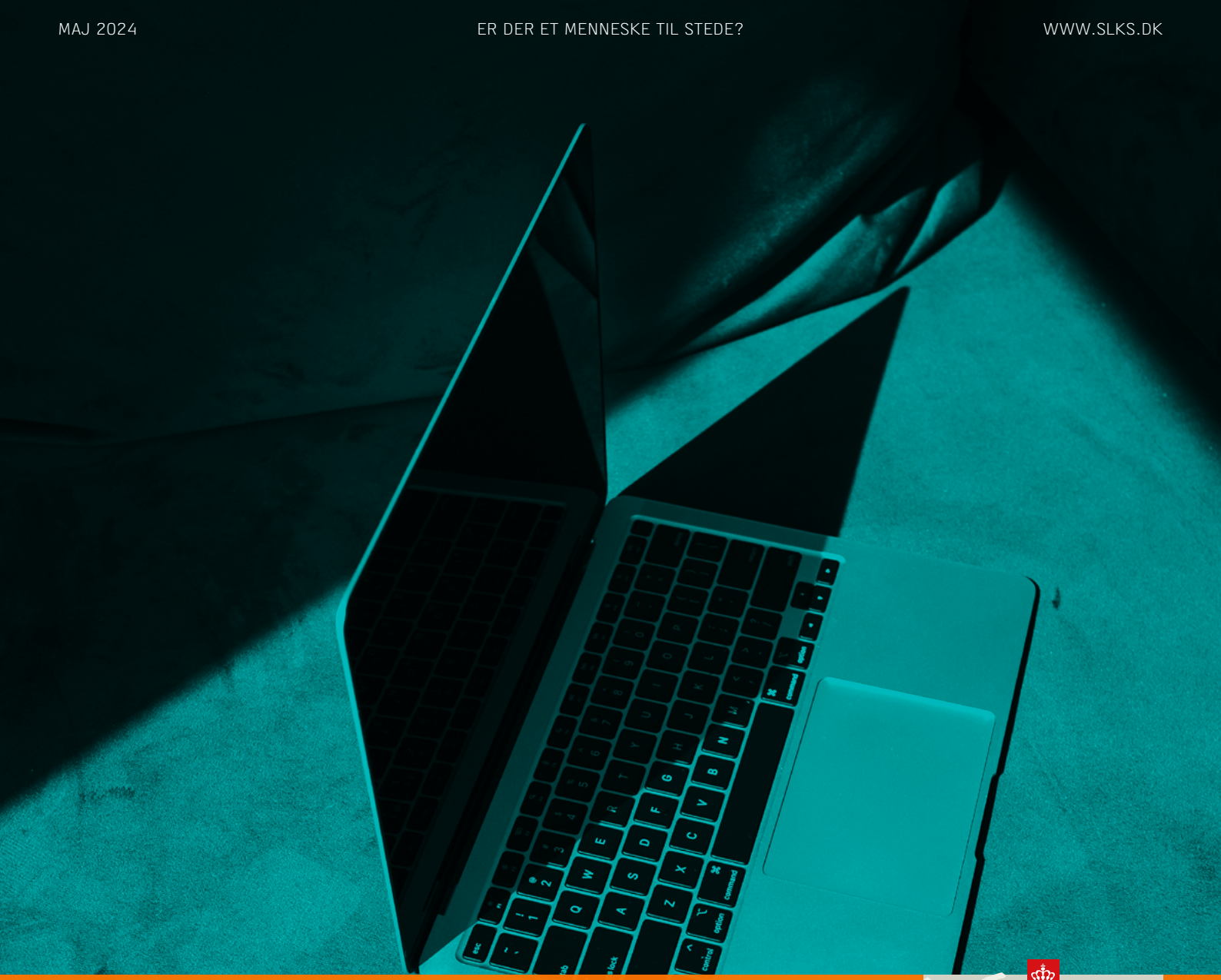


CENTER FOR
SOCIALE MEDIER
TECH OG DEMOKRATI

KULTUR MINISTERIET
SLOTS- OG KULTURSTYRELSEN

ER DER ET MENNESKE TIL STEDE?

EN KOMPARATIV ANALYSE AF STORE SOCIALE MEDIERS INDHOLDSMODERATORER

INDHOLD

INDLEDNING 3

OM UNDERSØGELSEN 4

UNDERSØGELSENS HOVEDRESULTATER 5

1 INDHOLDSMODERATION: HVORDAN OG HVORFOR 7

2 FLERE PLATFORME HAR INGEN ELLER FÅ DANSKTALENDE MODERATORER 9

3 STOR VARIATION I INDHOLDSMODERATORER PÅ TVÆRS AF EU 12

4 FORSKELLIGARTEDE OPGØRELSER KOMPLICERER SAMMENLIGNING 17

REFERENCER 19

INDLEDNING

Den offentlige, demokratiske samtale foregår i stigende grad på sociale medier. Enorme mængder indhold deles hver dag på platformene, heriblandt indhold, der kan være skadeligt for særligt børn og unge eller indhold, som kan bruges til at forvrænge og vildlede i samfundsdebatten gennem spredning af mis- og desinformation. Det kan i sidste ende påvirke valgresultater og underminere tilliden til demokratiske institutioner med potentielt vidtrækkende konsekvenser for vores samfund og demokrati til følge. Derfor er det vigtigt, at indhold på platformene modereres.

I det største valgår i historien, hvor mere end halvdelen af jordens befolkning skal til stemmeurnerne, er det afgørende med et særligt fokus på at sikre integriteten af frie og demokratiske valg. Adgangen til troværdig information og et stærkt svar mod spredningen af mis- og desinformation er vigtige forudsætninger i den forbindelse.

Det er en stor og vigtig opgave, der blandt andre er lagt i hænderne på tech-giganterne. En opgave der, hvis den misbruges eller ikke løftes, kan sætte ytringsfriheden under pres og svække demokratiet.

Det er afgørende for den demokratiske kontrol med tech-giganterne og deres indholdsmoderation, at offentligheden og den enkelte bruger gives et indblik i tech-giganternes maskinrum på en gennemsigtig og ensartet måde. Men der mangler gennemsigtighed med, hvordan der modereres, og hvem der modererer det indhold, som brugerne møder online.

De største tech-giganter er underlagt en række forpligtelser til at offentliggøre informationer om deres indholdsmoderation. De skal bl.a. rapportere om deres moderationspraksisser og deres indholdsmoderationshold, herunder moderatorernes kvalifikationer og sproglige ekspertise, som kan have betydning for platformes evne til at moderere indhold på tværs af sprog og kulturelle kontekster i EU.

Disse rapporter danner udgangspunkt for denne komparative analyse af tech-giganternes indholdsmoderation, som rejser spørgsmålet: Er der et menneske til stede?

OM UNDERSØGELSEN

I denne analyse undersøger Slots- og Kulturstyrelsens Center for sociale medier, tech og demokrati (CSTD) otte store sociale mediers gennemsigtighedsrapporter. Undersøgelsen ser nærmere på dele af platformenes indholdsmoderation og sammenligner deres hold af moderatorer og deres sprogkompetencer på tværs af EU.

De otte udvalgte store sociale medier er **Facebook**, **Instagram**, **LinkedIn**, **Pinterest**, **Snapchat**, **TikTok**, **X** (Twitter) og **YouTube**.

De otte sociale medier er udvalgt, fordi de er blandt de 23 store platforme og søgemaskiner, der på grund af deres størrelse er forpligtet til at skulle leve op til særligt strenge krav om gennemsigtighed i forordning om digitale tjenester (DSA'en). Det er derudover særligt på disse platforme, borgere deltager i den digitale offentlige samtale, dvs. debatterer, udtrykker deres holdninger, deler indhold og opsøger viden. Tjenesterne spiller derfor en central rolle i digitale demokratiske processer.

Dermed er store handelsplatforme, pornografiske videodelingsplatforme, samt store søgemaskiner og vidensplatforme fravalgt denne undersøgelses fokus.

Undersøgelsen baserer sig på en komparativ læsning af de otte udvalgte platformes gennemsigtighedsrapporter, og sammenligner de afrapporterede værdier på tværs af platformene. Rapporteringsperioderne i de udvalgte gennemsigtighedsrapporter varierer fra tre til seks måneder og løber inden for tredje kvartal 2023 til første kvartal 2024. De afrapporterede tal er indsamlet d. 27. maj 2024.

Undersøgelsen baserer sig på data om, hvad platformene selv rapporterer i deres gennemsigtighedsrapporter og giver derfor udelukkende indblik i platformenes selvrapporterede tilgang og ressourcer til at moderere indhold i EU.

UNDERSØGELSENS HOVEDRESULTATER

FLERE PLATFORME HAR INGEN ELLER FÅ DANSKTALENDE MODERATORER

- LinkedIn og X har ingen dansktalende moderatorer. Facebook, Instagram og Pinterest har under 10.
- Antallet af moderatorer er sjældent proportionelt med antallet af brugere.

DER ER STOR VARIATION I INDHOLDSMODERATORER PÅ TVÆRS AF EU

- Moderatorer hos LinkedIn og X dækker omkring en tredjedel af EU's 24 officielle sprog.
- Snapchat og Pinterest dækker omkring halvdelen af EU's 24 officielle sprog.
- Kun Facebook og Instagram har moderatorer, som kan dække alle EU's 24 officielle sprog
- De mest udbredte modersmål i Nord- og Østeuropa er dårligere dækket af platformenes indholdsmoderatorer end udbredte modersmål i Vest- og Sydeuropa.

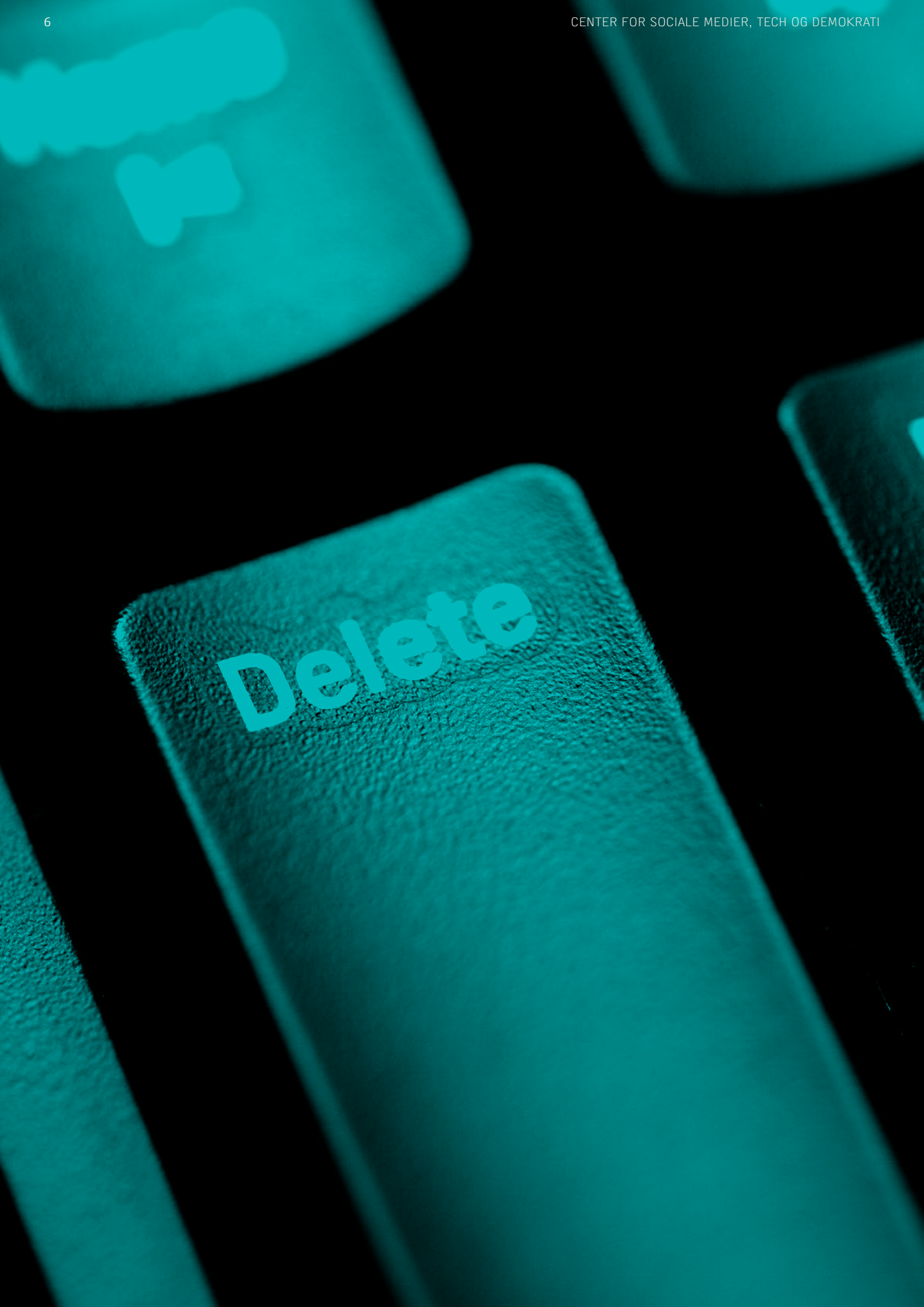
PLATFORMENES FORSKELLIGARTEDE OPGØRELSER KOMPLICERER SAMMENLIGNING

- Platformene opgør ikke på en ensartet måde, hvor mange unikke moderatorer de har tilknyttet, eller hvor mange moderatorer som tælles med flere gange, fordi de taler flere sprog.
- Platformene opgør ikke, hvor stor en andel af indholdet på de respektive sprog, der bliver modereret.
- Platformenes rapporteringsperioder er forskellige, hvilket vanskeliggør sammenligning på tværs.

ANBEFALINGER TIL TECH-GIGANTERS GENNEMSIGTIGHEDS-RAPPORTER

For at sikre, at tech-giganternes gennemsigtighedsrapporter i fremtiden giver bedre indsigter i platformenes indholdsmoderation og kan sammenlignes på tværs, anbefaler Center for sociale medier, tech og demokrati:

- Minimumskrav for hvor mange moderatorer tech-giganterne har i de lande, de driver forretning i.
- Fælles og sammenlignelige opgørelsesmetoder til opgørelse af moderatorer, så antallet bedre kan sammenlignes på tværs af platforme.
- Tydelige og præcise beskrivelser af, hvordan platformene indholdsmodererer.


Delete

1

INDHOLDSMODERATION – HVORDAN OG HVORFOR

”Every day, we remove millions of violating pieces of content and accounts on Facebook. In most cases, this happens automatically, with technology to detect, restrict, and remove content and accounts that may go against our Community Standards, Advertising Standards, and Commerce Policies. In other cases, our technology selects content for human review. Our review teams review a blend of user reports and content surfaced by our technology […].”

Facebooks gennemsigtighedsrapport for 2024

Når platformene begrænser indhold eller sanktionerer de konti, som deler indhold, kaldes det indholdsmoderation. Når platformene finder indhold, som skal modereres, kan det gøres på flere måder. I forskningen på området skelner man typisk mellem:

- **Hård** indholdsmoderation, såsom at fjerne indhold eller blokere konti, som har delt indhold.
- **Blød** indholdsmoderation, såsom at give indhold et mærkat som advarer om, at indholdet er vildledende eller at begrænse indholdets rækkevidde og spredning[1].

Indholdsmoderation er i dag et komplekst system, hvor platformene benytter både automatiserede værktøjer og moderatorer og får hjælp fra brugere til at rapportere indhold som, de mener, overskrider gældende retningslinjer.[2] Automatiserede værktøjer står for langt størstedelen af moderationen af de store mængder indhold, som dagligt deles på platformene, og indholdsmoderatorer assisterer særligt ved tvivlsspørgsmål og er med deres beslutninger med til at træne algoritmerne.

BOKS 1

EU ÅBNER SAG MOD X FOR UTILSTRÆKKELIG INDHOLDSMODERATION PÅ PLATFORMEN

I december 2023 åbnede Europa-Kommissionen en sag mod X for at undersøge, om platformen har overtrådt de nye EU-regler, som gælder for særligt store platforme. Kommissionen mistænker bl.a. X for ikke at have gode nok mekanismer til at modere indhold og dermed ikke i tilstrækkelig grad er i stand til at mindske risikoen for, at ulovligt indhold og informationsmanipulation spredes på platformen.[3]

Efter Elon Musks overtagelse af X i 2022 kunne en række medier rapportere, at en stor del af platformens tilknyttede moderatorer blev fyret.[4] Samtidig lancerede X, at platformens egne brugere fremover skulle stå for at faktatjekke hinanden via funktionen “Fællesskabsnoter”, hvor brugere kan skrive noter, som sættes i bunden af et opslag, fx om at indholdet er falsk. Hvilke noter, der vises i bunden af et opslag, afgøres ved, at brugerene stemmer på noterne. Det er siden blevet mødt med kritik, at indholdsmoderation skal forvaltes af platformens brugere på og ikke professionelle indholdsmoderatorer. Det danske faktatjek-medie TjekDet har bl.a. kunnet dokumentere, hvordan langt de fleste noter, der markerer indhold som mis- og desinformation, aldrig bliver vist på platformen.

BOKS 2 **EU STILLER NYE KRAV TIL TECH-GIGANTERNES GENNEMSIGTIGHED UNDER DSA'EN**

Den nyligt tiltrådte EU-forordning om det indre marked for digitale tjenester (DSA) har pålagt onlinetjenester i EU en række forpligtelser for bl.a at fremme gennemsigtighed og ansvarlighed.[5] Der gælder særlige forpligtelser for onlineplatforme og onlinesøgemaskiner, der har mere end 45 mio. aktive brugere i EU om måneden. Forpligtelserne omfatter bl.a. krav om, at platformene to gange årligt offentliggør rapporter (gennemsigtighedsrapporter) om bl.a. tjenesternes indholdsmoderation.

Gennemsigtighedsrapporterne skal bl.a. indeholde oplysninger om, hvordan platformene indholdsmodererer, f.eks.

- oplysninger om, hvilke foranstaltninger deres praksis for indholdsmoderation består af,
- antallet af påbud, de har modtaget fra bl.a. administrative myndigheder,
- oplysninger om, hvor meget indhold der er fjernet,
- nøjagtigheden og fejlprocenten i deres automatiserede indholdsmoderationssystemer, og
- indholdsmoderationsholdets kvalifikationer og sproglige ekspertise.

De nye regler for at fremme gennemsigtighed og ansvarlighed omfatter også visse forpligtelser om at meddele brugere, hvorfor deres indhold er blevet fjernet, eller hvorfor adgangen til en konto er blevet begrænset, ligesom brugerne skal have mulighed for at anfægte disse og tilsvarende afgørelser.

Hertil kommer, at Kommissionen har iværksat en gennemsigtighedsdatabase, som indsamler og offentliggør begrundelser for begrænsning eller fjernelse af indhold eller brugere for at gøre det muligt at kontrollere de beslutninger om indholdsmoderation, der træffes af onlinetjenesterne. DSA-gennemsigtighedsdatabasen er offentligt tilgængelig og kan findes **her**.[6]

HVILKET INDHOLD SKAL PLATFORMENE MODERERE?

De nye EU-regler for digitale tjenester (DSA'en) skal bl.a. være med til at sikre, at borgerne ikke udsættes for ulovligt indhold på online platforme og søgemaskiner. Reglerne forpligter derfor platformene til at begrænse udbredelsen af ulovligt indhold og til at fjerne indholdet, så snart de bliver opmærksomme på, at indholdet eksisterer.

Ulovligt indhold omfatter alt indhold, der ikke er i overensstemmelse med EU-retten eller en medlemsstats lovgivning. Det kan være alt fra indhold med seksuelt overgrebsmateriale mod børn, salg af ulovlige eller forfalskede produkter eller ulovlig deling af private billeder.

Hertil skal platformene vurdere og forsøge at minimere "systemiske risici". Eksempler på systemiske risici kan være, hvis tjenesterne anvendes til at udbrede eller forstærke misvisende eller vildledende indhold. Udover risikovurderinger ift. at bekæmpe bl.a. spredningen af desinformation, er platformene også forpligtet til at gennemføre risikobegrænsende foranstaltninger – fx i form af indholdsmoderation.

Tech-giganternes har herudover egne regler, community guidelines, for hvilken slags indhold, de mener, brugerne må dele på deres platforme.

Med EU's skærpede krav til tech-giganternes gennemsigtighed under forordning om det indre marked for digitale tjenester stilles også krav til de gennemsigtighedsrapporter, platformene er forpligtet til at udgive, og som denne rapport undersøger. Eksempler på disse krav er beskrevet i Boks 2 ovenfor.

2

FLERE PLATFORME HAR INGEN ELLER FÅ DANSK-TALENDE MODERATORER

De store sociale medier opgør i deres gennemsigtighedsrapporter, hvor mange af de officielle sprog i EU deres moderatorer dækker. Rapporterne giver dermed også et indblik i, hvor mange af platformenes moderatorer der taler dansk.

LINKEDIN OG X HAR INGEN DANSKTALENDE MODERATORER

Ud af de otte store sociale medier, som inkluderes i denne undersøgelse, har to platforme, LinkedIn og X, ikke tilknyttet dansktalende moderatorer, jf. tabel 1. Pinterest har én moderator, som taler dansk, og TikTok har med sine 27 dansktalende moderatorer flest moderatorer med danske sprogkundskaber tilknyttet.

Der er med andre ord stor forskel på, hvor mange moderatorer der er hos de otte platforme, der kan varetage moderation af indhold på dansk.

Det er værd at bemærke, at nogle platforme opgør én moderator, som kan tale flere sprog, for hvert sprog moderatoren taler. Derfor vidner tallene ikke nødvendigvis om, hvor mange *unikke* moderatorer hver platform har, men de giver bl.a. et indblik i, hvor mange forskellige *sprog* platformenes indholdsmoderatorer dækker – og ikke dækker.

Tabel 1
Oversigt over indholdsmoderatorer med danske sprogkunskaber på store sociale medier

Platform	Antal moderatorer med danske sprogkundskaber
LinkedIn	0
X (Twitter)	0
Pinterest	1
Facebook*	6
Instagram*	6
Snapchat	15
YouTube	18
TikTok	27

Anm.: *Facebook og Instagram, der begge ejes af Meta, opgør antal af moderatorer samlet for begge platforme.
Kilde: Tech-giganternes gennemsigtighedsrapporter, udgivet apr. 2024.

ANTALLET AF MODERATORER ER SJÆLDENT PROPORTIONELT MED PLATFORMENS STØRRELSE

Vurderingen af, om tech-giganternes moderatorer på et givent sprog er tilstrækkeligt, må bl.a. ses i forhold til, hvor mange brugere platformen har i det pågældende land. Sammenlignes antallet af brugere med antallet af moderatorer, er der store forskelle mellem platformene.

Af de seks platforme, som har dansktalende moderatorer tilknyttet, er YouTube den mest udbredte platform i Danmark. Platformen har et antal moderatorer, som svarer til, at der er én moderator pr. ca. 290.000 brugere af tjenesten i Danmark jf. tabel 2. Til sammenligning har TikTok et antal moderatorer, som svarer til én moderator til ca. 52.000 aktive brugere.

Mens Instagram og Facebook opgør et antal moderatorer, som svarer til én moderator pr ca. 630.000 og 730.000 aktive brugere på platformen, skal det bemærkes at Meta kun opgør et samtalet antal moderatorer på tværs af begge platforme. Derfor vil hver af de to platforme i praksis forventeligt have færre moderatorer til at moderere indhold.

Tabel 2
Antal månedlige aktive brugere på store sociale medier i Danmark pr. moderator med danske sprogkundskaber

Platform	Antal moderatorer med danske sprogkundskaber	Aktive månedlige brugere i Danmark**	Antal månedlige aktive brugere i Danmark pr. moderator med danske sprogkundskaber
LinkedIn	0	1.400.000	Ingen moderatorer
X (Twitter)	0	750.744	Ingen moderatorer
Pinterest	1	1.200.000	1.200.000
Facebook*	6	4.400.000	733.333
Instagram*	6	3.800.000	633.333
Snapchat	15	2.677.066	178.471
YouTube	18	5.200.000	288.889
TikTok	27	1.400.000	51.852

Anm.: *Facebook og Instagram, der begge ejes af Meta, opgør antal af moderatorer samlet for begge platforme, **Her medregnes platformenes egne opgørelser af brugere eller konti, som har vores logget ind med en konto på platformen. Det faktiske antal af danskere, som anvender platformene, kan variere.
Kilde: Udregninger foretaget af Center for sociale medier, tech og demokrati baseret på tal fra tech-giganternes gennemsigtighedsrapporter der er udgivet apr. 2024.

Antallet af aktive brugere på en platform afgør ikke nødvendigvis, hvor meget og hvilken slags indhold, der deles på platformen, men det vidner om, at platformenes evner til at moderere indhold på det danske sprog prioriteres forskelligt, og at antallet af moderatorer sjældent er proportionelt med antallet af brugere.

Platformene måler og opgør tal for gennemsnitlige aktive månedlige brugere forskelligt, og der mangler i flere gennemsigtighedsrapporter tilstrækkelig information omkring, hvordan disse tal er målt.* Hvis en bruger anvender flere forskellige konti på en platform, kan de fx risikere at bliver talt med flere gange. Derfor skal der i denne sammenligning mellem platformene tages forbehold for de forskelligartede opgørelsesmetoder platformene imellem og den usikkerhed, det medfører. For at sikre det bedst mulige sammenligningsgrundlag i denne analyse, er der her afrapporteret platformenes opgørelse af, hvor mange brugere, der i gennemsnit har været logget ind med en konto, og har været aktive på platformen i en given periode (fx inden for den seneste måned).

I DSA'en lægges op til, at brugere, som ikke er logget ind eller ikke har en konto, men som blot besøger platformen, medregnes i aktive månedlige brugere af platformen, da denne type af brugere kan udgøre en væsentlig del af brugergruppen. Derfor er et forventeligt, at det faktiske antal af brugere, som anvender platformene, kan variere fra tech-giganternes egne afrapporterede tal.

* Platformene Pinterest, Instagram, Facebook og Snapchat opgør deres aktive månedlige brugere, som brugere eller konti, som er logget ind og er aktive på platformen. LinkedIn og YouTube opgør ikke et samlet tal for aktive månedlige brugere, men opgør 1) aktive brugere, som er logget ind på platformen og 2) besøg på deres platforme fra brugere, som ikke er logget ind. Her afrapporteres tal for aktive brugere, som har været logget ind på platformene. Selvom X opgør aktive månedlige brugere som både aktive brugere der er logget ind sammenlagt med aktive brugere ikke er logget ind, afrapporteres her kun tal for aktive brugere, der er logget ind på platformen, for at sikre bedst muligt sammenligningsgrundlag platformene imellem.

3

STOR VARIATION I INDHOLDSMODERATORER PÅ TVÆRS AF EU

EU er et område med stor sprogdiversitet, og dansk er blot ét af de 24 officielle EU-sprog, som tech-giganterne er forpligtet til at rapportere om.

MODERATORER HOS LINKEDIN OG X DÆKKER KUN OMKRING EN TREDJEDEL AF EU'S 24 OFFICIELLE SPROG

Kun Facebook og Instagram har moderatorer, som kan dække alle EU's 24 officielle sprog, jf. tabel 3. TikTok og YouTube har moderatorer på deres platforme, som kan tale 22 ud af de 24 officielle EU-sprog.

Til sammenligning dækker Snapchat og Pinterest ca. halvdelen og LinkedIn og X kun ca. en tredjedel af EU's 24 officielle sprog med deres moderatorers sprogkundskaber.



Tabel 3
Antal af moderatorer tilknyttet store sociale medier som har sprogkundskaber inden for de officielle EU-sprog

Sprog	Facebook	Instagram	TikTok	YouTube	Snapchat	Pinterest	X (Twitter)	LinkedIn
bulgarsk	20	20	43	19	0	0	0	0
dansk	6	6	27	18	15	1	0	0
engelsk	98	98	2334	17507	1148	365	1152	1570
estisk	3	3	7	7	0	0	0	0
finsk	15	15	34	24	6	0	0	0
fransk	211	211	650	439	228	15	31	58
græsk	23	23	65	45	0	0	0	0
irsk	39	39	0	0	0	1	0	0
italiensk	164	164	439	229	14	3	23	1
kroatisk	19	19	17	34	0	0	0	0
lettisk	2	2	10	7	0	1	0	0
litauisk	6	6	9	14	0	0	0	0
maltesisk	1	1	0	0	0	0	0	0
nederlandsk	52	52	162	132	20	2	8	1
polsk	66	66	201	353	4	0	10	0
portugisisk	45	45	172	370	16	10	34	25
rumænsk	35	35	136	93	5	0	1	0
slovakisk	12	12	38	11	0	0	0	0
slovensk	7	7	39	7	0	0	0	0
spansk	147	147	515	675	72	20	40	25
svensk	42	42	111	37	24	1	0	0
tjekkisk	18	18	57	73	0	0	0	0
tysk	223	223	837	352	73	17	23	61
ungarsk	24	24	47	52	0	0	0	0
Antal sprog, der dækkes af moderatorers sprogkundskaber**	**24**	**24**	**22**	**22**	**12**	**11**	**9**	**7**

Anm.: *Facebook og Instagram, der begge ejes af Meta, opgør antal af moderatorer samlet for begge platforme, **Hvis en tjeneste ikke har opgjort nogle moderatorer, som taler et specifikt sprog, antages det, at tjenesten har 0.
Kilde: Tech-giganternes gennemsigtighedsrapporter, udgivet apr. 2024

DE NORDISKE SPROG I EU DÆKKES ENS AF INDHOLDSMODERATORERS SPROGKUNDSKABER

De nordiske sprog dækkes i store træk ens på tværs af platformene, hvor mellem 2-3 platforme ikke har nogen moderatorer, som taler disse sprog, jf. tabel 4. Antallet af aktive brugere på platformene pr. moderator som taler landets mest udbredte modersmål er relativt ensartet for de tre sprog. De største forskelle ses hos Instagram og Facebook, der underprioriterer dansk i forhold til de øvrige nordiske sprog.

Samlet set er de tre nordiske sprog dårligst dækket af moderatorer ved X, LinkedIn, og Pinterest , sammenlignet med de øvrige platforme.

** Rapporterne, som denne undersøgelse bygger på, opgør alene tal for EU-lande. I det følgende dækker termen "de nordiske sprog" således over dansk, svensk og finsk, men ikke norsk og islandsk

Tabel 4
Sammenligning mellem Danmark, Sverige og Finland: Antal moderatorer samt månedlige aktive brugere på store sociale medier i hvert land pr. moderator, som taler landets mest udbredte modersmål.

	Danmark	Sverige	Finland
Platform	Antal brugere pr. moderator med danske sprogkundskaber (samt samlet antal moderatorer)	Antal brugere pr. moderator med svenske sprogkundskaber (samt samlet antal moderatorer)	Antal brugere pr. moderator med finske sprogkundskaber (samt samlet antal moderatorer)
X (Twitter)	Ingen moderatorer	Ingen moderatorer	Ingen moderatorer
LinkedIn	Ingen moderatorer	Ingen moderatorer	Ingen moderatorer
Pinterest	1.200.000 (1 moderator)	1.900.000 (1)	Ingen moderatorer
YouTube	288.889 (18)	281.081 (37)	254.166 (24)
Instagram*	733.333 (6)	171.429 (42)	213.333 (15)
Facebook*	733.333 (6)	171.429 (42)	213.333 (15)
Snapchat	178.471 (15)	182.355 (24)	285.491 (6)
TikTok	51.851 (27)	29.729 (111)	47.058 (34)

Anm.: *Facebook og Instagram, der begge ejes af Meta, opgør antal af moderatorer samlet for begge platforme
Kilde: Beregninger foretaget af CSTD pba. tech-giganternes gennemsigtighedsrapporter udgivet apr. 2024.

DE MEST UDBREDTE MODERSMÅL I NORD- OG ØSTEUROPA ER DÅRLIGERE DÆKKET AF PLATFORMENES INDHOLDSMODERATORER END I VEST- OG SYDEUROPA

Ikke alle lande og geografiske områder på tværs af EU dækkes lige godt af indholdsmoderatorernes sprogkundskaber.

Alle otte platforme har moderatorer ansat, som taler de mest udbredte modersmål i alle vest- og sydeuropæiske lande i EU (Belgien, Frankrig, Holland,

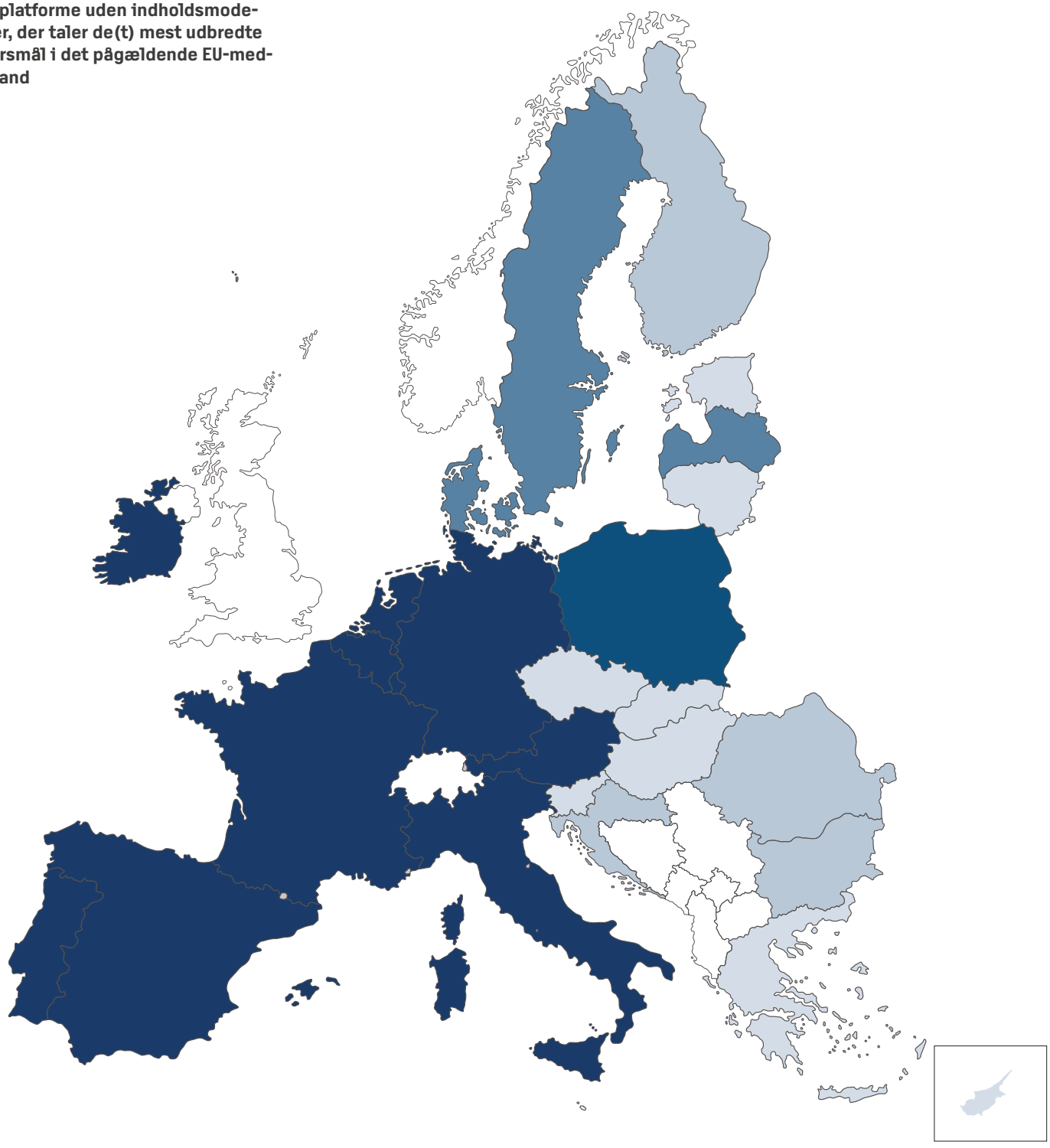
Italien, Irland, Spanien, Portugal og Østrig), med undtagelse af Cypern, Grækenland og Malta jf. figur 1.

Til sammenligning, er særligt de mest udbredte modersmål i de nord- og østeuropæiske lande i EU dårligt dækket af platformenes indholdsmoderatorers sprogkundskaber, hvor minimum to platforme ikke har moderatorer, som taler de pågældende landes mest udbredte sprog.

For Estland og Litauen, samt Tjekkiet, Ungarn, Slovenien og Slovakiet gælder det, at fire ud af otte platforme ikke har nogen moderatorer, som taler landets mest udbredte modersmål.

Cypern, Grækenland og Malta, hvor de mest udbredte modersmål er græsk og maltesisk, er ligeledes dårligt dækket af platformenes moderatorers sprogkundskaber, hvor kun hhv. fire og seks platforme har moderatorer, som taler sproget.

Figur 1
Antal platforme uden indholdsmoderatorer, der taler de(t) mest udbredte modersmål i det pågældende EU-medlemsland



Anm.: I Luxemborg er det mest udbredte modersmål Luxemborgish. Da dette ikke er et af de 24 officielle sprog, regnes der her med de næst mest udbredte modersmål nemlig fransk og tysk.
Kilde: Tech-giganternes gennemsigtighedsrapporter udgivet apr. 2024


Ctrl

4

FORSKELLIGARTEDE OPGØRELSER KOMPLICERER SAMMENLIGNING

Selvom gennemsigtighedsrapporterne giver indsigter i, hvilke sprog platformenes moderatorer taler, viser gennemgangen af rapporterne, at tech-giganterne ofte måler og afrapporterer disse tal forskelligt. Det kan derfor i praksis være svært at sammenligne platformenes tilgang til indholdsmoderation på de forskellige sprog meningsfuldt:

Det er ikke tydeligt, hvor mange unikke moderatorer alle platforme har tilknyttet, og hvor mange moderatorer som tælles med flere gange, fordi de taler flere sprog.

Det er uklart, hvor mange timer en moderator skal dedikere til at moderere indhold på et bestemt sprog, før de medregnes. Fx medregner YouTube i deres rapport moderatorer, som på platformen har gennemset mindst 10 videoer i rapporteringsperioden på seks måneder, mens andre platforme blot oplister moderatorer, som kan tale et givent sprog.

Platformene tæller og opgør aktive månedlige brugere forskelligt. Hvor nogle platforme skelner mellem aktive brugere, som er logget ind eller ikke er logget ind på tjenesten, opgør andre tjenester ét samlet tal. Det gør det vanskeligt at sammenligne, hvor udbredte tjenesterne er i et bestemt land, og dermed om der er tilstrækkeligt med moderatorer på et givent sprogområde.

Platformene opgør ikke i deres rapporter, hvor meget indhold der deles eller bliver modereret på forskellige sprog på platformen. Det gør det svært at sammenligne platformenes faktiske moderation af indhold på forskellige sprog.

Platformene rapporterer over forskellige perioder, som varierer mellem to til seks måneder.

Den manglende standardisering for, hvordan tech-giganterne skal tælle og opgøre tal i deres gennemsigtighedsrapporter, gør det vanskeligt at sammenligne de forskellige platforme og mindsker den reelle gennemsigtighed i platformenes arbejde med indholdsmoderation.[7]

Derfor kan denne undersøgelse, der tager udgangspunkt i tech-giganternes afrapporterede antal moderatorer på papiret, ikke vurdere platformenes egentlige prioritering af arbejdet med at moderere indhold. Undersøgelsen kan kaste lys over, hvorvidt de har moderatorer og dermed forudsætninger for at moderere indhold på et bestemt sprogområde.

Først og fremmest viser undersøgelsen, hvor mange sprogområder, der ikke er dækket af moderatorer fra de otte udvalgte platforme.

Undersøgelsen peger dertil på et overordnet behov for, at tech-giganterne i fremtiden giver øget indblik i deres indholdsmoderation på en gennemsigtig og standardiseret måde, så grundlaget for fremtidige undersøgelser er ensartet.

I december 2023 har Europa-Kommissionen udgivet et udkast til en skabelon for tech-giganternes fremtidige gennemsigtighedsrapporter, som fremadrettet skal sikre, at rapporterne bedre kan sammenlignes, og at tech-giganterne lever op til deres rapporteringsforpligtigelser.[8] Der lægges bl.a. her op til, at tech-giganterne skal rapportere mere detaljeret end nu om moderatorer, som har de nødvendige sprogkundskaber til at moderere på forskellige sprog, ligesom der lægges op til ensartede rapporteringsperioder og opgørelser over, hvilket indhold, der fjernes.

I tillæg til dette, foreslår Center for sociale medier, tech og demokrati en række anbefalinger til øget gennemsigtighed i tech-giganternes arbejde med indholdsmoderation, jf. Boks 3:

Det er ikke afgørende, om en given platform har fem eller ti dansktalende moderatorer, hvis det ikke er tydeligt, hvor meget tid disse moderatorer bruger på moderationsarbejdet, og om de samtidig er angivet som moderatorer for andre sprog. Derfor skal der stilles strengere krav til rapporteringen, hvilket allerede er en del af kommissionens arbejde. Selvom gennemsigtigheds-rapporterne er et skridt i den rigtige retning, så giver de endnu ikke den nødvendige indsigt i tech-giganternes moderationspraksis for at borgere, virksomheder og myndigheder kan have en oplyst samtale om tech-giganternes påvirkning af samfundet og den demokratiske samtale online.

BOKS 3
ANBEFALINGER TIL ØGET GENNEMSIGTIGHED OG STANDARDISERING I TECH-GIGANTERNES GENNEMSIGTIGHEDSRAPPORTER

For at sikre at gennemsigtighedsrapporterne, muliggør for myndigheder, aktører og interesserede borgere at forstå og sammenligne tech-giganternes arbejde om bla. indholdsmoderation, er det nødvendigt at stille mere specifikke krav til rapporternes indhold. Derfor anbefaler Center for sociale medier, tech og demokrati følgende krav til rapporterne:

- Minimumskrav for hvor mange moderatorer tech-giganterne har i de lande, som de er til stede i.
- Fælles og sammenlignelige opgørelsesmetoder til opgørelse af moderatorer, så antallet bedre kan sammenlignes på tværs af platforme.
- Tydelige og præcise beskrivelser af, hvordan tech-giganterne indholdsmodererer.
- Opgørelser over, hvor stor andel af det indhold, der er lagt op inden for perioden, der henholdsvis anmeldes, mærkes og fjernes på forskellige sprog.
- Tech-giganternes tidligere (og ikke kun nyeste) gennemsigtighedsrapporter og afrapporterede tal skal være tilgængelige.
- Krav til ensartede rapporteringsperioder.

REFERENCER

1 Sharevski et al. 2022: “Misinformation Warnings: Twitter’s Soft Moderation Effects on COVID-19 Vaccine Belief Echoes”

2 Drolsbach og Pröllochsouek 2023: “Content Moderation on Social Media in the EU: Insights From the DSA Transparency Database”

3 Europa-Kommissionen 2023: ”Press release: Commission opens formal proceedings against X under the Digital Services Act”. Udgivet d. 18. december 2023

4 Tjekdet 2024: ”Moderatorerne er fyret: Nu skal brugerne kontrollere brugerne på X”

5 Europa-Parlamentets og Rådets forordning (EU) om et indre marked for digitale tjenester (DSA’en)

6 Europa-Kommissionens DSA Transparency Database

7 Global Witness 2023: “How Big Tech platforms are neglecting their non-English language users”

8 Europa-Kommissionen 2023: “Digital Services Act – transparency reports (detailed rules and templates)”

OM CENTER FOR SOCIALE MEDIER, TECH OG DEMOKRATI

Center for sociale medier, tech og demokrati er etableret som en del af Medieaftalen 2023-2026. Centret er placeret i Slots- og Kulturstyrelsen, som i forvejen yder faglig rådgivning til Kulturministeriet på medie- og tech-området, og som beskæftiger sig med digitale medier, bl.a. gennem sekretariatsbetjening af Medierådet for Børn og Unge.

Centrets arbejdsopgaver omfatter bl.a. at bidrage med viden om brugernes mentale trivsel, om tech-giganternes betydning og konsekvenser for danske medier, og om hvilken indvirkning spredning af mis- og desinformation på digitale platforme har på den demokratiske samtale.

Læs mere om centret og dets øvrige udgivelser **her**.