Volume 4, Number 1, Spring 2020

Indholdspoliti

Den private udøvende magt på internettet

Emma Aagaard Helt[*]

Online content is constantly moderated by commercial entities. Now, The European Commission encourage an even stronger moderation in close collaboration between online service providers and so-called trusted flaggers – third parties with specialized expertise in identifying illegal content on the platforms. This article initially examines who these so-called trusted flaggers are, and then addresses the legal challenges spawned by EU's choice to privatize law enforcement on the internet. The author claims, that the use of trusted flaggers is problematic for numerous reasons. For one, the Danish Public Administration Act prohibits the delegation of any decisive authority related to citizens, to a private operator. Private operators may prepare, analyze, and advice. But the decisions must made by the public authority itself. Secondly, trusted flaggers enjoy a high degree of contractual freedom, and are not subject to regulation and control the way traditional law enforcement agencies, such as the police, are. That way, they themselves become sole responsible for balancing our most fundamental rights. Also, this privatization may fuel national differences within the EU, and ultimately challenge a unified EU strategy. Finally, there is a tangible risk that close collaboration between online service providers and trusted flaggers will lead to excessive moderation of online content. Moderation based on company policies and commercial interests; to the extent that freedom of speech is restricted. That way, the internet becomes subject to censorship. The author concludes that the European Commission has set an unclear non-binding recommendation, with potentially far reaching consequences.

[*] Jurastuderende, Københavns Universitet [emmahelt@gmail.com]

Introduktion

Du bruger din frokostpause på Facebook. Pludselig fanges dine øjne af et opslag, som er delt mere end 20.000 gange. Du klikker og læser. Opslaget hænger en navngiven person ud med vrede udtalelser, skældsord og trusler. Du liker og deler det med en kollega. Inden du ved af det, har historien spredt sig til hele kontoret. I forarges. Både du og kollegaerne har læst de hadefulde ord og deler en helt klar opfattelse af, hvad synderen er for en person. I virkeligheden har I kun hørt den halve sandhed. I ved faktisk ikke, hvad der er sandt og falsk, men tager de skrevne ord for gode varer – og scroller så videre.

Dette er blot ét eksempel på de digitale udfordringer, der følger med vores øgede anvendelse af internettet. Internettet der i dag anvendes af halvdelen af verdens befolkning1 og i stadig flere dele af vores hverdagsliv – både privat og på arbejde. De digitale medier er blevet en integreret del af medielandskabet og vores hverdag. Og Facebook er nærmest blevet vores fælles digitale forsamlingshus, hvor alle ytringer kan fremsættes, deles og spredes. På den ene side har internettet givet alle en megafon til offentlighedens ører og dermed styrket ytringsfriheden. På den anden side oplever vi alle, at tonen er blevet hård og nedladende. Det er langt lettere at ytre sig hårdt, hadsk og urigtigt bag skærmens anonymitet – og dermed også konsekvensløst. Det skaber usikkerhed og mistillid i befolkningen, hvilket på sigt kan få store konsekvenser. Men hvordan stopper man deling af hadsk, urigtigt eller ulovligt indhold på et globalt medie som internettet?

I forsøget på at give et svar på det spørgsmål og imødegå problematikken, udsendte EU-kommissionen i september 2017 en meddelelse om bekæmpelse af ulovligt indhold på nettet.[2] Meddelelsen, der er fulgt op af Henstilling (EU) 2018/334,[3] optegner retningslinjer og principper for, hvordan man kan intensivere kampen mod ulovligt indhold, herunder indhold der tilskynder til

[1] Henrik og Helle Stub, 'Snart bliver internettet for hele verden', *Videnskab.dk* (15. januar 2018), <https://videnskab.dk/teknologi-innovation/snart-bliver-internettet-for-hele-verden> besøgt den 30. august 2019.

[2] Commission, 'Tackling Illegal Content Online' (Communication) COM(2017) 555 final (28. september 2017).

[3] Commission Recommendation (EU) 2018/334 on measures to effectively tackle illegal content online [2018] OJ L63/50.

had – også kaldet "hate speech". For at komme problemet til livs, indeholder meddelelsen og henstillingen vigtige nyskabelser, idet de begge opfordrer til at give tredjeparter – såkaldte troværdige meddelere – en privilegeret mulighed for at markere, vurdere og indberette ulovligt, digitalt indhold. I retsakterne opfordrer Kommissionen både onlineplatforme og leverandører af hostingtjenester til at arbejde tættere sammen med nationale, kompetente myndigheder og såkaldt troværdige meddelere – "trusted flaggers", som er offentlige eller private specialiserede enheder med ekspertviden om, hvad der udgør ulovligt indhold. Disse enheder skal vurdere, hvornår noget falder under kategorien "ulovligt indhold" og derefter træffe beslutninger om mulige foranstaltninger. En opgave, som oprindeligt varetages af offentlige myndigheder, men som løses af platformene selv – nu med hjælp fra trusted flaggers. Dermed har opgaven bevæget sig endnu et led væk fra de offentlige myndigheder.

Det rejser imidlertid en række retlige spørgsmål, når vigtige offentlige arbejdsopgaver privatiseres – i dette tilfælde ved at lade private aktører overtage den initierende del af retshåndhævelsen. Det er disse spørgsmål, som undersøges kritisk i nærværende artikel. Først og fremmest rejser sig spørgsmål af materielt retlig karakter i forhold til, hvem der besidder den rette kompetence til at kunne varetage opgaver på vegne af offentlige myndigheder og onlineplatforme. Desuden opstår med trusted flaggers en række mere generelle retlige spørgsmål. Det synes for det første problematisk, at håndhævelsen bliver privatiseret, fordi den derved falder uden for de offentlige administrative regler. For det andet er det problematisk, ud fra hensynet om legitimitet og gennemsigtighed, at trusted flaggers' arbejde baseres på en privatretslig aftale (frivillig eller kontraktlig) mellem dem og onlineplatformene/hostingstjenesteyderne. For det tredje synes det problematisk, at private aktører kommer til at agere en slags politi – eller anklagemyndighed – i komplicerede sager, når lovgivningen på de områder ofte ligger i et retligt krydsfelt mellem nationale bestemmelser og grundlæggende menneskerettigheder. Afgørelserne beror på et skøn. Et skøn, som meget vel kan blive pejlemærke for håndtering af fremtidige sager af lignende karakter.

Trusted flaggers tillægges i givet fald et stort medansvar for at håndhæve og ikke mindst "udfylde" lovgivningen på de digitale platforme, og de bliver af den grund nemt et decideret indholdspoliti på internettet. Traditionelt set tilhører

dette ansvarsområde politiet, der som en del af den udøvende magt, har til opgave at føre kontrol med, at lovene overholdes og skride ind over for lovovertrædelser ved efterforskning og retsforfølgning. EU-kommissionen understreger i meddelelsen, at de privatejede onlineplatforme spiller en central rolle i samfundet, idet de giver brugerne adgang til væsentlig information. Af den grund bærer de et betydeligt samfundsmæssigt ansvar for at sikre, at de ikke bliver fora for eksempelvis hadefulde ytringer mod individer. Endvidere skal de sikre, at de samme rettigheder, som folk har offline også gør sig gældende online – nøjagtigt som De Forenede Nationers Menneskerettighedsråds slog fast i 2016.[4] For at sikre den hurtigste og mest effektive håndhævelse, opfordrer EU-Kommissionen derfor til, at en del af ansvarsområdet uddelegeres til private aktører – de såkaldte trusted flaggers.

Formålet med denne artikel er at bidrage med en kritisk analyse af opfordringen fra EU, og undersøge hvilke problematikker der ligger i brugen af trusted flaggers. Artiklen kan kritiseres for at "male fanden på væggen", da den alene behandler ikke-bindende lovgivning. Men tværtimod. Det kan nemlig være startskuddet til en langt mere omfattende regulering af området. Vi er – uden tvivl – udfordret af ulovligt indhold der deles på internettet og har derfor et behov for at komme problemet til livs. Omvendt er det et principielt problem, at EU skaber retningslinjer for brugen af internettet, som risikerer at undergrave vores europæiske grundværdier og den digitale "frihed". Desuden ligger der et problem i at lade private institutioner stå for implementering og efterfølgelse af disse retningslinjer.

Artiklen er struktureret på følgende måde: I afsnit 2 redegøres for begrebet "hate speech". I afsnit 3 beskrives, hvilken retlig betydning en meddelelse og henstillingen fra EU-kommissionen kan have i dansk ret. I afsnit 4 redegøres der for det nye indholdspoliti, herunder hvem de er, og hvad de er. I afsnit 5 gennemgås de væsentligste problemstillinger. I afsnit 5.1 undersøges det ud fra et forvaltningsretligt synspunkt, hvilke konsekvenser det har, når retshåndhævelsen privatiseres. I afsnit 5.2 undersøges, hvilket regelsæt der gælder for trusted flaggers. I afsnit 5.3 behandles problemstillingerne omkring national

[4] United Nations, General Assembly, 'The promotion, protection and enjoyment of human rights on the Internet', 27. juni 2016, A/HRC/32/L.20.

lovgivning og den menneskeretlige kurs i EU. Afsnittet undersøger – særligt med fokus på hate speech – hvordan ytringsfriheden fortolkes på national plan, da det er centralt for at forstå, om trusted flaggers harmonerer med idéen om en fælles vision i EU. I afsnit 5.4 præsenteres begrebet "virksomhedscensur". Her forsøges det at overføre forbuddet mod censur i grundloven fra offentlig- til privatretlig regi. Derigennem vil jeg problematisere og diskutere, at indholdspolitiet (private aktører) handler ud fra egne normer, egne fortolkninger af lovgivningen og egeninteresser. I afsnit 5.5 behandles begrebet retssikkerhed. Slutteligt behandles bekymringen for en (mulig) endnu større regulering kort i afsnit 5.6.

1. Teori og metode

Forslaget fra EU-kommissionen blev i september 2017 fremlagt i en meddelelse om bekæmpelse af ulovligt indhold på nettet[5] (og desuden præsenteret i en pressemeddelelse)[6] og efterfølgende fulgt af Henstilling (EU) 2018/334 fra marts 2018 om foranstaltninger til effektiv bekæmpelse af ulovligt indhold på internettet.[7] Der er ikke udarbejdet specifikke retningslinjer for, hvem de private aktører kan være, hvorfor der heller ikke sket en udvælgelse af trusted flaggers. Af den grund er både viden og litteraturen om opfordringen naturligvis sparsom. Kun en enkelt artikel[8] – der fortsat er under udarbejdelse – fra november 2018, er kommet denne artikel forkøbet, men ellers har det ikke været muligt at finde hverken forskningsbaseret materiale eller andre artikler, som behandler nærværende problemstilling.

[5] Meddelelsen (n 3).

[6] Europa-Kommissionen, 'Sikkerhedsunionen: Kommissionen skærper indsatsen i kampen mod ulovligt indhold på internettet' (Pressemeddelelse, 28. september 2017) <https://europa.eu/rapid/press-release_IP-17-3493_en.htm> besøgt den 20. august 2019.

[7] Henstilling (n 4).

[8] Sebastian Felix Schwemer, 'Trusted notifiers and the privatization of online enforcement', (University of Copenhagen and Danish Internet Forum (DIFO), november 2018).

Artiklen kaster således et helt nyt og første lys på, hvilke problematikker en større udbredelse[9] af trusted flaggers kan få på komplekse juridiske områder som disse. Specielt set i en dansk kontekst. Af hensyn til artiklens omfang har det kun været muligt at behandle én af de utallige digitale udfordringer, der eksisterer, og som retskilderne fra EU søger at imødegå: hate speech. Når problemstillingerne med trusted flaggers eksemplificeres, vil det derfor være med fokus på den gråzone, der opstår mellem informations- og ytringsfriheden og strafbare udtalelser, hvor fænomenet "hate speech" ligger placeret. Området er valgt, da der, modsat digitale udfordringer som deling af terrorvideoer og børnepornografisk materiale, ikke ligger en klar linje for håndtering af denne type digital udfordring. Artiklen vil endvidere koncentrere sig om de tilfælde, hvor arbejdet som trusted flaggers tildeles private aktører, og hvor arbejdet består i at fjerne hele eller dele af hjemmesider. Artiklen behandler ikke de tilfælde, hvor internetdomæner lukkes.

2. Hate speech

Fænomenet hadeytringer eller hate speech har de facto altid eksisteret. Går man knapt hundrede år tilbage i historien, er antisemitismen[10] måske det mest åbenlyse eksempel herpå. Selve termen "hate speech" er af nyere dato og er i dag karakteriseret ved at være mundtlige eller skriftlige udsagn eller i en vis udstrækning kropssprog,[11] der angriber, truer eller fornærmer en person eller en gruppe. Fænomenet optræder på mange forskellige måder, og de nedladende ytringerne kan fremsættes på baggrund af en lang række årsager, herunder hudfarve, seksualitet, køn, religion, etnisk oprindelse, handicap, politiske overbevisning, moralske opfattelse, beskæftigelse eller lignende.

---

[9] European Commission, 'EU Internet Forum: Bringing together governments, Europol and technology companies to counter terrorist content and hate speech online' (Pressemeddelelse, 3. december 2015), IP/15/6243.

[10] Antisemitisme var kendt før nazisternes tid, men jødehadet var en stor del af Hitlers ideologi, da han overtog magten i 1933.

[11] Rigsadvokaten Informerer, nr 18/2006 (senest opdateret 23/10-08) 11, vedrørende bødeforelæg til en person (overtrædelse af straffelovens § 266 b), der var fremkommet med »abelyde« på et værtshus.

Hate speech kan enten være rettet mod enkeltpersoner eller hele grupper. Fx på Facebook, hvor der chikaneres eller diskrimineres i kommentarfeltet til et opslag – både i lukkede grupper eller på egne profiler. Det kan være ytringer, der opfordrer til forbrydelse, jf. TfK 2017.725. Det kan være systematisk mobning af skolebørn, også kaldet trolling. Det kan være falske oprettede profiler på de sociale medier. Listen over eksempler er blevet uendelig, og fænomenet er stærkt stigende. Meget tyder på, at internettet er en væsentligt katalysator hertil.[12]

I 2016 fremlagde EU-Kommissionen og fire store IT-virksomheder, Facebook, Microsoft, Twitter og YouTube derfor adfærdskodekset "Code of conduct on countering illegal hate speech online",[13] som en reaktion på den store udbredelse af hadtale på onlineplatforme. Adfærdskodekset signalerer et kollektivt ansvar for at fremme den digitale ytringsfrihed, og har desuden til formål at hjælpe brugere med at anmelde ulovligt hadesprog på de sociale platforme. Parterne blev enige om, at platformene skal vurdere de fleste brugeranmeldelser indenfor 24 timer, overholde både EU's og national lovgivning om hadtale samt fjerne det ulovlige indhold, såfremt det overtræder de nævnte regler.

Adfærdskodekset er et eksempel på, at der i dag allerede findes systemer og modeller for fjernelse af ulovligt indhold, hvor private parter vurderer lovligheden af det indhold, som brugerne anmelder. EU-Kommissionens meddelelse om bekæmpelse af ulovligt indhold på nettet fra september 2017 og henstillingen fra marts 2018 om foranstaltninger til bekæmpelse ulovligt indhold på online platforme skal ses i forlængelse af den strategi, som gav anledning til adfærdskodekset.

Hate speech, som retlig term, er imidlertid ukendt i dansk strafferet. Da fænomenet omfatter forskellige typer af udsagn, har flere af straffelovens bestemmelser betydning for, hvilke indgreb der kan iværksættes overfor den

---

[12] Vincent F. Hendricks, 'Tillid i en digital virkelighed', i Henriette Korf, Anna Vibe Onsberg Hansen, Anders Young Rasmussen og Merete Arentoft (red), *Når forbrydelser bliver digitale: En antologi om IT-kriminalitet og adfærd på internettet* (Det Kriminalpræventive Råd 2016) 74–77.

[13] European Commission, 'Code of conduct on countering illegal hate speech online' (30. juni 2016) <https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combatting-discrimination/racism-and-xenophobia/countering-illegal-hate-speech-online_en> besøgt den 1. september 2019.

uønskede kommunikation. Overfor ytring- og informationsfriheden i EMRK[14] artikel 10 står straffelovens § 266 b om racistiske udtalelser og reglerne om freds- og ærekrænkelser i straffelovens kapitel 27, foruden § 136 om offentlig tilskyndelse til forbrydelse. Fænomenets brede definition og vage placering i strafferetten, gør det svært at komme problemet til livs på national såvel som international plan. Vurderingen af, hvornår ytringer betegnes som hadefulde og ulovlige, afhænger af den betydning, som selve ytringen og ytringsfriheden tillægges i den nationale retskultur.[15]

Hate speech er med andre ord et diffust begreb, som modsat eksempelvis ulovlig deling af terrorvideoer og børnepornografisk materiale, lægger op til komplekse, juridiske afvejninger, der virker bekymrende at skulle overlade til private tredjeparter. Disse bekymringer behandles senere i nærværende artikel.

3. EU-retskilders betydning i dansk ret

Henstillinger og udtalelser er ikke-bindende retsakter (sekundære retsregler), jf. traktaten om den Europæiske Unions Funktionsmådes (TEUF) artikel 288, stk. 5, som EU ifølge traktaten om den Europæiske Unions (TEU) art. 5, stk. 1 og 2, har kompetence til at udstede. EU-domstolen har dog fastslået i sin praksis, at begge typer retsregler kan have nogle retlige virkninger. Domstolen fastslog eksempelvis i sagen *C-317-320-08, Alassini,* at selvom henstillinger og udtalelser ikke er rettigheder, som private kan påberåbe sig ved en national domstol, "kan de dog ikke anses for at være ganske uden retsvirkninger". Det betyder, at nationale retsinstanser, ifølge EU-domstolen, skal "tage hensyn til henstillingerne", eksempelvis hvis de har til formål at udfylde bindende EU-retlige bestemmelser, jf. præmis 40 i afgørelsen. Det samme gør sig gældende med meddelelser fra EU-kommissionen, som trods deres ikke-bindende karakter, ifølge Karsten Engsig Sørensen og Jens Hartig Danielsen,[16] alligevel i

[14] Council of Europe, Convention for the Protection of Human Rights and Fundamental Freedoms, 1950, ratificeret ved Lov nr 285 af 29. april 1992 (Den Europæiske Menneskerettighedskonvention).

[15] Trine Braumbach og Peter Blume, 'Had på nettet – en retlig udfordring' [2009] Juristen 7.

[16] Karsten Engsig Sørensen og Jens Hartig Danielsen, *EU-retten* (7. udg., Jurist- og Økonomforbundets Forlag 2019) 111–112.

et eller andet omfang må binde den institution, som har udstedt retsakten. Det skaber ifølge forfatterne en berettiget forventning om, at den pågældende institution i praksis vil anvende samme fortolkning, når den håndhæver reglerne.

Konkret for artiklens omdrejningspunkt vil det betyde, at Kommissionens meddelelse og henstilling kan og typisk vil udgøre et væsentlig inspirationsgrundlag, hvis der i medlemsstaterne skal udarbejdes retsakter på området. Endvidere spiller det en rolle i forhold til, hvordan virksomheder sikrer ordentlig adfærd og håndhævelse af lovene på deres platforme.[17] Det er af den grund interessant at undersøge, hvilke problematikker opfordringen fra EU har. For selvom de nævnte EU-retskilder i sin rene form ikke udgør bindende lovgivning, kan de at få stor betydning hos de privatejede platforme.

4. Trusted flaggers – det nye indholdspoliti

Et større fokus på god onlinekultur kommer – foruden af ovennævnte adfærdskodeks – også af den strategi, der blev lagt for Kommissionen i 2014. Europa-kommissionens snarligt aftrædende formand Jean-Claude Juncker fremhævede på daværende tidspunkt[18] "det digitale indre marked", som én af de i alt ti politiske prioriteringer for årene 2015-2019. Fokus i denne periode er at skabe et erhvervsklima, som tilpasses de digitale muligheder. Målet er at skabe en bedre og mere sikker onlinehandel samt gøre digitaliseringen til drivkraft for vækst i unionen. Der er et ønske i EU, om at digitaliseringen udnyttes fuldt ud, men samtidig at de kriminelle konsekvenser af teknologiens udvikling stoppes gennem et samarbejde på tværs af geografiske grænser.

Konsekvenserne er som nævnt allerede ved at være legio,[19] og derfor mener EU, at onlineplatformene, som den primære adgangskilde til digitalt indhold (for de fleste internetbrugere) bærer et betydeligt samfundsmæssigt ansvar med hensyn til at beskytte brugerne og samfundet som helhed. Onlineplatforme formidler adgang til digitalt indhold, eksempelvis gennem hostingtjenester, som

[17] Det fremgår ligeledes af EU-kommissionens meddelelse fra september 2017.
[18] Daværende formand for Europa-kommissionen, Jean-Claude Junckers, politiske retningslinjer, 15. juli 2014.
[19] Meddelelsen (n 3) 2.

giver mulighed for upload af indhold fra tredjeparter.[20] Disse hostingtjenester kan eksempelvis være onlinemarkedspladser, videodelingsplatforme, sociale netværk og bloggingwebsteder.[21] I EU-kommissionen meddelelse fra september 2017[22] opfordres onlineplatformene til at forhindre, at deres infrastruktur og virksomhed anvendes til at udøve krænkende aktiviteter og forbrydelser (gennem onlinetjenester). Kommissionen lægger op til større selvregulering og opstiller i meddelelsen ikke-bindende retningslinjer og principper for, hvordan onlineplatformene i samarbejde med medlemsstaterne, de nationale myndigheder og andre relevante interessenter kan intensivere kampen mod ulovligt indhold.

Kommissionens vigtigste forslag er, at platformene 1) opretter automatiske værktøjer, der kan forhindre, at tidligere fjernet indhold optræder igen, 2) opretter specifikke tidsrammer for fjernelse af indholdet, 3) skaber automatiske påvisningsteknologier, der kan indberette ulovligt indhold samt 4) opbygger et tættere samarbejde med betroede meddelere – "trusted flaggers". Denne artikel har fokus på sidstnævnte.

4.1 Henstilling (EU) 2018/334

Kommissionen lovede i meddelelsen at overvåge fremskridtene og vurdere, om der er behov for yderligere foranstaltninger for at sikre hurtig og proaktiv påvisning og fjernelse af ulovligt indhold på nettet, herunder mulige lovgivningsmæssige foranstaltninger.[23] 1. marts 2018 fulgte Kommissionen op med Henstillingen (EU) 2018/334 om foranstaltninger til effektiv bekæmpelse af ulovligt indhold på nettet. Henstillingen indeholder operationelle foranstaltninger og beskyttelsesforanstaltninger, som virksomheder og medlemsstater skal træffe for i endnu højere grad at bekæmpe ulovligt indhold, der i henstillingens pkt. 14 er defineret som "enhver form for indhold, som ikke

[20] ibid 5.
[21] ibid.
[22] ibid 3.
[23] European Commission, Fact Sheet, 'Stepping up the EU's efforts to tackle illegal content online' (1. marts 2018) <https://europa.eu/rapid/press-release_MEMO-17-3522_en.htm> besøgt den 29. august 2019.

er i overensstemmelse med EU-retten eller med medlemsstaternes lovgivning, uanset lovgivningens art eller særlige genstand".

Henstillingen søger at udmønte de politiske retningslinjer, som blev fremlagt af Kommissionen i meddelelsen fra september 2017. De søger at præcisere de (forskellige) mekanismer til indberetning af ulovligt indhold, der skal indføres, samt hvordan anmeldelser af lovligt indhold skal behandles (sikkerhedsforanstaltninger). Men hvor meddelelsen var rettet mod onlineplatformene, retter henstillingen sig i stedet mod medlemsstaterne og udbyderne af hostingtjenesterne. Kommissionen udvider således deres løsninger for, hvordan der bedst holdes justits på de digitale medier, ved nu at tildele hostingtjenesterne en del af ansvaret. Kommissionen underbygger denne udvidelse ved at fastslå, at "hostingtjenesteyderne spiller en særlig vigtig rolle i bekæmpelse af ulovligt indhold på nettet, da de lagrer oplysninger, der leveres af deres brugere og på disses anmodning…"[24] Henstillingen fastslår desuden, at såfremt det er hensigtsmæssigt, kan henstillingen også anvendes på andre berørte udbydere af internettet.[25] Der er i henstillingen ikke blevet uddybet, eksemplificeret eller lavet vejledning til, hvem disse "hostingtjenesteydere" er (foruden nedenstående i afsnit 4.2). Det efterlader en usikkerhed for medlemsstaterne og de berørte tjenester, der selv får til opgave at definere ansvarsfordelingen ud fra skøn.

Hostingtjenesterne tilskyndes imidlertid til at anvende automatiske metoder til identificering og fjernelse af ulovligt indhold, til at træffe en række gennemsigtighedsforanstaltninger og særligt at arbejde sammen indbyrdes med medlemsstaterne og med trusted flaggers, hvilket ifølge Kommissionen vil bidrage til hurtigere bekæmpelse af ulovligt indhold.[26]

4.2 Karakteristik af indholdspolitiet

Trusted flaggers eller troværdige meddeleler er personer eller enheder, herunder ikke-statslige organisationer og erhvervsorganisationer, med særlig ekspertise og et ønske om på frivillig basis at påtage sig et vist ansvar i forbindelse med

[24] Henstilling (n 4), præambelbetragtning nr 15.
[25] ibid.
[26] Henstilling (n 4), præambelbetragtning nr 25, 29.

bekæmpelse af ulovligt indhold på nettet.[27] De er udpeget specifikt af hostingtjenesteyderne til at bistå med særlig ekspertise og et særligt ansvar med henblik på bekæmpelse af ulovligt indhold på nettet.[28] Det er således op til tjenesteyderne selv at vurdere, hvilke personer eller enheder de anser for egnet til at udføre deres aktiviteter på en omhyggelig og objektiv måde og med respekt for de værdier, som Unionen bygger på, jf. TEU artikel 2. Tjenesteyderne opfordres til at offentliggøre klare og objektive kriterier for bestemmelse af, hvilke personer eller institutioner de anser for troværdige meddelere. Disse kriterier er endnu ikke at finde.

Kommissionen uddyber i henstillingen vigtigheden af, at samarbejdet mellem nationale myndigheder,[29] hostingtjenesteyderne og trusted flaggers styrkes grundet det høje antal anmeldelser om ulovlig aktivitet på internettet. De understreger, at anmeldelser fra meddelerne skal tillægges høj prioritet, og at der skal udvises en passende grad af tillid til deres korrekthed. Det betyder, at tjenesteyderne (i højere grad) bliver forpligtet til skulle overvåge – blandt andet med hjælp fra tredjeparter – deres egne medier og støvsuge dem for ulovligt indhold. Det er på sin vis en ændring af de fritagelsesregler, der følger af E-handelsdirektivet,[30] som ellers har til formål at sikre europæiske udbydere (mellemænd) mod ansvar for den information, der udveksles på deres tjenester.[31] Mellemmænd er fysiske eller juridiske personer, der stiller tjenester til rådighed, som giver internetbrugere mulighed for at udveksle information.[32] På EU-plan sætter E-handelsdirektivet de overordnede retlige rammer for fjernelse af ulovligt indehold, og indeholder vigtige regler om ansvarsfrihed for "mellemmænd".

[27] Henstilling (n 4), præambelbetragtning nr 29.

[28] ibid artikel 1, stk 4, litra g.

[29] ibid artikel 1, stk 4, litra j: "De kompetente myndigheder er udpeget af medlemsstaterne i overensstemmelse med deres nationale lovgivning til at udføre opgaver, som omfatter bekæmpelse af ulovligt indhold på nettet, herunder retshåndhævende myndigheder og forvaltningsmyndigheder, der har til opgave at håndhæve den gældende lovgivning på visse særlige områder, uanset lovgivningens art eller særlige genstand."

[30] Europa-Parlamentets og Rådets direktiv 2000/31/EF af 8. juni 2000 om visse retlige aspekter af informationssamfundstjenester, navnlig elektronisk handel, i det indre marked (Direktivet om elektronisk handel) [2000] EUT L178/1.

[31] Henrik Udsen, *IT-RET* (4. udg, Ex Tuto 2019) 279.

[32] ibid 272.

Ifølge direktivets artikel 14, kan disse "mellemmænd" ikke drages til ansvar for information, der lagres efter anmodning fra tredjemand, forudsat en række betingelser er opfyldt. Det følger desuden af artikel 15, at medlemsstaterne ikke må pålægge "mellemmændene" en forpligtelse til at overvåge den information, der florerer på deres platforme (aktiv overvågning), medmindre de får konkret viden eller bevidsthed om den ulovlige aktivitet. Det er interessante problemstillinger – om direkte modsætninger i EU-retsakterne – som ikke behandles yderligere i nærværende artikel.

Trusted flaggers får til opgave at holde øje i al den information, der florerer på hostingtjenesteydernes informationstjenester. Det indhold, som de mener ikke er i overensstemmelse med EU-retten eller den berørte medlemsstats lovgivning, skal de indberette til hostingtjenesterne med en anmodning om, at tjenesteyderne hurtigt får fjernet eller deaktiveret adgangen til indholdet. Trusted Flaggers kommer dermed til at fungere som indholdspoliti – en slags privatiseret retshåndhæver på internettet. De kommer til at sidde med en lang række opgaver, der vedrører alt fra at mindske udbredelse af terrorrelateret propaganda, over deling af videoer, hvor børn seksuelt bliver misbrugt, til sager om krænkelse af intellektuelle ejendomsrettigheder (IP) og hate speech. De tillægges dermed et stort ansvar for at vurdere, hvad der udgør ulovligt digitalt indhold, og derigennem også udfylde lovens rammer på de mere "uklare" retsområder, som eksempelvis hate speech.

Trusted flaggers har, som beskrevet, ikke en direkte adgang til at fjerne ulovligt indhold, men kommer til at have en væsentligt funktion i den initierende del af retshåndhævelsen. I lyset af trusted flaggers' status om "troværdig meddeler" og Kommissionen anbefaling om, at deres indberetninger bør gives høj prioritet og anerkendes som korrekte,[33] bliver deres anmeldelser afgørende for, hvad der fremover vil blive opfattet som ulovlig aktivitet på internettet, og dermed bør fjernes fra hjemmesider. De sætter altså en form for standard for praksis.

[33] Henstilling (n 4), præambelbetragtning nr 29.

5. Privat myndighedsudøvelse

En myndighed er en samfundsinstitution, som forvalter og som har autoritet til at udøve magt på et bestemt område. Det kan for eksempel være statslige myndigheder, herunder ministerierne, der under en ministers ledelse har ansvaret for at forberede og gennemføre lovgivning. Det kan også være politiet, der som allerede nævnt har ansvaret for føre kontrol med, at lovene overholdes og at skride ind over for lovovertrædelser. Fælles for de offentlige myndigheder er, at de demokratiske organisationer (i hvert fald i Danmark) nyder en generel tillid fra borgerne, og at de har opnået legitimitet, idet deres arbejde udgør en væsentlig del af staten.[34] Niveauet af tillid er helt specielt i Danmark. Seneste tryghedsundersøgelse fra 2016, foretaget af European Social Survey,[35] viser, at danskernes tillid til politiet og domstolene ligger i toppen, når man sammenligner med andre europæiske lande. Vi – som samfund – stoler på vores offentlige myndigheder, og tror i store træk på, at særligt politiet arbejder ud fra en objektiv tilgang.

Denne grundfæstede opfattelse af, at politiet besidder legitimitet, fordi de, som en del af staten, sørger for, at lovene overholdes, er ikke ensbetydende med, at andre aktører, der udøver samme form for arbejde, vil opnå samme legitimitet. Vores tillid bygger på en klar tiltro til, at politiet behandler identiske tilfælde ens og dermed overholder grundprincippet om, at alle er lige for loven.

I EU's fælles forsøg på at intensivere kampen mod ulovligt indhold, herunder hate speech, ligger en forventning fra EU-kommissionen om, at privatejede hjemmesider ikke blot følger deres egne retningslinjer, men ligeledes overholder de offentlige love, herunder Den Europæiske Menneskerettighedskonvention (EMRK).[36] Friheden til selv at vælge indhold er ikke absolut, selvom

---

[34] Justitsministeriets forskningskontor, 'Tryghed og holdning til politi og retssystem' (januar 2016) <http://www.justitsministeriet.dk/sites/default/files/media/Pressemeddelelser/pdf/2015/ESS%2520rapport%25202015%2520%25282%2529.pdf> besøgt den 1. september 2019.

[35] ibid.

[36] Council of Europe, Convention for the Protection of Human Rights and Fundamental Freedoms, 1950, ratificeret ved Lov nr. 285 af 29. april 1992 (Den Europæiske Menneskerettighedskonvention).

hjemmesiderne er private, og det er således ikke tilladt at ignorere lovgivningen. Med argumentet, om at trusted flaggers besidder en større ekspertise indenfor det lovgivningsmæssige område end hostingtjenesteyderne, opfordrer Kommissionen aktørerne til at anvende dette nye indholdspoliti.[37] En udvikling der skal ses i lyset af, at hverken politiet eller anklagemyndigheden selv har ressourcerne til at efterforske og retsforfølge alle dem, som oploader eller deler ulovligt indhold.[38]

Fokus i dette afsnit er at gennemgå nogle af de problemstillinger der ligger i, at trusted flaggers kommer til at blive en privat myndighedsudøver, og hvad der kan tale imod, at disse private enheder er repræsentative i udøvelsen af deres arbejde.

5.1 Forvaltningsret

Den første problemstilling er hele spørgsmålet, om hvorvidt det er tilladt at overlade retshåndhævelse til selvstændige private enheder. Ud fra et forvaltningsretligt synspunkt vil der overordnet set ske ekstern delegation af forvaltningsretlig kompetence, fordi private aktører pålægges en forpligtelse til at håndhæve loven på digitale medier. Retshåndhævelsen delegeres dermed fra politiet til hostingtjenesterne og videre til trusted flaggers. Det er ikke en ukendt tendens i dag,[39] og det er indiskutabelt en nødvendighed at bede de digitale medier og platforme om hjælp, hvis ulovligt indhold hurtigt og effektivt skal fjernes. Ikke desto mindre rokker det med den klassiske magtfordelingslære.

Tager man et kig på forvaltningsloven, eksisterer der ikke en hovedregel om, at delegation kræver hjemmel, men omvendt gælder der heller ikke en hovedregel om, at alle former for ekstern delegation frit kan foretages. Dette skyldes, at delegationsreglerne ikke i alle tilfælde er klare, hvorfor vurderingen af om der er hjemmel til at foretage ekstern delegation, må træffes fra område til område.[40]

Delegation til private af kompetence til at træffe afgørelse i forhold til borgerne, jf. forvaltningslovens § 2, stk. 1, kræver klar og udtrykkelig

[37] Fact Sheet (n 22).
[38] Pressemeddelelsen (n 6).
[39] Fx Code of Conduct (n 9).
[40] Niels Fenger, *Forvaltningsret* (Jurist- og Økonomforbundets Forlag 2019) 141.

lovhjemmel.[41] Dette følger blandt andet af Justitsministeriets udtalelse i 1991, i forbindelse med bevarelse af en række spørgsmål til lovforslag L 135 om ændring af færdselsloven, hvortil de svarede, at kommunerne ikke kunne (videre)delegere kompetencen til at foretage parkeringskontrol til private, idet der var tale om en myndighedsopgave. Derimod antages det i dag, at der ikke gælder et almindeligt krav om lovhjemmel, når det drejer sig om delegation til private i forberedelsen af afgørelsessager.[42] Retsstillingen blev fastslået i FOB 2013-9,[43] hvor Ombudsmanden konkluderede, at dele af sagsforberedelsen i afgørelsessager kunne overlades til private, såfremt en række krav er opfyldt. Det er interessant at dykke ned i denne sag, fordi grænsen mellem, hvornår arbejdet blot består i at udarbejde dele af sagsforberedelsen kontra, hvornår der er tale om at træffe endelig afgørelse, ikke er klar. Trusted flaggers kommer til at blive en del af sagsbehandlingerne i afgørelsessager[44] (særligt den initierende del), men grænsen mellem ovenstående har betydning for, hvordan anmeldelser, leveret af trusted flaggers, skal håndteres og behandles.

FOB 2013-9 vedrørte det daværende Miljøklagenævns aftale (i 2009) med Kammeradvokaten, om bistand til behandlingen af nogle klagesager over kommunale afgørelser om husdyrsbrug. Det fremgår af udtalelsen, at Kammeradvokatens bistand blandt andet bestod i gennemgang af sagens akter og udarbejdelse af udkast til afgørelsen, herunder rådgivning om juridiske spørgsmål. Udkastet sendte de til Miljøanklagenævnet, som foretog de fornødne rettelser. Derefter blev det tilrettede udkast af Kammeradvokaten sendt videre til partshøring, hvorefter Kammeradvokaten selv foretog de fornødne rettelser herefter. Miljøanklagenævnet læste slutteligt udkastet til afgørelsen, og kunne blot underskrive, hvorefter sagen blev afsluttet. Ombudsmanden konkluderede i udtalelsen, at selvom der ikke var hjemmel i lovgivningen, til at Miljøklagenævnet kunne overlade opgaver i forbindelse med behandlingen af

[41] ibid 143.
[42] ibid 143.
[43] FOB 2013-9, 'Lovlig delegation af sagsforberedelse til advokatfirma' (28. juni 2013) <http://www.ombudsmanden.dk/find/udtalelser/beretningssager/alle_bsager/2013-9/dokument/> besøgt den 1. september 2019.
[44] En afgørelse i forvaltningslovens forstand er en retsakt (forvaltningsakt), dvs. en udtalelse, der går ud på at fastsætte, hvad der er eller skal være ret i et foreliggende tilfælde.

konkrete sager på området til private, var det i overensstemmelse med de almindelige forvaltningsretlige principper om delegation af myndighedsudøvelse til private.

Det er vigtigt at understrege, at ovennævnte ikke 1:1 kan overføres på forholdet mellem de delegerende aktører og trusted flaggers. Der er ikke tale om direkte myndighedsudøvelse efter bogens regler; både fordi begge parter er private og fordi, trusted flaggers ikke tager den endelige beslutning, om hvorvidt indholdet skal fjernes. Alligevel kan man med en vis forsigtighed bruge konklusionen i ombudsmandsudtalelsen på ovennævnte forhold.

Arbejdsdelingen mellem hostingtjenesterne og trusted flaggers er endnu ikke blevet klarlagt på området for håndtering af ulovligt, digitalt indhold herunder for håndtering af hate speech. Af den grund er det usikkert, hvilken rolle trusted flaggers får, og hvilken værdi deres anmeldelser tillægges. Det må dog forventes, at denne ikke bliver ubetydelig. Det blev gjort klart i henstillingen, jf. redegørelsen i afsnit 4.2, at anmeldelser fra meddelerne skal gives høj prioritet, og at der skal udvises en passende grad af tillid til deres korrekthed. Ligeledes er hovedargumentet for at anvende trusted flaggers, at deres ekspertise rækker ud over, hvad hostingtjenesteyderne besidder. Det må ud fra konklusionen i ombudsmandsudtalelsen antages, at delegation af sagsbehandlingsskrift, der har væsentligt indflydelse på både forvaltningsprocessen og på forvaltningsaktens indhold, ikke er udelukket. Den vurdering som trusted flaggers foretager af det digitale indhold, og som eventuelt senere bliver indberettet til hostingtjenesteyderne, må således godt have afgørende indflydelse på afgørelsens endelige udfald. Den udvikling må siges at stå i modstrid med udgangspunktet om, at der efter forvaltningslovens § 2, stk. 1, kræves klar og udtrykkelig lovhjemmel, når der delegeres til private af kompetence til at træffe afgørelse i forhold til borgerne. Denne "nye linje"[45] og brede fortolkning af delegationsforbuddet er også at se i nyere praksis. Senest kan nævnes højesteretsdom UfR 2017.3389 H, hvor det ikke fandtes i strid med lovgivningen, at SKAT, i deres afgørelse om afskedigelse af en medarbejder, havde henvist til en privat konsulentvurdering, som SKAT i partshøringsbrevet erklærede sig enig i. Domstolen fandt frem til, at SKAT hverken formelt eller

[45] Michael Gøtze, 'Når offentligt bliver privat' [2018] Juristen 32.

reelt kunne anses for at have delegeret sin kompetence til at afskedige deres medarbejder, selvom deres afgørelser i høj grad byggede på konsulentens vurdering.

Set ud fra synspunktet om effektivisering, er brugen af trusted flaggers – og ekstern delegation i det hele taget – et gode og et led i en effektiv udnyttelse af offentlige midler. Man kan måske gå så langt som at sige, at det er en modernisering af delegationslæren. Fra en formalistisk sort på hvidt-forvaltningsret, hvor det offentlige stod med ansvaret, jf. lovens ordlyd, til en mere nytænkende forvaltningsret, hvor opgaver overlades til dem med den rette – og måske bedste – ekspertise på området. Men selvom denne såkaldte "nye linje" er effektiv, er den ikke uden retssikkerhedsmæssige problemer (behandles yderligere i afsnit 5.5).

Fra en borgers synsvinkel virker det usikkert og uigennemsigtigt, at en del af magtudøvelsen på de digitale medier skal varetages af ikke-synlige og måske ej heller identificerbare private aktører. Aktører, som ikke hører direkte under vores demokratisk kontrollerede myndigheder. Her kan der argumenteres for, at det grundlæggende problem eksisterer såvel når hostingtjenesterne som når trusted flaggers står for retshåndhævelsen. En væsentlighedsvurdering, der ikke synes ubetydelig. For hvem har ansvaret for, at der føres retlig kontrol med aktørerne og delegationsordningerne?

Det er vigtigt at huske på, at delegationen til trusted flaggers består i magtudøvelse – et kerneområde inden for retlig forvaltningsvirksomhed. I forhold til hate speech betyder det delegation af bemyndigelse, til blandt andet at kunne gøre indgreb i den enkelte borgers kernerettighed; ytringsfriheden. Af den grund bør der være en klar fælles opfattelse blandt borgere og staten, om at de aktører, som skal kunne gøre indgreb i en så fundamental rettighed, besidder en stor legitimitet. Deres afgørelser skal opfattes som værende korrekte og ud fra en objektiv antagelse bygge på vedtagne fælles interesser og værdier.[46] Selvom det er ønsket fra EU, at den digitale håndhævelse i højere grad tillægges private aktører, og delegationsforbuddet i dansk ret må siges at være blødt op, stiller forvaltningsretten alligevel krav til, hvordan aftalegrundlaget for samarbejdet mellem hostingtjenesteyderne og trusted flaggers skal udformes. Det bør måske

[46] Henstilling (n 4), kapitel 2, pkt 27.

ikke være op til tjenesteyderne selv at vurdere, hvem de finder egnet til udførelse af dette stykke arbejde.

5.2 Regelsættet

Den anden problemstilling er regelsættet. Trusted flaggers får, som nævnt ovenfor, til opgave at anmelde indhold, der ikke stemmer overens med EU-retten eller medlemsstaternes lovgivning. Det bliver, for så vidt angår deling af terrorrelateret indhold eller deling af børnepornografisk materiale, nemt for meddelerne at identificere indholdets ulovlighed. På andre og mindre 'præcise' retsområder, som eksempelvis hate speech, vil påvisningen afhænge af meddelernes ekspertise og dennes fortolkning af lovgivningen. Det betyder helt grundlæggende, at der kan opstå situationer, hvor trusted flaggers beslutter, at indhold, eksempelvis hele eller dele af hjemmesider, skal fjernes, uden at dette nødvendigvis begrundes i konkrete retsregler. I stedet begrundes det i en række privatretlige retningslinjer eller brugsvilkår udstedt af hostingtjenesteyderne[47] eller af trusted flaggers selv. Der kan på den ene side argumenteres for, at hjemmesider selv bør være herre over, hvad de ønsker at formidle på deres private sider. Omvendt blev det i Danmark, ved Højesterets afgørelse U.2010.2221 H, som forpligtede Telenor til at hindre adgang til hjemmesiden www.thepiratebay.org, understreget, at der på internettet ikke er ubegrænset ytringsfrihed. Internettet er private rum, der stilles til rådighed af private onlinetjenester, som nu får til opgave at holde justits.

Én af problemstillingerne ved at lade private onlineplatforme stå for håndhævelsen af loven på egne tjenester blev allerede i 2014 fremhævet af Europarådets menneskerettighedskommissær. I "The rule of law on the Internet and in the wider digital world" udtalte Kommissæren, at "(...) private entities can impose (and be "encouraged" to impose) restrictions on access to information without being subject to the constitutional or international law constraints that apply to state limitations of the right to freedom of expression". Opgaver som at identificere og fjerne ulovligt indhold og generelt håndhæve loven på digitale platforme ligger traditionelt set placeret i offentlige myndigheders sfære. Men hvor offentlige myndigheder er underlagt en række

[47] Meddelelse (n 3) 16.

juridiske sikkerhedskrav, i udførelsen af dette arbejde (fx de forfatnings- såvel som forvaltningsretlige principper), nyder private aktører en større kontraktlig frihed. I Danmark indeholder forvaltningsloven eksempelvis regler for sagsbehandling, som giver borgere rettigheder og beføjelser i forbindelse med myndigheders behandling af en sag. Eksempler herpå er reglerne om sagsoplysning, partsaktindsigt og begrundelse. Ekstern delegation af retshåndhævelsen vil af den grund medføre, at man som borger ikke har samme rettigheder, som hvis sagen blev behandlet af en offentlig myndighed.

Henstillingen fra marts 2018 indeholder i stedet en række beskyttelsesforanstaltninger og garantier som eksempelvis gennemsigtighed, nøjagtigt og retfærdighed[48], der skal skabe sikkerhed i anmeldelses- og indgrebsordningerne. Disse foranstaltninger forpligter både de private såvel som offentlige meddeleler, men de er langt mindre klare og præcise, som de juridiske sikkerhedskrav de offentlige myndigheder er bundet af.

En anden problemstilling ved at overlade retshåndhævelsen til hostingtjenesteyderne selv og deres trusted flaggers er lovgivningens regler (eller mangel på samme) for indgreb i brugers ytrings- og informationsfrihed. Det fremgår eksplicit af henstillingen, at onlineplatformene er forpligtet til at tage hensyn til "brugeres grundlæggende rettigheder og legitime interesser"[49] og desuden den centrale rolle, de selv spiller med hensyn til at fremme offentlig debat samt formidling og modtagelse af oplysninger i overensstemmelse med lovgivningen.[50] Problemet heri er blot, at hvor staten er underlagt strenge regler for indskrænkning af ytrings- og informationsfriheden, jf. EMRK art. 10, stk. 1, 2. pkt., gør det sig ikke gældende for de private aktører. Dette vil blive behandlet senere i artiklen.

Det er et helt tydeligt budskab i henstillingen, at retshåndhævelsen på de digitale medier i høj grad skal privatiseres – men retningslinjerne synes usikre. Trusted flaggers bliver ikke én fælles og synlig enhed som eksempelvis det danske politi. Politiet, vi kender alle fra gaden, iført lyseblå skjorter og i hvide patruljebiler. I stedet bliver de en masse små selvstændige enheder, placeret under private eller offentlige virksomheder. Man vil som bruger ikke kunne genkende

[48] Henstilling (n 4), præambelbetragtning nr 18–20.
[49] ibid, præambelbetragtning nr 13.
[50] ibid.

dem – om man så gik forbi dem på gaden. Manglen på fælles forståelse og retningslinjer kan medføre situationer, hvor digitalt indhold, der tilskynder til had, ikke behandles ens. Denne problematik vil også blive behandlet nedenfor.

5.3 Menneskeretlig kurs

Henstillingen fra marts 2018 er et forsøg fra EU på at give hostingtjenesteyderne mere ansvar over det brugergenerede indhold på deres tjenester. Heri ligger en anselig opgave for udbyderne i at afgøre, hvornår det digitale indhold enten forbyder sig imod eller er beskyttet af vores lovgivning, herunder af bestemmelserne i EMRK[51] og vores nationale straffelov.[52] Gør man eksempelvis indgreb i personers udtalelser eller holdninger på de sociale medier, kan det retlig betyde et indgreb i ytrings- og informationsfriheden, med begrundelse i beskyttelsen af retten til privatliv eller de nationale strafbestemmelser – og vice versa.

Hate speech, der kan komme til udtryk på et væld af forskellige måder, skaber retlige konflikter mellem de enkelte af vores grundlæggende rettigheder samt mellem de grundlæggende rettigheder og vores nationale straffelovgivning. Imidlertid anerkendes ytrings- og informationsfriheden i dag som en af grundstenene i et demokratisk samfund[53] og som den helt fundamentale politiske menneskerettighed.[54]

EMRK artikel 10 beskytter den formelle og materielle ytringsfrihed, modsat Grundlovens § 77, der kun beskytter den formelle. Ytringsfriheden efter artikel 10 er imidlertid ikke absolut. Efter bestemmelsens stk. 2 kan der gøres indgreb i ytringsfriheden, hvis indgrebet har hjemmel i lov og er nødvendigt i et demokratisk samfund af hensyn til et af de legitime hensyn, der er opregnet i stk. 2. Ifølge praksis fra Den Europæiske Menneskeretsdomstols (EMD) falder hate speech desuagtet ikke under artikel 10, idet udsagn af en sådan karakter anses for

[51] Council of Europe, Convention for the Protection of Human Rights and Fundamental Freedoms, 1950, ratificeret ved Lov nr 285 af 29. april 1992 (Den Europæiske Menneskerettighedskonvention).
[52] Lovbekendtgørelse nr 1156 af 20. oktober 2019 (Straffeloven).
[53] Trine Baumbach, 'Racismebestemmelsens usikre grænser', [2014] Juristen 47.
[54] Trine Baumbach, *Medieret – frihed og ansvar* (Karnov Group 2017), 42.

at være ytringer, som fornægter konventionsforudsætningerne (demokratisk styreform, retsstatsprincippet og en vilje til at overholde menneskerettighederne) og konventionens bærende værdi (princippet om den menneskelige værdighed).[55]

EMD behandler i stedet forholdene efter EMRK artikel 17, om misbrug af rettigheder.[56] Dette hænger nøje sammen med artikel 10 stk. 2, som eksplicit påpeger, at udøvelsen af frihedsrettigheden medfører "pligter og ansvar", der nok ikke anses for at være efterkommet, hvis ens kommunikation udmønter sig som hate speech. Ifølge Trine Baumbach findes der ikke en universel definition af hate speech.[57] Alligevel kan fænomenet – den egentlige hate speech omfattet af artikel 17 – ud fra EMD's praksis med nogenlunde sikkerhed defineres som "ytringer, der direkte opfordrer til vold, eller er stærkt voldsforherligende, samt ytringer, der lægger befolkningsgrupper for had".[58] Ytringer af mindre grov og hadefuld karakter falder derimod under EMRK artikel 10, stk. 1, og indgreb i sådanne skal således retfærdiggøres efter stk. 2.

At bekæmpe hadefulde ytringer på nettet indeholder derfor en problemstilling om, i hvilken udstrækning og på hvilken måde der kan ske en begrænsning af den materielle ytringsfrihed begrundet i, at den hadefulde ytring krænker de berørte personers privatliv og integritet på en samfundsmæssigt uønskelig måde.[59]

I Danmark er der opstillet materielle begrænsninger for ytringsfriheden i eksempelvis straffelovens § 266 b, der beskytter mod hån og uberettigede beskyldninger (nærmere kendt som racismeparagraffen) samt i straffelovens kapitel 27 om freds- og ærekrænkelser. Straffelovens § 266 b kriminaliserer langt fra alle former for hate speech, men grundet artiklens begrænset længde vil fokus være på denne. Bestemmelsen har til formål at beskytte befolkningsgrupper mod racistiske udtalelser, der krænker gruppens menneskelige værdighed – en

[55] ibid 71.
[56] ibid samt *Perincek v Switzerland* ECHR 2015-IV 181 (om fornægtelse af det armenske folkedrab mv) og *Pavel Ivanov v Russia* App no 35222/04 (EMD, 20. februar 2007) (om stærkt antisemistiske udsagn omfattet af artikel 17).
[57] Baumbach 2017 (n 55) 72.
[58] ibid.
[59] Braumbach og Blume (n 16) 8.

beskyttelsesinteresse, der er omfattet af de legitime hensyn efter EMRK artikel 10, stk. 2.[60] Da straffelovens § 266 b kriminaliserer handlinger, som borgere har en menneskeret ret til, skal den fortolkes i lyset af ytringsfriheden[61] og den praksis, som EMD har fastlagt.[62]

Ikke desto mindre skal der – som skitseret ovenfor – foretages en række afvejninger af ytringens karakter. Og det giver anledning til nationale forskelle. De materielle begrænsninger, som staterne kan fastsætte, afspejler den betydning, som ytringen tillægges i de respektive retskulturer. I Danmark er vi eksempelvis forpligtet til at sikre, at hate speech straffes i overensstemmelse med hjemlen i straffelovens § 266 b, såfremt tiltalte objektivt og subjektivt har realiseret gerningsindholdet.[63] En svær afvejning, der også giver anledning til problemer for vores egne domstole, jf. eksempelvis U 2003.2435 V (om politiske udtalelser på et landsmøde, der efterfølgende gav debat, om hvorvidt domstolenes retsanvendelse er for mekanisk således, at udtalelser er for generaliserende),[64] U 2012.2361 H (om manglende fortsat til udbredelse) og U 2014.73 V (om meget generaliserende truende og nedværdigende udtalelser, der var omfattet bestemmelsen, selvom de indgiv i saglig debat).

De nationale afvejninger, altså om den pågældende ytring falder under EMRK artikel 10 eller artikel 17 samt under en af vores nationale materielle begrænsninger, bliver essentiel i trusted flaggers arbejde. De får til opgave at vurdere, hvornår ytringer og udsagn har en vis nytteværdi for samfundet, og dermed frit kan offentliggøres, kontra hvornår ytringer overtræder den nationale straffelovgivning (som godt nok skal fortolkes i lyset af EMRK) og dermed bør fjernes. Hertil kommer desuden, at EMD har fastslået, at staten ikke er berettiget til at tage alle udsagn for pålydende eller forstå alle udsagn bogstaveligt.[65] Der er altså lagt op til vurderinger, som kan spænde bredt – både på grund af de forskelligartede nationale retskulturer, men særligt også indbyrdes mellem de

[60] Baumbach 2017 (n 55) 262.
[61] ibid 223.
[62] Baumbach 2014 (n 54) 47.
[63] ibid 54.
[64] Baumbach 2017 (n 55) 263.
[65] *Gül and Others v Turkey* App no 35071/97 (EMD, 4. december 2003) og desuden Baumbach (n 55) 271.

forskelligartede hostingtjenester, der som nævnt skal holde justits med det indhold, som florerer på deres medie(r).

Kommissionens anbefaling om at anvende trusted flaggers i forsøget på at skabe den mest effektive og samordnet tilgang til bekæmpelsen af ulovligt indhold er derfor ikke uden komplikationer. Særligt taget i betragtning af internettets og hostingtjenesternes globale virke.

Det er et indbygget problem i vores overnationale menneskerettigheder, at der opstår statslige forskelle i håndhævelsen af EU-lovgivningen. Dét, at vi nationaliserer vores internationale regler,[66] skaber egne erfaringer, interesser og udsigtspunkter, som kan risikere at medføre, at vi vil se andre staters værdier som trusler mod vor egne på trods af, at de tager afsæt i samme menneskerettigheder. Et digitalt og globalt problem som hate speech er derfor svært at bekæmpe i fællesskab, også selvom EMD har forsøgt at skabe harmoni mellem reglerne. Disse forhold gør det usikkert at overlade rollen som samfundets indholdspoliti til mere eller mindre tilfældige private aktører.

5.4 Virksomhedscensur, transperens og over-fjernelse

Det voksende fokus på onlineplatforme og hostingtjenesteyders ansvar for ulovligt digitalt indhold, har – foruden EU-kommissionens opfordring om trusted flaggers og den frivillige aftale om "Code of conduct" – også ført til andre tiltag. Eksempelvis indførte Tyskland i 2018 en ny medielov (NetzDG), som pålægger sociale medieplatforme at fjerne ulovligt indhold inden for 24 timer. Men selvom tiltagene prøver at bidrage positivt til bekæmpelse af de digitale udfordringer, kritiseres de også for at legitimere privat censur.

Fænomenet privat censur opstår af den redaktionelle frihed, som medierne har til selv nøje at udvælge og formidle det indhold, de privatejede platforme ønsker. Grundlovens § 77 og EMRK artikel 10 hjemler begge et censurforbud, hvorefter myndigheder ikke må foretage en forhåndskontrol af de ytringer og den information, der offentliggøres. Endvidere indeholder EMRK artikel 10 som nævnt en ret til at modtage information. Men forbuddet forbyder alene

[66] Anthea Roberts, *Is international Law International*? (Oxford University Press 2017) 325.

offentlig forhåndskontrol, hvorfor kun staten og ikke private er omfattet. Der er ganske enkelt en stor frihed for både onlineplatformene og hostingtjenesteyderne til selv at definere deres platform og deres ansigt ud af til. Af den grund opstiller virksomhederne forventeligt nok en række retningslinjer for, hvilket indhold de ønsker at formidle, og hvilket de må fjerne. Det kan føre til situationer, som nævnt i afsnit 5.2, hvor indhold fjernes, ikke fordi det direkte strider imod lovgivningen, men fordi det strider imod tjenesternes politik og retningslinjer. Og det er herigennem, der opstår privat – og fuldt lovlig – censur. Opfordringen til at anvende trusted flaggers kan risikere at "fodre" dette fænomen – ud fra andre og udokumenterede retningslinjer. Hvor privat censur sker på baggrund af én aktørs udvælgelse, kan man fremadrettet forestille sig, at digitalt indhold fjernes, fordi det ikke harmonerer med en intern strategi, som hostingtjenesteyderne og trusted flaggers har fastsat. Denne form for strategi er udtryk for censur på baggrund af en (måske uofficiel) virksomhedsstrategi – heraf ordet lovlig "virksomhedscensur".

Kommissionen har som nævnt anbefalet hostingtjenesteydere at gøre brug af trusted flaggers til identificering og indberetning af ulovligt indhold, der strider mod EU-retten eller den berørte medlemsstats lovgivning. Det betyder, at hostingtjenester både har adgang til at fjerne indhold, som ikke lever op til deres egne regler og retningslinjer (virksomhedsstrategi) – uagtet at indholdet er lovligt, og nu også får til opgave til holde justits med det indhold, som florerer på deres medie, og som ikke lever op til lovgivningens krav. En form for dobbeltrolle, der virker bekymrende.

Onlineplatforme og hostingtjenester spiller i dag begge centrale roller den offentlige debat – især dennes demokratiske kvalitet og pluralisme. Vi mennesker er sociale dyr, som er enormt påvirkelige overfor det indhold, der florerer online, og vi danner meninger ud fra, hvad andre deler, "liker" og kommenterer.[67] Af den grund er det ikke underordnet, hvilket og i hvor høj grad digitalt indhold fjernes. For den almindelige bruger af de digitale platforme, bliver lovens grænser for henholdsvis lovligt og ulovligt indhold vanskelig at gennemskue. De private aktører får med deres "dobbeltrolle" rig mulighed for at forme og sætte grænserne for vores ytringsfrihed og for ulovligt, digitalt indhold. Uagtet lovgivningens ord,

[67] Hendricks (n 13).

skaber tjenesternes fjernelses- og blokeringsprocedure de uskrevne regler for, hvordan brugere af tjenesterne må og skal opføre sig. Det er her vigtigt at understrege, at man har mulighed for at få afprøvet beslutninger om fjernelse af hele eller dele af hjemmesider ved domstolen. Men det kræver, at der er kendskab til baggrunden fra fjernelsen.

Hostingtjenesterne bør derfor i høj grad være transparente i deres samarbejde med meddelerne og særligt gennemsigtige i forhold til, hvad de foretager sig, lige fra udvikling af retningslinjer til indsigt i, hvordan de håndhæves. Transparens kan også være med til at sikre, der ikke sker en form for "over-fjernelse". En over-fjernelse af digitalt indhold, som begrænser den reelle ytringsfrihed. For hvis alt digitalt indhold bliver underlagt en lang række interne retningslinjer, vil den digitale ytringsfrihed indsnævres – langt mere end den allerede er i forvejen. Det er ikke utænkeligt, at platformene ud fra et konkurrencesynspunkt (ønsket om at bevare deres stilling på det digitale marked) og tjenesteydernes større medansvar for at bekæmpe ulovligt digitalt indhold samt delegationen af retshåndhævende kompetence til trusted flaggers, alle kan medføre en øget over-fjernelse af digitalt indhold.

Det er problematisk, hvis private aktører både får til opgave at sætte de overordnede rammer for internettet – i forhold til ønskeligt indhold, tonen og de generelle retningslinjer – og samtidig får ansvaret for at retshåndhæve deres egne såvel som de offentlige regler på internettet. På den måde bliver den lovgivende, udøvende og dømmende magt de facto til én aktør. Og det hele vil ske uden indblanding fra "rigtige" myndigheder og retshåndhævere.

5.5 Retssikkerhed

Der er mange årsager til, at ulovligt indhold – herunder hate speech – er udfordringer, som skal håndteres. Det er dog en nødvendighed, at midlerne til bekæmpelse af de digitale udfordringer ikke går på kompromis med hverken retssikkerheden eller andre grundlæggende rettigheder.

Retssikkerhed kan overordnet forklares som det grundlæggende element i en demokratisk retsstat, der sikrer private mod vilkårlige, ulovhjemlede og

uforudsigelige indgreb fra statens side.[68] Retssikkerheden hænger nøje sammen med retsstatsprincippet, der i henhold til artikel 2 i TEU er en af EU's grundlæggende værdier. Det er princippet om, at både EU og alle EU-landene er styret af et lovkompleks (regelsæt og procedurer), der er vedtaget efter fastlagte processer i stedet for ved skønsmæssige beslutninger eller ad hoc-afgørelser. Det er helt essentielt for begge principper, at der er lighed for loven og forudsigelighed i afgørelser, så borgerne kan have tillid og tiltro til systemet. Trusted flaggers' afgørelser udfordrer både disse grundlæggende principper og generelt den traditionelle juridiske metode, som bygger på fastlæggelse af jus og retsfaktum, hvorefter der foretages en subsumption. Deres arbejde formodes i højere grad at være baseret på egen fortolkning af lovgivning, egen virksomhedsstrategi og ud fra tidligere afgørelser af pågældende type, mere end det baseres på fastlæggelse af regelsættet og sagsoplysning. Rammerne for den demokratiske debat og vores ytringsfrihed overlades på den måde til en "sort boks", hvor ukendte vurderingskriterier henholdsvis indskrænker eller liberaliser den enkeltes frihed til at debattere og ytre sig.

Det må almindeligvis antages, at det på ingen måde må forringe borgerens retsstilling, at opgaven som digitalt politi bliver uddelegeret til private aktører. Men selvom Kommissionen anfører "højere kvalitetsmeddelelser" og "hurtigere nedtagninger", som fordele ved betroede anmeldere, præciserer de ikke yderligere, hvad der ligger i disse kvalitetsforbedringer.[69] Kommissionen har primært fokus på effektiviseringsstrategier til bekæmpelse af de digitale udfordringer – men det er ikke nok. De retssikkerhedsmæssige problematikker skal nøje overvejes og gennemtænkes, særligt hvis trusted flaggers skal besidde en legitimitet svarende til det rigtige politi og opnå samme tillid fra borgerne.

[68] Institut For Menneskerettigheder, 'Lov og Ret' <https://menneskeret.dk/emner/lov-ret> besøgt den 3. september 2019.
[69] Meddelelsen (n 3) 8.

5.6 Bindende lovgivningsmæssige foranstaltninger på området for hate speech

Artiklen har i det foregående koncentreret sig om de to ikke-bindende EU-retsakter, og hvilke retlige problematikker disse medfører. Afslutningsvist vil der kort knyttes et par ord til, hvilke problematikker der kan opstå ved, at Kommissionen, som nævnt i meddelelsen fra september 2017, supplerer retningslinjerne med lovgivningsmæssige foranstaltninger[70] – og dermed bindende retsakter – på området for hate speech.

Der er ingen tvivl om, at der er brug for effektive midler, der kan forhindre eller i det mindste reducere deling af ulovligt og krænkende indhold. Spørgsmålet er bare – hvilke midler?

Lovgivningsmæssige foranstaltninger skaber klare retningslinjer og måske en mere ensartet håndhævelse. Det vil endvidere harmonere med den virkelighed, vi normalt lever i – altså når vi ikke er online på internettets mange platforme, hvor ytringsfriheden også møder materielle begrænsninger (fx i straffeloven). Internettets utallige digitale udfordringer har potentiale til at skade menneskets viden, holdninger og den generelle tillid til hinanden og samfundet.[71] I større skala kan det potentielt give demokratiske udfordringer og allerværst være demokratiundergravende. Af den grund er samfundet forpligtet til at reagere.

Det er naivt at tro, at det hele nok skal blive bedre – altså med mindre vi vil skyde en hvid pil efter demokratiet. Men det er imidlertid svært at forestille sig, hvordan reglerne skal udformes. Internettet er globalt med den konsekvens, at ytringerne på nettet er underlagt mange jurisdiktioner. De ytringer, der krænker personer her i landet, opfattes ikke nødvendigvis ikke som værende krænkende endsige strafbare uden for Danmarks grænser, jf. ovenstående. Så hvis vi i endnu højere grad nationaliserer vores internationale regler, risikerer vi i endnu højere grad at komme til at se andre staters værdier som trusler mod vor egne.

Det gavner desuden ikke demokratiet og ytringsfriheden, hvis staten og de retshåndhævende myndigheder begynder at sondre mellem tilladelige og ikke-tilladelige ytringer, og begynder at stille for rigide krav til både afsender og modtager af disse. Ud fra en mere filosofisk tilgang kan dette også begrundes

[70] Meddelelsen (n 3) 3, 21.
[71] Hendricks (n 13) 77.

med, at for at vi som mennesker kan danne os oplyste og velbegrundet meninger, har vi måske også brug for at møde usandheder, radikale holdninger og andet 'provokerende' materiale. Altså til en vis grad. Der er behov for, at problemer og emner sættes under debat – og det skal ikke kun være efter de retningslinjer, som magthaverne sætter.

Slutteligt kan der også sættes spørgsmålstegn ved, om det overhovedet er en offentlig opgave at intervenere i borgeres indbyrdes verbale konflikter. Måske en del af reguleringen skal ligge hos internetbrugerne og deres reflekterede opfattelse af den digitale virkelighed. Det kan overvejes om de humanistiske virkemidler skal tages i brug, og det i stedet handler om at danne brugerne.[72] En national handleplan for digital dannelse, kan også være en måde, hvorpå der skabes en bedre digital infrastruktur og rarere debatkultur.[73]

6. Konklusion

En styrket indsats mod digitale udfordringer, som eksempelvis den udbredte brug af hate speech, er en nødvendighed. Der er behov for effektive midler, der kan forhindre, reducere eller helt fjerne ulovligt og krænkende indhold fra både enkeltpersoner, organisationer og platformes hjemmesider. Forslaget fra EU-kommissionen om uddelegeringen af ansvar til private aktører som retshåndhævere på de digitale medier synes derfor umiddelbart både tidssvarende og effektivt.

Der er ingen tvivl om, at hostingtjenesteyerne bør bære en langt større del af ansvaret for bekæmpelsen af den omfattende spredning af ulovlig indhold. Internettet er blevet et uundværligt redskab i hverdagen. Platformene har gennem de tjenester, de stiller til rådighed, muliggjort en hidtil uset adgang til information og udveksling af meningen og har åbnet op for helt nye markedsmuligheder for virksomheder. De spiller en central rolle for den offentlige debat og kan være med til at styrke ytringsfriheden – kvaliteten og pluralismen i den demokratiske debat. Imidlertid har det også givet anledning til

[72] ibid.

[73] Man har forsøgt at sætte digital dannelse på skoleskemaet i de danske folkeskole, jf <https://www.uvm.dk/folkeskolen/laering-og-laeringsmiljoe/it-i-undervisningen/digital-dannelse> besøgt d. 15. september 2019.

lang række bekymringer, da flere bruger platformene som fora til deling af ulovligt indhold, hvad end det er terrorrelateret indhold, hate speech eller andet. Det har fået flere aktører herunder FN og EU til at slå fast, at hvad der er ulovligt offline også er ulovligt online.

Det større samfundsansvar, der tillægges hostingtjenesteyerne for beskyttelse af brugere og samfundet som helhed, mod de, som anvender deres tjenester til forbrydelser, vil efter anbefalingerne skulle deles med trusted flaggers. Det synes fordelagtigt at lade private aktører overtage en del af retshåndhævelsen på de digitale medier, fordi de som udgangspunkt besidder en større ekspertise på området. Artiklen har dog vist, at dette forslag ikke er uden retsmæssige bekymringer.

Traditionelt set er retshåndhævelsen placeret hos de offentlige myndigheder, hvor politiet er den myndighed, som har ansvaret for at føre kontrol med, at lovene overholdes og at skride ind over for lovovertrædelser. Trusted flaggers bliver modsat hertil aktører placeret under private enheder. Således bliver de langt mindre kontrolleret og kommer til at nyde en stor kontraktlig frihed i valg af metoder og grad af åbenhed. Trusted flaggers vil ikke blive underligt samme forvaltningsregler som politiet, selvom de får væsentlig indflydelse på forvaltningsprocessen (hvilket ifølge Ombudsmandens praksis er tilladt, jf. FOB 2013-9.) Det i sig selv synes ikke tilstrækkeligt, eftersom forvaltningslovens § 2 hjemler et forbud mod ekstern delegation til private af kompetence til at træffe afgørelse i forhold til borgere. Med det in mente betyder det, at såfremt Henstilling (EU) 2018/334 var bindende i alle sine enkeltheder, vil den være vanskelig at implementere i dansk lovgivning (trods EU-rettens forrang).

Henstillingen anbefaler, til trods for at trusted flaggers ikke nødvendigvis er offentlige myndigheder[74], at aktørerne respekterer vores fundamentale rettigheder herunder ytringsfriheden. Henstillingen tildeler dermed trusted flaggers offentligretlige forpligtelser, men uden at gå så langt som til at forpligte til en overholdelse af alle forvaltning – såvel som forfatningsretlige regler. En gennemført respekt for de fundamentale rettigheder vil dog kræve en mere ensartet menneskeretlig kurs og forståelse herfor – hvilket nok er en utopi. Stater

[74] Artiklen har kun haft fokus på de tilfælde, hvor trusted flaggers er private aktører.

nationaliserer til en vis grad de internationale regler og forpligtelser, og det kan betyde, at staterne indbyrdes kommer til at modarbejde hinanden.

Det er heller ikke utænkeligt, at der kan opstå så tæt samarbejde mellem hostingtjenesteyderne og trusted flaggers, at de gennem interne retningslinjer og egen fortolkning af vores menneskerettigheder overregulerer og over-fjerner digitalt indhold. Måske endda i så høj grad, at de begynder at forme og sætte grænserne for vores ytringsfrihed og for ulovligt, digitalt indhold – der måske nok er lovlig, men stadig uacceptabel. I forlængelse heraf ønskes langt flere og klarerer regler for, hvordan retssikkerheden opretholdes. Her tænkes i forhold til kontrollen af trusted flaggers og transparensen i deres arbejde. Offentlig kontrol og indsigt i deres vurderinger og fjernelse af ulovligt indhold vil ydermere øge deres legitimitet og troværdighed. Det samme vil en lovfæstet forpligtelse til at begrunde enhver fjernelse af indhold.

Det er klart, at ulovligt indhold på internettet ikke må underminere borgernes tillid og tiltro til det digitale miljø og demokratiet generelt. Ej heller må det true de økonomiske muligheder. Det digitale indre marked, der også er en del af EU-kommissionens strategier for perioden 2015-2019, skal styrkes. Det må dog være et ufravigeligt krav, at der skal findes passende beskyttelsesforanstaltninger, som hverken går på kompromis med retssikkerheden eller vores lovgivning – herunder menneskerettighederne. Der skal findes metoder, som tilpasses de nationale regler og som samtidigt sikrer, at bekæmpelsen af de digitale udfordringer sker på baggrund af en fælles strategi og vision fra EU. Internettet er globalt – og derfor skal metoderne også fungere globalt. Fordelene og ulemperne ved privatiserede håndhævelsesmekanismer skal nøjes afbalanceres – og effektivitetshensynet må ikke uden videre trumfe.

Alene på baggrund af de udvalgte behandlede betænkeligheder – en brøkdel af alle de retsmæssige konflikter, må det konkluderes, at EU-kommissionen har opstillet uklare, ikke-bindende anbefalinger. Henstillingen, som vejledning for både hostingtjenesteyderne og trusted flaggers selv, synes både usikker og ugennemtænkt og efterlader desuden mange tvivlsspørgsmål.