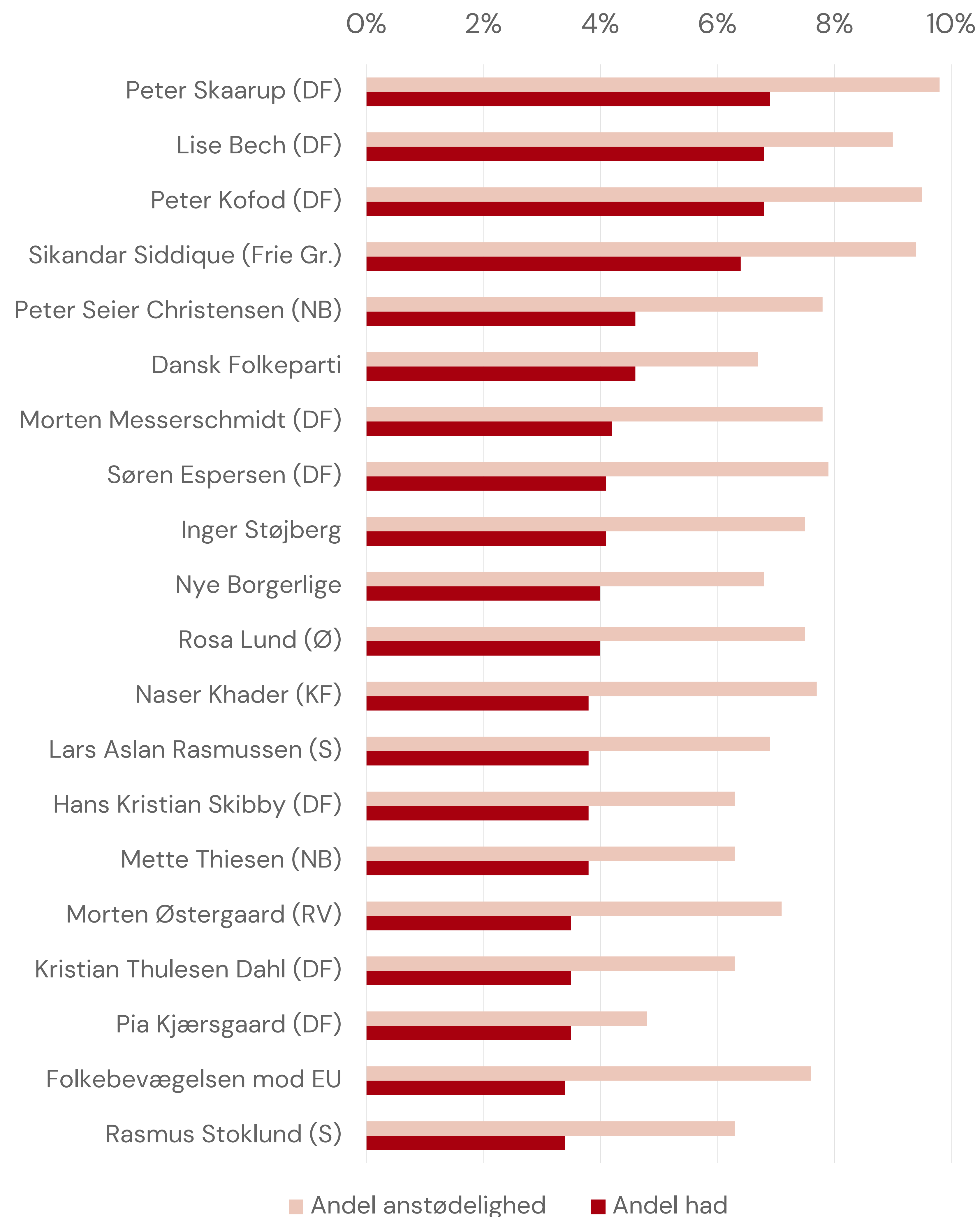


# Had og anstødelighed på politikernes sider

På figuren til højre er angrebene delt op på, om de er hadefulde eller anstødelige. Vi har sorteret politiker- og partisiderne efter, hvor der er den største andel af had. Figuren ligner meget figuren fra før, hvilket er udtryk for, at de politikere, der har procentvist flest angreb i kommentarsporene, også er dem, hvor den hadefulde tone fylder mest. Der er blot et enkelt nyt navn på listen, nemlig Pia Kjærsgaard, der er sprunget op blandt de 20 politikere og partier med størst andel af had i kommentarsporet.

Peter Skaarup, Peter Kofod, Lise Bech og Sikandar Siddique ligger fortsat øverst på listen over politikere og partier med størst andel af hadefuldt sprog i kommentarsporet. Hos Peter Skaarup er 7 pct. af kommentarerne hadefulde, og 10 pct. er anstødelige. Hos Kofod, Bech og Siddique udgør hadtale mellem 6-7 procent, og mellem 9-10 pct. er anstødeligt sprog. Dykker vi ned i kommentarerne, kan vi se, der er forskel på, hvem de hadefulde og anstødelige angreb retter sig imod. Mens nogle politikere i høj grad er "værter" for hadefulde angreb, oplever andre i større eller mindre grad at være "genstand" for hadefulde angreb.

Had og anstødelighed på politiker- og partisiderne





## Eksempler på angreb i Skaarups, Kofods og Bechs kommentarspor:

### *Etnicitet*

Det lort hører ikke til her, havde det været i deres egen land, så var de sat i kassen. Islam er en sygdom der breder sig som kræft hvis det ikke bliver bekæmpet. Forbyd Islam i Danmark, og send alle pædofil-profetens følgere ud af landet. Islamfrit Europa!

den hatte neger skal bare ud og ned i jobcentret med ham og tag din søster med og alt jeres øgle yngel

Hjem med hele skidtet. hører ikke hjemme i EU

Lad dem rådne op i deres jordhuler

### *Politiske modstandere og andre*

En tyv der stjæler igen og igen, han burde have fingerne amputeret.

Hvad med en jagttid på veganer. de skulle sku have et skud salt lige i røven. Tag deres mobil fra dem, og giv dem et lag tæsk.

Brug kræfterne på noget i har indflydelse på DF.... Det der er skønne spildte kræfter..... Danmark bliver fyldt med de hjernevaskede muslimske psykopater der har halshugget kristne og andet godtfolk i Syrien og Irak.... Samtidig skal vi betale for deres usle feltadresser af koner og deres afkom..... Kvalmende... og lad vær at brug kræfter på at overbevise mig om at det ikke kommer til at ske..... Og hvem kommer til at betale med død og ødelæggelse når en af kulturberigerne får held til at lave en terroraktion et sted i Danmark, og nej Peter Kofod, det bliver ik dig..... Det bliver Hr og Fru jensen..... Føj for helvede....

Pernille Skipper måtte man såmænd godt for min skyld gi' et spark i røven. Hun opfordrer jo ligefrem til voldelig adfærd. Sikke en "Luder"

Løgnagtig skide landsforrædder

Hjernedøde svenskere !

## Angreb i DF'ernes kommentarspor

Angrebene i DF-politikernes kommentarspor ligner generelt hinanden. Til venstre har vi udvalgt nogle eksempler fra de tre DF-politikere med de største andele af angreb og had i kommentarsporene. Et fællestræk er, at mange af angrebene retter sig mod muslimer. Når angreb retter sig mod personer med beskyttede karakteristika, karakteriserer vi det som hadtale. Så når Facebookbrugere betegner en religiøs gruppe eller etnicitet som f.eks. "hele skidtet", "det lort", "en sygdom" eller "hjernevaskede psykopater" er der tale om hadefuldt sprog.

Foruden angreb mod muslimer og indvandrere, finder vi også angreb mod andre politikere og politiske modstandere. For eksempel mod Pernille Skipper (Ø), som bliver kaldt en "luder", Morten Østergaard (RV), der bliver kaldt "løgnagtig skide landsforræder" og mod veganere, der skal jages og have tæsk. Eller mod de "hjernedøde svenskere", der åbner op for at tage flere migranter.

Når angrebene baserer sig på køn ("luder"), nationalitet ("hjernedød svensker") eller politisk overbevisning ("veganer"), betegner vi dem som hadtale. Når angrebene retter sig imod offentlige personer (f.eks. tidligere formand for Radikale Venstre Morten Østergaard) eller kriminelle generelt (f.eks. tyve), betegner vi dem som anstødelige.

## Eksempler på angreb i Sikandar Siddiques (FG) kommentarspor:

Luk røven siddique. du er den største racist mod etniske danskere, gid du snart må blive smidt ud fra borgen

GUD GØR VI SQ DA EJ. Men du skal være velkommen til at gå og tænke på holocaust hele dagen din abe

Du gift for det danske demokrati, med din Islam udbredelse. Godt ske der er religions frihed men ikke lighed.

Godt man kan slå lyden fra denne grønne racistiske marsmand...denne racist er jo fuldstændig uden for dansk pædagogisk rækkevidde...ydermere har racistens intelligensniveau for længst... nået bunden...omtrent lige der hvor kloakken...starter sit udløb !!

Danskerne Har bare fået nok af jer, og giver igen af samme skuffe. Og Så syntes jeg ærligt talt at du snart skal holde din åndsvage kæft lukket, der kommer kun lort ud af den. I muslimer har ødelagt Danmark, Så jeg siger også fuck koranen

Din Store Møghund!!Du er Danmarks stærkeste mand!! Du er eneste nulevende levende abort der selv er kravlet op af skraldespanden!

Hold nu din kæft skidekande du er nul og ingenting din kamel abort

nu er morten messerschmidt europæer og ikke en sandkasse rotte med en forkvaklet hjernedød tankegang som dig der ikke engang hører til i hverken danmark el europa

Hjem til jordhulen med dig

## Tonen er særligt hård, når man selv har muslimsk baggrund

Mens angrebskommentarerne på de tre DF'eres Facebooksider i høj grad er rettet mod muslimer og politiske modstandere til DF, så er angrebene på Sikandar Siddiques Facebookside i højere grad rettet mod ham selv og hans etniske ophav.

Siddique får at vide, at han er en "sandkasse rotte", en "abe", en "møghund" og "gift for demokratiet", og at han skal tage "hjem til jordhulen". Flere af angrebene er rettet imod Siddiques etnicitet eller religiøse overbevisning, og derfor klassificerer vi dem som hadefulde.

Det er bemærkelsesværdigt, at mens angrebene på de tre DF'eres Facebooksider i høj grad kom fra personer, der var enig i DF's politik, så er angrebene på Siddiques side fra personer, der er imod ham.

Der er også andre politikere med muslimsk baggrund i Folketinget – hvorfor er det lige Siddiques Facebookside, der scorer så højt på angreb og had? Et bud kunne være, at Siddique skriver om temaer som religion, diskrimination, racisme, hadforbrydelser og værdier. Det er netop emner, som ofte får bølgerne til at gå højt i kommentarsporene.



### Angrebskommentarer hos Rosa Lund (Ø):

Total hjernedød, intet sans, ingen erfaring, fuldstændig mærkelig budskaber der overhovedet ikke har noget med samfundet eller dens nuværende problemer. Er du bare dum ?

Luk nu røven so. De skal sgu ikke til Danmark.. De her deres eget land og de valgte selv at tage hjem som iss terrorister og sidder de en fangelejr...Hvad fanden kommer det os ved.Ikke en skid.Hvis du syntes det er så synd for dem.. Så flyt ned til dem..Du har intet at sige. Du er jo selv kriminelle.

Man skal satme være fuld for at sprænge på dig ....på kanten til udpumpning og ...så ville JEG være krænket ...du er så symmetrisk skæv i kraniet at en skulptør skulle lave det makværk ville blive fyret på stedet...AF KOMMUNEN ! Selv indmaden på 4 sal er en gang datooverskredet gullash suppe ...karma is a bitch Rødbedehjelm... hop i rensningsanlægget på Amager og befri verden for en klaphat som sig ....BA OUT

### Angrebskommentarer hos Lars Aslan Rasmussen (S):

Lars, skulle vi så ikke følge den til dørs og få smidt så mange islamister som muligt ud af vores land? Hvorfor skal vores liv forpestes og hvorfor skal vi overhovedet bruge tid på de idioter? Væk med dem og vi kan komme videre med vores liv. De skulle aldrig have været her alligevel....

Tak Lars Aslan. Vi mangler nogle politikker som giver kontant afregning. Håber de er identificeret og kan sendes hjem! Hvad laver de her?? Unge demonstranter har betalt og stadig betaler stor pris med livet som indsats. og de idioter sidder her og udnytter ytrings og frihed!

### Angrebskommentarer hos Naser Khader (KF):

Hvis bare alle indvandre var som Naser Khader, så var de velkommen og de ville ikke give problemer, men man skulle tro at det er det muslimske affald de sender til Danmark.

Naser, du er som en lille hund der slikker røven på din master. Vi skal nok lade være med at smide dig ud af danmark. Du slikker og slikker og slikker, men ingen viser sig ansigt. Du er så langt fra hvad dansk er. Du er araber og du kommer aldrig til at blive danskere. Selv hvis du slikker aller danskeres røvhuller.

## De sprængfarlige ordførerskaber

Flere af politikerne på listerne over, hvor der er størst andel af angreb og had i kommentarsporene, er integrations-, udlændinge-, indfødsrets og/eller retsordførere. Det gælder f.eks. Rosa Lund (Ø). Flere af angrebene i Lunds kommentarspor er rettet mod hende og Enhedslistens politik, og i nogle tilfælde går de også på hendes køn, hvorfor vi vil klassificere dem som hadtale.

Hos Naser Khader (KF), der er retsordfører, og hos Lars Aslan Rasmussen (S), der bl.a. er indfødsretsordfører, er kommentarerne i overvejende præget af opbakning til deres linje. Khader og Rasmussen har begge selv muslimske aner, men i modsætning til kommentarerne på Siddiques Facebookside, er de fleste angreb ikke rettet mod Khader og Rasmussen selv. De er typisk mod de mennesker og grupper, som Khader og Rasmussen skriver om i deres opslag. Det gælder dog ikke alle kommentarer, f.eks. beskyldes Khader for at være en "lille hund, der slikker røv" og aldrig bliver dansk.

Hos udlændinge og integrationsordfører Rasmus Stoklund (S) er der både en del angreb, som er rettet mod Stoklund og Socialdemokratiets "slappe udlændingepolitik", og en del angreb, der retter sig mod islamister, kriminelle udlændinge og muslimer generelt samt politiske modstandere til Socialdemokratiet.

## Eksempler på angreb i Pernille Skippers (Ø) kommentarspor:

Hjerneblæste kvindemenneske tag skyklapperne af kan hun slet ikke se hvad der sker i vores engang fredelige land skal vi have endnu mere kriminalitet voltægter overfald osv

Du er bare den størst luder på to ben vi skal ikke have flere af sådan nogen ind du ødelægger landet så ud med de alle

Det er dog noget af en modbydelig svada den mægkælling kommer med. Hun er en sæk der ikke jan tåle at få en smule magt.

Lille dumme tøs ,det kan være man er nød til at rejse væk så man slipper for at høre og se på jeres ævl hele tiden Det er sku da os selvstændige som kommer til at betale for jeres sindssyge politik ,og ja jeg er også ude af Danmark men tager selvfølgelig de test som jeg skal ,føj

Kæft en rundetrunte, kæft hun er frastødende, klamme so

Flet dog næbbet Skipper, falske kvindemenneske

du syg i skallen kælling, hvergang der er nogle fejlfarver der ynker det mindste så får de i hovede og røv, imens der sidder mennesker på vores plejehjem der er lænket til deres senge smurt ind i urin og lort. dem skider i på, send forhenvede de perler hjem så vores egne kan blive passet på hvem lider mest i dette samfund

Hun ter sig som en Sindssyg kælling!!! Hun tænker ikke på de mange tabte arbejdspladser

Du er jo helt væk i roen, kvindemenneske

Klamme kælling, du er en taber

Pernille. Du er ikke bare en ordinær racist mod danskerne, dit folk. Som du, til trods for at være jurist, ikke kan ligestille folkeretsligt, juridisk m. Verdens øvrige folk. Du er en ualmindelig ækel racist. En skam for dit folk og en gemen landsforræder.

Føj!

## Rundetrunter, ludere og sindssyge kællinger i politik

Vi ser flere eksempler på, at kvindelige politikere, der blander sig i debatter om f.eks. integration og ligestilling, får mange hadefulde angreb med på vejen i kommentarsporene.

Et godt eksempel på en sådan politiker er Enhedslisten Pernille Skipper, som var politisk ordfører fra 2016–2021 og således dækkede en bred vifte af emner. På Pernille Skippers Facebookside finder vi mange eksempler på hadefuld tale, fordi angrebene bruger nedsættende ord om hendes køn. Pernille Skipper bliver bl.a. kaldt "hjerneblæst kvindemenneske", "rundetrunte", "klamme so", "lille dumme tøs", "luder", "sæk" og "møgkælling". Pernille Skipper ligger nr. 45 på listen over politikere og partier med den største andel af angreb i kommentarsporet. 7 pct. af kommentarerne på Skippers side er ifølge algoritmen et angreb.

Vi finder også grove angreb hos mandlige politikere, men når vi dykker ned i udsnit af kommentarerne er disse sjældent rettet mod det maskuline køn eller baseret på køn. Angrebene vil derfor ofte blive kategoriseret som anstødeligt sprog fremfor hadefuld tale – med mindre at angrebene er rettet mod andre former for beskyttede karakteristika. Det kunne være etnicitet, som vi så det hos Sikandar Siddique, eller handicap, som vi skal se på næste side.



## Eksempler på angreb i Kristian Hegaards (RV) kommentarspor:

Du er en dumt og ulækker lille spadser

Synes du sku hænge dig dit misfoster

Pudsigt når en hjerneskadet kalder en mand i kørestol for idiot. Men han ved jo nok ikke hvor hjernen sidder...

Han må være dybt debil, hvis han ikke kender retsstat og retsprincipperne!

Du er sgu for retarderet at høre på.... tag dig sammen

Var han født i førerens tid var han smidt på møddingen sammen med natpotten arme lille.. mund hans mor ikke fortryder at ha født han

Din lille lede undermåler – du vil hellere støtte islamisk pædofili – end at hædre en ægte dansk værdikriger – som kæmper for at holde islam væk fra danmark. Dit lille elendige afskum – jeg håber fandeme at der en dag kommer én fra ""fredens religion"" og pifter din rullestol. Syge, syge, radikale, radigale udskud

Kunne du ikke bare køre i havnen. Radikale udskud. I er bare så syge i potten. I er ikke til gavn for nogen som helst danskere. Håber I ryger under spærregrænsen, ved næste valg. I er godt nok det værste parti i folketinget. Danskerne, skal ikke betale til jer.

Skub ham forhelvede i havnen radikale møjdyr

Fucking store multi idiot

Din lille rotte. må du få kløe i røven for den kan du vel ikke nå

Håber sku det bliver din tur en dag din skøre spasser

## Et handicap fungerer som skydeskive

Kristian Hegaard er Radikale Venstres ordfører for bl.a. integration, udlændinge og retspolitik. Hegaard er desuden født med en muskelsygdom, der gør, at han sidder i kørestol – og dét giver genlyd i kommentarsporene på hans Facebookside.

Vi finder flere hadefulde angreb, som er rettet mod hans handicap. Kristian Hegaard får i kommentarsporet at vide, at han er en "lille spadser", et "misfoster", "dybt debil", "retarderet", en "lille led undermåler", en "multi idiot" og et "udskud", og mens nogen opfordrer ham til at køre i havnen, skriver andre, at han "i førerens tid" var blevet "smidt på møddingen sammen med natpotten".

På Kristian Heegaards Facebookside er knap hver 10. kommentar et angreb, og i vores data ligger han som nr. 25 på listen over politikere og partier med den største andel af angreb i kommentarsporet. Det er ikke alle angreb, der er rettet mod ham selv. En stor del af angrebene på hans side er rettet mod andre personer eller grupper, han skriver om i sine opslag.

Men det er iøjnefaldende, hvordan det at repræsentere en minoritet i dansk politik bliver en skydeskive for en række hadefulde angreb i kommentarsporene på Facebook.

**5**

**5**

**5**

**5**

**Angreb på  
mediernes  
Facebooksider**

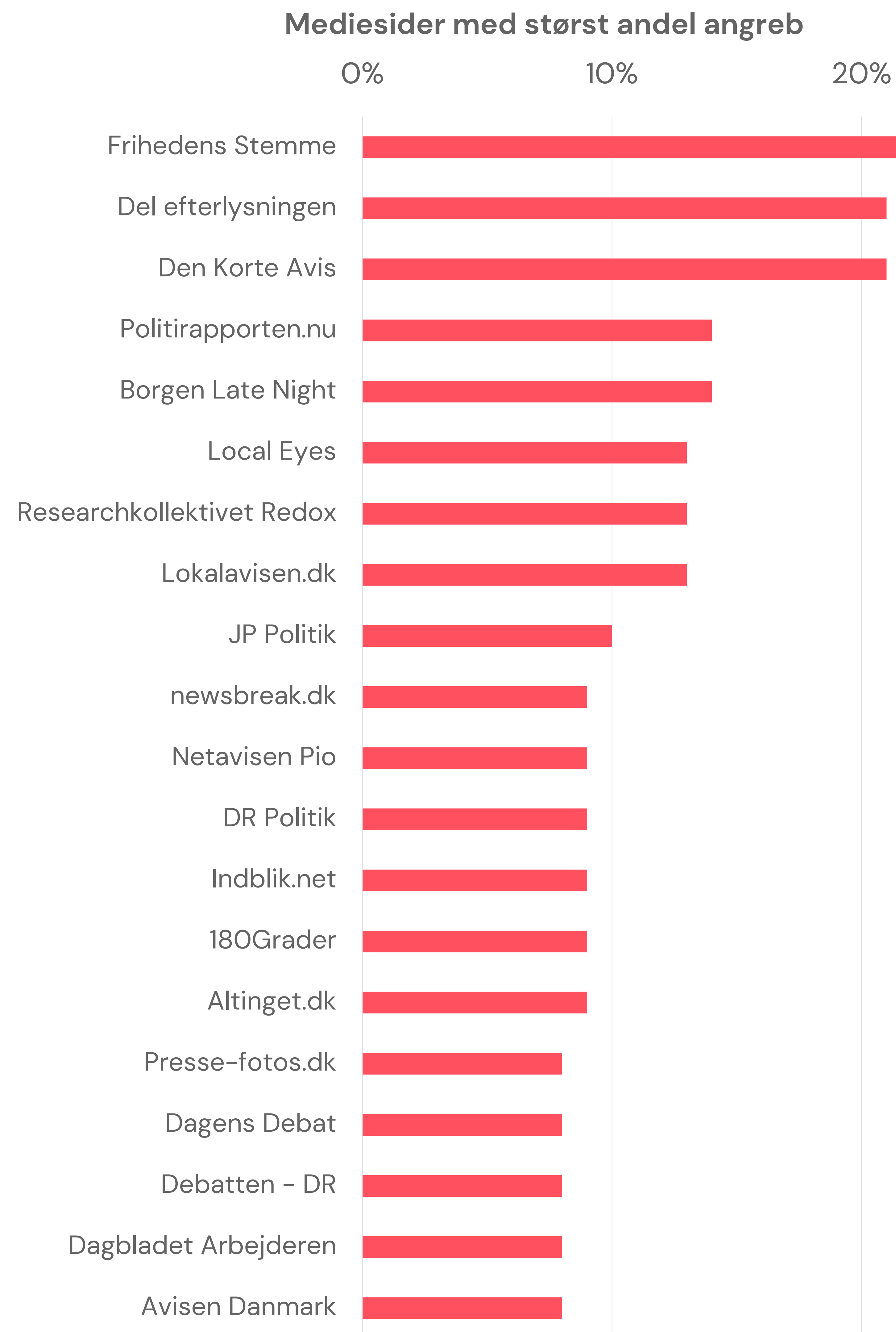


# Medier med størst andel angreb

På listen over de mediesider med den største andel af angreb i kommentarsporene, ser vi både små alternative medier og store landsdækkende medier. Vi finder både højreorienterede medier som Den Korte Avis, Indblik.net og 180Grader og venstreorienterede medier som Research-kollektivet Redox, Netavisen Pio og Dagbladet Arbejderen.

Frihedens Stemme er det medie, som relativt set har flest angreb i kommentarsporene. Her er knap hver fjerde kommentar (24 pct.) et angreb. Frihedens Stemme rapporterer efter eget udsagn om "friheden" og "trusler mod friheden" i Danmark og deler mange opslag om islam og muslimsk indvandring. Dernæst følger Del efterlysningen og Den Korte Avis, hvor godt hver femte kommentar (21 pct.) er et angreb.

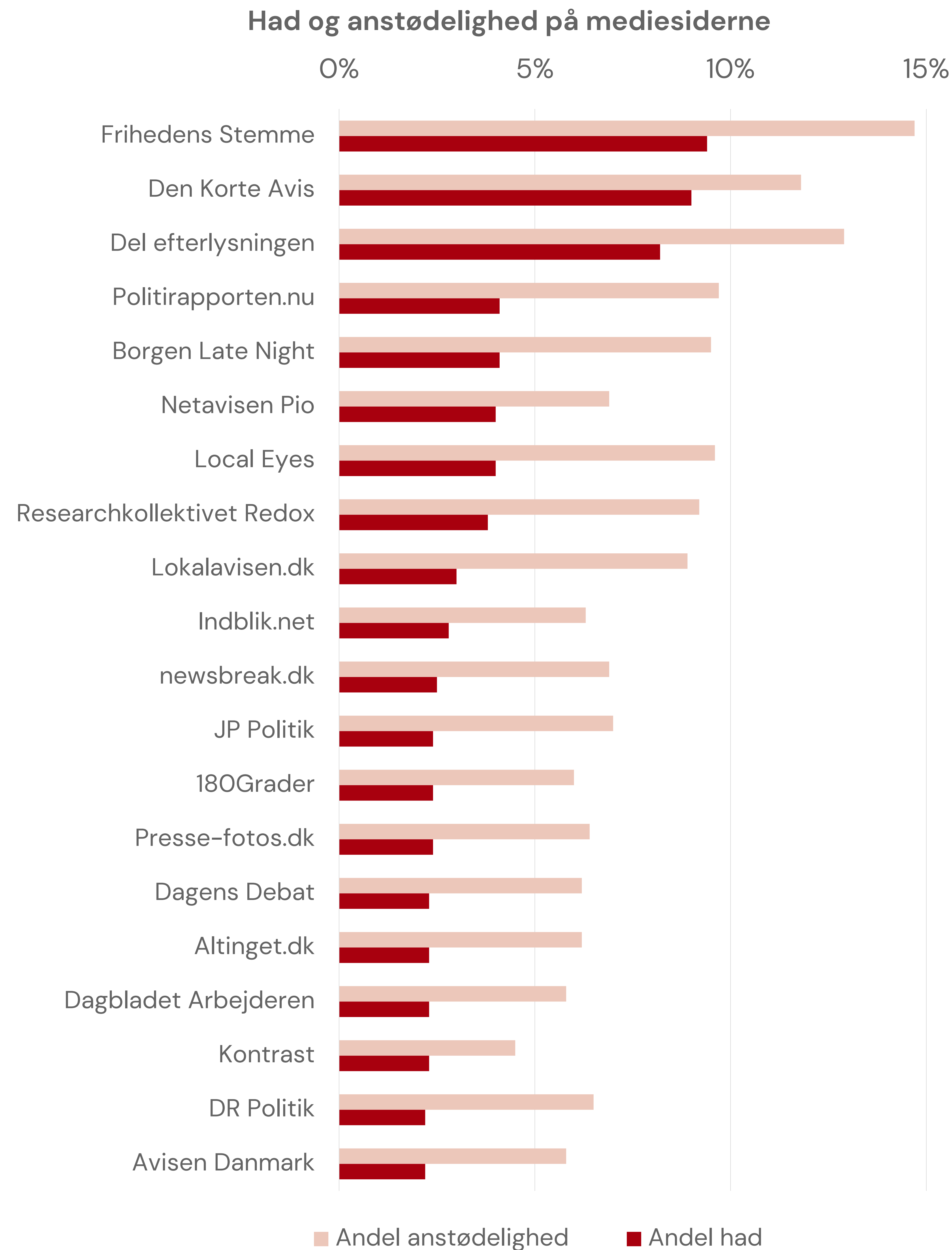
Der er flere medier på listen, som fokuserer på kriminalitetshistorier. Hos Del efterlysningen kan man dele mediets opslag om forskellige former for kriminalitet og efterlyste personer. Kriminalitetshistorier er også omdrejningspunktet for Politirapporten.nu, Local Eyes og Presse-fotos.dk.



# Had og anstødelighed på mediernes sider

På figuren til højre har vi delt angrebene op på, om de er hadefulde eller anstødelige. Ligesom hos politikerne er der et stort sammenfald mellem de mediesider, der relativt set har flest angreb i kommentarerne, og de mediesider, hvor vi ser den største andel af hadefuldt sprog. Når vi sorterer efter andelen af had i stedet for andelen af angreb, er der et enkelt nyt medie, der kommer op på listen, nemlig Kontrast, som kalder sig "Danmarks borgerlige medie". Netavisen Pio er kravlet højere op på listen, mens DR Politik ligger lavere målt på had end på angreb.

Frihedens Stemme er igen det medie, der topper listen, og her er næsten hver 10. kommentar, der bliver skrevet i kommentarsporene, hadefuldt. Herefter følger Den Korte Avis, hvor 9 pct. af kommentarerne er hadefulde, og hos Del efterlysningen er lidt over 8 pct. af kommentarerne hadefulde. På de tre sider er mellem 12 og 15 pct. af kommentarerne anstødelige.





Eksempler på angreb i kommentarspor hos Del efterlysningen, Politirapporten.nu, Local eyes og Presse-Foto.dk

*- Kriminelle generelt -*

Syge klamme fucking nederen retarderet lorte undskyldning af et menneske!!

Lol, "fængsel" i danmark. Det er jo ikke et fængsel, det er en hyggeklub med god mad og playstation i cellerne. Skyd ham og smid kadaveret i vejkanten, han fortjener ikke engang en begravelse den lille søn af en tør kusse.

Lade mig slå dem alle ihjel , de Svin skal død

*- Kriminelle indvandrere/udlændinge -*

Lad dem komme ud på en øde ø. Giv dem en masse våben og lad dem skyde sig selv.

Igen og igen Sådan nogle indvandre god at de blev godt og grundig skambidt. Men omvendt var det synd at hunden skulle smage på sådan noget lort

Østeuropæisk mand, et forbandet møgdyr, som aldrig skulle have haft muligheden for at komme til Danmark

Det kan jo kun være de primitive huleboere igen

*- Dyremishandling-*

Mon ikke den spinkle fyr er bedst tjent med at melde sig selv til Politiet hurtigt, og inden resten af Danmarks hesteejere finder ud af hvem den spinkle modbydelige, afstumpede, perverse, sexueltafvigende og mentalt syge fyr er ...!?

Fucking snotabe, du skulle være kastet ud fra et fly

## Små og store forbrydelser bringer sindet i kog

Flere af de medier, som har en stor andel angreb i kommentarsporene, skriver om kriminalitet. Hos Del efterlysningen, Politirapporten.nu og Local Eyes kan man læse om alt fra tricktyveri og biltyveri til voldtægt og overfald. Presse-Foto.dk laver opslag med billeder af bilbrande, demonstrationer, hærværk og i det hele taget "fotoegnet kriminalitet".

Der er især en anstødelig tone i debatten, når der er tale om grov kriminalitet som voldtægt, misbrug af børn, drabsforsøg mv. Men også demonstrationer og bilræs kan antænde angreb i kommentarsporene. Nogle gange tipper det over i had, når angrebet f.eks. benytter sig af nedsættende ord om handikappede ("retardereret lorte undskyldning af et menneske") eller går på etnicitet. Der er rigtig mange hadefulde kommentarer, som retter sig mod gerningsmændenes etnicitet, hvis de er udlændinge eller indvandrere – eller hvis det antages, at de er det.

Endelig får dyremishandling en særlig hård medfart i kommentarsporene. En nyhed om en dreng, der har kastet en kat op i luften, får i kommentarsporet følgende reaktion: "Fucking snotabe, du skulle være kastet ud fra et fly".



## Eksempler på angreb i kommentarspor hos Research-kollektivet Redox, Netavisen Pio og Dagbladet Arbejderen

RAKKERPAK hele bundtet. Nationalister, racister og nazister, – andet er det ikke.	NB er i sin grundsubstans nazistisk, nok også derfor Pernille Vermund hellere ser den syge mand – Paludan i folketinget fremfor nogen anden, under alle omstændigheder er stranken af rotter den samme.
Det er cool nok at han laver pussygrab den gris. Øhhh jeg husker det ikke. Typisk gris men han ender jo nok med et job til et par mill . Den gris. Han skal ned med nakken den gris. Skide seriekrænker svin	Skidtspanden Nye Borgerlige er en højreradikal loge og et samlingspunkt for nazister og andet utøj fra den aller yderste højrefløj
Fucking nazisvin..	Knippel liderlige møgsvin de ku ha prøvet at svinge dem jo
"Møgdyr! Min fulde foragt for den zionistiske forbrydelser-stat israel!"	Et formynderisk socialdemokratisk parti støttet af stalinister og klimatosser. Det skal nok gå godt.
En pædofil, en fascist og så det bimlende hjernedøde kvindemenneske til Trine Bramsen – hvad kan gå galt?	luk røven gamle nar...

## Den politiske venstrefløj

Blandt de medier, som har størst andel af angreb og hadefuldt sprog, i kommentarsporene, finder vi både højre- og venstrefløjen repræsenteret. På venstrefløjsmedierne Research-kollektivet Redox, Dagbladet Arbejderen og Netavisen Pio finder vi især angreb rettet mod højrefløjen. Nye Borgerlige bliver kaldt racister og nazister og sammenlignet med rotter, og Morten Østergaard (RV) bliver kaldt et "skide seriekrænker svin" i kommentarer til #metoo-debatten.

I kommentarsporene bliver det tydeligt, at den dominerende politiske overbevisning blandt følgerne hos de tre medier varierer, og at der også er folk, der er grundlæggende uenig med flertallet af følgerne på siden, der blander sig. På Netavisen Pio, hvis ideologi flugter med Socialdemokratiets, bliver der bl.a. skrevet om et "formynderisk socialdemokratisk parti støttet af stalinister og klimatosser" i kommentarsporet, og et opslag om Mogens Lykketoft får meldingen: "luk røven gamle nar...".

Redox og Dagbladet Arbejderen ligger længere ude på venstrefløjen, og her får socialdemokrater og "andre socialister" også indimellem en hård medfart. F.eks. mødes et opslag om regeringens tryghedspakke af en kommentar om en "hjernevasket befolkning" og "facistisk stasi Danmarks regering".



## Eksempler på angreb i kommentarspor hos Den Korte Avis, Indblik.net, 180Grader og Kontrast

Diktator, narcissist, løgner og massemorder af mink

Det vi da gi Danmark dødsstøddet.. Så HELL NO. De aber skal fandeme ikke i regering. De skal HELT ud af folketinget

Så kan de syge hjernedøde politikere lære det.. Dem og deres forpulede afgifter på ALT. Snart kommer der vel målere op, der så kan måle gassen i den pyt man skår og så er der fandeme også afgift på det. Jeg håber sku det lykkedes for smuglerne at få en helvedes masse cigaretter uset over grænsen..

De har selv valgt side, så de må rådne op, der hvor de er. .. DER ER SGU INGEN DER VIL SAVNE DEM. .

Har vi ikke for længe føjet alle de dumme perkere

Det er nord Koreanske forhold der er indført i Danmark. Hun er sgu godt nok en magtsyg kælling. Godt hun ikke kan bestemme hvad vi spiser hos os selv. ENDNU !!!de tager alt selvbestemmelse fra danskerne.

Alternativt parti Uffe Elbæk og hans bøsse venner hører ikke til i Folketinget de er kun ude på at ødelægge Danmark med Mellemøstlige tilstande og en etnisk udrensning af Danmark befolkning

Signe Munk er syg i sit lille socialistiske hovede!

Smid dog det yngel ud med kloak vandet , endnu et tilfældet af venligeboerne og nogle partier

Kujoner.. ALTID skal de svin være flere om EEN person. NOSSELØSE aber. Op i træerne med dem.

## Den politiske højrefløj

På højrefløjsmedierne Den Korte Avis, Indblik.dk, 180Grader og Kontrast er angrebene især rettet mod rød blok. Særligt regeringens coronahåndtering får en overhaling i kommentarsporene, og Mette Frederiksen kaldes bl.a. en "magtsyg kælling" og en "Diktator, narcissist, løgner og massemorder af mink". Enhedslistens politikere betegnes som "fæle socialister" og "aber", mens Signe Munk (SF) kaldes "syg i sit lille socialistiske hoved" i forbindelse med en udtalelse om elbiler.

Ligesom med venstrefløjsmedierne er der også forskel på højrefløjsmediernes placering i det politiske spektrum og deres fokusområder. Mens vi hos Indblik.net, 180Grader og Kontrast finder anstødeligt og indimellem hadefuldt sprog inden for forskellige aktuelle sager – fra Black Lives Matter og præsidentvalg til klima og corona – handler de hadefulde og anstødelige kommentarer hos Den Korte Aviser primært om muslimer, flygtninge og indvandrere. Angreb rettet mod indvandrere og muslimer dukker også regelmæssigt op hos de øvrige højrefløjsmedier, særligt hos 180Grader, men her fylder andre højrefløjsdebatter også.



## Eksempler på angreb i kommentarsporene hos JP Politik, Altinget.dk, Debatten – DR og DR Politik:

Skrigeskinke over alle skrigeskinke.  
Forfærdelige pigebarn.

Gud fader et pattebarn ..... pas din  
skole , snottøs.

Hende Samira den Über krænkede  
Radigale psykose. Hun stopper  
aldrig før alle mand undskylder  
overfor hvad vi kunne finde på i  
hendes verden.. Sindssyge islamist.

A: "pia er danmarks nazi dame..  
sådan har det altid været" hvis man  
på intet plan kan nikke  
genkendende til A ... så er man fuld  
af lort lol.

pia lignte en dreng før i tiden, så hun  
blev mobbet, så hun mobbed  
tilbage .. og at spille den slags spil,  
har eskaleret siden. At hun nu ligner  
en semi nuttet uglemor.. er så hvad  
det er. Temmelig nemt shit.. barn  
mobbes barn slår tilbage og går for  
langt. \*dette\* er pia .... :v ked af at  
sige det sådan her, men hellere  
dette, end NAZI SHIT Satans griske  
liderlige fjols. Skrub ud og få dig et  
job din idiot.

Shhhhh din møj hvalp , endnu en  
klovn fra L.A , sæt du dig nu ned  
skrammens hjørne og vær helt stille.

Kuk kuk idiot.. Klimatosse.. Send alle  
flygtninge ud, det er der problemet  
er, men i skjuler det bag jeres  
miljøpis..

En af de korrupte Demo rotter

Du har fuldstændig ret Lars Boje de  
radikale venstre er totalt blæst i  
bøtten håber danskerne husker det  
til næste valg og ikke stemme de  
fucking halvhjerner ind de er kun til  
for de skide muslimer og de har  
bevist rigeligt hvad de er

Den latterlige gamle grå gris der i  
årevis har svinet og været ligeglad  
med vores verden HOLDA KÆFT EN  
EGOIST. gamle grå gris

## Hos de store landsdækkende medier går angrebene i mange retninger

Selvom de landsdækkende medier relativt set ikke har lige så mange som angreb, som flere af de mindre, alternative medier, er flere af de landsdækkende medier kravlet op på listen over de 20 mediesider med størst andel angreb i kommentarsporene.

Fælles for de landsdækkende medier med størst andel angreb på siderne er, at de alle har politik som omdrejningspunkt: JP Politik, Altinget.dk, DR's Debatten og DR Politik.

I modsætning til venstre- og højrefløjsmedierne skyder angrebene på de landsdækkende mediers Facebooksider i mange retninger, men oftest er angrebene rettet mod politiske personer og andre offentlige personer. Derfor vil vi kategorisere størstedelen af angreben på siderne som anstødeligt sprog, fordi det at være en offentlig person ikke er et beskyttet karakteristika. Men vi ser også eksempler på, at angrebene går på køn, alder eller etnicitet. F.eks. når formanden for den unge formand for klimaorganisationen Friday for Future beskrives som "et pattebarn" og en "snottøs"; når nyvalgt 27-årige formand Alex Vanopslag (LA) kaldes "møghvalp"; når Pernille Skipper (Ø) kaldes "skrigeskinke" og "forfærdeligt pigebarn", og når Samira Nawa (RV) anklages for at være en "sindssyg islamist". I disse tilfælde har vi at gøre med hadtale.





**Sprængfarlige  
emner i den  
offentlige debat**

# Værdipolitik sætter debatten i kog

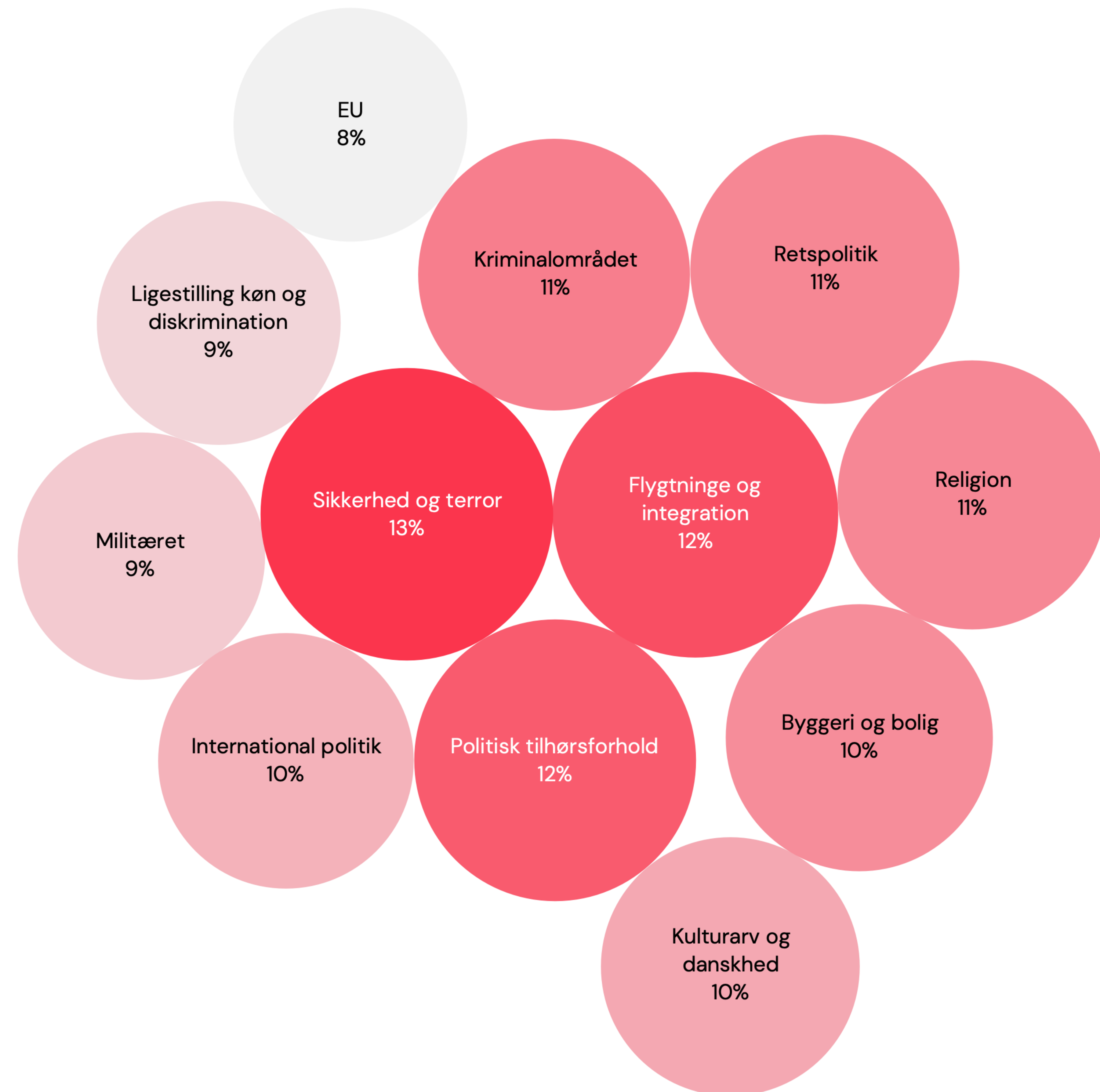
Når vi ser på de knap 3,3 millioner angrebskommentarer på mediernes og politikernes Facebooksider, er det tydeligt, at det er nogle bestemte debatområder, der især antænder den anstødelige og hadefulde tone. Måske har læseren allerede fået færten af en del af de fremhævede temaer.

Den kategori, der indeholder den største andel angreb, er *sikkerhed og terror*. Der er desuden rigtig mange angreb, der er knyttet til en form for identitetsmarkør. *Religion, politisk tilhørsforhold, køn* samt *kulturarv og danskhed*. Kategorien *flygtninge og integration* handler i høj grad om etnicitet, ligesom angrebskommentarerne i kategorien *byggeri og bolig* især handler om udsatte boligområder/ghettoer – og dermed også er knyttet op på etnisk herkomst.

Endelig er der nogle kategorier, som peger i retning af internationale forhold og udenrigsforhold: International politik, militær og EU.

## Temaerne i de hårde ringhjørner af den offentlige debat

For at kortlægge de mest hårde og hadefulde emner i den offentlige debat, har vi analyseret angrebskommentarerne ved hjælp af en emnemodel. Emnemodellen består af 2.600 søgeord fordelt på flere niveauer af over- og underemner. Figuren herunder viser i overordnede kategorier, hvilke emner der oftest forbindes med angreb. Bag hvert af emnerne gemmer sig en liste af underemner, og under dem gemmer sig lange lister af søgeord, der er relateret til de pågældende underemner.



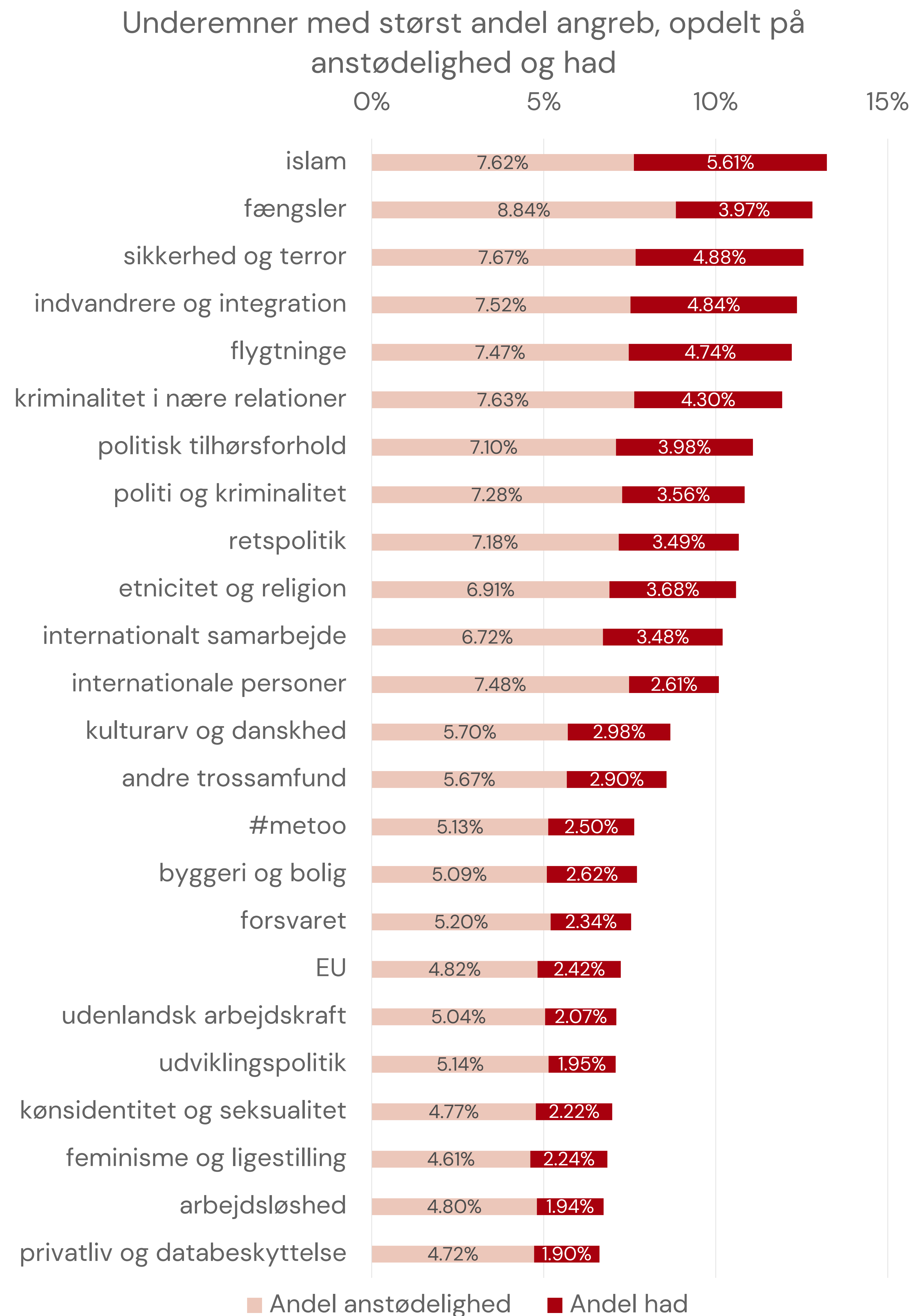


# Emner, der emmer af angreb og had

I figuren til højre har vi kigget på andelen af angreb inden for de mere specifikke underemner. Figuren er sorteret efter, hvilke emner der har den største andel af angreb på mediernes og politikernes sider. Det ses samtidig, hvor stor en andel af angrebene, algoritmen klassificerer som hadtale.

Emnet *Islam* indeholder den største andel af angreb. Mere end hver 8. kommentar om islam (13 pct.) er et angreb ifølge algoritmen, og heraf kan cirka halvdelen (6 pct.) karakteriseres som hadefulde. Højt på listen ligger også: *Indvandrere og integration*, *flygtninge* samt *etnicitet og religion*, der ligesom *islam* ofte handler om muslimer.

En stor andel af kommentarerne om kriminalitet indeholder angreb. *Fængsler* samt *sikkerhed og terror* er de kriminalitetsemner, der tiltrækker den største andel angreb. Går vi længere ned ad listen, finder vi en stor andel angreb i en række emner om internationale forhold, ligesom *#metoo* og ligestilling er tydeligt repræsenteret på angrebslisten. I bunden af er emner som hobby, forbrug, ferie, fysiske sygdomme, kæledyr, scenekunst og litteratur og graviditet.





## Eksempler på angreb om islam, indvandrere og integration, flygtninge samt etnicitet og religion:

De skal alle sammen ud af landet – de er som kræft celler der giver metastaser – jeg kom ud af Iran for at undgå at se på dem men de spreder sig rundt i verden – Religion ødelægger vores smukke verden og det er en skam!

De skulle have en KUGLE lige imellem øjnene hvis de gik over den danske/tyske grænse... Sådan nogle taliban krigere

DF. Uden hjerte, i aner intet om islam, men hvad i selv tror den er. DF er en respektløs parti. Selvfølgelig dem med et lille hjerne vil aldrig noglesind forstå hvad islam er.

Så DF er for gamle racister. NB for rige racister. Og SK for krigsliderlige racister. De har meget at vælge imellem på det sorte højre

Fuck de lorte PERKEXX giv dem en fucking kugle i panden. De kan for fanden da ikke komme her til hvores land hvores Danmark og tro de skal bestemme. Nej nu er det fandme nok. Ikke nok med at de vil have deres fucking lorte moské her i hvores land men nu vil de også ændre på hvores DANSKE historie. De kan fandme få en kugle i panden eller skride hjem til deres eget lorteland. Og bliver du offended af mit opslag er du ikke en rigtig dansker for det er dem der har invaderet Hvores Danmark

Du er et klamt menneske Morten Østergaard. Sælger de danske værdier for ussel mammon! FØJ

Hvorfor er der så mange højredrejede psykopater der skriver herinde? Kan vi ikke få ryddet op i tråden og få sendt lortet tilbage hvor det kommer fra? Jeg følger ikke Enhedslistens Facebookside for at glo på nazist-kommentarer.

## Islam og muslimer lagt for had i debatten

Inden for emnerne *islam, indvandrere og integration, flygtninge* samt *eticitet og religion* finder vi en stor andel angrebskommentarer. En del af dem er hadefulde, fordi angrebene netop går på beskyttede karakteristika: Etnicitet og religion.

I nogle af de grove kommentarer bliver muslimer sammenlignet med kræftceller og får at vide, at de "lorte perkexx" skulle have en kugle for panden. Vi ser jævnligt, at Facebookbrugere anvender dæknævne – formentlig for at gøre det vanskeligere for Facebooks algoritmer at opdage den hadefulde tale. Det kunne være "perkexx" i stedet for "perker" eller "musse" i stedet for "muslimer". Vores algoritme har gennemskuet flere af disse dæknævne. Debatsporet bliver særligt hårdt, når debatten går på kriminelle muslimer, islamister, radikaliserede og fundamentalistiske bevægelser som Hizb ut-Tahrir.

Foruden muslimer og islam retter angrebene sig mod f.eks. radikale eller venstreorienterede politikere og partier, der står for en mere moderat kurs i udlændingepolitikken. Der er også angreb, som retter sig imod politikere fra DF og Nye Borgerlige, og indimellem finder vi angreb, der retter sig mod de mennesker, der skriver hadefuldt om muslimer og indvandrere i kommentarsporet: "Hvorfor er der så mange højredrejede psykopater der skriver herinde? (...) Jeg følger ikke Enhedslistens Facebookside for at glo på nazist-kommentarer."



## Eksempler på angreb om fængsler, sikkerhed og terror, kriminalitet i nære relationer, politi og kriminalitet, retspolitik

Hvad FANDEN sker der for OS  
Danskere : at SKYDE på hinanden på  
ÅBEN gade / by = bare FORDI man  
IKKE får sine egen VILJE ligesom  
SMÅ baby eller børn " som IKKE har  
fået : mad, rent ble på, sutte,  
modermælkserstatning og  
MARIEKIKS / så gør man FANDME i  
for små SKO = når man er SÅ  
narcissistisk, psykopater eller  
STJERNEPSYKOPATER,,...!!!

Normalt i den kultur at kvinder bliver  
voldtaget og undertrykt. De er sygt  
hjernevaskede og kan ik selv se det

Ussel mentalt forstyrret terrorist  
væsen

hvorfor tager man ikke kontant  
hjælpen fra dem , alm mennesker  
arbejder om fredagen , de er jo  
nazister fatter ikke hvad de laver i  
Danmark og det bliver værre og  
værre , bare se rundt omkring i vest  
europa , måske er det Putin der skal  
rede os en dag, for vi kan åbenbart  
ikke se vores fjender lige uden foran  
vores hoveddør

Det var fedt han fik tæsk af politi lad  
os få mere af det

Hun skal da også brændemærkes.  
Hun har løjet

Tænk engang af de gider og løbe  
efter det fede svin . Han stikker af  
fra dem der prøver og beskytte ham  
. Lad ham være, lad ham løbe så  
tosset han vil, klamme svin

Seksuelle overgreb er for livstid for  
offeret. Jeg kan godt forstå hvis  
faderen havde gjort det. Det må  
være hårdt at se sine børn lide, fordi  
en nar skulle tilfredsstille sig selv.  
idiot! 11 år i fængsel, resten i  
helvedet håber jeg!!!!

Satans abeyngel, de hører ikke  
hjemme i et civiliseret land!!!

## Ingen nåde, hvis du er på kant med loven

Kriminalitet kan få folk til tasterne og vække den hårde tone i kommentarsporene. Emnet *fængsler* ligger særligt højt målt på angreb, hvilket bl.a. skyldes søgeord som prøveløsladelser, varetægtsfængslinger og fængselsdomme for kriminelle. Militante grupper som Islamisk Stat og Al-Qaeda fremkalder rigtig store andele angreb og had, samt sikkerhedsmæssige trusler som terrorisme, missiltest og selvmordsbombere. I disse tilfælde er omkring 15-30 pct. af kommentarerne angreb ifølge algoritmen. Derudover er kriminalitet i nære relationer genstand for en stor andel angreb. Her er det især søgeord som voldtægt, social kontrol og vold mod kvinder, hvor mellem ca. 10-15 pct. af kommentarerne er angreb. Når det kommer til kriminalitet generelt, er det især hjemmerøverier og overfald, der hidser danske Facebookbrugere op, men der bliver også reageret hårdt på f.eks. hærværk.

Mange af angrebene om kriminalitet definerer vi som anstødelige, fordi de ikke retter sig mod beskyttede karakteristika. Men der er også mange angreb, som benytter en nedsættende retorik om kliniske diagnoser (f.eks. psykopati), eller som går på etnicitet og religion. Endnu engang træder muslimer frem som en gruppe, der er særligt lagt for had, når der er kriminalitet indblandet.



## Eksempler på angreb om internationalt samarbejde, internationale personer, EU, udenlandsk arbejdskraft og udviklingspolitik:

Han slikker r.. på EU, og påstår, at det vil kunne stoppe social dumping  
Han er så hjernelam, at det gør helt ondt...

Forbyd Islam og problemerne med indvandring forsvinder....Hvad skete der med vores retsforbehold....I er dagligt skyld i vold mod danskere med jeres naive tilgang til Islamisk indvandring

Kofoed du er en usling, en plattenslager da du var gruppeformand for Df i retsudvalget, ufatteligt At du ikke er blevet anholdt og spærret ind paa sikret afdeling, du respekterer ikke vore grundlovssikrede rettigheder og vore retssikkerhedslove, Du er en tynd gang havregrød der burde have ti spark i røven,-FØJ

Skulle den lille mongol ikke bare passe sin skole!

Magen til møg fiske skal man lede længe efter

Mutter Merkel er Europas undergang. Hun burde spærres inde eller smides i et mørkt huld

slut, med det løgn tøse diktator trump

Narcissisme, løgn på løgn på løgne, på løgne og atter løgne. Den farligste mand nogensinde næst efter hitler.

Læger uden moralske grænser. Er er en menneskesmugler organisation, de vil aldrig få en krone af mig igen.  
De har jo sejlet tusindvis af økonomiske migranter over Middelhavet og kaldt det at rede dem

Menneskesmugler, væk med ham

Du har solgt Danmarks velfærd til sydeuropa taburetliderlige heks. Men sådan er det vel hvis man vil ned til brüssels senere i livet , så skal man sælge en stump af Danmark. Snablen i kassen og ødslen med andres penge er du god til. fy for fanden.

## Den store verden vækker store følelser

Den internationale scene, storpolitiske forhold og EU er også emner, der kan antænde kommentarsporene.

EU-politiske emner får hug i kommentarsporene, bl.a. retsforbeholdet, som (igen) kobles til indvandring. Vi finder ligeledes nogle angreb om østeuropæere, udenlandsk arbejdskraft og social dumping. Mange af angrebene vedrørende EU retter sig mod politikere. F.eks. "diktatoren Merkel" og danske politikere, der beskæftiger sig med EU. Peter Kofod (DF) kaldes en "usling" og "plattenslager", men Lars Løkke Rasmussen (V) er så "hjernelam, at det gør helt ondt".

Vi finder flere angreb, der drejer sig om storpolitik. Tidl. præsident Trump kaldes f.eks. en "tøse diktator" og "den farligste mand næst efter Hitler". Selvom der er relativt få kommentarer om Tyrkiets præsident Erdogan, er tonen i disse kommentarer meget hård – og kobles ofte til indvandring.

Endelig finder vi en del angreb, der knytter sig til internationale spørgsmål som klimakrise og nødhjælp. Vi ser bl.a. eksempler på nødhjælpsarbejdere og læger uden grænser, der anklages for at være "menneskesmuglere". Også klimaaktivisten Greta Thunberg er genstand for en hel del angreb, som for eksempel: "Skulle den lille mongol ikke bare passe sin skole!"



## Eksempler på angreb om #metoo, kønsidentitet og seksualitet, feminisme og ligestilling:

KÆRE MAGT-LIDERLIGE-KVINDER, I er igang med at brænde alle broerne bag jer. I tilsidesætter naturlige kraftfulde drifter til fordel for jeres eget ego's skyld!! Anerkend nu at mænds (og kvinders) liderlighed og drifter ikke kan styres og kontrolleres af love og paragraffer og af hensigtserklæringer. HVEM FANDEN AF JER NYDER IKKE AT BLIVE KURTISERET AF DET MODSATTE KØN MED ET KÆRLIGT KLAP I RØVEN OG EN FRÆK BEMÆRKNING! (...)

De eneste hekse i hele den debat er [navn], [navn] og [navn]

Hetro sexuelle Skal da også have et flag op... og en parade hvor vi klæ'r os latterligt.. og økonomisk støtte skal vi da også ha... stop nu det Bride pis... Det fylde alt for meget. Og deres flag skal squatte være der... Øh men der er selvfølgelig og mange homoer ansat derinde (...)

Det her blir din politiske død din hygleriske, nosseløse forræder

Hold nu kæft kælling, du er fuld af lort der er ingen der gider at røre dig. er det der for du skaber dig sådan?

STOP. Ikke flere vanvittige kvindemennesker det er til at blive idiot af

Nogle forsvarsadvokater er luderkarle ,du kan betale dem til at mene hvad som helst,og forsvare en hvilken som helst ugerning...(..)

Nej, de krænkelser hystestader skal sgu stoppes nu, deres selvfede egoismen hører ingen steder hjemme. Udstil kvinderne også, så de også bliver stillet til regnskab.

Ekstremistisk feminister og transkønnede er til skyld for dette! Og White knights som ham der! Det fuldstændigt absurd, at børn skal ud i det her! (...) Vi er biologisk set han- eller hunkøn. Al den snak om Nonbinær og rend mig [Navn]! (...) Mænd kan i forvejen ikke finde ud af at være mænd, og når de kan, så de voldtægtsforbrydere. (...) Stop nu, det børnemishandling at fastholde små børn i det her pjat! (...)

## Hekse brændes på bålet i kommentarsporene

Flere af de emner, som popper op på listen, når vi ser på, hvad der især antænder angreb og had i kommentarsporene, handler om køn, ligestilling, seksualitet og feminisme.

#Metoo fylder meget – særligt holdningen om, at #metoo er gået over gevind, og vi ser eksempler på, at kvinder bliver kaldt "hystader", "kællinger" og "vanvittige kvindemennesker". I disse tilfælde er der tale om hadefuldt sprog, fordi angrebene går på køn, som er et beskyttet karakteristika. Angrebene er bl.a. rettet mod nogle af de kvinder, som er stået frem i #metoo-debatten, og mod kvindelige politikere fra forskellige partier, særligt Radikale Venstre.

Vi finder en del angreb rettet mod homoseksuelle, nonbinære og transkønnede. Debatten om, hvorvidt børn skal kunne gennemføre kønsskifte, har især antændt angreb i kommentarsporene.

Der er også angreb på mænd. Ofte benytter disse angreb mere kønsneutralt nedsættende ord som f.eks. "idiot". Desuden ser vi eksempler på, at mænd nedgøres gennem nedsættende ord om kvinder ("tøsedreng", "luderkarl") eller homoseksuelle ("svans", "bøsserøv"). Men der er enkelte eksempler, som går direkte på det mandlige køn, f.eks. "Nosseløs forræder" og "vatpik".



**Hadet er  
koncentrert om  
bestemte  
grupper**

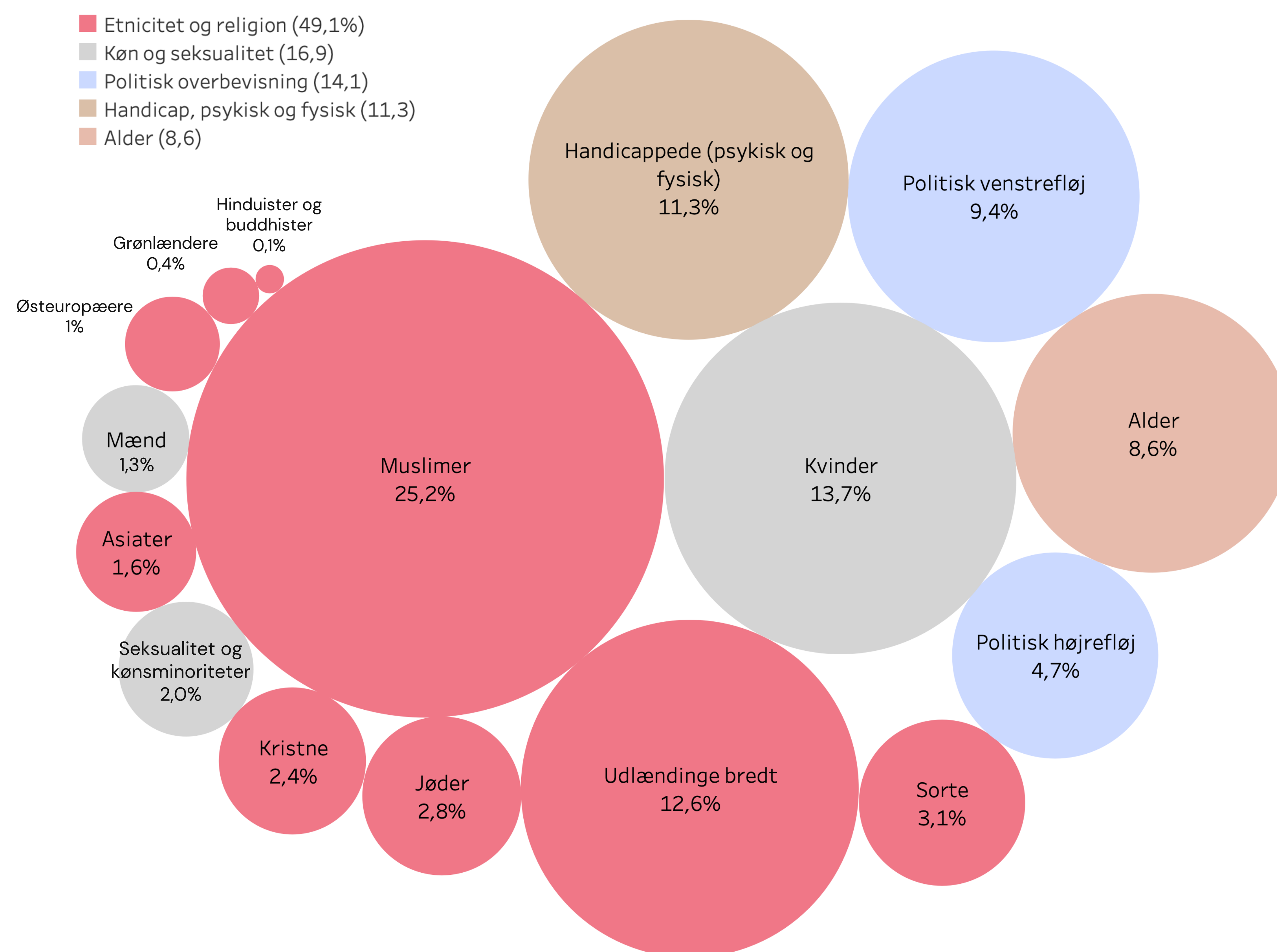


# Hvem er de udsatte grupper i debatten?

I et forsøg på at kvantificere, hvilke samfundsgrupper der er udsat for angreb i den offentlige debat, har vi udviklet en separat søgenøgle. Med søgenøglen forsøger vi at tælle, hvor mange gange forskellige betegnelser for beskyttede karakteristika og beskyttede grupper bliver nævnt i angrebskommentarerne. I alt har vi listet 255 ord, der enten er neutrale (f.eks. homoseksuel, flygtning, liberal) eller nedsættende (svans, stodder, væmligboer) for beskyttede karakteristika. Lidt mere end 600.000 gange bliver betegnelserne for beskyttede karakteristika nævnt i de 3,3 mio. angrebskommentarer.

Opgørelsen giver langt fra det fulde overblik over, hvor mange af angrebene der retter sig mod bestemte beskyttede grupper. Det er f.eks. ikke muligt at identificere gruppetilhørsforhold i kommentarer som "de skulle skydes", ligesom forekomsten af neutrale betegnelser som "ældre" ikke nødvendigvis betyder, at angrebet er rettet mod den pågældende gruppe. Figuren til højre bekræfter vores indtryk af, hvilke beskyttede grupper der oftest bliver genstand for angreb i debatten.

Totalt benævnelser fordelt på beskyttede grupper



**NB:** For mænd og kvinder har vi kun inkluderet nedsættende betegnelser, da neutrale betegnelser gav for mange falske positive. Desuden har vi fra start valgt, at nedsættende brug af kliniske betegnelser som "psykopat" eller fornærmelser som "retarderet" tæller som et angreb mod fysisk og psykisk handicappede. Nedsættende brug af ord som "svans" tæller som et angreb baseret på seksuel orientering. Ord som "olding", "møgunge" og "gammel nar" er eksempler på betegnelser, der indikerer angreb baseret på alder. Angreb baseret på politisk overbevisning tæller neutrale ideologiske begreber, men også nedsættende slang som "nazister" og "kældermennesker" om højrefløjen eller "væmmeligboer", "feltmadras" og "radigale" om venstrefløjen.



**Afsluttende  
bemærkninger**



# Konklusion:

## Nu har vi kortlagt debatten i sin helhed: Vi har et problem

Vi er lykkedes med at udvikle en sprogalgoritme på dansk, der med hidtil uset præcision og genkaldelsesgrad tillader os at kortlægge andelen af anstødelige og hadefulde indlæg i den offentlige debat i sin helhed, forstået som mere end 63 mio. offentlige indlæg over de sidste 2 år.

Kortlægningen viser, at efter moderation fra både Facebook og sideindehaverne udgør angreb mere end 5 pct. af den offentlige debat mellem 2019 og 2021, svarende til 3,3 mio. kommentarer. Heraf er 900.000 af indlæggene hadefulde. Angreb i debatten er altså et alvorligt og udbredt problem. Og når vi ser på angrebene indhold, må vi konstatere, at flere af dem er ganske voldsomme. Det gælder både de anstødelige og de hadefulde kommentarer.

Selvom vores kortlægning med en stikprøve tilbage til 2015 viser et lille fald i andelen af angreb, der får lov at blive stående i debatten, viser andre undersøgelser, at flere danskere i dag end for fem år siden afholder sig fra at deltage i debatten på grund af tonen. Her er det vigtigt igen at nævne, at tonen sagtens kan være blevet hårdere, både i det offentlige kommentarspor og i folks private indbakker. Denne kortlægning medtager ikke kommentarer, som modereres væk, umiddelbart efter de publiceres og inkluderer heller ikke private indlæg. Et angreb når ofte frem til sin modtager på få minutter og bliver i modtagerens bevidsthed længe efter, at det modereres væk. Derudover er der en anden oplagt grund til, at særligt nogle danskere opfatter debatten som meget voldsomt.

## Angrebene fordeler sig ikke ligeligt imellem os

Om vi kigger kvantitativt eller kvalitativt; om vi tæller ord eller begreber; eller om vi forsøger at afdække temaer, så er konklusionen klar: Etniske minoriteter og særligt muslimer er det hyppigste mål for angreb og had. Selv når vi dykker længere ned i andre emner i vores opgørelser over de mest hadefulde debatter, er angrebskommentarerne relateret til udlændinge. Under temaet byggeri og boglig er det debatten om ghettoer, der er hadefuld. Når det handler om kriminalitet, er der en forholdsvis stor debat om kriminelle med anden etnisk baggrund. International politik handler om grænserne, ligesom sikkerhed og terror ofte handler om islamiske organisationer, Syrienskrigere, mv.

Knap 50 pct. af de angrebskommentarer, hvori det er muligt at identificere en betegnelse for beskyttede karakteristika ved hjælp af søgenøgler, er relateret til etnicitet, nationalitet og religion. Det næstmest udsatte beskyttede karakteristika er køn, og her er det især kvinderne, der står for skud.

Derudover fordeler angrebene sig heller ikke lige mellem siderne. Nogle medier og politikere har markant flere angreb på deres sider. Og bevæger vi os ud på politiske fløje og fløjmedierne, stiger andelen af angreb og had. For politikerne er det dog meget forskelligt, i hvilken grad angrebene er "med" eller "mod" dem.

# Konklusion fortsat:

## Nye metoder skal hjælpe os med at værne om det digitale medborgerskab

Formålet med at kortlægge og analysere Danmarks Digitale Medborgerhuse er at blive klogere på de tendenser og dynamikker, der hhv. svækker og styrker det digitale medborgerskab. Derudover er formålet at udvikle nogle værktøjer, der gør os bedre i stand til at måle det digitale medborgerskab og sige noget kvalificeret om emnet.

Som vores metodiske gennemgang viser, er algoritmen ikke perfekt, men den leverer hidtil uset gode resultater. Og vigtigst gør den det muligt for os at få et overblik over de enorme datamængder, som i dag udgør den offentlige debat. Når man forsøger at finde 3,3 mio. nåle i en høstak af 63 mio. strå, er det godt at have automatiseret hjælp. Vi opfordrer ikke til, at hverken medier eller politikere fyrer deres sociale medie-managere og hyrer en algoritme i stedet. Men vi tror på, at det stærkeste værn mod udemokratiske strømninger i vores digitale fællesskaber skal bygges gennem et samarbejde mellem mennesker og teknologi.

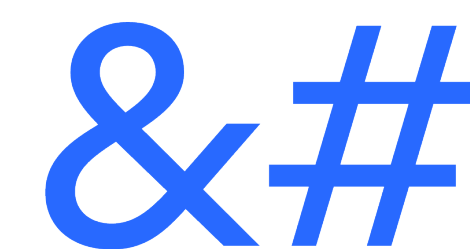
Som samfundsengagerede aktører kan vi heller ikke lade være med at spekulere på, hvor god en algoritme man kunne bygge til moderatører, hvis man som Facebook har det samlede modererede indhold tilgængeligt at træne sine modeller på.

Vores algoritmen er god og "intelligent". Den har fået en god forståelse for angrebssemantik og fanger forskelligartede indlæg fulde af stavefejl, slang og angreb, som vi aldrig har præsenteret den for.

Algoritmen er primært blevet så god, fordi der har været ressourcer og tid til at træne den på et hidtil uset stort annoteret datasæt. Dette har kun været muligt på grund af den seriøse opbakning fra TrygFonden, som skal have stor tak for at prioritere den digitale tryghed som indsatsområde.

I kortlægningen af angreb og hadtale har vi stået på skuldrene af mange andres undersøgelser, definitioner og erfaringer. I projektets anden fase glæder vi os til at rette blikket mod al den anerkendelse, der også bliver afsendt og modtaget i den offentlige debat på Facebook.

Stay tuned.

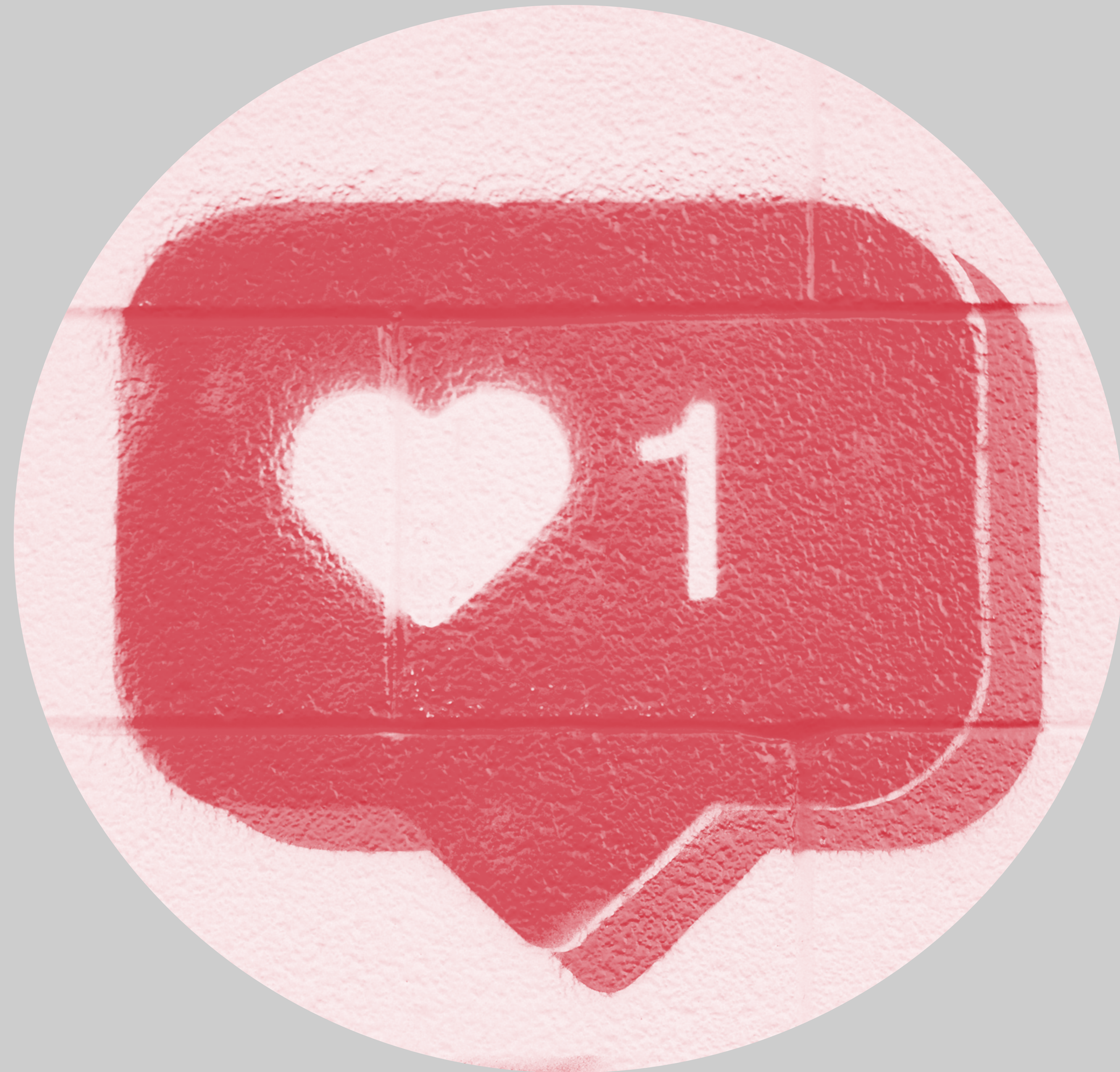


Analyse & Tal

Med støtte fra  
**TrygFonden**



# En stor tak til:



## **Projektets annotører:**

Ida Marcher Manøe  
Julie Enevoldsen  
Nikolaj Meldgaard Christensen  
Naja Bau Nielsen

## **Projektets advisory board:**

Andreas Birckbak, Associate Professor, TANTlab, AAU  
Bolette Sandford Pedersen, Professor, Center for Sprogteknologi, KU  
Leon Derczynski, Associate Professor, Computer Science, ITU  
Marianne Rathje, Seniorforsker, Dansk Sprognævn  
Michael Bang Petersen, Professor, Institut for Statskundskab, AU  
Rasmus Rønlev, Adjunkt, Center for Journalistik, SDU

## **Vores samarbejdspartnere hos TrygFonden:**

Anders Hede, Forskningschef  
Christoffer Elbrønd, Projektchef  
Christian Nørr, Dokumentarist  
Peter Pilegaard Hansen, Projektchef

## **Open-source teknologi vi står på skuldrene af:**

The Danish Gigaword Project; Leon Derczynski, Manuel R. Ciosici  
Ælectra; Malte Højmark-Bertelsen

## **Holdet hos Analyse & Tal:**

Daniel Juhász Vigild, Praktikant  
Edin Lind Ikanovic, Partner  
Jakob Bæk Kristensen, Partner  
Mikkeline Thomsen, Partner  
Ronnie Brandt Taarnborg, Partner  
Sarah Steinitz, Partner  
Tobias Bornakke, Partner





**BILAG**



# Referencer:

1. Facebook Audience Insights, danskerne: <https://www.Facebook.com/ads/audience-insights/people?act=54864242&age=18-&country=DK>
2. Facebook Audience Insights, aktivitet: <https://www.Facebook.com/ads/audience-insights/activity?act=54864242&age=18-&country=DK>
3. Slots og Kulturstyrelsen (2020) Nyheder, baggrund og breaking news – Hvilken rolle spiller de sociale medier i danskernes nyhedsforbrug, <https://mediernesudvikling.kum.dk/index.php?id=39274>
4. Astrid Skov Andersen & Maja Langberg (2021): "Nogle personer tror, at de gør verden til et bedre sted ved at sende hadbeskeder, siger ekspert". TV 2. <https://nyheder.tv2.dk/samfund/2021-03-13-nogle-personer-tror-at-de-goer-verden-til-et-bedre-sted-ved-at-sende-hadbeskeder>
5. Institut for Menneskerettigheder (2017): "Hadefulde ytringer på Facebook". <https://menneskeret.dk/udgivelser/hadefulde-ytringer-paa-Facebook>
6. Institut for Menneskerettigheder (2017): "Hadefulde ytringer på Facebook". <https://menneskeret.dk/udgivelser/hadefulde-ytringer-paa-Facebook>
7. Analyse & Tal og Common Consultancy (2020): "Digital Ytringsfrihed og sociale medier – en undersøgelse af hvilke kommentarer der bliver slettet fra fem danske nyhedsmediers Facebooksider". <https://strapi.ogtal.dk/uploads/daaa0500998540ca9e1eea5bd3c71128.pdf>
8. Laurits Hjortkjær Henningsen (2021): "Voldtægtskultur og kvindehad på Facebook". Politiken. <https://politiken.dk/kultur/medier/art8139946/Voldt%C3%A6gtskultur-p%C3%A5-Facebook-Danske-m%C3%A6nd-dyrker-kvindehad-i-private-grupper>
9. Mark Zuckerberg (2021): "A Blueprint for Content Governance and Enforcement". Facebook blogpost. <https://www.Facebook.com/notes/751449002072082/>
10. Valerio Basile, Cristina Bosco, Elisabetta Fersini, Debora Nozza, Viviana Patti, Francisco Rangel, Paolo Rosso, Manuela Sanguinetti (2019) SemEval-2019 Task 5: Multilingual Detection of Hate Speech Against Immigrants and Women in Twitter, <https://www.aclweb.org/anthology/S19-2007.pdf>
11. Gudbjartur Ingi Sigurbergsson, Leon Derczynski (2019) Offensive Language and Hate Speech Detection for Danish, <https://arxiv.org/abs/1908.04531>
12. Marcos Zampieri, Shervin Malmasi, Preslav Nakov, Sara Rosenthal, Noura Farra, Ritesh Kumar (2019) SemEval-2019 Task 6: Identifying and Categorizing Offensive Language in Social Media (OffensEval)
13. Marcos Zampieri, Preslav Nakov, Sara Rosenthal, Pepa Atanasova, Georgi Karadzhov, Hamdy Mubarak, Leon Derczynski, Zeses Pitenis, Çağrı Çöltekin (2020) SemEval-2020 Task 12: Multilingual Offensive Language Identification in Social Media (OffensEval 2020), <https://arxiv.org/pdf/2006.07235.pdf>
14. Alexandra, B., & Ralf, S. (2009). Rethinking Sentiment Analysis in the News: from Theory to Practice and back.
15. Al Kuwatly, H., Wich, M., & Groh, G. (2020). Identifying and Measuring Annotator Bias Based on Annotators' Demographic Characteristics. *Proceedings of the Fourth Workshop on Online Abuse and Harms*, 184–190. <https://doi.org/10.18653/v1/2020.alw-1.21>

# Bilag: Hvor gode er vores algoritmer (nørdudgaven)



## Detektion af angreb vs. Ikke angreb:

Precision score:	0.8389
Recall score:	0.8295
F1 score:	0.8341



## Kategorisering af angreb i anstødeligt sprog vs. hadtale:

Precision score:	0.8198
Recall score:	0.8145
F1 score:	0.8171



Analyse & Tal  
Analyse & Tal  
Analyse & Tal  
Analyse & Tal  
Analyse & Tal

# Analyse & Tal f.m.b.a.

Analyse & Tal er et digitalt analysebureau med kontorer i København, Århus og Oslo. Vi er specialiserede i at analysere dét, der kan være svært at analysere. Vi skaber nye indsigter ved at kombinere klassiske metoder med nye digitale.

Analyse & Tal har eksisteret siden 2014 og tæller i dag 18 faste medarbejdere. Vores drøm er at skabe et mere demokratisk og retfærdigt samfund. Derfor har vi valgt at organisere os som et medarbejderejet kooperativ.

Vi er stolte af at investere vores overskud i udviklingen af nye metoder, projekter og i demokratiseringen af vores samfund som helhed.