

# **"ONE CUT HERE, ONE CUT THERE"**

## **ANALYSE AF TIKTOKS TIL DIG-ALGORITME**

En dansk undersøgelse af, hvad TikTok eksponerer brugere med interesse for selvskade, selvmord og psykisk lidelse på deres Til dig-side.

# Indholds- fortegnelse

Indledning	3.
Opsummering	4.
Fremgangsmåde	5.
Resultater	6.
Diskussion	9.
Konklusion	10.
Anbefalinger	11.
Noter	12.
Bilag	13.



# Indledning

”One cut here, one cut there, no one will see, and no one will care”. Sådan lyder en kunstigt genereret stemme over en grynet video af en ung kvinde med et udtryksløst ansigt. Et swipe og næste video i feedet viser en hospitalsseng med teksten ”pov: you were too honest with the nice psychologist”.

Det er eksempler på indhold, børn og unge med interesse for selvska- og depressionsrelateret indhold bliver eksponeret for på TikTok.

Flere undersøgelser peger på, at børn og unge mistrives. Af og til i så voldsom grad, at de selvskader eller forsøger selvmord.

Fra 2009 til 2019 er omfanget af 15–24-årige, der selvskader, steget med 58%.<sup>1</sup> For 15–19-årige er selvmordsraten steget siden 2011,<sup>2</sup> og antallet af selvmordsforsøg blandt børn i alderen 0–14 år steget fra 835 i 2020 til 1454 i 2021.<sup>3</sup>

Sideløbende har der været en voksende opmærksomhed på sociale mediers rolle og det indhold og de netværk om selvska- og selvmord, som eksisterer på sociale medier.

For hvad er det, børn møder, når de kommer på et socialt medie som TikTok? Og hvad betyder TikToks algoritmer for dem, der i forvejen udviser interesse for selvska- og depressionsrelateret indhold?

TikTok er på få år blevet et meget populært

medie – både på verdensplan og i Danmark. I Danmark er de ugentlige brugere i alderen 15–29 år steget fra 3% i 2019 til 41% i 2022. Selvom TikTok har en aldersgrænse på 13 år, er det de 9–14-årige, der udgør den mest aktive gruppe af brugere i Danmark.

Da mediet ikke benytter sig af aldersverifikation, er det nemt at omgå. I gennemsnit bruger de danske børn på 9 til 14 år 1 time og 18 minutter dagligt på TikTok.<sup>4</sup>

I denne analyse ser vi nærmere på, hvilket indhold, sårbare profiler på TikTok bliver anbefalet.

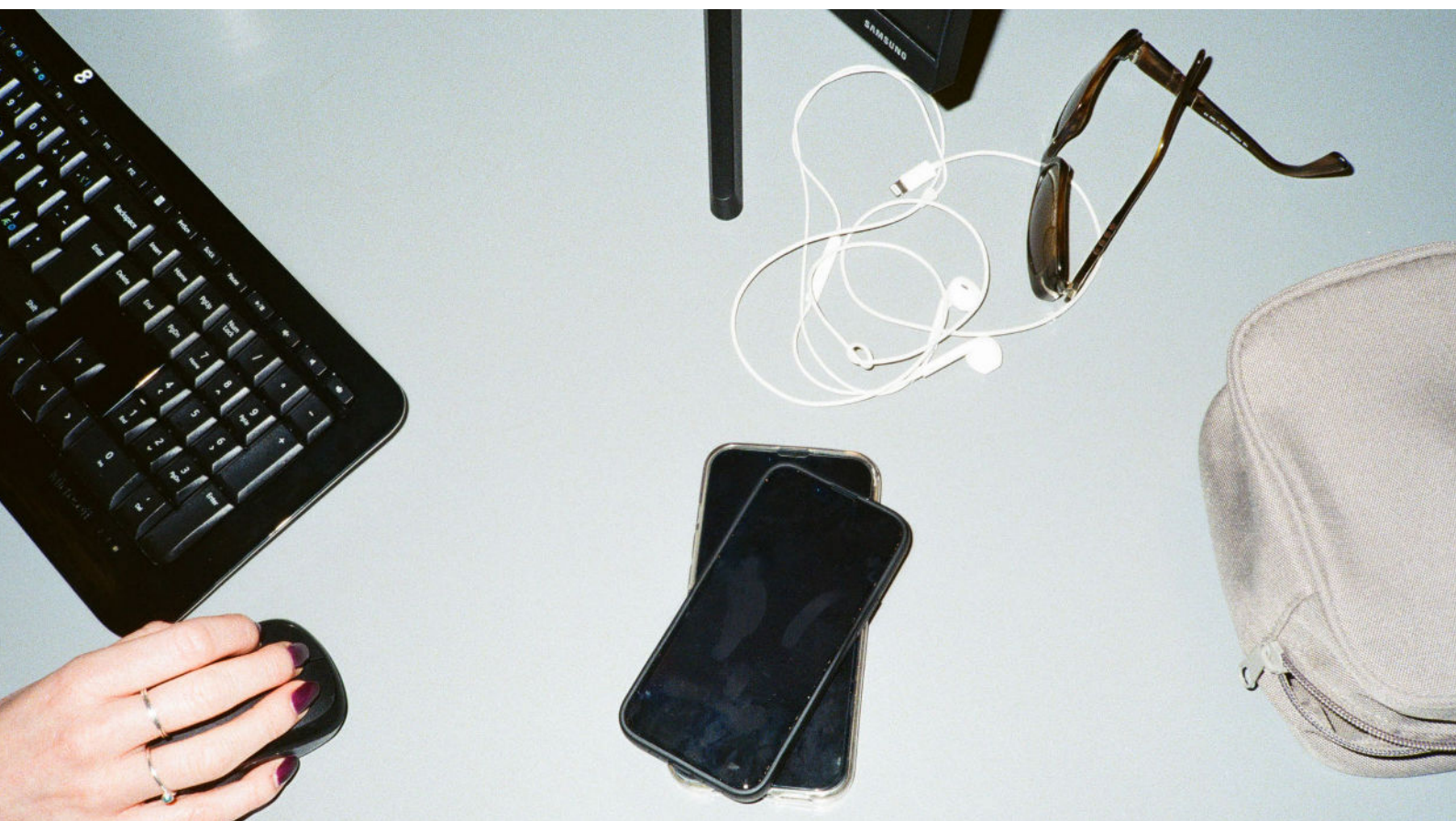
Hvorvidt TikTok spiller en rolle i den store mistrivsel blandt børn og unge kan vi ikke med nærværende undersøgelse sige noget om – derimod undersøger vi det univers, børn og unge med interesse for selvska- og selvmord og psykiske lidelser bliver præsenteret for på tjenesten. Den gruppe af brugere kalder vi i analysen for sårbare brugere.

Lignende undersøgelser er tidligere foretaget af blandt andet den amerikanske NGO Center for Countering Digital Hate samt Amnesty International.<sup>5</sup> Arbejdet med selvska- og selvmordsrelateret indhold ligger desuden i forlængelse af en rapport, vi i marts 2023 udgav om digital selvska- og selvmord, som bl.a. viste, at 2% af voksne danskere indenfor det seneste år har delt dokumentation af fysisk selvska- og selvmord.<sup>6</sup>



# Opsummering

- \_01** TikToks algoritme er designet til at give mere af den type indhold, som den enkelte bruger interagerer med, også selvom det er indhold om selvskade, selvmord, psykisk sygdom og svære følelser rettet mod individet selv.
- \_02** Hvor hurtigt en sårbar brugers Til dig-side ender med en høj frekvens af depressions- og selvskaderelateret indhold, afhænger af, hvornår brugeren for første gang eksponeres for en video, som falder inden for denne kategori.<sup>7</sup>
- \_03** Ud af fire sårbare profiler har to af profilerne allerede efter et kvarter et feed, der i intervaller eksponerer brugeren for mellem 70-80% depressions- og selvskaderelateret indhold. De to andre profiler er på samme tidspunkt få minutter fra at ende i et lignende feed.
- \_04** Efter første eksponering bliver de sårbare brugeres Til dig-side dog alle præget af en høj frekvens af selvskade- og selvmordsrelateret indhold – i kortere sekvenser på op til 80-100%.
- \_05** Det depressions- og selvskaderelaterede indhold, forsøgets brugere er blevet eksponeret for på TikTok, består kun sjældent af ord som selvskade og selvmord. Brugere er heller ikke blevet eksponeret for visuelle gengivelser af udført selvskade. Derimod refereres til fænomenerne gennem omskrivninger, der dog ofte er ret entydige.



# Fremgangsmåde

## Til dig-siden

På TikTok er Til dig-siden eller For You-siden der, hvor brugerne bruger langt størstedelen af deres tid. Det er en helt central del af oplevelsen af at være på TikTok, som løbende tilpasses den enkelte bruger. Det er samtidig det første, nyoprettede brugere eksponeres for. TikTok oplyser selv, at algoritmen blandt andet baserer deres anbefalinger på følgende input:

- Brugerens interaktioner såsom delinger, likes, kommentarer, hvem brugeren følger, og indhold, brugeren selv skaber.
- Videoinformation som billedtekster, lyd og hashtags.
- Brugerindstillinger og enhedens indstillinger som sprog, område og enhedens type.<sup>8</sup>

Til dig-siden er bygget op, så den kræver mindst mulige valg af brugeren. I stedet tilpasser den sig selv til det, den opfanger som brugerens interesse. Det vil sige, at du ikke behøver lede efter profiler eller selv opsøge det indhold, du leder efter. TikTok består dog også af en række andre funktioner udover For You-siden, som også har betydning for brugeroplevelsen – herunder søgefunktion, mulighed for likes, kommentarer med videre.

## Opsætning af telefoner

Vi benytter os af rensede telefoner med nye simkort. Alle telefoner er nye smartphones, som ikke har været anvendt før, som er stillet til rådighed af Telia. Det samme gør sig gældende for de anvendte simkort. Ved opsætning af telefonerne vælger vi ikke at

oplyse køn og sætter fødselsdato tilsvarende den brugerprofil, vi ønsker at oprette med telefonen – dvs. alder på enten 13 eller 18 år. Sprog sættes til dansk og lokation indstilles automatisk på telefonen. Vi opretter en ny Google-konto med dertilhørende gmail til hvert simkort.

## Oprettelse af profiler

For at oprette en profil på TikTok, kræves følgende oplysninger om dig: Alder, telefonnummer eller e-mail og brugernavn. Lokationen henter den selv gennem dine indstillinger. Benytter man en rensede telefon med en anonym email, er det altså de tre informationer, TikToks algoritme har til at tilpasse indhold til dig, før tjenesten har indsamlet data om din adfærd på platformen.

Hver bruger oprettes på en rensede telefon med indstillinger på dansk med googlekontoen tilknyttet det nye telefonnummer og en anonym email. De sårbare profiler giver vi et brugernavn, der angiver en interesse i selvskade.

Vi opretter i alt 6 profiler: To kontrolprofiler med alder på hhv. 13 og 18 år, samt to sårbare profiler på 13 år og to sårbare profiler på 18 år.

- Brugernavne for sårbare profiler: Selvskader + tilfældigt tal
- Brugernavne for kontrolgruppe: Tilfældig kombination af tal og bogstaver

## Adfærd for sårbare profiler

Sårbare profiler dækker i dette tilfælde over kunstige profiler, vi har oprettet til forsøg-

get.<sup>9</sup> Ved oprettelse på TikTok angiver vi ikke kategorier af indhold, som profilen er særligt interesseret i.

Profilerne er defineret af en interesse for selvskaade, selvmord, psykiske lidelser og spiseforstyrrelser. Herunder angst, depression dårligt selvværd og andre svære følelser knyttet til sig selv.

For profilernes adfærd på platformen betyder det, at når en video, der omhandler selvskaade, selvmord, psykiske lidelser og spiseforstyrrelser, herunder angst, depression dårligt selvværd og andre svære følelser knyttet til sig selv, identificeres på Til dig-siden, afspilles den i 10 sekunder og markeres med et "like". For alle andre videoer inkl. reklamer scrolles umiddelbart videre.

Profilerne liker kun indhold, der har et dystert og destruktivt fokus på den type indhold – dvs. den ikke liker for eksempel

bedringshistorier, motivation til at få det bedre med videre.

Profilerne liker ikke videoer, der drejer sig om lidelse, der relaterer sig til relationer – som for eksempel til ekskæreste, forældre eller skolekammerater.

## Adfærd for kontrolgruppe

Kontrolgruppen scroller i ensartet hastighed på 10 sekunder pr. video i indhold på dansk og engelsk. Videoer på andre sprog scroller vi umiddelbart videre fra.

## Optagelse og kategorisering af videoerne

Vi benytter telefonens skærmoptagelsesfunktion til at optage den første time, profilerne er på Til dig-siden. Bagefter ser vi indholdet igennem og registrerer de videoer, som den sårbare profil er sat til at like.

# Resultater

## Hvilket indhold bliver vi præsenteret for?

Det selvskaade- og selvmordsrelaterede indhold, de sårbare brugere i forsøget er blevet eksponeret for, benytter sig sjældent direkte af ord som selvmord, suicide mv. I stedet refereres der til fænomenerne – ofte ret åbenlyst – ved hjælp af mere dagligdags termer med videre.

På tekst bliver selvmord og selvskaade blandt andet refereret til med termer som for eksempel "depression will probably win", "I'm ready to grow my wings" eller "no one gives a shit abt my mental health till it started showing on my skin". Ofte koblet med

mørke billeder, en hospitalsseng, et grædende ansigt eller lignende.

Nogle gange fremgår referencen til selvskaade, selvmord, psykiske lidelser mv. ikke visuelt gennem hverken billeder eller tekst. Det kan for eksempel være en video af et umiddelbart udtryksløst ansigt – men hvor en voice over fortæller "one cut here, one cut there, no one will see, and no one will care" eller musik spiller med sangteksten "pill diet, pill diet (...) if they say to kill yourself, then you will try it".

I andre tilfælde kræver det mere forhåndsviden at afkode. Når for eksempel der står "opening the back of ur phone case after a bad day" på en video, skal man kende til



trenden med at gemme barberblade i telefoncovers for at forstå, at det er en reference til selvskaede.

Omskrivningerne er sandsynligvis en måde at omgå TikToks moderation. Moderationen fungerer, når det kommer til visuelle gen- givelser af selvskaede. De sårbare brugere bliver ikke eksponeret for blodige billeder eller andre måder at skade sig selv eller begå selvmord.

## Fra første like til et univers af svære følelser

I figuren på næste side kan vi se, at fre- kvensen af selvskaede- og selvmordsrelate- ret indhold stiger i takt med, at vi scroller og liker denne type indhold. Hver streg i figuren er en video med selvskaede- og selv- mordsrelateret indhold. X-aksen viser antal videoer, som vi har swipet igennem.<sup>10</sup>

Vi kan se fra figuren, at de to kontrolprofiler af og til bliver præsenteret for selvskaede- og selvmordsrelateret indhold, men da vi ikke liker videoerne, stiger frekvensen ikke.

De fire sårbare profiler viser derimod, hvad der sker, når algoritmen udleder, hvilken type indhold personaen interagerer med – den giver mere af det samme: Flere videoer om selvskaede, selvmord, psykisk sygdom og svære følelser.

Det er forskelligt fra sårbar profil til sårbar profil, hvor hurtigt Til dig-algoritmen udle- der personaens præferencer. Dette hænger formodentligt sammen med, hvornår per- sonaen første gang bliver eksponeret for selvskaede- og selvmordsrelateret indhold: I starten af en nyoprettet TikTok-profil prø- ver Til-dig-algoritmen forskellige indholds- typer af, og personaliseringen kan først begynde, når den er i stand til at udlede, at brugeren i højere grad interagerer med en bestemt type indhold.

I ”profil 18 år sårbar 2” og ”profil 13 år sår- bar 1” bliver de hurtigere eksponeret for selvskaede- og selvmordsrelateret indhold, hvilket har den effekt, at disse profiler også hurtigere får en Til-dig-algoritme, hvor frekvensen af selvskaede- og selvmordsrela- teret indhold er høj.

Hvor hurtigt en sårbar brugers Til dig-side ender med en høj frekvens af selvskaede- og selvmordsrelateret indhold, afhænger altså af, hvornår brugeren for første gang eks- poneres for en video, som falder inden for denne kategori. Efter første eksponering bliver alle de sårbare brugeres Til dig-side præget af en høj frekvens af selvskaede- og selvmordsrelateret indhold.

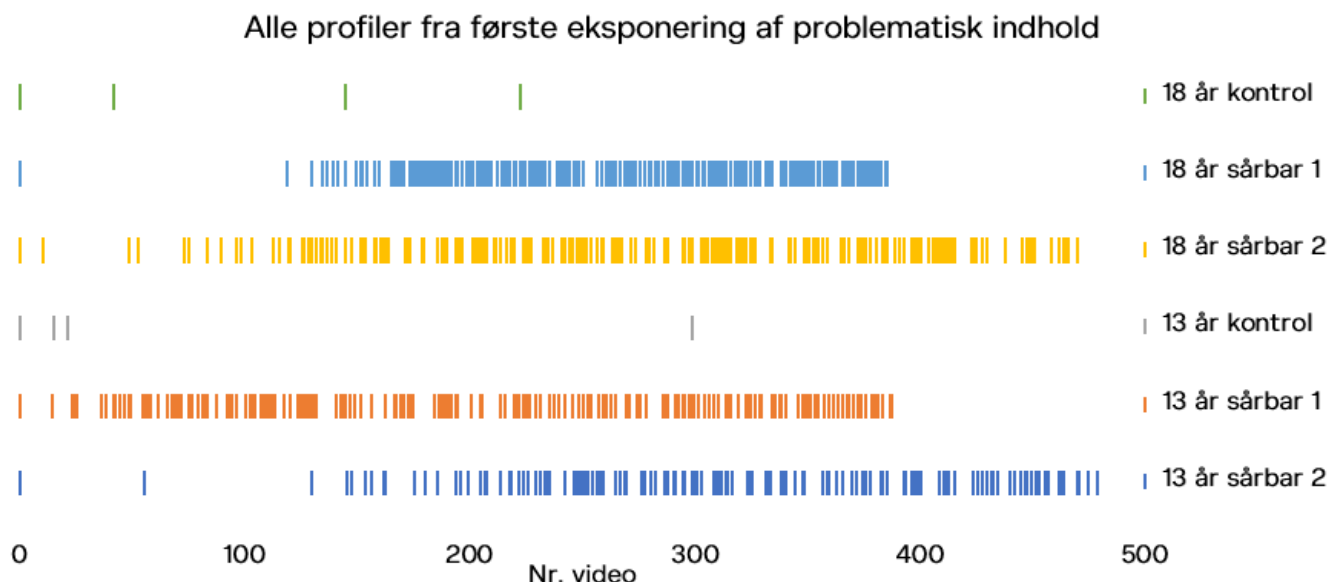
## Men hvor stor en andel af det indhold, man eksponeres for af Til dig-algoritmen, kan karakteriseres som selvskaede- og selv- mordsrelateret?

Vi har udregnet et glidende gennemsnit af andelen af selvskaede- og selvmordsrelate- ret indhold.<sup>11</sup> Figureerne for de fire profiler er lagt i bilag.

Ved første øjekast ser vi samme mønster som ved foregående figur, at tidspunktet for, hvornår Til dig-algoritmen begynder at eksponere brugeren for selvskaede- og selv- mordsrelateret indhold, varierer. Derudover får vi et yderligere indblik i, hvor stor en andel af det samlede indhold, som brugerne møder, er selvskaede- og selvmordsrelateret i takt med, at vi scroller.

Den maksimale andel af selvskaede- og selv- mordsrelateret indhold, som de fire profiler eksponeres for i en sekvens på 10 videoer, er henholdsvis 90%, 80%, 100% og 90% af indholdet.

## Frekvens af selvskade- og selvmordsrelateret indhold opgjort fra første video



Frekvensen af selvskade- og selvmordsrelateret indhold er ikke konsekvent så høj, i visse perioder falder andelen af selvskade- og selvmordsrelateret indhold helt ned 10-20% igen – men vores indsamlede data viser, at algoritmen vender tilbage til at udvælge og præsentere indhold i store mængder om selvskade, selvmord, svære følelser og psykisk sygdom. Også til profiler, der angiver en alder på tretten år.

Forskellene mellem profilerne på 18 år og 13 år er så små, at vi ikke kan udlede noget entydigt ud fra disse forskelle pba. dette datagrundlag. Dog er der indikationer på, at profilerne på 18 år får en anelse mere problematisk indhold præsenteret af algoritmen.

**For de fire sårbare profiler kommer frekvensen af indhold relateret til selvskade, selvmord, psykiske lidelser og svære følelser knyttet til sig selv i perioder op på**

# 80-100%



## Hvornår tager algoritmen fat?

En anden vigtig dimension vi kan måle TikToks Til dig-algoritme på, er hvor hurtigt den tager fat.

Som tidligere nævnt er denne dimension afhængig af, hvornår brugeren første gang bliver eksponeret for problematisk indhold, da algoritmen først der kan udlede, at brugeren er interesseret i denne type indhold.

Man kan derfor ikke opstille en fast regel for, hvor hurtigt brugeren ender i et feed, der består af en høj grad af problematisk indhold, men figur 1 viser, at det på tværs af de fire profiler, går enormt stærkt.

To af profilerne har allerede efter et kvarter et feed, der i intervaller eksponerer brugeren for mellem 70–80% problematisk indhold. De to andre profiler er på samme tidspunkt få minutter fra at ende i et lignende problematisk feed.

## Diskussion

Vores resultater viser, at Til dig-algoritmen kan eksponere voksne og børn for en bekymrende stor mængde indhold omhandlende selvskaade, selvmord, psykisk sygdom og svære følelser.

En legitim indvending mod disse resultater er, at vores forsøg mangler ekstern validitet, dvs. der kan sættes spørgsmålstegn ved, hvorvidt disse resultater kan generaliseres til den virkelige verden og en organisk brug af TikTok: Få, hvis nogen, TikTok-brugere liker kun én type indhold ligesom vores personaer gjorde. Til dette vil vi indvende, at dette eksperiment nødvendigvis ikke kun har værdi, hvis resultaterne illustrerer, hvordan de fleste brugere af TikTok får de-

Disse resultater er tankevækkende, set i lyset af mængden af tid danske børn bruger på TikTok dagligt: Danske børn fra 9 til 14 år anvender i gennemsnit TikTok 1 time og 18 minutter hver dag.

Et barn, der opretter TikTok for første gang kan derfor allerede efter et kvarter risikere at være have et Til dig-feed, der hovedsageligt består af problematisk indhold. Og barnet vil kunne blive eksponeret for selvmord, selvskaade og psykisk sygdom en hel time mere den dag – og det er altså første dag med TikTok installeret på telefonen.

Næste dag hvor samme barn åbner telefonen, vil det møde samme feed, idet algoritmen ikke starter forfra. En ny dag med potentielt 1 time og 18 minutters indhold om selvskaade, selvmord, psykisk sygdom, angst og depression er startet.

res indhold kurateret.

I stedet kan eksperimentet forstås, som en form for "red teaming" af Til dig-algoritmen. Hvad er algoritmen villig til at eksponere både børn og voksne for af selvskaade- og depressionsrelateret indhold for at maksimere tiden og opmærksomheden, der bruges på platformen?

Her er vores resultater entydige; algoritmen har få eller ingen sikkerhedsmekanismer ift. mængden af selvskaade- og selvmordsrelateret indhold, som den vil eksponere brugere for. Så spørgsmålet er ikke, hvorvidt disse resultater afspejler, hvordan de fleste børn får kurateret deres indhold på deres

Til dig-side, men hvorvidt vi skal acceptere, at det overhovedet er muligt for børn og voksne at ende med et feed, hvor der præsenteres en stor mængde af potentielt skadeligt indhold.

## Konklusion

Vores undersøgelse viser, at TikToks algoritme præsenterer brugere for mere af den type indhold, som brugeren bruger tid på og interagerer med. Og det gælder også, når der er tale om dystert og destruktivt indhold, der bl.a. omhandler selvskade, selvmord, psykisk sygdom og svære følelser rettet mod individet selv.

Indholdet, der omhandler selvskade, selvmord, psykisk sygdom og svære følelser mod sig selv, består sjældent af blodige selvskadebilleder eller direkte referencer til selvmord. Henvisningerne til for eksempel selvmord og selvskade optræder i TikToks indhold indirekte, men tydeligt. På tekst kan det for eksempel være, at et forestående selvmord beskrives med ordene "depression will probably win" eller "I'm ready to grow my wings".

Efter første eksponering for en depressions- eller selvskaderelateret video ledes brugeren mod et feed, som i perioder består af 80-100% depressions- og selvskaderelateret indhold. Da det varierer, hvor hurtigt en bruger eksponeres for en depressions- eller selvskaderelateret video, varierer det også, hvor hurtigt en sårbar bruger Til dig-side ender med en høj frekvens af depressions- og selvskaderelateret indhold.

Ud af fire sårbare profiler har to af profilerne allerede efter et kvarter et feed, der i intervaller eksponerer brugeren for mellem 70-80% depressions- og selvskaderelateret

Og skal børn overhovedet, uanset mængden, eksponeres for indhold, der romantiserer selvmord?

indhold. De to andre profiler er på samme tidspunkt få minutter fra at ende i et lignende feed.

Endelig ser det ikke ud til, at Til dig-algoritmen har sikkerhedsmekanismer, der forhindrer brugere i at blive eksponeret for en høj grad af indhold, der omhandler selvskade, selvmord, psykisk sygdom og svære følelser. Selvom de lykkes med at fjerne visuelle repræsentationer af selvskade og direkte referencer til selvmord, ender vores profiler alligevel i et feed, der er stærkt præget af fænomenerne.

Ud fra vores undersøgelse ser det ikke ud til at alder spiller en stor rolle i forhold til det indhold, som brugeren eksponeres for. Disse resultater er tankevækkende, set i lyset af mængden af tid danske børn bruger på TikTok dagligt: Danske børn fra 9 til 14 år anvender i gennemsnit TikTok 1 time og 18 minutter hver dag. Et barn, der opretter TikTok for første gang kan derfor allerede efter et kvarter risikere at være have et Til dig-feed, der hovedsageligt består af problematisk indhold. Og barnet vil kunne blive eksponeret for selvmord, selvskade og psykisk sygdom en hel time mere den dag – og det er altså første dag med TikTok installeret på telefonen. Næste dag hvor samme barn åbner telefonen, vil det møde samme feed, idet algoritmen ikke starter forfra. En ny dag med potentielt 1 time og 18 minutters indhold om selvskade, selvmord, psykisk sygdom, angst og depression er startet.

# Anbefalinger

- \_01** TikTok og andre sociale medier pålægges af EU at indføre aldersverifikation. Selvom aldersgrænsen på TikTok er 13 år, er der ingen sikkerhedsmekanismer, der forhindrer yngre børn i at oprette sig som bruger.
- \_02** TikTok bør kuratere et særligt feed til børn og unge. I kombination med aldersverifikation vil en særlig Til dig-algoritme for børn og unge under 18, som stiller større sikkerhedskrav til indholdet, mere effektivt kunne beskytte børn for negativ påvirkning.
- \_03** TikTok og andre sociale medier skal pålægges af EU at implementere "guardrails". Guardrails er mekanismer, der sikrer, at algoritmen opererer etisk forsvarlig, sikkert og lovligt, og som minimerer risikoen for eksponering af skadeligt indhold og misbrug af data (Privacy and Safety by Design).

Anbefalingerne støttes af Børns Vilkår, Center for Digital Pædagogik, Dansk Stalking Center, Joan-Søstrene, Mediesundhed for børn og unge, Red Barnet, Sex og Samfund, Søstre mod vold og kontrol og UNICEF.

**BØRNS  
VILKÅR**

**SØSTRE**  
MOD VOLD OG KONTROL

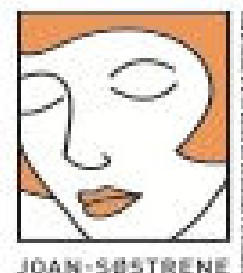
**SEX**  
SAMFUND



**Center for  
Digital Pædagogik**



**Red Barnet**



# Noter

1 <https://ugeskriftet.dk/videnskab/selvskade-er-et-stigende-problem-men-der-er-hab-for-ude>

2 <https://selvmordsforskning.dk/wp-content/uploads/sites/2/2022/06/Nr-46-Selv-mord-2011-2021.pdf>

3 <https://nyheder.tv2.dk/samfund/2022-12-27-eksperter-bekymrede-over-udvikling-i-selv-mordsforsoeg-blandt-boern-her-er-deres-raad>

4 <https://www.dr.dk/om-dr/fakta-om-dr/medieforskning/medieudviklingen/2022/tik-tok-tordner-frem>

5 [https://counterhate.com/wp-content/uploads/2022/12/CCDH-Deadly-by-Design\\_120922.pdf](https://counterhate.com/wp-content/uploads/2022/12/CCDH-Deadly-by-Design_120922.pdf) samt <https://www.amnesty.org/en/wp-content/uploads/2023/11/POL4073502023ENGLISH.pdf>

6 <https://digitaltansvar.dk/wp-content/uploads/2023/03/Digital-vold-i-Danmark.-Digital-selvskade.pdf>

7 Sårbare brugere er brugere med interesse for selvskade, selvmord, psykiske lidelser og svære følelser koblet til en selv.

8 <https://newsroom.tiktok.com/en-us/how-tiktok-recommends-videos-for-you>

9 Det vil sige en form for "research persona" som beskrevet af Bounegru, Devries og Weltevrede: [https://link.springer.com/chapter/10.1007/978-981-19-2476-7\\_5](https://link.springer.com/chapter/10.1007/978-981-19-2476-7_5)

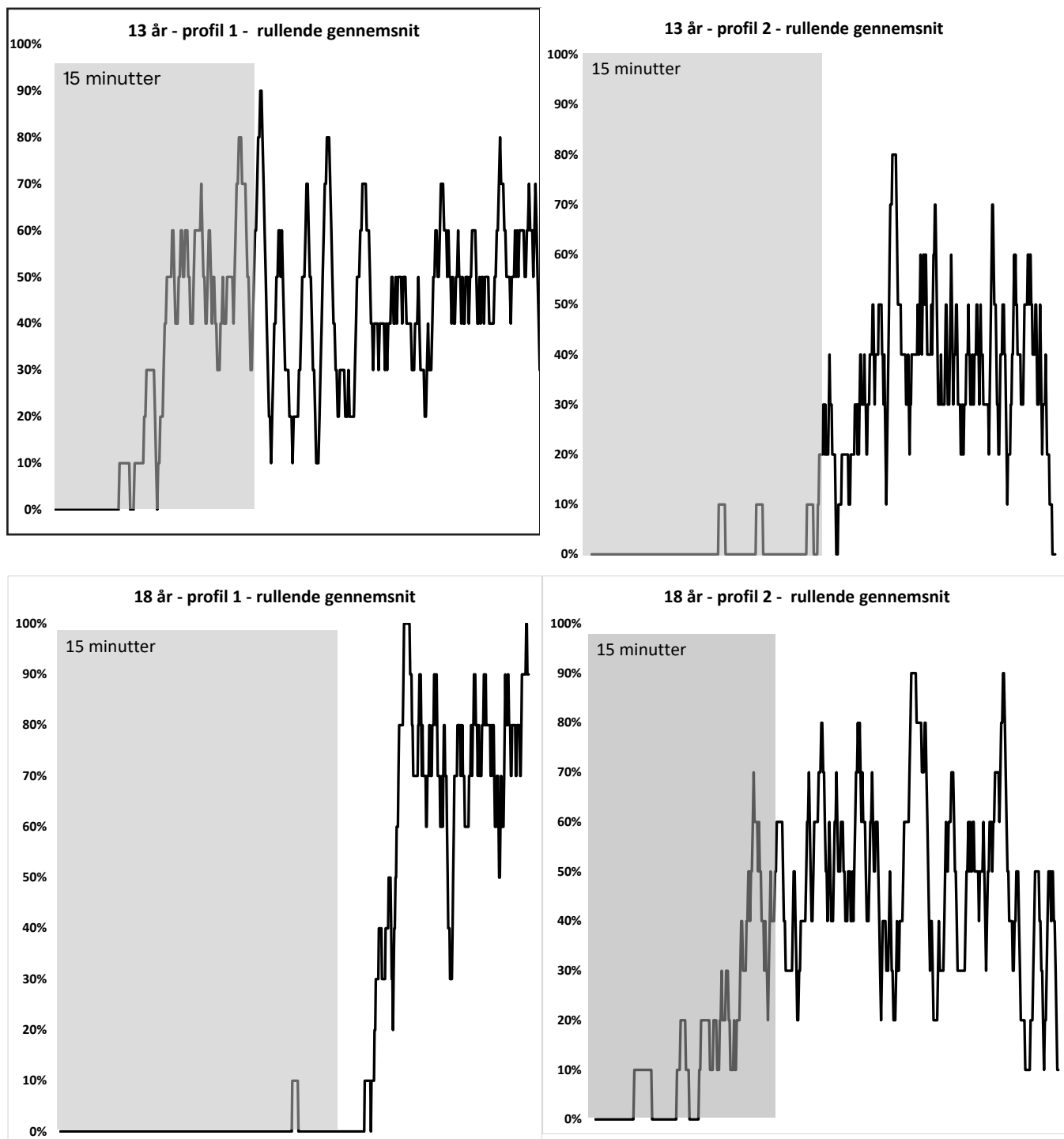
10 Det er forskelligt fra profil til profil, hvor mange videoer vi swiper igennem på en time. Dette skyldes, at vi afspiller videoer med problematisk indhold i 10 sekunder, inden vi swiper, mens vi swiper hurtigt videre for alle andre videoer; dette har den konsekvens, at for de profiler, hvor vi hurtigt eksponeres for mange videoer med problematisk indhold, når vi ikke at swipe igennem lige så mange videoer i alt.

11 Det glidende gennemsnit er udregnet for et interval på 10 videoer, dvs. vi tager andelen af problematisk indhold for video 1-10, dernæst for video 2-11, så 3-12 etc.



# Bilag

## Andel problematisk indhold over antal videoer



# /Digitalt Ansvar

Udgivet af Digitalt Ansvar, december 2023

Ansvarshavende redaktør: Ask Hesby Krogh

Redigering: Anne Tscherning Larsen og Ask Hesby Krogh

Tekst og analyse: Anne Tscherning Larsen, Asger Nim og Asta Iris Rohde

Telia har doneret simkort og telefoner til forsøget.

Projektet er muliggjort med midler fra Medieaftalen 2023–2026 samt midler fra Oak Foundation Denmark.