

APRIL 2019

NÅR DIGITALE PLATFORME SKAL BEGRÆNSE YTRINGSFRIHEDEN

EU lægger op til, at digitale platforme som fx Google, YouTube og Facebook skal bruge automatiserede indholdsfiltere for at undgå ulovligt indhold på nettet. Men der er en række menneskeretlige udfordringer ved en automatisk begrænsning af brugernes ytrings- og informationsfrihed.

EU er i gang med at indføre to nye regelsæt, som lægger op til, at digitale platforme som Google, Facebook og YouTube i øget omfang skal gøre brug af automatiserede indholdsfiltere for at overholde lovgivningen. Det drejer sig om en forordning om forebyggelse af terrorrelateret onlineindhold, og et direktiv om ophavsret på det digitale indre marked.

Fælles for begge regelsæt – som beskytter meget forskelligartede samfundsinteresser – er, at digitale platformen får adgang til at bruge ”automatiserede værktøjer” eller ”indholdsgenkendelsesteknologier” som led i en forpligtelse til at fjerne ulovligt indhold.

Regelsættene supplerer gældende EU-regulering som for eksempel e-handelsdirektivet, som forpligter virksomheder til at fjerne ulovligt indhold, når de bliver gjort opmærksomme på det.

Den øgede brug af automatiserede filtre fører en række menneskeretlige udfordringer med sig.

AUTOMATISEREDE INDHOLDSFILTRE

Automatiserede indholdsfiltere er et maskinlæringsprogram, hvor algoritmer

automatisk identificerer og fjerner – eller hindrer upload af – tekst eller billeder på nettet. Det kan fx dreje sig om ulovligt indhold.

Indholdsfilterne analyserer en stor mængde data og bliver på den baggrund i stand til at finde flere og flere sammenhænge og mønstre i dataene. Filterne lærer løbende, og de kan med tiden træffe såkaldte intelligente beslutninger. Det vil sige, at et filter, der skal identificere og fjerne ulovligt materiale, med tiden vil identificere og fjerne andre typer materiale end det, filtret oprindeligt var programmeret til.

I visse tilfælde vil filtrets resultater efterfølgende blive gennemgået af en person, der skal afklare, om indholdet reelt er ulovligt og skal fjernes. I andre tilfælde er der ingen menneskelig kontrol med indholdsfilteret.

UPRÆCIS FILTRERING

Algoritmerne, som ligger til grund for filterne, er begrænsede i deres forståelse af konkrete omstændigheder og sammenhæng. Det kan føre til en upræcis sortering. Indholdsfiltere kan både risikere at fjerne materiale, der ikke er ulovligt, og de kan overse ulovligt indhold.

Indholdsfiltere kan med andre ord komme til at fjerne både for lidt og for meget.

MANGLENDE GENNEMSIGTIGHED

Det er desuden et menneskeretligt problem, at filtrenes "beslutninger" foregår i det skjulte, og at det derfor ikke er muligt at få indblik i dem. Indholdsfilternes sorteringer er reelt umulige at gennemskue.

UPROPORTIONALE BEGRÆNSNINGER

Ytrings- og informationsfriheden er blandt andet beskyttet i Den Europæiske Menneskerettighedskonventions artikel 10 og grundlovens § 77. Begrænsninger i ytringsfriheden skal være fastlagt ved lov, tjene et legitimt formål og være proportionale.

Det er staten, der skal beskytte borgernes ytrings- og informationsfrihed. Staten har derfor en forpligtelse til at sikre, at private aktørers praksis ikke fører til ulovlige begrænsninger i borgernes ytrings- og informationsfrihed.

De automatiserede indholdsfiltere skal med andre ord sikre, at ytrings- og informationsfriheden ikke indskrænkes mere end strengt nødvendigt.

De to udkast til regelsæt fra EU henviser til, at brugen af indholdsfiltere skal være "effektiv", "passende" og "forholdsmæssig", men giver ingen konkrete anvisninger til, hvordan platformene skal sikre dette.

FN's specialrapportør for ytringsfrihed har ligeledes fremhævet, at de nye EU-regler ikke forholder sig til, hvordan platformene skal sikre en korrekt balance mellem på den ene side ytrings- og informationsfriheden og på den anden side kravet om at fjerne ulovligt indhold.

KRAV OM EFFEKTIVE KLAGEMULIGHEDER

Meget taler derfor for, at staten skal påtage sig et større ansvar for at sikre ytrings- og informationsfriheden, hvis den indfører regulering, der medfører brug af automatiserede indholdsfiltere. Herudover er det vigtigt at få sikret, at brugerne har effektive klagemuligheder.

Et effektivt klagesystem vil som minimum indebære, at brugerne:

- bliver informeret, når de har fået indhold fjernet eller blokeret
- oplyses om grundlaget for beslutningen
- har mulighed for at gøre indsigelser indenfor en rimelig tidsfrist.