



Brottssoffermyndigheten

# Näthat och demokratiskt deltagande

– en kunskapsöversikt

Måns Svensson  
Oscar Björkenfeldt  
Fredrik Åström  
Karl Dahlstrand

# Näthat och demokratiskt deltagande

– en kunskapsöversikt

Måns Svensson  
Oscar Björkenfeldt  
Fredrik Åström  
Karl Dahlstrand



## **Brottsoffermyndigheten**

Brottsoffermyndigheten

Box 470, 901 09 Umeå

Telefon: 090-70 82 00

[registrator@brottsoffermyndigheten.se](mailto:registrator@brottsoffermyndigheten.se)

[www.brottsoffermyndigheten.se](http://www.brottsoffermyndigheten.se)

© Brottsoffermyndigheten 2021

ISBN: 978-91-983020-5-9

Grafisk form: Matador Kommunikation

Tryck: Taberg Media Group, 2021

# Förord

Rätten att uttrycka åsikter, att kritisera och att ifrågasätta är en del av en levande demokrati. Ett demokratiskt samhälle måste vara öppet för olika röster och ge alla möjlighet att komma till tals. Demokratien är beroende av att människor deltar och engagerar sig i samhällsfrågor. Med rätten till yttrandefrihet kommer också ett ansvar. Var och en bör skydda, värna och bevara yttrandefriheten genom att respektera andra och bidra till en samhällsdebatt fri från hot och hat. Tystas människor tystnar demokratin.

Den digitala utvecklingen gör det möjligt för fler att göra sin röst hörd. Det är positivt men för också med sig utmaningar. Att föra fram och sprida hot och hat har blivit enklare. Uttalanden kan få snabb spridning i sociala medier och på så sätt få större genomslagskraft i dag än tidigare. Det kan missbrukas av personer som inte vill bevara det demokratiska samhället.

Personer som engagerar sig i samhällsfrågor är särskilt utsatta för hot och hat. Det gäller till exempel förtroendevalda, journalister, konstnärer, opinionsbildare, forskare och representanter för det civila samhället. Utsatthet kan också vara kopplat till exempelvis en persons kön, könsidentitet, hudfärg, sexuella läggning eller ålder. När kvinnor och hbtqi-personer får hot och hat riktat mot sig har det ofta att göra med kön och sexualitet.

På uppdrag av Brottsoffermyndigheten har forskare vid Lunds universitet och Högskolan i Halmstad genomfört en enkätundersökning som visade att mer än hälften av de tillfrågade har känt sig kränkta av sådant som andra skriver om dem på nätet. Särskilt utsatta är unga och personer med utländsk bakgrund, och följderna blir omfattande nivåer av självcensur bland dessa grupper. Det innebär att de anpassar hur de uttrycker sig för att undvika hot och hat, eller helt avstår från att publicera och delta i samhällsdebatten. Det demokratiska samtalet begränsas och viktiga röster riskerar att tystna.

Även denna kunskapsöversikt är ett resultat av samarbetet mellan Brottsoffermyndigheten, Lunds universitet och Högskolan i Halmstad. Den belyser olika frågor om näthat och demokratiskt deltagande, såsom vilka områden som genererar hot och hat, vilka samhälls- och yrkesgrupper som är särskilt utsatta och vilka konsekvenser näthat leder till. Genom att uppmärksamma frågan om hot och hat på internet vill Brottsoffermyndigheten belysa vilka konsekvenser det kan få för vår demokrati. Myndigheten vill också synliggöra de personer som utsätts för näthat och bidra till en levande samhällsdebatt.

Umeå, juni 2021

Annika Öster

Generaldirektör, Brottsoffermyndigheten

# Innehåll

<b>1. Inledning</b> .....	<b>6</b>
<b>2. Metod</b> .....	<b>7</b>
<b>2.1 Systematisk litteraturoversikt</b> .....	<b>7</b>
(a) Planering av studien .....	8
(b) Söka, identifiera och organisera artiklar .....	8
(c) Extrahera och värdera materialet .....	9
<b>2.2 Nationell kunskapsöversikt</b> .....	<b>10</b>
<b>2.3 Bibliometrisk analys</b> .....	<b>10</b>
<b>3. Resultatredovisning</b> .....	<b>12</b>
<b>3.1 Bibliometrisk analys</b> .....	<b>12</b>
Forskningsfält .....	12
Forskningsinriktning .....	13
<b>3.2 Systematisk litteraturoversikt</b> .....	<b>15</b>
Det postfaktiska samhället .....	15
Populism .....	19
Näthat .....	21
Näthat mot yrkesgrupper med viktiga funktioner i demokratin .....	24
Juridik och yttrandefrihet .....	27
Tekniska lösningar för att motverka näthat och desinformation .....	29

<b>3.3 Nationell kunskapsöversikt .....</b>	<b>30</b>
Introduktion till de rättsliga ramarna avseende näthat .....	30
Digitaliseringens inverkan på det demokratiska samtalet .....	33
Digitala påverkansförsök på svenska val .....	34
Näthat som fenomen .....	34
Allmänhetens utsatthet för näthat .....	35
Särskilt utsatta samhällsgrupper .....	36
Särskilt utsatta yrkesgrupper .....	39
Rättens hantering av näthat .....	44

## **4. Sammanfattning och avslutande reflektioner .....**

**46**

### **4.1 Systematisk kunskapsöversikt .....**

**46**

Det postfaktiska samhället .....	46
Populism .....	47
Näthat .....	47
Näthat mot yrkesgrupper med viktiga funktioner i demokratin .....	47
Juridik och yttrandefrihet .....	48
Tekniska lösningar för att motverka näthat och desinformation .....	48

### **4.2 Kunskapsläget om förhållandet i Sverige .....**

**49**

### **4.3 Avslutande reflektioner kring strategier för att motverka problemen .....**

**50**

## **Referenslista .....**

**52**

# 1. Inledning

Denna kunskapsöversikt utgör en underlagsrapport inom ramen för regeringens uppdrag till Brottsoffermyndigheten att förbereda en informationsinsats mot näthat kopplat till demokratiskt deltagande och delaktighet (Ku2020/01463/MD). Syftet är att förebygga och motverka förekomsten av näthat i det demokratiska samtalet, samt att bidra till att stödja de som riskerar att drabbas av näthat. Mot denna bakgrund har Brottsoffermyndigheten gett i uppdrag till Lunds universitet och Högskolan i Halmstad att ta fram en underlagsrapport i form av en kunskapsöversikt. Projektgruppen som har genomfört kunskapsöversikten består av Måns Svensson, professor i rättssociologi vid högskolan i Halmstad (projektledare), Oscar Björkenfeldt, doktorand i rättssociologi vid Lunds universitet, Fredrik Åström, docent och bibliometriker vid Lunds universitet och Karl Dahlstrand Fil. Dr. i rättssociologi vid Lunds universitet.

Näthat ska i det här sammanhanget ses som ett paraplybegrepp för olika slags kränkningar, trakasserier eller hot som utförs på nätet, till exempel via sociala medier eller e-post. Med det här synsättet kan handlingar som faller innanför lagens gränser såväl som olagliga handlingar förstås som näthat.

Mer specifikt avser kunskapsöversikten att behandla frågor som relaterar till:

- Näthatets omfattning
- Vilka områden som särskilt genererar näthat
- Om det finns särskilt utsatta samhällsgrupper
- Om det finns särskilt utsatta yrkesgrupper
- Hur utsattheten yttrar sig
- Vilka konsekvenser näthatet medför
- Vilka möjligheter till stöd och hjälp det finns för de som utsätts idag
- Vilka metoder som är effektiva för att ge stöd, hjälp och skydd till personer som utsätts för näthat

## 2. Metod

Kunskapsöversikten inkluderar följande tre delar:

- En systematisk litteraturöversikt över den engelskspråkiga vetenskapliga och referentgranskade litteraturen med fokus på näthat och demokrati
- En bibliometrisk analys av den litteratur som identifierades genom den systematiska litteraturöversikten
- En mer ostrukturerad kunskapsöversikt, baserad på nyckelordsbaserade sökningar efter svenskspråkiga utredningar, rapporter och liknande, så kallad grå litteratur

Styrkan hos den systematiska litteraturöversikten är att man på ett systematiskt och transparent vis kartlägger det internationella vetenskapliga kunskapsläget. Publikationerna som söks fram med den här metoden är referentgranskade och håller hög vetenskaplig kvalitet. Emellertid rör denna litteratur sällan de specifika svenska förhållandena. För att kompensera för denna brist genomfördes även den mer ostrukturerade kunskapsöversikten. Den bibliometriska analysen ger förutsättning att förstå fältet på en övergripande och orienterande nivå, inte minst visuellt.

### 2.1 Systematisk litteraturöversikt

Den systematiska litteraturöversikten är en metod genom vilken man samlar in, analyserar och sammanställer vetenskapliga studier inom ett visst område, eller utifrån en viss frågeställning. Således kan den systematiska litteraturöversikten beskrivas som ett sätt att sammanställa kunskapsläget inom ett visst område utifrån en strukturerad och systematisk insamling och granskning av vetenskapliga studier.

Insamlingen av de vetenskapliga publikationerna har gjorts genom en sökning, innefattande en kombination av nyckelord, i Web of Science, som är en av de mest etablerade och erkända samlingarna av databaser (se mer nedan). Systematiken i den här studien tar sin utgångspunkt i de tre steg som Tranfield et al. beskriver i *Towards a Methodology for Developing Evidence-Informed*



Management Knowledge by Means of Systematic Review (2003):

(a) Planering; (b) söka, identifiera och organisera publikationer; samt (c) extrahera och värdera. Vid varje steg krävs det ett antal övervägande som har att göra med såväl sökkriterier som frågor rörande vilka publikationer som ska inkluderas respektive exkluderas. Nedan redovisas hur Tranfields (et al.) metod har applicerats inom ramen för denna studie.

### **(a) Planering av studien**

Kunskapsöversikten fokuserar på frågor som rör näthat och demokratiskt deltagande. Av särskilt intresse är vilka områden som genererar hot och hat, vilka samhälls- och yrkesgrupper som är särskilt utsatta, vilka konsekvenser näthat leder till, och vilka strategier som används för att motverka näthat, samt för att stödja och hjälpa de som utsätts. Om möjligt ska studien kunna belysa eventuella begränsningar av specifika gruppers möjligheter (exempelvis baserat på könstillhörighet och etnicitet) att fritt delta i det offentliga samtalet, samt identifiera tematiska områden för vilka frågeställningarna är relevanta.

Viktiga avgränsningar utgörs av det faktum att det systematiska sökandet efter vetenskapliga publikationer i den här delstudien begränsas till vetenskapliga tidskrifter vilka är skrivna på engelska. De publikationer som finns på svenska är i regel inte vetenskapliga, det vill säga de är inte peer-review granskade och har därmed inte genomgått den process som garanterar vetenskaplig kvalitet. Ett viktigt undantag utgörs av doktorsavhandlingar vilka uppfyller vetenskapliga krav men som ändå faller utanför begränsningarna för den här studien. Med detta sagt har delstudie två till syfte att fånga sådan litteratur som faller utanför den systematiska litteraturoversikten.

### **(b) Söka, identifiera och organisera artiklar**

Internets och sociala mediers inverkan på hur vi kommunicerar i dagens samhälle har inneburit en stor omvälvning för materialiserandet av demokratiskt deltagande. De frågor som relaterar till näthatets framväxt rör ett flertal olika forskningsfält. De frågor som hanteras är, exempelvis, hur människor förhåller sig till objektiva fakta i en digital tid, hur en hatisk samtalston främjas av populistiska samhällsströmmar och algoritmiska logiker, hur privata plattformsföretag förhåller sig till offentligt ansvar, samt de yttrandefrihetsutmaningar som de nya teknikerna medför. Ambitionen med denna översikt är att skapa en så bred bild som möjligt av var forskningen avseende näthat och demokratiskt deltagande står.

Med tanke på att studieobjektet innefattar forskning från ett brett register av vetenskapliga discipliner hade det varit otillräckligt att söka efter publikationer i disciplinspecifika databaser. Av denna anledning användes Web of Science (WoS), som är en databas (inkluderande bland annat 12 000 vetenskapliga tidskrifter av högsta kvalitet) publicerad av Thomas Reuters, vilken förtecknar internationell forskningslitteratur, framför allt tidskriftsartiklar på engelska. Fördelen med denna databas är att informationen är innehållsmässigt rik och möjliggör avancerade bibliometriska analyser.

Vid sökningen av de vetenskapliga publikationerna användes Boolean sökmotetik, vilket är en typ av sökkommandon (AND, OR och NOT) som kan användas för att utöka eller begränsa en sökning. Här kombinerade vi olika nyckelord som tillsammans genererade en söksträng för att söka i publikationernas titel och abstrakt. Den givna söksträngen för denna litteraturöversikt redovisas i metodbeskrivningen av den bibliometriska analysen nedan. För att få fram de mest relevanta nyckelorden arbetades olika förslag igenom mellan de involverade forskarna.

### **(c) Extrahera och värdera materialet**

För att identifiera forskningslitteraturen användes följande söksträng i sökfältet Topic i WoS (ett sökfält som inkluderar publikationernas titlar, sammanfattningar, samt de nyckelord som används för att beskriva publikationens innehåll), kombinerad med avgränsningar för publikationsår, språk och dokumenttyp:

```
(hate OR threat* OR harassment* OR bullying* OR troll*) AND (self-censorship* OR "freedom of speech" OR "freedom of expression" OR silencing* OR censor* OR democra*) AND (internet* OR cyber* OR digital* OR cyber* OR social media* OR online comments*) AND PY=2010-2020 AND LANGUAGE: English AND DOCUMENT TYPES: (Article OR Book OR Book Chapter OR Letter OR Note OR Proceedings Paper OR Review)
```

Sökningen resulterade i 610 publikationer, men en genomgång av sökresultaten visade att många publikationer låg utanför fokus för analysen. Efter att icke-relevanta träffar eliminerats återstod 160 publikationer, för vilka data från WoS laddades ner för bearbetning och analys.

Eftersom den andra screeningfasen minskade antalet publikationer till en hanterbar mängd genomfördes därefter en granskning av artiklarnas fullständiga text. Detta steg innebar alltså en mer noggrann analys av studiernas relevans för det slutliga urvalet. I denna fas exkluderades ytterligare 31 publikationer, vilket innebar att 129 artiklar återstod till det slutliga urvalet.

## 2.2 Nationell kunskapsöversikt

För den nationella kunskapsöversikten tillämpades ett mindre systematiskt tillvägagångssätt än för den systematiska litteraturöversikten. Som tidigare nämnts rekommenderas det ofta att systematiska sökningar kompletteras med mer osystematiska manuella sökningar; i vårt fall har detta resulterat i ytterligare ett 60-tal relevanta, men ej referentgranskade, publikationer. De manuella sökningarna har gjorts genom användande av olika söktermer i sökmotorer såsom Google Scholar men framför allt genom sökningar på myndigheters hemsidor. Vi har även analyserat flera rapporters referenslistor för att på så sätt finna ytterligare litteratur. Utöver detta har vi dessutom varit i kontakt med personer som besitter god kunskap på området som ett sätt att identifiera studier som vi eventuellt annars missat. Majoriteten av de publikationer som innefattas i den nationella kunskapsöversikten är myndighetsrapporter, avhandlingar, bokkapitel, statliga utredningar och rapporter från civilsamhällesorganisationer.

## 2.3 Bibliometrisk analys

För att få en överblick över den forskningslitteratur som behandlar frågor om näthat och demokrati har vi valt att kombinera den systematiska forskningsöversikten med bibliometriska analyser. Bibliometri handlar i grunden om kvantitativa och statistiska analyser av texters och textsamlingars egenskaper. Genom att studera statistiska samband mellan begrepp som används inom ett visst forskningsområde kan man kartlägga de olika teman som behandlas inom området. Genom att analysera till exempel de tidskrifter som forskningen publicerats i och som artiklarna refererar till går det att se inom vilka forskningsfält ett visst forskningsområde studeras, och i vilken utsträckning dessa forskningsfält relaterar sina resultat till andra fältstudier av ett visst ämne. Vi får alltså en bred överblick över de teman som behandlas inom ramen för denna specifika frågeställning, samt en uppfattning om de forskningsområden inom vilka frågan studeras.

Den bibliometriska analysen bygger på data från WoS, som är internationell och utan ämnesmässiga begränsningar; och även har den fördelen att referenslistorna för de publikationer som ingår i databaserna indexeras, såväl som data om själva publikationen i sig. Samtidigt bör man också vara medveten om databasernas begränsningar: i huvudsak är det artiklar i vetenskapliga tidskrifter som indexeras, och det är framför allt engelskspråkig litteratur vi finner i databaserna. Detta gör att till exempel forskning som publiceras i monografier eller som kapitel i antologier och på svenska sannolikt inte kommer att vara representerad i analyser som bygger på WoS-data.

För att bearbeta data och utföra analyserna användes dels Bibexcel (<https://bibliometrie.univie.ac.at/bibexcel/>), ett program för bibliometriska analyser där man bland annat kan renodla informationen från WoS för att analysera inte bara specifika fält i databaserna, till exempel titel, författare eller citerade referenser, utan också delar av specifika fält, till exempel tidskriftsnamnen för den citerade referensen. Dels användes VOSviewer (<http://www.vosviewer.com>), ett program där man kan skapa visualiseringar baserade på nätverksanalyser. I den bibliometriska analysen inkluderades de 160 publikationer som återstod efter den initiala screeningen.

För att identifiera inom vilka forskningsfält man studerar frågan om näthat och demokrati användes co-citeringsanalys på tidskriftsnivå. Tanken är att man studerar den litteratur som forskningen använt genom att analysera referenslistorna och antar att artiklar eller – i detta fall – tidskrifter som citeras tillsammans har ett ämnesmässigt samband. Om man gör dessa analyser baserat på hundratals eller tusentals artiklar med tiotusentals eller hundratusentals referenser, så bildar samförekommande citerade artiklar eller tidskrifter kluster som kan sägas representera olika forskningsinriktningar eller forskningsfält.

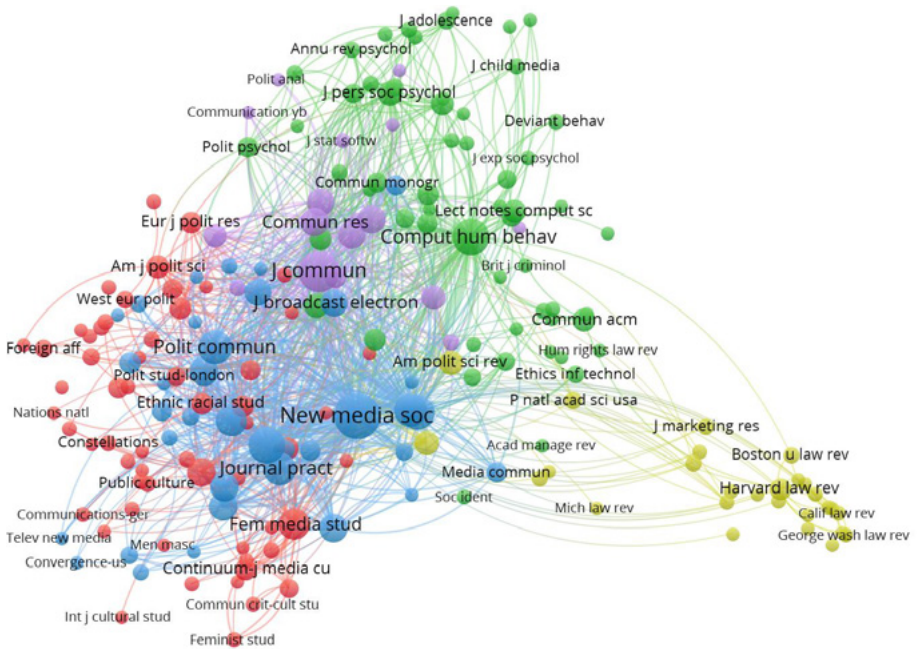
Analysen bygger alltså på hur ofta citerade tidskrifter förekommer tillsammans i referenslistorna för de artiklar som identifierades i sökningen. För att sedan gå vidare och identifiera inte bara inom vilka forskningsfält som frågor om näthat och demokrati studeras, utan också vad det är man forskar om, gjordes en motsvarande analys som – till skillnad från analysen av forskningsfält – inte bygger på samförekomster av referenser utan istället på samförekomster av de termer som förekommer i publikationerna. Från artiklarna som identifierats i WoS-sökningen hämtades de nyckelord som beskriver artiklarnas innehåll. Liksom i den tidigare analysen grupperas de begrepp som förekommer tillsammans i dokumenten.

## 3. Resultatredovisning

### 3.1 Bibliometrisk analys

#### Forskningsfält

Denna analys identifierar alltså tidskrifter som citeras tillsammans i de 160 utvalda publikationerna och grupperar dem tillsammans för att representera olika forskningsfält. Storleken på noderna i nätverket visar hur ofta tidskrifterna citeras, och styrkan hos sambanden mellan tidskrifterna markeras av hur nära de förekommer på kartan, av länkar mellan tidskrifterna, samt av kluster (representerade genom olika färger) definierade av statistiska analyser av samband mellan olika enheter (Figur 1). Kartan har två huvuddelar: ett litet kluster på den högra sidan bestående av juridiska tidskrifter och, på den vänstra sidan, ett antal kluster av tidskrifter med anknytning till samhälls- och beteendevetenskaperna. I den nedre delen av den vänstra sidan hittar vi två kluster där det röda klustret framför allt består av statsvetenskapliga tidskrifter tillsammans med tidskrifter med anknytning till studier av genus- och etnicitetsfrågor eller kritiska kulturstudier, medan det blå klustret framför allt representerar kommunikationsvetenskap och journalistik. I den vänstra sidans övre del finns ett kluster som domineras av psykologi och beteendevetenskap, men som också innehåller tidskrifter med koppling till människa-datorinteraktion, det vill säga beteendevetenskapliga perspektiv på datoranvändning. I mitten av den vänstra delen finns också ett mindre, lilafärgat kluster, som innehåller en blandning av de ämnen som finns representerade i de tre andra klustren på den vänstra sidan.



**Figur 1.** Co-citeringsanalys av de 197 citerade tidskrifter som förekommer tillsammans minst tre gånger.

Även om vi i klusteranalysen finner fyra distinkta ämnesinriktningar representerade (statsvetenskap och kritiska kulturstudier, kommunikationsvetenskap, psykologi och beteendevetenskap samt juridik) kan man notera att det också finns starka länkar mellan tre av de fyra forskningsfälten. Detta tyder på en kommunikation mellan fälten, där till exempel forskning inom kommunikationsvetenskap också tar till sig av det som kommer fram i statsvetenskaplig forskning. Den forskning som ter sig mest isolerad i förhållande till de andra fälten är juridiken, men detta kan åtminstone till någon del förklaras av traditioner kring vilken annan forskning man refererar till, och hur man refererar, inom juridiken.

### Forskningsinriktning

Denna analys bygger på en statistisk behandling av de nyckelord som används för att beskriva innehållet i de inkluderade publikationerna. På liknande sätt som i föregående figur markerar storleken på nyckelorden hur ofta de förekommer, medan närhet och kluster representerade av färger baseras på hur ofta nyckelorden förekommer tillsammans (Figur 2).



nyckelord som participation, deliberation, user comments och civility eller incivility. I den nedre delen finner vi det blå och lila klustret där forskningen mer verkar inrikta sig på ”den mörkare delen” av frågor om tillgång till offentligheten genom sociala medier, där vi finner nyckelord kopplade till authoritarianism, power och fake news.

Näthatsklustren på den högra sidan visar dels på forskning om frågor kring nätmobbning av barn i den övre delen av kartan, dels om trollande, trakasserier och våld riktat mot kvinnor och journalister. Lite mindre kopplat till positionen på kartan och i klustren finner vi också att forskningen identifierar viktiga teman i frågor om yttrandefrihet (freedom of expression), censur och propaganda.

## 3.2 Systematisk litteraturoversikt

De 129 artiklar som inkluderades i det slutliga urvalet för den systematiska litteraturoversikten studerades i fulltext och kategoriserades därefter i olika forskningsteman utifrån undersökningens forskningsfokus och studieobjekt.

Följande teman har identifierats för denna litteraturoversikt (antalet artiklar anges inom parentes):

- Det postfaktiska samhället (26)
- Populism (13)
- Näthat (30)
- Näthat mot yrkesgrupper med viktiga funktioner i demokratin (18)
- Juridik och yttrandefrihet (30)
- Tekniska lösningar för att motverka näthat och desinformation (12)

Nedan följer en genomgång av respektive forskningstema.

### **Det postfaktiska samhället**

– 26 granskade publikationer

På ett generellt plan beskriver publikationerna under det första forskningstemat olika infallsvinklar på det postfaktiska samhället (post-truth society), det vill säga ett tillstånd där den offentliga debatten till stor del är bortkopplad från objektiva fakta och där människor ger efter för känsloladdade argument snarare än sanningar (Dahlgren 2018; Hannan 2018; Barton 2019; Estelles & Castellvi 2020; Yerlikaya 2020). En stor mängd publikationer beskriver falska nyheter (fake news) som en av de största drivkrafterna bakom



den postfaktiska utvecklingen (Marres 2018; Barton 2019; Miletskiy Cherezov & Strogetskaya 2019; Duffy, Tandoc & Ling, 2020; Lee & Hosam, 2020; Schaewitz et al. 2020; Yerlikaya 2020). Trots de båda begreppens symbiotiska förhållande är det viktigt att skilja dem åt. Medan det postfaktiska samhället är en referens till ett epokskifte där fakta förlorat status, är falska nyheter snarare en slags bedräglig mediaprodukt vars syfte är att medvetet vilseleda läsare genom direkta lögner – och inte att förväxla med felaktig rapportering eller opinionstexter (Dahlgren 2018; Barton 2019).

Ett antal studier påvisar att falska nyheter sprids under strategiska former, ofta till utvalda användare på sociala medier med omfattande kontaktnät, för att accelerera distributionen och påskynda debatten kring det aktuella ämnet (Barton 2019; Schaewitz et al. 2020). Beträffande motivet bakom falska nyheter är det en vanlig missuppfattning att spridningen bedrivs enbart utifrån ideologiska skäl. Forskningen visar emellertid att rent ekonomiska motiv också spelar en central roll. Uppseendeväckande (falska) nyheter genererar trafik som i sin tur genererar annonsintäkter (Barton 2019). Med andra ord, ju mer lockande och intressant en artikel tycks vara desto fler läsningar kommer den få och desto mer pengar kommer den dra in. Även om detta framstår som exploaterande, är samma typ av ekonomisk ackumulering drivande för nästan alla stora nyhetsorganisationer – skillnaden ligger i bristen på sanning och avsikten att vilseleda konsumenterna.

I frågan om falska nyheter har man, under senare år och i synnerhet efter det amerikanska presidentvalet 2016, flitigt uppmärksammat inverkan på den allmänna åsiktsbildningen av automatiserade konton (sociala botar). Inom detta forskningsfält finns det omfattande empiri som stödjer antagandet att sociala botar används på ett strukturerat sätt för att uppnå stor spridning (Shao et al. 2018; Frost 2020), vilket många gånger också är ett effektivt tillvägagångssätt. Av en av artiklarna framgår det exempelvis att närvaron av sociala botar på sociala medier faktiskt påverkar hur olika ämnen diskuteras. Ross et al. (2019) visar att i redan åsiktpolariserade onlinemiljöer räcker det med ett relativt litet antal sociala botar för att svänga åsiktsbildningen. Samma studie drar slutsatsen att åsikter propagerade av sociala botar snabbt utvecklas till majoritetsuppfattningar som följaktligen tystar åsikter som upplevs vara i mindretal. Liknande resultat går att finna i andra studier som framhäver att sociala botar regelbundet påverkar åsiktsbildningen på nätet, ofta genom spridning av falska nyheter (Shao et al. 2018; Ross et al. 2019; Frost 2020).

Av forskningen framgår det vidare att effektivare användning av maskin-inlärningsmetoder för att upptäcka och blockera sociala botar (Shao

et al. 2018), samt för att minska effekten av så kallade filterbubblor (Bozdog & van den Hoven 2016) kan ha viss positiv effekt för att minska spridningen av desinformation. Samtidigt påpekas det att sådana åtgärder inte ska ses som den ultimata lösningen av den mer djupgående problematiken med desinformation och det framväxande postfaktiska samhället.

Mot bakgrund av ovanstående beskriver Dahlgren (2018) hur de postfaktiska tendenserna i dagens digitala samhälle är ett resultat av överflödet av snabb information och en politisk kultur som kännetecknas av misstro. Detta har resulterat i att antaganden om den sociala verkligheten och förutsättningarna för legitim kunskap ofta undergrävs. Dahlgrens (2018) hållning stöds av flera forskare som även belyser att denna utveckling till del påskyndats av ideologiska idéströmningar som bidragit till ett undergrävande av demokratiska förutsättningar rörande ansvarstagande och civilt engagemang (Dahlgren 2018; Marres 2018; Sehl, Simon & Schroeder 2020; Sussman 2020; Yerlikaya 2020). Den organiserade spridningen av falska nyheter, kombinerad med sociala medieplattformars algoritmbaserade infrastruktur och deras utvidgade inflytande som informations- och nyhetsförmedlare, har således haft flera negativa följder för den fria opinionsbildningen i västerländska demokratier (Hemsley et al. 2018; Jameel, Ali & Malik 2019; Yerlikaya 2020).

I förhållande till den ovan angivna problematiken lyfter forskningen fram den drastiska omställningen av medielandskapet som ett allvarligt hot mot demokratin. Åtskilliga studier pekar på att det växande utbudet av digitala nyhetssidor har decentraliserat den journalistiska publiceringsprocessen till förmån för en ny våg av medborgarjournalistik. Det moderna medielandskapet innehåller en uppsjö av nyhetsmedier som inte följer professionella journalistiska normer och etiska principer. Den nya medielogiken har resulterat i att skillnaden mellan åsikter och fakta ofta suddas ut, vilket i sin tur har haft betydande konsekvenser för den traditionella journalistikens auktoritet, trovärdighet och status (Dahlgren 2018; Lu & Liu 2018; Luhrmann et al. 2018; Barton 2019; Stetka Mazak & Vochocova 2020). Sammantaget skapas goda förutsättningar för desinformation och näthat (Estelles & Castellvi 2020). Med andra ord har det stora utbudet av digitala nyhetsmedier lett till allt mer polariserade digitala miljöer där nyhetskonsumenter serveras "sanningar" som stödjer deras egna åsikter (Kligler-Vilenchik, Baden & Yarchi 2020). Det framgår vidare att misstro mot det så kallade etablissemanget, det vill säga etablerade medier, politiker och experter, har förstärkts i och med utbredningen i samhället av internet och i synnerhet sociala medier (Dahlgren 2018). I relation till detta påpekar forskningen att det pågår en allt mer distinkt autokratisering, där valdeltagande visserligen är intakt eller till

och med ökar, men där mediefrihet, yttrandefrihet och rättsstaten på olika sätt undergrävs i flera länder (Lu & Liu 2018; Luhmann et al. 2018). Således kan man konstatera att flera länder nu står inför ett historiskt demokratiskt dilemma, då misstron mot demokratiska institutioner fördjupas samtidigt som basen för etablerade och legitimerade fakta destabiliserats (Dahlgren 2018; Luhmann et al. 2018; Barton 2019). Det råder stor enighet inom forskningen om att det postfaktiska samhället, spridandet av falska nyheter och bristen på kritiskt tänkande är ett allvarligt hot mot demokratin.

Hannan (2018) menar att det polariserade medielandskapet har bidragit till en hänsynslös samtalsretorik som i grunden har förändrat det politiska språkbudet. Sociala mediernas diskursiva utrymme, som ofta saknar etablerade normer för sanning, logik, bevis och hövlighet, har resulterat i en slags vilda västern för kommunikation där ociviliserat språkbruk, delvis inspirerat av ledande populistiska politiker, normaliseras (Hannan 2018; Valdez, Burbach & Ziefle 2018; Lee & Hosam 2020).

Som ett svar på den alltmer uppmärksammade spridningen av hat och desinformation på sociala medier har flera sociala medieplattformar successivt utvecklats som verktyg för att identifiera och flagga inlägg med vilseledande och falsk information. Sådana metoder har en viss påvisad positiv betydelse: exempelvis visar en studie att falska nyheter på sociala medier upplevs som mindre korrekta när människor får en allmän varning (Clayton et al. 2020). Dock menar en större majoritet forskare att sådana åtgärder enbart har marginell effekt (Barton 2019; Ross et al. 2019; Valdez & Ziefle 2019). Flertalet forskare hävdar att tyngdpunkten på åtgärderna snarare bör läggas på utbildning och träning i kritisk reflektion. Med andra ord; falska nyheter är inte det egentliga problemet utan det verkliga problemet är snarare att människor tror på dem. För att få bukt med denna problematik föreslås ett brett utbud av utbildningsinsatser, inklusive verktyg för att kritiskt analysera innehållet i sociala medier och öka förståelsen för de maktstrukturer som ligger bakom desinformationen (Beam, Hmielowski & Hutchens 2018; Dahlgren 2018; Schaewitz et al. 2020). Här argumenterar flera studier att utbildningssystemet har misslyckats med att hålla jämna steg med digitaliseringens explosion av information och de tekniska förändringarna inom nyhetsmedia (Dahlgren 2018; Barton 2019).

## Populism

– 13 granskade publikationer

Det digitala medielandskapet har medfört nya förutsättningar som bland annat har skapat grogrund för hatisk och populistisk retorik (jfr Ekman 2015; Alvares & Dahlgren 2016; Bonikowski 2017; Jakubowicz 2017; Davis 2020; Mott & Cockayne 2020). Populism sägs ofta vara ett svårdefinierat begrepp och innebörden i begreppet varierar (Bonikowski 2017). Men i allmänhet beskrivs populism på ett eller annat sätt som politik baserad på ett motstånd mot samhällets elit där ”vanliga människor” ses som den enda legitima källan till politisk makt (Alvares & Dahlgren 2016; Bonikowski 2017; Davis 2020). Vilka som anses tillhöra eliten varierar mellan olika samhällen. Oftast pekas förtroendevalda politiker ut som de främsta företrädarna för eliten, men även myndighetsrepresentanter, journalister, opinionsbildare och akademiker beskrivs som den elit som motståndet riktar sig mot (Bonikowski 2017).

Forskningen pekar på att även populism med etno-nationalistiska inslag är vanlig (jfr Ekman 2015; Alvares & Dahlgren 2016; Bonikowski 2017; Evolvi 2019). Denna typ av populism riktar sig inte enbart mot samhällets ”elit” utan även mot minoritetsgrupper som utpekas som syndabockar i förhållande till den samhällsutveckling som hotar ”det vanliga folkets” ställning. Det finns ett samband mellan invandringskritiska åsikter och förakt mot etablerade medier och politiker (Watson & Riffe 2013; Ekman 2015; Alvares & Dahlgren 2016). Invandring målas upp som ett hot mot västerländska samhällen (Ekman 2015; Evolvi 2019).

Ekonomiska, politiska och kulturella förändringar har lett till att ”vanligt folk” upplever ett hot (Alvares & Dahlgren 2016; Bonikowski 2017; Davis 2020). Det är vanligt att främlingsfientliga ståndpunkter utgår från förespråkandet av västerländska värderingar och rätten till yttrandefrihet (Ekman 2015; Evolvi 2019). Olika studier pekar på att det algoritmbaserade medielandskapet medfört nya förutsättningar för att organisera och sprida den här typen av budskap (Alvares & Dahlgren 2016; Mott & Cockayne 2020). Det finns empiriskt stöd för att nätverk som sprider den här typen av idéer är organiserade, dynamiska och effektiva. Bland annat används satir som ett verktyg för att normalisera rasistiska och främlingsfientliga budskap (Schwarzenegger & Wagner 2018; Sponholz & Christofoletti 2019), samt för att berättiga hot och inspirera till hatkampanjer (Fangen & Holter 2020). Det globala mönstret indikerar att det nya medielandskapet bidrar till populism som banar väg för organiserade digitala attacker mot förtroendevalda, journalister, kvinnor och minoriteter runt om i världen (Schwarzenegger & Wagner 2018).

De två forskarna Fangen och Holter (2020) genomförde nyligen en intervjustudie i Norge med personer vars rasistiska, diskriminerande eller hatiska kommentarer tagits bort från nyhetssidors kommentarsfält. Resultatet av deras studie bekräftar de populistiska strömningarnas betydelse i det att respondenterna beskrev sig själva som marginaliserade sanningskrigare mot de lögner som sprids av korrupta politiker och media.

Davis (2020) använder begreppet den anti-offentliga sfären (anti-public sphere) som utgångspunkt för att analysera den växande mängden hatiska uttryck på nätet. Anti-offentliga sfärer definieras som digitala miljöer som befinner sig i opposition mot det rådande samhället: exempelvis vit makt-hemsidor, anti-klimatvetenskapliga forum, militanta mansrörelser, invandringskritiska Facebook-grupper). Dessa anti-offentliga sfärer uppstår genom att företrädare för olika aparta åsikter har kunnat använda det nya medielandskapet för att bryta sin relativa isolering och nå fram till likasinnade. Davis visar hur nya grupperingar uppstår. Människor med åsikter som tidigare avskrevs som marginella, extrema och betydelselösa hittar sina sammanhang och kan mobilisera gemensamma aktioner. Davis hävdar att det inte längre är gångbart att betrakta sådana diskursiva formeringar som anti-moderna och barbariska. I själva verket är de etablerade i den moderna och post-industriella offentligheten (Davis 2020), både online och offline (Mott & Cockayne 2020). Sociala medier har bidragit till att förstärka och normalisera ett språkbruk som bygger på rasism, sexism och främlingsfientlighet (Klein 2012; Ekman 2015; Jakubowicz 2017; Schwarzenegger & Wagner 2018). Sökmotorer, politiska bloggar, nyhetskanaler och sociala medier bidrar till att falska och kränkande påståenden legitimeras och ges trovärdighet (Klein 2012). Andra forskare understryker att det strukturella förtryck som nu tar fart på nätet har funnits långt innan den nya tekniska utvecklingen (Mott & Cockayne 2020).

När det gäller forskarnas tankar om hur man kan motverka framväxten av populism och de problem i form av exempelvis främlingsfientlighet som den ger upphov till så liknar svaren de som gavs i förhållande till de postfaktiska strömningarna. Vikten av utbildningsinsatser som från ung ålder fokuserar på kritiskt tänkande understryks (Splitter 2019). Olika former av moderering och städning i sociala medier och kommentarsfält kan möjligen ha en viss positiv effekt men kommer inte att lösa den grundläggande problematiken (Fangen & Holter 2020).

## Näthat

– 30 granskade publikationer

De publikationer som analyseras inom ramen för den här studien är alla på engelska och det begrepp som vanligen används för att beskriva det fenomen som vi vill förstå är hate speech (jfr Fernandez Alvarez & Kouadio Alexis 2014; Eckert 2018; (Steiner 2015); Schmuck & Tribastone 2020; Tontodimamma et al. 2020). På svenska har vi valt att tala om näthat och detta kan ta sig uttryck genom olika slags kränkningar och hot av olika juridisk allvarlighetsgrad. Det övergripande forskningsområdet som rör näthat kan i sin tur delas in i olika inriktningar. Genom en bibliometrisk analys har Tontodimamma et al. (2020) identifierat tre dominerande inriktningar: (a) forskning som problematiserar förhållandet mellan näthat och yttrandefrihet; (b) forskning som utvecklar och testar olika automatiserade (AI- och maskininlärningsbaserade) tekniker för att upptäcka och ta bort näthat; och (c) forskning som studerar nätkränkningar i form av sexism och nätmobbning. De pekar också på två tillväxtfaser: den första 1992–2010 med en långsam ökning av antalet publikationer och den andra 2010–2019 med en omfattande forskningstillväxt.

Även efter 2019 har den snabba ökningen av forskningen på området fortsatt. På ett övergripande plan visar forskningen att de initiala förhoppningarna rörande internet och sociala mediers potential att främja en demokratisk utveckling i flera avseenden har visat sig vara överdrivna eller till och med felaktiga. Exempelvis visar omfattande forskning på hur människor ofta undviker att engagera sig i diskussioner på nätet på grund av risken att utsättas för näthat (Gearhart & Zhang 2014; Lumsden & Morgan 2017; jfr Nadim & Fladmoe 2019; Syahputra 2019; Porten-Chee, Kunst & Emmer 2020). I synnerhet gäller detta för politiska diskussioner på sociala medier. Det framkommer även att många ser det som lönlöst att diskutera seriösa ämnen på sociala medier på grund av den omfattande åsiktpolariseringen och trångsyntheten gentemot meningsmotståndare (Kruse, Norris & Flinchum 2018). Det utbredda näthatet har skapat en kollektiv passivitet som innebär att människor väljer att avstå från att delta i politiska samtal (Syahputra 2019; Khan & Roshan 2020). Även Gearhart och Zhang (2014) visar hur samtalstonen på sociala medier gör att människor ofta avstår från att yttra sig i olika frågor.

Graden av näthat har betydelse för i vilken utsträckning människor minskar sitt deltagande online. En allmänt otrevlig ton har en viss effekt medan hotfulla och hatiska kommentarer i högre grad leder till att människor drar sig undan ett aktivt deltagande i diskussioner (Rossini 2020).

Flera studier har fokuserat på HBTQ-personers och kvinnors deltagande online mot bakgrund av det ökande sexistiska näthatet. Detta kan analyseras under tre dominerande teman: (a) att skrämma, svartmåla och skambelägga kvinnor; (2) att cementera kvinnors underlägsenhet och heteronormativitetens överlägsenhet; och (3) att förespråka patriarkala samhällsstrukturer (Sobieraj 2018). Sexistiska yttranden har normaliserats i många etablerade digitala miljöer och sexistiska kommentarer är ofta grafiska och innehåller grova hot (jfr Lumsden & Morgan 2017; McLaughlin 2019; Nadim & Fladmoe 2019; Sanders 2019; Slater 2019; Waldman 2019; Chadha et al. 2020). En rad studier visar att kvinnor, som ett resultat av detta, upplever sig begränsas i det offentliga (digitala) samtalet. Detta uttrycks både av kvinnor som själva blivit utsatta för näthat och av kvinnor som inte utsatts men som är medvetna om riskerna på nätet (Fernandez Alvarez & Kouadio Alexis 2014; Lumsden & Morgan 2017; Eckert 2018; Sobieraj 2018; McClintock 2019; Nadim & Fladmoe 2019; Kanai & McGrane 2020).

Det är tydligt att kvinnor som uttrycker sympati för feministiska åsikter eller multikulturalism är särskilt utsatta. Exempelvis visar en av artiklarna, som behandlar kvinnliga bloggare som skriver om feminism, familj och mödravård i Tyskland, Schweiz, USA och England, att över 70 procent blivit utsatta för allvarliga hot (bland annat hot om dödligt våld och våldtäkt). På ett liknande sätt visar Duguay et al. (2020) att queerpersoners möjlighet att vara aktiva på Tinder, Instagram och Vine är kraftigt begränsad och att de ofta använder olika former av självzensur som strategi för att undvika trakasserier. Samma studie visar även att nätbolagens användarregler och anmälningsfunktioner inte innebär något reellt skydd mot trakasserier.

Även om det inte råder något tvivel om att sexistiskt näthat begränsar kvinnors möjligheter att fritt uttrycka sig på nätet är det mindre klart om kvinnor, på ett övergripande plan, utsätts för mer näthat än män. Detta beror på att befintlig forskning i hög grad begränsats till mindre befolkningsmängder eller utvalda sociala medier. Mot denna bakgrund undersökte Nadim och Fladmoe (2019) skillnaderna mellan kvinnors och mäns utsatthet genom två storskaliga befolkningsundersökningar i Norge. Deras resultat var något överraskande, då män uppvisade mer erfarenhet av näthat än kvinnor. Dock anser författarna att resultatet kan förklaras med att män engagerar sig mer i diskussioner på nätet än vad kvinnor gör. Detta stöds till viss del av att män i huvudsak utsätts för attacker förknippade med deras åsikter eller attityder, medan kvinnor framför allt utsätts för sexistiska attacker. Resultaten demonstrerar samtidigt att sexistiska trakasserier i högre grad leder till självzensur.

En annan grupp som har identifierats som särskilt utsatt för näthat är unga personer. Under de senaste åren har forskningen på området haft ett uttalat fokus på nätmobbning, som kortfattat beskrivs som en digital förlängning av traditionell fysisk mobbning (Fernandez Alvarez & Kouadio Alexis 2014; Kokkinos et al. 2016; Chadha et al. 2020). Flera studier kommer fram till att flickor och unga kvinnor utsätts mer än pojkar och unga män, samt att sexistiska kommentarer är vanliga.

Det visar sig också att inte många unga söker socialt eller rättsligt stöd när de utsätts för trakasserier. Orsaken är att utsattheten ofta är förknippad med känslor av skam (Fernandez Alvarez & Kouadio Alexis 2014). I stället utvecklar de som utsätts olika strategier för att handskas med näthatet (Byrne 2020). Exempelvis visar en studie att kvinnliga universitetsstudenter ofta använder sig av olika defensiva strategier såsom att normalisera trakasserier och att undvika riskfyllda digitala miljöer (Chadha et al. 2020). Likt vuxna undviker unga personer att diskutera politiska frågor på sociala medier då det upplevs som riskfyllt (Ekström 2016; Powers Koliska & Guha 2019).

Vad gäller psykologiska konsekvenser av näthat finns det starka indikationer att utsatthet kan bidra till psykisk ohälsa i form av oro, ångest och stress (Gearhart & Zhang 2014; Eckert 2018; Duguay, Burgess & Suzor 2020). Det framkommer även att känslan av utsatthet kan dröja kvar under längre perioder. Detta är särskilt tydligt i de fall där en person utsätts på grund av grupp- eller könstillhörighet. Exempelvis visade en studie att unga muslimers exponering för invandringskritiska och rasistiska budskap bidrog till en långvarig (2–3 veckor) försiktighet på nätet och en allmän obehagskänsla av diskriminering (Schmuck & Tribastone 2020).

Olika lösningar för övervakning och moderering av diskussioner på nätet har tagits fram och används i dag flitigt av såväl mediebolag och plattformar som enskilda publicister. Strategin bygger på en idé om att det är ett fåtal väldigt aktiva deltagare i debatter och diskussioner som sätter agendan och avskräcker andra från att delta (Porlezza 2019; Porten-Chee, Kunst & Emmer 2020; Wright, Trott & Jones 2020). Det finns emellertid inte mycket stöd i forskningen för idén om att moderering (manuell eller automatiserad) är ett effektivt instrument. Däremot beskrivs counter-speech eller collective civil moderation, vilket i korta drag innebär strategiska och samordnade insatser som syftar till att förbättra samtalskvaliteten på nätet, som mer lovande (Friess, Ziegele & Heinbach 2020; Jääskeläinen 2020). Även om det inte är sannolikt att sådana insatser kommer leda till att näthatet upphör menar flera forskare att de kan ha en normbildande effekt samt uppmuntra en mer civiliserad och



konstruktiv samtalston (Friess, Ziegele & Heinbach 2020; Porten-Chee, Kunst & Emmer 2020). En studie som berör mobbning mot homosexuella stödjer detta. Studien visade att positiva kommentarer avseende samkönade äktenskap kunde dämpa hatiska kommentarer och generera en mer öppen och tolerant samtalston (Gearhart & Zhang 2014).

### **Näthat mot yrkesgrupper med viktiga funktioner i demokratin**

– 18 granskade publikationer

Majoriteten av publikationerna på det här området fokuserar på den ökade utsattheten bland offentliga yrkesgrupper, såsom politiker och journalister (jfr Binns 2017; Dorostkar & Preisinger 2017; Gardiner 2018; Scaramuzzino 2020; Waisbord 2020), men även andra yrkesgrupper, såsom akademiker, avhandlas (Barlow & Awan 2016; Marwick 2020).

Även på det här området visar forskningen att det finns tydliga mönster av sexism, rasism och främlingsfientlighet kopplade till hatet. Personer som i sin professionella roll är positivt inställda till feminism, invandring och jämställdhet är särskilt utsatta. Akademiker och journalister som skriver om feminism och antirasism drabbas ofta av näthat i organiserad form. Kvinnor och personer med utländsk bakgrund utsätts för hot och hat som är kopplat till deras könstillhörighet och etnicitet (jfr Barlow & Awan 2016; Koirala 2020; Sobieraj et al. 2020; Wagner 2020).

Gardiner (2018) visar, genom en analys av kommentarer under tidningen The Guardians nyhetsartiklar som modererats bort på grund av trakasserier eller desinformation, att artiklar skrivna av kvinnliga journalister eller journalister med utländsk bakgrund var överrepresenterade när det gällde förekomsten av hot och hat. Kvinnor som i sin journalistiska gärning bevakar mansdominerade sektorer eller företeelser visade sig vara mest drabbade. Hatet mot dessa innehöll en hög grad av sexism och angrepp på ett personligt plan.

Tromble och Koolehar visar i en studie (2020) att kvinnliga politikere utrymme att interagera med allmänheten på Twitter försvåras som ett resultat av sexistiska inlägg och hatiska responser.

Det är tydligt att akademiker (Barlow & Awan 2016; Marwick 2020), journalister (jfr Binns 2017; Koirala 2020; Scaramuzzino 2020) och politiker (Theocharis et al. 2016; Wagner 2020) som arbetar med feministiska och antirasistiska frågor kontinuerligt motarbetas genom mer eller mindre organiserat hat. Wagner (2020) betonar exempelvis att näthatet har en negativ inverkan på kvinnliga politikere demokratiska deltagande och att det finns en

organiserad avsikt att få kvinnor och minoritetsgrupper att känna sig otrygga på sociala medier. Samtidigt visar forskningen att sociala medier är ett viktigt verktyg för dessa yrkesgrupper (Theocharis et al. 2016; Marwick 2020). Att helt avstå från yrkesrelaterad interaktion med sociala medier är således ofta inte ett alternativ. I regel är det ett förhållandevis litet antal personer som står för en stor andel av näthatet (Hellmueller et al. 2020).

Utsatthet för näthat är vidare starkt förknippad med offentlig exponering: ju mer högprofilerad eller omtalad en person är, desto större är risken att denna blir utsatt (Gorrell et al. 2020). Trots att fokus har speciellt riktats mot kvinnliga politikers, journalisters och akademikers utsatthet för sexistiskt näthat, framkommer det i vissa studier att män utsätts för en större omfattning av trakasserier och attacker (ibid). Andra studier visar att det inte finns någon skillnad mellan mäns och kvinnors utsatthet; exempelvis visar en av artiklarna som analyserar Twitterdata från USA, Nederländerna och Storbritannien att det inte förekommer någon olikhet i tonen mellan kommentarer skickade till kvinnliga och manliga politiker. I alla tre länderna var dessutom könsrelaterade förolämpningar ovanliga och i Nederländerna var tonen mot kvinnliga politiker mer positiv än mot män (Tromble & Koole 2020). Det är dock viktigt att påpeka att dessa studier (Gorrell et al. 2020; Tromble & Koole 2020) använder Twitterdata från en relativt kort tidsperiod, sex respektive två veckor.

När det kommer till näthatets konsekvenser inriktar sig en övervägande del av forskningen på personliga konsekvenser, det vill säga hur det psykologiska välbefinnandet påverkas. Det står klart att näthatet har stor negativ inverkan på både individens privatliv och yrkesliv. Det finns begränsad forskning som undersöker konsekvenser ur ett samhällsperspektiv, exempelvis förekomsten av själv censur som leder till åsiktsförflyttningar. Själv censur förekommer bland både journalister, politiker och akademiker (Barlow & Awan 2016; Hess & Waller 2020; Scaramuzzino 2020). Ett betydande antal personer i dessa yrkesgrupper justerar sina uttalanden eller undviker att uttrycka sig kring specifika ämnen för att undvika hot och hat. Exempelvis visar en nyligen genomförd svensk studie att tre av fyra journalister anpassar hur de skriver som ett direkt resultat av näthat (Scaramuzzino 2020). Vidare påvisas att ett antal personer har lämnat sitt yrke och att ett något större antal har övervägt detta. Med andra ord står det klart att näthat leder till själv censur bland personer med samhällsviktiga och opinionsbildande yrken, samt att det finns en medvetenhet om vilka ämnen som genererar mest hat. Dock har själv censuren inte problematiserats i förhållande till demokratins förutsättningar i någon betydande utsträckning. Det finns emellertid ett intressant undantag där själv censuren som fenomen belyses ur ett makroperspektiv. Waisbord (2020)

visar hur det ökade näthatet mot yrkesgrupper med viktiga demokratiska funktioner, i detta fall journalister, kan förstås utifrån tre förklarande faktorer: (1) tillgången till digitala medel; (2) en framväxande hatkultur; och (3) ökande populism och den medföljande demoniseringen av traditionella nyhetsmedier. Waisbord menar att denna utveckling har skapar en så kallad mobbcensur – en ny form av censur som skiljer sig från statlig censur i att den pådrivs av civila medborgare som i allmänhet agerar inom lagens gränser. Mobbcensur kan ses som ett resultat av en underifrån kommande vigilantism som syftar till att disciplinera och tysta åsikter i den offentliga debatten. Detta antagande är något som implicit stöds av flera andra forskare som menar att populistiska rörelser har ett tydligt syfte att kontrollera den offentliga diskussionen genom att angripa fakta, nyheter och information som inte återspeglar deras politiska agenda (Barlow & Awan 2016; Meza, Vincze & Mogos, 2018; Sobieraj et al. 2020). Forskarna är i det närmaste eniga i att varna för allvarliga konsekvenser för demokratiska samhällen (Turner 2018). Dock poängteras att det finns ett stort behov av mer forskning på området.

Även om det generellt sett råder brist på empiriska studier som undersöker strukturella former av självcensur så finns det i forskningen målande exempel. I en studie av Hess och Walter (2020) beskrivs hur en lokaltidning i en mindre stad aktivt valde att undvika rapportering kring nyheten om bygget av en ny moské för att undvika rasistiskt näthat. Lokaltidningens beslut motiverades utifrån en moralisk kompass avseende huruvida viss journalistik skulle generera lämpligt eller olämpligt diskursivt beteende, snarare än idéer om journalistikens roll i samhället. I det här fallet ersattes opartisk rapportering till följd av ett beslut att motverka spridningen av hat, rasism och intolerans i lokalsamhället. Hess och Walter (2020) menar att den här typen av självcensur sannolikt inte är unik.

Teknologiska strategier (till exempel moderering och blockering av vissa användare) ses som något som tillfälligt kan lindra skadorna från näthat som riktas mot olika professioner. Men även här efterfrågas utbildningsåtgärder för att angripa problematikens grund (Ziegele et al. 2018). Samtidigt påpekas att sociala medieplattformar bör ta mer ansvar för att motverka den negativa utvecklingen. Flera av artiklarna belyser även att det finns ett stort behov för arbetsgivare att genomgå utbildning och utforma strategier för att hjälpa de som utsätts inom sitt yrke (Sobieraj et al. 2020; Wagner 2020).

## Juridik och yttrandefrihet

– 30 granskade publikationer

Den granskade litteraturen utgår från en problembild där digitalisering och sociala medier har skapat nya förutsättningar och utmaningar för yttrandefriheten (Bangstad 2014; Bakalis 2018; Celik 2018; Edwards et al. 2020; Lowe 2020; Tenove 2020). Inte minst har detta undersökts ur ett regleringsperspektiv. En av de avgörande frågorna handlar om huruvida näthatets negativa konsekvenser är så omfattande att det motiverar en rättsligt reglerad begränsning av yttrandefriheten. Forskningen på området är ofta komparativ och jämför amerikansk och europeisk reglering och rättslig praktik. En vanlig reflektion är att begreppet näthat är svårdefinierat och att beskrivningen av de kriterier som krävs för att begränsa yttrandefriheten varierar (Bakalis 2018; Cavaliere 2019). Det går att identifiera en debatt mellan två läger vad gäller balansen mellan yttrandefrihet och näthat. Många forskare drar slutsatsen att den reglering som syftar till att minska näthatets konsekvenser bör stärkas kraftigt (jfr Bangstad 2014; Bakalis 2018; Balkin 2018; Miklaszewska 2019; Lowe 2020). Andra menar att inskränkningar av yttrandefriheten kan leda till maktmissbruk som hindrar människor från deltagande på nätet (Poyhtari 2014; Berkey 2017; Penney 2017; Jones 2018). Flertalet forskare hävdar emellertid att problematiken kräver väl balanserade avvägningar där eventuella ingrepp i yttrandefriheten balanseras mot den eventuella nyttan (se exempelvis Bakalis 2018; van Benthem 2018; van Dijk 2020). En vanlig slutsats är att lagstiftningen i dag är fragmenterad och inte ger tillräckligt skydd till offren (Bakalis 2018).

Den yttrandefrihet som i stor utsträckning tog sin form omkring 1900-talets början anses av flera vara otillräcklig för dagens digitala samhälle. Balkin (2018) för fram att det krävs lagstiftning som kan fånga olika typer av hot och hat på nätet, exempelvis hot mot individer och grupper, men även lagar som kan förhindra radikaliserings- och organiserade former av näthat. Näthatet utgör ett allvarligt problem för den fria åsiktsbildningen och dagens lagstiftning snarare skyddar de som trakasserar och hotar än de som faktiskt utsätts (Guiora & Park, 2017).

I några av de granskade studierna har empiriska undersökningar genomförts avseende internetanvändares attityder till reglering av hot och hat på sociala medier. Dessa studier visar att normer kring yttrandefrihetens betydelse har förändrats under de senaste åren. Det finns en känsla av apati och cynism beträffande sociala medieföretags hantering av näthat och deras vilja att bidra till en respektfull digital samtalston (Celik 2018; Johnson 2018; Mac

Sithigh 2020). En av studierna visar att det finns en utbredd uppfattning att utövarna av olika former av näthat bör lagföras och att det inte är rimligt att näthat försvaras av yttrandefrihetsprinciper (Celik 2018). Den sortens attityder och uppfattningar är mer utbredda bland personer som tillhör olika minoritetsgrupper. Kvinnor anser oftare än män att sociala medieplattformar bör ha en skyldighet att skydda deras rätt att vara fredade på sociala medier (Bangstad 2014; Celik 2018; Johnson, 2018; Enarsson & Lindgren 2019; Erdosova 2019). Bristande skydd medför att människor undviker att dela innehåll som handlar om deras trosuppfattning, etniska bakgrund, sexuella läggning och politiska åsikter (Celik 2018).

Annan forskning pekar ut unga personer som särskilt sårbara och att socio-ekonomiskt utsatta grupper drabbas i större utsträckning än andra (Celik 2018). De sociala medieplattformarna bör ges ökat ansvar (Celik 2018; Johnson 2018) för att värna rätten för alla att uttrycka sig utan att drabbas av hat. Även i den här typen av studier understryks vikten av utbildningsinsatser i digital kompetens och kritiskt tänkande (Celik 2018; Johnson 2018; Edwards et al. 2020). Flera forskare menar att det utöver utbildningsinsatser krävs striktare och tydligare lagstiftning på området (Erdosova 2019).

Bland annat Liang (2015) argumenterar utifrån att det skett ett gradvis paradigmskifte där censurproblematiken inte längre enbart handlar om staters ingrepp i yttrandefriheten genom lagar och andra maktmedel. Den nya situationen innebär att censur utövas av populistiska rörelser, politiska ledare och individer på nätet, som använder det utrymme som yttrandefrihetslagarna ger för att förtrycka meningsmotståndare, ofta genom rasistiska och sexistiska attacker. Demokratiska samhällen måste försvara utsatta grupper mot de personer och rörelser som begränsar grundläggande rättigheter genom att sprida hot och hat på nätet (Guiora & Park 2017; Celik 2018; Limba & Sidlauskas 2018; Enarsson & Lindgren 2019; Erdosova 2019).

Flera forskare pekar på att de sociala medieplattformarna bör ta ökat ansvar för upprätthållandet av demokratin (Wagner 2014; Milosevic 2016; Einwiller & Kim 2020; Rogers 2020; van Dijck 2020). Det är allmänt erkänt i forskningen att privata plattformsföretag, som inte sällan sätter kommersiella intressen före offentlig samhällsnytta, påverkar samhällets struktur och demokratiska processer (Einwiller & Kim 2020; van Dijck 2020). Såväl statliga som internationella aktörer sätter allt kraftigare press på sociala medieföretag att ta bort olagligt och olämpligt innehåll (Kasakowski et al. 2020).

Trots detta menar van Dijck (2020) att det fortfarande finns ett behov av att stater och internationella organisationer, i ännu högre grad än vad som redan gjorts, utvecklar gemensamma värden och normer kring det digitala offentliga rummet (se även Lysaker & Syse 2016). Med andra ord, om internet ska kunna vara ett öppet och demokratiskt forum krävs det insatser från ett brett spektrum av aktörer, alltifrån internationella organisationer (till exempel EU och FN) till enskilda stater, myndigheter och civilsamhället (Wagner 2014; Cohen-Almagor 2015, 2017; Miklaszewska 2019; van Dijck 2020).

### **Tekniska lösningar för att motverka näthat och desinformation**

– 12 granskade publikationer

Den här forskningen behandlar olika automatiserade funktioner som med hjälp av maskininlärningstekniker och algoritmer motverkar spridningen av näthat och desinformation. Denna litteratur är av förklarliga skäl tekniskt orienterad och understryker vikten av att förstå den digitala kodens normerande karaktär. I det digitala samhället utgör digital kod ett viktigt regleringsinstrument.

Under de senaste åren har flera studier publicerats som beskriver hur man med hjälp av avancerade maskininlärningstekniker och artificiell intelligens försöker möta de utmaningar som uppstår i takt med framväxten av desinformation och näthat. Inom ramen för denna litteraturöversikt har vi identifierat ett antal studier som presenterar lösningar för att motverka spridningen av näthat och nätmobbning på sociala medier (Van Royen et al. 2017; Artime et al. 2020; Muneer & Fati 2020; Ullmann & Tomalin 2020), samt tekniska metoder för att skilja mellan faktabaserade och falska nyheter (de Oliveira, Medeiros & Mattos 2020; Singh, Ghosh & Sonagara 2021).

Det har även gjorts försök att klassificera hatiskt, intolerant och sexistiskt språkbruk på sociala medier i syfte att minska utsattheten bland kvinnor och HBTQ-personer (Dias Oliva, Antonialli & Gomes 2020; Pamungkas, Basile & Patti 2020; Rodriguez-Sanchez, Carrillo-de-Albornoz & Plaza 2020), samt system som kan identifiera och blockera sociala botar i syfte att minska spridningen av desinformation (Pennekamp et al. 2019; Lundberg & Laitinen 2020; Monakhov 2020).

Denna översikt har inte för avsikt att redogöra för de ingående tekniska resonemangen som den här forskningen presenterar, utan nöjer sig med en övergripande reflektion om att detta forskningsområde är relativt ungt men att snabba framsteg görs.

### 3.3 Nationell kunskapsöversikt

I det här avsnittet kommer vi att behandla svenskspråkig och rapportbaserad litteratur (bland annat så kallad grå litteratur). Följande teman kommer att behandlas.

- Introduktion till de rättsliga ramarna avseende näthat
- Digitaliseringens inverkan på det demokratiska samtalet
- Digitala påverkansförsök på svenska val
- Näthat som fenomen
- Allmänhetens utsatthet för näthat
- Särskilt utsatta samhällsgrupper
- Särskilt utsatta yrkesgrupper
- Rättens hantering av näthat

Dessa teman speglar översiktligt den tillgängliga kunskapen på området avseende förhållanden i Sverige. Eftersom de publikationer som använts för den här delen av kunskapsöversikten inte nödvändigtvis är vetenskapliga kan de inte heller sökas fram på det systematiska sätt som gällde för den internationella studien. Det finns helt enkelt ingen samlad databas för den här typen av litteratur, där sådana systematiska studier kan genomföras. Som ett resultat blir denna del av kunskapsöversikten inte heltäckande, utan bygger snarare på ett urval.

#### **Introduktion till de rättsliga ramarna avseende näthat**

Som det ofta brukar framhållas är näthat inget rättsligt begrepp eller specifikt brott utan det är ett samlingsbegrepp som beskriver när någon kränker, trakasserar eller utsätter någon annan person för hot via internet, till exempel på sociala medier eller med e-post. Näthatsbrott kan således begås genom flera olika brottsliga gärningar, men ofta nämns brotten olaga hot, ofredande, ärekränkning (förtal och förolämpning), olaga integritetsintrång, kränkande fotografering, hets mot folkgrupp och sexuellt ofredande. Eftersom näthat också har relevans för demokratiskt deltagande och delaktighet aktualiseras även den grundlagsreglerade yttrandefriheten. De rättsliga förutsättningarna och ramarna rörande samlingsbegreppet näthat som här kommer att kortfattat presenteras tar endast hänsyn till svensk gällande rätt.

Näthat handlar från rättslig synvinkel i regel om att traditionella brott, ofta handlingar kriminaliserade i brottsbalkens (BrB) 4–6 kapitel där brott mot

frihet och frid, ärekränkingsbrott och sexualbrott beskrivs, begås mot någon annan genom digitala kontakter och meddelanden eller genom att uppgifter publiceras och sprids online. Speciellt i BrB 4 kap. finns flera av de gärningar kriminaliserade som brukar inräknas i näthat. I kapitlet återfinns exempelvis olaga hot (BrB 4:5), ofredande (BrB 4:7), olovlig identitetsanvändning (BrB 4:6b) och olaga integritetsintrång (BrB 4:6c). Därtill återfinns samlingsbrotten olaga förföljelse, eller så kallad stalkning (BrB 4:4b), olaga integritetsintrång (BrB 4:6c), samt grov fridskränkning och grov kvinnofridskränkning (BrB 4:4a). Kapitlet innehåller även brottet dataintrång, även kallat facerape (BrB 4:9c), som kan aktualiseras vid näthat eller nätkränkningar. I BrB 5 kap. återfinns ärekränkingsbrotten förtal (BrB 5:1) och förolämpning (BrB 5:3) som ofta aktualiseras när någon sprider känsliga personuppgifter eller hatar eller kränker någon i meddelanden eller liknande. I BrB 6 kap. beskrivs sexualbrotten, där det främst är sexuellt ofredande (BrB 6:10) som förknippas med näthat, även om flera andra brott i kapitlet kan genomföras online och aktualiseras vid spridning av känsliga personuppgifter som ett led i att hata eller kränka annan.

Vad som delvis utmärker flera av de ovan nämnda brotten är att straffskalorna har relativt låga maximum. Därtill är brotten inte kriminaliserade innan fullbordat brott föreligger enligt BrB 23 kap. (försök, förberedelse, stämpling och medverkan till brott). Båda dessa faktorer kan tänkas påverka rättsväsendets förmåga att utreda och lagföra gärningspersonerna. Näthat i form av näthat kan dock tänkas aktualiseras vid stämpling och anstiftan, det vill säga det straffrättsliga ansvaret för att psykologiskt förmå någon annan att begå eller försöka begå ett brott. På liknande sätt kan näthat tänkas aktualiseras vid brottet övergrepp i rättssak: hot om våld för att någon har anmält ett brott eller vittnat i en rättegång eller försök att hindra en person att anmäla och vittna (BrB 10:17). Detsamma gäller vid hedersrelaterade brott som kan inbegripa olika repressalier såsom hat och hot (BrB 29:2 p. 10). En annan straffskärpningsregel, som kanske oftare brukar förknippas med näthat, är när ett motiv för ett brott har varit att kränka en person, en folkgrupp eller en annan sådan grupp av personer på grund av något av följande: ras, hudfärg, nationellt eller etniskt ursprung, trosbekännelse, sexuell läggning, könsöverskridande identitet eller uttryck eller annan liknande omständighet (BrB 29:2 p. 7). Med hatbrott avses även brotten hets mot folkgrupp (BrB 16:8) och olaga diskriminering (BrB 16:9), vilka båda är att anses som brott mot allmän ordning, där även brott som uppvigling, barnpornografibrott och olaga våldsskildring ingår. Brottet hets mot folkgrupp, som alltså innebär att i det offentliga sprida uttalanden som hotar eller uttrycker missaktning för en



eller flera av de i lagen utpekade folkgrupperna, är även ett grundlagsreglerat tryckfrihets- och yttrandefrihetsbrott vilket aktualiserar den grundlagsreglerade yttrandefriheten. Även om de traditionella analoga brotten i hög grad tillämpas på nyare digitala fenomen såsom näthat har en del försök gjorts för att anpassa brottskatalogen till de aktuella problemen. Relativt nyligen gjordes en sådan bred översyn av flera av de brott som beskrivits ovan för att modernisera och anpassa det straffrättsliga skyddet för den personliga integriteten (Justitiedepartementet 2016).

Näthat riskerar att begränsa det demokratiska deltagandet och delaktigheten varför vi nedan kommer att beröra de grundlagsreglerade fri- och rättigheter som är särskilt relevanta. Även om näthat många gånger straffrättsligt begås av enskilda gärningspersoner och är riktat mot andra enskilda så kan brottslighetens sammantagna effekt skada demokratin och grundläggande fri- och rättigheter i samhället och det finns således även en rad relevanta bestämmelser av statsrättslig karaktär att beakta.

Enligt regeringsformen (RF) har varje medborgare rätt att i tal, skrift eller bild meddela upplysningar och uttrycka tankar, åsikter och känslor (RF 2:1). Bland de grundläggande fri- och rättigheterna återfinns en rad olika opinionsfriheter såsom yttrandefrihet, informationsfrihet och tryckfrihet. Det finns dock en rad angivna förutsättningar för begränsningar av dessa fri- och rättigheter, men detta får endast göras för att tillgodose ändamål som är godtagbara i ett demokratiskt samhälle. Begränsningen får vidare bara vara en sistahandslösning och får inte utgöra ett hot mot den fria åsiktsbildningen eller enbart vara på grund av politisk, religiös, kulturell eller annan sådan åskådning (RF 2:21). Rikets säkerhet, folkförsörjningen, allmän ordning och säkerhet, enskildas anseende, privatlivets helgd eller förebyggandet och beivrandet av brott kan vidare vara anledningar att begränsa yttrandefriheten och informationsfriheten, samtidigt som begränsningar rörande politiska, religiösa, fackliga, vetenskapliga och kulturella angelägenheter ska undvikas (RF 2:23). Tryckfrihetsförordningen (TF) inkluderar ett antal gärningar som anses vara tryckfrihetsbrott och som delvis överlappar med ovan beskrivna bestämmelser i brottsbalken: här återfinns exempelvis brotten olaga hot, förtal, förolämpning, hets mot folkgrupp, övergrepp i rättssak och hot mot tjänsteman. Det finns således ett samband mellan brottskatalogerna i BrB och TF och det mer detaljerade innehållet anpassas efterhand till teknik- och samhällsutvecklingen (Utredningen om ett modernt och starkt straffrättsligt skydd för den personliga integriteten 2016). Utmaningen ligger i att avväga integriteten och det straffrättsliga skyddet mot yttrandefriheten. Trenden är dock att lagstiftaren delvis stärker skyddet mot näthat genom att reglera, och därmed i någon mån

begränsa, yttrandefriheten. Tryck- och yttrandefriheten är som redan nämnts en relativ frihet som kan begränsas vid särskilda behov. 2018 års tryck- och yttrandefrihetskommitté (2020), som är en parlamentariskt sammansatt kommitté för att se över regleringen på området i ett antal hänseenden, beaktar även olika hot mot den fria åsiktsbildningen, även om det inte är ett centralt tema för kommittén. Yttrandefriheten är samtidigt även skyddad i Europakonventionen. Enligt artikel 8 om rätt till skydd för privat- och familjeliv har det enligt Europadomstolens praxis även ansetts vara statens ansvar att säkerställa straffbestämmelser och utredningsmöjligheter avseende olika former av näthat. Samtidigt finns det ett ökat behov av att stärka det rättsliga skyddet för journalister. Enligt kommittédirektiv (Justitiedepartementet 2020) ska en utredare se över det straffrättsliga skyddet för vissa samhällsnyttiga funktioner och några andra straffrättsliga frågor. Genom tidigare utredningar har lagstiftaren även skärpt lagstiftningen för att motverka näthat: exempelvis har lagen om ansvar för elektroniska anslagstavlor (1998:112) ändrats för att tydliggöra skyldigheten att ta bort vissa brottsliga meddelanden från till exempel kommentarfält på bloggar (Justitiedepartementet 2016).

### **Digitaliseringens inverkan på det demokratiska samtalet**

Internet och sociala medier har skapat nya förutsättningar för populistiska och antidemokratiska rörelser att sprida information och åsikter, mobilisera grupper och skapa engagemang. Grupper som tidigare givits begränsat utrymme i den offentliga debatten får nu ökat utrymme genom den nya tekniken och det nya medielandskapet (Sundell 2018; Kaminski 2020). Åsikter som tidigare isolerats till slutna sammanhang har förflyttats till kommersiella sociala medier; där rasistiskt, sexistiskt och främlingsfientligt språkbruk successivt normaliserats (Bladini 2017; Nationell satsning på medie- och informationskunnighet och det demokratiska samtalet 2020). Förändringar i medielandskapet och nyhetskonsumtionen har inneburit att etablerade nyhetsmedier förlorat en stor del av sina annonsintäkter till plattformsföretagen. Bristen på publicistiskt ansvar bland de många digitala nyhetsmedierna har i sin tur inneburit att de fundamentala banden mellan massmedia, allmänintressen och demokrati har försvagats (Medieutredningen 2016; Myndigheten för samhällsskydd och beredskap [MSB] 2016; Carlsson (red) 2018).

Svenska myndigheter belyser här att det nya digitala medielandskapet har banat väg för spridningen av desinformation och illvillig propaganda med syfte att underminera och störa det demokratiska samtalet (MSB 2019; Giandomenico & Linderstål 2020). Sammanfattningsvis är det tydligt att

den digitala medieekologin både har polariserat mediekonsumtionen och förenklat spridandet av antagonistiska åsikter. Detta har i sin tur genererat starka åsiktpolariseringar online, vilket i kombination med att det är lätt att kommunicera på sociala medier skapat en produktiv miljö för personer som vill uttrycka hatiska och hotfulla åsikter om andra.

### **Digitala påverkansförsök på svenska val**

Inom ramen för den här kunskapsöversikten har vi funnit ett fåtal studier som undersöker digital påverkan kopplad till det svenska valet 2018. Vad som är framträdande i den här litteraturen är sättet på vilket den demonstrerar de svenska alternativmediesajternas inflytande över det digitala politiska samtalet. Exempelvis visar en studie utförd av Oxford Internet Institute (OII) att svenska användare av sociala medier delade en stor mängd artiklar – avsevärt fler än i andra europeiska länder – från högervidna alternativa nyhetsmedier i anslutning till valet 2018 (till exempel Samhällsnytt och Fria Tider) (Hedman et al. 2018). Liknande resultat framkommer i rapporten Smearing Sweden: International Influence Campaigns in the 2018 Swedish Election, där man identifierar ett relativt isolerat högernätverk i Sverige vars mest konsekventa desinformationskampanj relaterade till anklagelser om valbedrägerier (Colliver et al. 2018).

Därutöver visar en studie genomförd av Totalförsvarets forskningsinstitut (FOI) på uppdrag av Myndigheten för samhällsskydd och beredskap att botar stod för en ansenlig mängd av det publicerade innehållet på Twitter som var kopplat till valet 2018 och svensk politik under tidsperioden 5 mars till 30 september (Fernquist et al. 2018). Studiens innehållsanalys visar att kritik mot invandring och negativa konsekvenser av invandring var ett ämne som diskuterats flitigt på Twitter i samband med valet.

### **Näthat som fenomen**

Ett antal källor intresserar sig för vilka som står bakom näthatet. Dock framförs det att det är svårt att ge en tydlig bild av den typiska näthataren (Dunkels 2016). Detta beror på att näthat är ett brett fenomen som förekommer i många olika former och sammanhang. Exempelvis skiljer sig ungas utsatthet från vuxnas i den avseendet att unga personer ofta känner förövaren, eller åtminstone vet vem denna är (Friends 2017), medan det är vanligare att identiteten på den som utsätter vuxna personer för näthat är okänd för den utsatte – detta gäller i synnerhet för personer som offentligt engagerar sig i olika samhällsfrågor (Brottsförebyggande Rådet [Brå] 2015a; Bladini 2017; Civil Rights Defenders

2019). Flera källor pekar på att det främst är män som ägnar sig åt näthat, även om andra publikationer påvisar att till exempel universitetsstudenter och välutbildade kvinnor också är aktiva i sammanhanget (Dunkels 2016).

När det kommer till mer organiserade former av näthat, exempelvis mot förtroendevalda och journalister, är könsskillnaden tydlig. Baserat på förtroendevaldas egna utsagor är förövaren i första hand en förargad manlig medborgare, eller rättshaverist, i 45–60 års åldern (Brå, 2019b). Dessa resultat stöds av en studie utförd av FOI där man undersökt vilka personer som uttrycker hat mot eller skriver nedsättande kommentarer om journalister på nätforumet Flashback (Fernquist et al. 2019). Majoriteten visade sig vara män i åldern 40–60 år.

Ett förhållandevis litet antal personer anses stå för en stor majoritet av hotfulla kommentarer riktade mot journalister (Kaati & Olsson 2018). Samtidigt är det vanligt att förtroendevalda, journalister, opinionsbildare och kulturpersoner utsätts för näthat av personer som tillhör eller sympatiserar med högerextrema grupper (MSB 2016; Myndigheten för kulturanalys 2016; Brå 2019; Fernquist et al. 2019). Något överraskande framkommer det även i Brås studie om hot och hat mot förtroendevalda att 16 procent av de som någon gång utsatts uppfattade att förövaren var en annan förtroendevald (Brå 2019c).

### Allmänhetens utsatthet för näthat

Sociala medier har på kort tid blivit ett självklart inslag i många människors vardag och brukas av en stor majoritet – 83 procent – av svenska internetanvändare (Internetstiftelsen 2019). Unga personer är i majoritet på sociala medier och föredrar plattformar såsom Snapchat och TikTok framför Facebook där i många fall deras föräldrar befinner sig (ibid). Mot bakgrund av att internet och sociala medier används av en stor majoritet av befolkningen är det inte förvånande att utsatthet för IT-relaterade brott också blivit allt vanligare. Enligt Brå (2016) ökade andelen personbrott med IT-inslag från 11 till 26 procent mellan åren 2006 och 2014. Ökningen innefattar framför allt en växande utsatthet för hotfulla och kränkande meddelanden på sociala medier. Däremot är det fortfarande en förhållandevis låg andel av befolkningen (2,5 procent) som utsatts för nätkränkningar av brottslig karaktär (Brå 2019d). Detta är dock en ökning från föregående år då andelen var 2,1 procent.

Det går dock att urskilja från litteraturen att risken för att utsättas för näthat ökar när man uttrycker sin åsikt angående en aktuell samhällsfråga online. Exempelvis visar en rapport från SOM-institutet att 5 procent av de tillfrågade respondenterna blivit utsatta för hot, trakasserier eller våld i samband med

att de uttryckte en åsikt gällande en aktuell samhällsfråga på internet, till exempel på sociala medier, bloggar eller diskussionsforum (Tipple 2019). I Internetsstiftelsens studie är siffran ännu högre, cirka 14 procent.

Brå har undersökt utsatthet för näthat och presenterat resultaten i olika rapporter. Enligt senaste Nationella trygghetsundersökningen (NTU) har 2,6 procent av respondenterna i åldern 16–84 år svarat ja på enkätfrågan ”Har någon i syfte att kränka eller skada dig spridit känsliga uppgifter, bilder, filmer eller kommentarer om dig på internet under förra året?” (Brå 2020a, s. 94). Utifrån självrapporterad utsatthet för nätkränkning under åren 2016–2019 är andelen utsatta män något större än andelen kvinnor (ibid). Unga kvinnor i åldersgruppen 16–19 år uppger dock högst utsatthet för nätkränkning (8,1 procent) (Brå 2020a, s. 95). Annars är könsfördelningen i fråga om utsatthet för nätkränkning och näthat relativt jämn, även om kvinnor använder sociala medier mer än män (Brå 2019b, s. 7). Tittar man på självrapporterad utsatthet för kränkningar via internet bland unga finns en svagt positiv trend. Enligt den årliga Skolundersökningen om brott (SUB) som undersöker utsattheten hos niondeklassare, uppger nästan en av fyra elever att de blivit utsatta för att någon har skrivit kränkande saker om dem på internet. Drygt en av fem niondeklassare uppger att de blivit utsatta för att någon har lagt upp bilder eller filmklipp på dem, som de inte vill ska spridas på internet, men bland flickorna uppger färre att de utsatts för att någon lagt upp oönskade bilder eller filmklipp på internet 2019 än 2015 (Brå 2020b).

## Särskilt utsatta samhällsgrupper

### Sexistiskt och rasistiskt näthat

Det är påtagligt att vissa samhällsgrupper är mer utsatta på nätet än andra. Det handlar framför allt om HBTQ-personer, personer med funktionsnedsättningar, personer tillhörande etniska minoriteter och unga personer (Bladini 2017; Statens medieråd 2019; Landsrådet för Sveriges ungdomsorganisationer [LSU] 2020). Trots att mycket tyder på att kvinnor i allmänhet utsätts mer än män, återfinns den största skillnaden i hur utsattheten yttrar sig. Kvinnor får motta näthat som i stor utsträckning är kränkande i fråga om deras könstillhörighet. Följaktligen framgår det i litteraturen att hot och hat ofta är kopplat till kön, makt och sexualitet. Exempelvis menar Totalförsvarets forskningsinstitut (2020) att Sverige kan vara det land i världen som har flest så kallade incels per invånare, det vill säga män som lever i ofrivilligt celibat och sprider hat mot kvinnor på internet.

Även om det sexistiska hatet är ett allmänt utbrett problem är utsattheten som störst bland personer som debatterar ämnen som feminism, genus, jämställdhet och rasism på nätet (Kulturdepartementet 2017). Det är således uppenbart att näthatet är förenat med sexistiska, rasistiska och främlingsfientliga diskurser. Sveriges Kvinnolobby har exempelvis i en ny rapport uppmärksammat att hat och hot mot kvinnor är ett växande problem och föreslår att kön bör införas som grund för beteckningen hat- och terrorbrott när motivet är könsbaserat (Sveriges Kvinnolobby 2021). Detta är något som även föreslås av Bladini (2017) som genomfört en kartläggning av den rättsliga regleringen kring hot och hat i Norden.

I ljuset av vad som beskrivs ovan har brottsanmälningar för nätrelaterade hatbrott, det vill säga olika brott som begås på grundval av gärningspersonens negativa inställning till en annan persons sexuella läggning, hudfärg eller religionstillhörighet, fördubblats mellan 2016 och 2018 (Brå 2019a). Samtidigt visar en studie av Delegationen mot segregation (2020) att boende i områden med socioekonomiska utmaningar i större utsträckning har avstått från att delta i olika politiska aktiviteter på grund av hot, trakasserier eller våld, i jämförelse med de som bor i områden utan socioekonomiska utmaningar och i riket i stort.

### **Näthat och nätmobbning mot unga personer**

För unga personer är sociala medier numera den vanligaste brottsplatsen för hot, kränkningar och trakasserier (Brå 2020b). Enligt Internetstiftelsen (2019) har tre av tio personer under 25 år någon gång drabbats av näthat, vilket är en ökning sedan 2016 (Statens medieråd 2019), och bland elever i årskurs 9 har drygt 42 procent av flickorna och drygt 30 procent av pojkarna utsatts för hot som skett via internet och sociala medier (Brå 2020b). Följaktligen kan man konstatera att flickor drabbas något mer än pojkar. Utsattheten skiljer sig dock mest i den meningen att hot och hat mot flickor ofta förknippas med oönskad spridning av bilder och sexuella trakasserier, medan pojkar i större utsträckning får motta dödshot eller hot om våld med vapen (ibid.). I en undersökning utförd av Friends (2017) uppgav exempelvis 20 procent av flickorna att de under det senaste året utsatts för sexuella trakasserier. Unga flickor är dessutom mer oroliga för att utsättas för näthat än vad pojkar är (Myndigheten för ungdoms- och civilsamhällsfrågor [MUCF] 2019). Vidare visar Friends (2017) rapport att 7 av 10 som utsätts för nätkränkningar har fått dem från någon de känner, eller vet vem det är. Det framkommer även att mer än hälften av de unga har sett vuxna bete sig illa mot andra på nätet.

Mot bakgrunden av ungas omfattande utsatthet på nätet har nätmobbning under en längre period undersökts flitigt inom den svenska kontexten. Nätmobbning kännetecknas, precis som mobbning utanför internet, i huvudsak av tre kriterier: avsikt att skada, repetition och obalans i maktförhållandet (Statens medieråd 2020). För nätmobbning finns det alltså en tydlig definition, vilket skiljer begreppet från näthat som innefattar ett bredare register av obehagliga händelser på nätet. Med andra ord är all nätmobbning också näthat – men allt näthat är inte nätmobbning (ibid.). I Folkhälsomyndighetens (2018) undersökning Skolbarns hälsovanor, som bygger på svar från 4 500 barn i åldrarna 11, 13 och 15 år, kan man se att mobbning generellt har ökat. Detta är ett anmärkningsvärt resultat då mobbningens omfattning varit densamma i Sverige sedan början av 1900-talet. Det ska dock påpekas att det är svårt att med säkerhet säga om internet och sociala medier är orsaken till denna ökning, i och med att nätmobbning ofta sker parallellt med fysiska former av mobbning (ibid.).

Det är vanligast att nätmobbning relaterar till den utsattes utseende och intressen (Friends 2017; Statens medieråd 2020). Samtidigt har nätmobbning ofta en koppling till diskrimineringsgrunderna kön, religion, sexualitet, könsidentitet och funktionsnedsättning (Statens medieråd 2020). Man har funnit att konsekvenserna av nätmobbning i huvudsak är desamma som vid mobbning utanför internet. Däremot har man sett att social ångest verkar vara en specifik riskfaktor i fråga om utsatthet för nätmobbning, samt att den psykiska hälsan avsevärt påverkas vilket leder till känslomässiga reaktioner såsom ilska, oro, rädsla, nedstämdhet och skam (ibid.). Exempelvis beskriver Maria Fridh, doktor i medicinsk vetenskap vid Lunds universitet, i sin avhandling, vars empiri bygger på enkätsvar från skolelever i årskurs nio, att det finns tydliga kopplingar mellan nätmobbning och psykisk ohälsa (Fridh 2018). Ett annat utmärkande drag är att den som utsatts ofta utvecklar en negativ kroppsuppfattning, det vill säga de är mer missnöjda med sin kropp och utseende (Statens medieråd 2020). Exempelvis uppgav nära hälften av de tillfrågade flickorna i en undersökning från Länsförsäkringar (2018) att sociala medier gör att de inte känner sig nöjda med sina kroppar. En annan undersökning från Lunds universitet visar att flickors självförtroende avsevärt försämras efter mellanstadiet, vilket förklaras av det ständiga flödet av bilder av ouppnåeliga kroppsideal på sociala medier (Hoff, Daukantaitė & Birgerstam 2020).

Det finns även indikationer på att förekomsten av grooming och sexuell exploatering på nätet ökar (Sveriges Kvinnolobby 2021). En undersökning från 2018 visar att hälften av alla 15-åriga flickor har blivit kontaktade i sexuella

syften av någon de tror är vuxen (Childhood 2018). Vidare har även, sedan lagen om olaga integritetsintrång infördes 2018, anmälningar av så kallad hämndporr och olovlig spridning av bilder på nätet ökat – även om polisen bedömer att mörkertalet fortfarande är mycket stort. År 2019 anmäldes 1 437 brott i form av olaga integritetsintrång varav endast 16 lagfördes (Brå 2019e). Smygfilmning och spridning av kränkande bilder på sociala medier utförs oftast av pojkar mot flickor under 18 år.

### **Särskilt utsatta yrkesgrupper**

Som angetts ovan är andelen enskilda personer som utsatts för näthat relativt låg; risken ökar dock om personen i fråga är ung, kvinna eller tillhörande en minoritetsgrupp. En annan starkt bidragande faktor förenlig med utsatthet är engagemang i det offentliga samtalet. Med andra ord, ju mer en person syns i offentliga sammanhang eller är aktiv på sociala medier, desto större är risken att hen utsätts för hot och hat. Därför uppmärksammas det i litteraturen att näthat i synnerhet riktas mot vissa yrkesgrupper: politiska förtroendevalda, journalister och opinionsbildare (så kallade influencers) och andra aktörer som aktivt deltar i det demokratiska samtalet. Även konstnärer, forskare, personer som är aktiva inom civilsamhället samt företrädare för rättsväsendet eller statlig och kommunal förvaltning utsätts i hög grad. Den nationella grå litteraturen behandlar således ett bredare urval av yrkesgrupper än vad som iakttagits under den internationella litteraturgenomgången. Nedan följer en beskrivning av omfattningen och konsekvenserna av näthat mot särskilt utsatta yrkesgrupper.

### **Politiskt förtroendevalda**

Hot och hat mot politiskt förtroendevalda har alltmer uppmärksamats i Sverige under de senaste åren. Bland annat har Brå sedan 2014 publicerat Politikernas trygghetsundersökning (PTU) där omfattningen och konsekvenserna av trakasserier, hot och våld mot förtroendevalda undersöks (Brå 2012, 2014, 2015b, 2017, 2019c). Den senaste undersökningen bekräftar att näthat mot förtroendevalda fortsatt utgör ett allvarligt hot mot det demokratiska samhället och, i jämförelse med de tidigare PTU-undersökningarna, ser den negativa utvecklingen inte ut att avta. Under 2018 har var tredje förtroendevald utsatts för trakasserier, hot och våld i samband med sitt politiska uppdrag. Man kan vidare se att oron för utsatthet är större hos kvinnor än hos män, samt att oron är större hos de som sedan tidigare blivit utsatta. Totalt sett är det en något större andel kvinnliga än manliga förtroendevalda som uppger att de blivit utsatta för trakasserier, hot eller våld (32 respektive 30 procent). Majoriteten av dem utsattes under 2018 vid



upprepade tillfällen: endast 16 procent uppgav att det enbart gällde ett enskilda tillfälle. Resultaten visar även att yngre förtroendevalda är mer utsatta än äldre, och att förtroendevalda med utländsk bakgrund är mer utsatta än personer med svensk bakgrund (Brå 2019c).

Det finns därutöver en stark korrelation mellan graden av offentlig exponering och utsatthet. Förtroendevalda som är mer aktiva på internet och sociala medier, eller mer omtalade i massmedia, utsätts oftare än de som inte är lika aktiva eller omskrivna. Vidare är storstadspolitiker, riksdagsledamöter och ordförande särskilt drabbade (Brå 2019c). Liknande resultat påvisas i den rapport om digitalt hat mot kommunpolitiker som genomfördes av dåvarande Sveriges Kommuner och Landsting (SKL) tillsammans med Totalförsvarets forskningsinstitut (FOI). Rapporten visar att 20 procent av alla kommentarer på öppna diskussionsforum och alternativa medier som handlar om en kommunstyrelseordförande innehöll någon form av hat, hot eller kränkningar (SKL 2018a). Samtidigt innehöll kommentarer som riktade sig mot kvinnliga kommunpolitiker grövre och i högre grad sexistiska inslag (ibid). Det är även vanligare att kvinnor uppger att de fått hotfulla e-postmeddelanden och utsatts för hot och hat på sociala medier, ofta i form av meddelanden innehållande sexuella anspelningar (Brå 2019c). Här efterfrågas att säkerhetspersonal i riksdagen och säkerhetsansvariga i kommuner och regioner samt inom politiska partier ökar kunskapen om könsbaserat hat, hot och trakasserier (Sveriges Kvinnolobby 2021).

Totalt uppger nästan var fjärde förtroendevald i Brås (2019c) undersökning att hot och hat, eller oron för att utsättas för det, påverkat dem i deras uppdrag. Andelen är större bland de förtroendevalda som redan har erfarenhet av att motta hot och hat (42 procent). Den vanligaste konsekvensen av utsatthet, eller oron för att utsättas, är att man undviker att engagera eller uttala sig i en specifik fråga. Hela 30 procent av de tillfrågade anger att de censurerat sig själva för att undvika trakasserier, hot eller våld. Därutöver har 18 procent av de som utsatts övervägt att lämna ett förtroendeuppdrag med anledning av hot och hat, och 5 procent har lämnat ett uppdrag (ibid). För att få en djupare förståelse av den beskrivna problematiken har SKL genomfört en intervjustudie med förtroendevalda med erfarenhet av näthat. Liksom i Brås undersökning poängterar intervjupersonerna att den ökade utsattheten har en negativ inverkan på det demokratiska samtalet, samtidigt som det i hög grad påverkar de förtroendevaldas privatliv (SKL 2018b).

## Journalister

Precis som förtroendevalda är journalister en yrkesgrupp som historiskt sett fått motta kritik och kränkningar i samband med sitt arbete. I och med sociala mediers ökade popularitet i samhället har även journalisters utsatthet för näthat uppmärksammats inom den svenska kontexten. År 2015 presenterade Brå en kunskapsöversikt om utsattheten inom yrkesgrupper som är viktiga för det demokratiska samtalet. I studien noterade Brå att en stor majoritet av svenska journalister någon gång blivit utsatta för kränkningar online (Brå 2015a). Liknande resultat påvisas i en undersökning genomförd av Institutionen för journalistik, medier och kommunikation (JMG) vid Göteborgs universitet, där 70 procent av de tillfrågade uppgav att de utsatts för nedsättande eller kränkande kommentarer, och 30 procent att de utsatts för hot, under de senaste tolv månaderna (Löfgren Nilsson 2017). Det är vanligast att hoten genomförs via e-post eller på sociala medier, följt av hotfulla telefonsamtal. Hot genom artikelkommentarer står för en mindre del och har kontinuerligt minskat i takt med att kommentarsfält stängs ner eller kraftigt modereras. I samma undersökning angav de tillfrågade att de upplevde att motivet bakom hatet var kopplat till negativa inställningar till personers hudfärg, nationalitet eller etniska bakgrund, samt religionstillhörighet. Samtidigt riktas hot och hat i synnerhet mot journalister som engagerar sig i frågor som rör immigration, flyktmottagning, integration, infrastruktur, utrikeskonflikter, religion och klimat (ibid.).

Som tidigare nämnts vittnas det i flera undersökningar om att kvinnor är något mer utsatta än män, både vad gäller specifika yrkesgrupper och bland allmänheten. Resultaten från JMG:s undersökning sammanfaller dock inte med detta antagande: tvärtom är män i högre grad utsatta än kvinnor. Samma resultat har kommit fram i en tidigare studie som genomförts vid Göteborgs universitet (Löfgren Nilsson 2013, 2017). Detta betyder dock inte att kvinnliga journalister inte är en utsatt grupp. En studie genomförd av branschorganisationen Tidningsutgivarna i Sverige, visar exempelvis att 72 procent av de tillfrågade kvinnliga chefredaktörerna, ledarskribenterna, krönikörerna och kommentatorerna har utsatts för hot eller trakasserier under de senaste tolv månaderna. I samma studie uppgav var tredje kvinnlig journalist att de funderat på att sluta med journalistiken till följd av sin utsatthet (TU-Medier i Sverige 2016). Liknande siffror redovisas i JMG:s undersökning där 25 procent av samtliga tillfrågade journalister uppgav att de någon gång har funderat på att lämna journalistiken (Löfgren Nilsson 2017).

Hot och hat förefaller även vara ett problem på organisationsnivå, i Utgivarnas och TSN Sifos undersökning av hot mot mediehus, redaktioner och medarbetare uppgav 42 av de 94 svarande redaktionerna att de hotats under det senaste året (Utgivarna 2016). Även i denna undersökning återfinns självcensurerande tendenser då medarbetare på 12 av de 42 hotade redaktionerna tvekat inför en publicering och medarbetare på 11 av de hotade redaktionerna har undvikit bevakningsområden eller arbetsuppgifter efter allvarliga hot.

Även om problematiken är utbredd är profilerade journalister i högre grad utsatta än andra journalister. Exempelvis visar en studie utförd av FOI, som fokuserar på hot och hat mot journalister på fyra av de mest profilerade högervridna diskussionsforumen, att det är ett fåtal journalister som ofta nämns. De mest nämnda journalisterna, främst ledarskribenter och politiska kommentatorer från både vänster- och högerkanten, förekom i drygt 10 procent av omnämningarna. Studien visar även att det är ett fåtal användare som står för 20–30 procent av det totala antalet hatfulla inlägg om journalister (Kaati & Olsson 2018).

Många svenska journalister upplever att hotet och trakasserierna syftar till att påverka deras yrkesutövning (Löfgren Nilsson 2017). I linje med en sådan observation kan förövarens handlande sägas vara strategiskt i bemärkelsen att hen vill åstadkomma en förändring i bevakningen kring vissa ämnesfrågor. En avgörande aspekt i relation till näthatets demokratiska konsekvenser är därför huruvida journalister faktiskt påverkas av sin utsatthet. Som vi tidigare angett rapporteras det att flera journalister överväger att lämna yrket på grund av hat och hot. Det framkommer även att mer än var fjärde utsatt journalist har undvikit att bevaka ett visst ämne eller fråga. Liknande siffror påvisas även bland journalister som någon gång undvikit att bevaka en viss individ eller grupp (ibid.). Med andra ord går det, precis som när det gäller förtroendevaldas utsatthet, att urskilja självcensurerande tendenser bland journalister. Det är däremot svårt att utifrån de undersökningar som gjorts dra mer konkreta slutsatser kring hur stor påverkan näthat har på den journalistiska rapporteringen och vilka eventuella konsekvenser detta har för det demokratiska samtalet.

### **Andra utsatta yrkesgrupper**

Utöver förtroendevalda och journalister vittnar andra aktörer med viktiga funktioner i demokratin om ökad utsatthet för nätrelaterat hot och hat. En studie genomförd av människorättsorganisationen Civil Rights Defenders

(2019) visar exempelvis att oberoende opinionsbildare (det vill säga frilansskribenter, debattörer, samt kultur- och civilsamhällesaktörer) som engagerar sig i frågor som rör jämställdhet, diskriminering, våldsbejakande extremism, migration och rasism ofta exponeras för hot och hat. Av omsorg om den egna säkerheten väljer därför många opinionsbildare att inte skriva eller prata om aktuella samhällsfrågor, och avstår inte sällan från att medverka i seminarier och andra offentliga engagemang. Likaså utsätts konstnärer, författare och andra kulturarbetare för hot och hat. En undersökning utförd av Myndigheten för kulturanalys (2016) visar att var tredje konstnär och författare någon gång varit utsatt och att var sjätte har drabbats av hot och hat under det senaste året, varav en stor del har skett över internet och sociala medier.

FOI genomförde nyligen en undersökning baserad på inhämtade data från diskussionsforumet Flashback, i syfte att undersöka hur det digitala hatet skiljer sig åt mellan kvinnliga och manliga journalister, politiker, influencers, komiker och musiker. I undersökningen framkom det att nästan hälften av alla individer (totalt 440) hade utsatts för hatfulla kommentarer, att kvinnor blir mer utsatta än män, samt att kvinnor får ta emot mer sexuella trakasserier och utseenderelaterade förolämpningar medan männen oftare förolämpas på grund av sitt yrke eller sin kompetens. Kvinnliga journalister och influencers var vidare den grupp som i störst utsträckning utsattes för hat (Fernquist et al. 2020).

En oroande utveckling är därutöver att hat och hot mot ungdomsrörelser ökat. Drygt hälften av de medverkande organisationerna i en rapport från Landsrådet för Sveriges Ungdomsorganisationer (LSU) uppgav att de utsatts för hot och hat. Nästan 20 procent av organisationerna har på olika sätt anpassat sin verksamhet på grund av hatet, exempelvis genom inställd medverkan på Almedalen och mässor, uteblivna debattartiklar och inlägg på sociala medier. Oroväckande i sammanhanget är att utsattheten i flera fall normaliseras – hatet blir en del av vardagen – något som kommer med uppdraget. Även här är unga kvinnor och personer med utländsk bakgrund särskilt drabbade (LSU 2020).

Ett annat resultat som dyker upp i den svenska litteraturen är att individer som saknar redaktionellt stöd eller en större organisation att vända sig till när de utsätts är särskilt sårbara (Civil Rights Defenders 2019). Beträffande opinionsbildare framhävs en allt större omfattning av organiserade former av hot- och hatdrev, både mot individer och grupper (Löfgren Nilsson 2017; Brå 2019c; Civil Rights Defenders 2019). Liknande resultat påvisas även bland forskare som ägnar sig åt frågor som kan uppfattas som politiskt kontroversiella (Bladini 2017).

## Rättens hantering av näthat

I litteraturen framhävs det att den ökade användningen av ny teknik och det eskalerande näthatet medför nya utmaningar för lagstiftaren, då rätten måste utvecklas på ett sätt som skyddar människor från olagliga former av hot och hat samtidigt som denna anpassning måste ske på ett sätt som inte i alltför stor utsträckning begränsar de positiva aspekterna av den nya tekniken (Enarsson 2017). Denna balansgång beskrivs ofta som paradoxal och ett vanligt antagande är att regler som skyddar människor från angrepp på nätet i sin tur begränsar friheten för andra. Detta är något som Márten Schultz, professor i civilrätt vid Stockholms universitet, bestämt bestrider. Enligt Schultz (2020) är inskränkningar i yttrandefriheten en självklarhet i demokratiska rättsstater i enlighet med grundprincipen att det föreligger en betydande risk att den starke tystar den svage i ett samhälle där dessa inskränkningar inte tillämpas (ibid.).

Flera undersökningar visar att staten inte förmår att tillhandahålla tillräckligt stöd åt brottsoffer som utsätts för näthat och att det stöd som finns är både svårtillgängligt och geografisk varierande (Bladini 2017). Att döma av personupplklaringsstatistiken är bristerna i rättens hantering av näthat markanta. Enligt Brå (2019a) upplklarades mindre än fem procent av anmälda nätrelaterade hatbrott under 2016. Den vanligaste orsaken till att så många anmälningar läggs ner rapporteras vara relaterade till utrednings- och bevissvårigheter.

Det framkommer vidare att kunskapen och resurserna avseende brott som begås online ligger långt efter brott som sker utanför internet (Sackemark, Schultz & Alderhorn 2015). Detta har även bekräftats av Polismyndigheten som medger att mer kunskap och resurser krävs för att få fler dömda (Polismyndigheten 2019). Samtidigt har organisationen Näthatsgranskaren (2019) riktat hård kritik mot Polismyndigheten angående deras arbete med demokrati- och hatbrott på internet: bland annat menar organisationen att Polisens arbete präglas av bristande engagemang och otillräcklig kunskap.

Vad gäller förtroendevalda, journalister, opinionsbildare och andra särskilt utsatta aktörer beskrivs förtroendet för rättsväsendets förmåga att hantera olagligt kränkningar på nätet som lågt (Löfgren Nilsson 2017; Brå 2019c; Civil Rights Defender, 2019; LSU 2020). En bekymmersam utveckling bland förtroendevalda är att många som utsätts för hot och hat online inte tror att en anmälan skulle leda till något eller att händelsen skulle betraktas som en småsak. Liknande mönster återfinns bland journalister: färre än var tredje journalist som utsatts för hot och hat vidtar någon åtgärd och endast tolv procent av dessa polisanmäler händelsen. Andelen polisanmälningar bland

journalister som utsatts har dessutom kraftigt minskat från cirka 30 procent år 2013 till 12 procent år 2017 (Löfgren Nilsson 2013, 2017). Ytterligare ett oroande resultat är att både journalister, förtroendevalda och andra opinionsbildare i flera fall anser att näthat är en del av jobbet. På samma sätt går det att konstatera att unga personer mycket sällan förlitar sig på att rättsväsendet ska erbjuda skydd och stöd vid utsatthet för nätrelaterade hot, trakasserier och kränkningar. Det finns också osäkerhet kring vilka lagar som gäller på nätet (Svensson & Dahlstrand 2014; Friends 2017).

Som ett resultat av att flera undersökningar, liksom utsagor från utsatta personer, visat att rättens hantering av olagliga hot och kränkningar på nätet är bristfällig har det under de senaste åren lyfts fram att det straffrättsliga skyddet är för svagt och straffvärdet för lågt. Detta gäller i synnerhet när det gäller utsatthet bland yrkesgrupper som är viktiga för det demokratiska samhället (Nationell satsning på medie- och informationskunnighet och det demokratiska samtalet 2020). Bland annat argumenterar Mårten Schultz (2020) för att hot och kränkningar som skadar demokratin ska erkännas som mer än ett brott mot en enskild individ, det vill säga omfattas av det straffrättsliga skyddet för det demokratiska samhället. Med andra ord anser Schultz att olagliga former av näthat mot offentliga opinionsbildare borde leda till strängare åtgärder än vad som sker idag. Ytterligare ett förslag som lyfts fram är att näthat med sexistiskt innehåll riktat mot kvinnor bör jämföras med våld mot kvinnor eftersom det finns betydande evidens att det på liknande sätt skapar rädsla och tystnad och begränsar kvinnors handlingsutrymme (Bladini 2017; Sveriges Kvinnolobby 2021).

Trots betydande brister i den rättsliga hanteringen av nätrelaterade kränkning brott har ett omfattande arbete bedrivits under de senaste åren för att komma till rätta med den nuvarande problematiken. Exempelvis har straffbestämmelserna beträffande olaga hot, ofredande, grovt förtal och förolämpningar förtydligats och moderniserats. Dessutom har skyddet för den personliga integriteten stärkts genom att bland annat spridning av vissa slags bilder eller andra uppgifter straffbeläggs (Justitiedepartementet 2016). Det har även införts en straffskärpning för olaga hot mot förtroendevalda (Justitiedepartementet 2018). Därutöver tillsatte regeringen i maj 2020 en särskild utredare för att bland annat analysera och ta ställning till hur man kan utforma ett förstärkt straffrättsligt skydd för vissa samhällsnyttiga funktioner och ingripa mot brott som begås mot någon som utövar sin yttrandefrihet, särskilt inom ramen för yrkesmässigt bedriven nyhetsförmedling eller annan journalistisk verksamhet, bör utformas (Justitiedepartementet 2020).

## 4. Sammanfattning och avslutande reflektioner

Den här sammanfattningen redogör kortfattat för innehållet i rapporten. Projektgruppen som har genomfört kunskapsöversikten består av Måns Svensson, professor i rättssociologi vid högskolan i Halmstad (projektledare), Oscar Björkenfeldt, doktorand i rättssociologi vid Lunds universitet, Fredrik Åström, docent och bibliometriker vid Lunds universitet och Karl Dahlstrand Fil. Dr. i rättssociologi vid Lunds universitet. I sammanfattningen redogörs för innehållet från de två delstudierna (a) en systematisk litteraturöversikt över den engelskspråkiga vetenskapliga och referentgranskade litteraturen med fokus på näthat och demokrati; och (b) en mer ostrukturerad kunskapsöversikt, baserad på nyckelordsbaserade sökningar efter svenskspråkiga utredningar, rapporter och liknande, så kallad grålitteratur. I den fullständiga rapporten redogörs även för resultatet av en tredje delstudie som bygger på en bibliometrisk analys på den litteratur som söktes fram inom ramen för den systematiska litteraturöversikten.

### 4.1 Systematisk kunskapsöversikt

I det här avsnittet redogörs översiktligt för de sex dominerande teman som identifierades i den internationella forskningen om näthat och demokratiskt deltagande. Ordningen mellan dessa teman är vald med utgångspunkt i att framställningen skall forma en logisk berättelse och skall inte ses som en rangordning av något slag. Den här delstudien genomfördes genom en systematisk sökning i forskningsdatabasen Web of Science där slutligen 129 vetenskapliga artiklar inkluderades.

#### **Det postfaktiska samhället**

– 26 granskade publikationer

Forskningen talar om ett postfaktiskt tillstånd (eller samhälle, post-truth society) där den offentliga debatten i hög utsträckning tar sin utgångspunkt i känsloladdade argument snarare än objektiva fakta och sanningar. Förenklat

diskuterar litteraturen inom det här temat hur ett, i hög grad algoritmberoende, informationsöverflöde förefaller att bidra till en åsiktpolarisering, där sökandet efter en gemensam verklighetsbeskrivning minskar i betydelse. Det digitala medielandskapet har bidragit till framväxten av ett postfaktiskt tillstånd i det offentliga samtalet där emotionellt styrda åsikter får en mer central betydelse, på bekostnad av mer faktabaserade och objektiva diskussioner. Forskningen pekar med andra ord på att fakta-begreppet är på väg att urholkas och förlora sin betydelse i en bred social och politisk mening.

### **Populism**

– 13 granskade publikationer

I förhållande till den postfaktiska samhällsutvecklingen fokuserar det här forskningstemat på den våg av populistiska rörelser som, under senare år, tagit allt större plats i västerländska demokratier. Här menar forskningen att antidemokratiska narrativ, kännetecknade av främlingsfientlighet, misstro mot etablerade medier, politiker och experter, har fått allt större fotfäste i det offentliga samtalet. Avsaknaden av en tillräcklig gemensam verklighetsuppfattning och en faktabaserad diskussion har skapat grogrund för populistiska politiska rörelser, som utnyttjar nätets möjligheter. Detta har i sin tur fört med sig en normalisering av emotionellt drivna hatiska och hotfullt uttryck.

### **Näthat**

– 30 granskade publikationer

Forskningen inom det här temat visar att det utbredda näthatet har skapat en kollektiv passivitet som innebär att människor ofta väljer att avstå från att delta i diskussioner kring aktuella samhällsfrågor. Kvinnor, unga människor och personer tillhörande olika minoritetsgrupper upplever i högre grad än andra att de är begränsade, när det gäller att uttrycka sina åsikter. Människor som tystas väljer sällan att söka hjälp utifrån – varken social eller rättslig. I stället tenderar utsatta att använda försvarsstrategier som går ut på att acceptera situationen (en slags normaliseringsprocess) och att undvika riskbeteenden (att uttrycka sina åsikter) och digitala riskmiljöer.

### **Näthat mot yrkesgrupper med viktiga funktioner i demokratin**

– 18 granskade publikationer

Inom det här forskningstemat fokuseras frågor som har att göra med att hot och hat mot yrkesgrupper med betydelse för demokratins upprätthållande – exempelvis journalister, politiker och akademiker – utgör ett utbrett problem. Även om hoten och hatet är riktat mot enskilda personer, så är det tydligt att det



är specifika perspektiv i det offentliga samtalet som angrips. Yrkesverksamma som lyfter fakta, kunskap, bevakning och debatt på ett sätt som uppfattas som att de bidrar till att stärka vissa perspektiv rörande feminism, invandring och jämställdhet är särskilt utsatta. Det finns allt starkare indikationer på att yrkesgrupper som har stor betydelse för demokratin i ökande utsträckning tillämpar olika former av självzensur för att hantera hot och hat. Inom forskningen på det här området råder det i dag en stor enighet om att den här utvecklingen utgör ett allvarligt hot mot den fria åsiktsbildningen och i förlängningen det demokratiska samhället.

### **Juridik och yttrandefrihet**

– 30 granskade publikationer

De rättsliga utmaningarna är framför all centerade kring balansen av olika skyddsvärda intressen. Detta avser i synnerhet det avvägande som har att göra med att avgöra om näthatets konsekvenser är så allvarliga att de motiverar begränsningar i yttrandefriheten. Även om det råder viss oenighet kring hur avvägningsfrågorna skall lösas, så är de flesta forskarna överens om att juridiken misslyckats med att anpassa sig till de utmaningar digitaliseringen medför. Bland annat pekar man på en risk att lagstiftningen idag tenderar att överfokusera skyddet av den traditionella yttrandefriheten på bekostnad av de som utsätts för hot och hat. Som ett resultat har en ny form av censurproblematik uppmärksammats där populistiska rörelser, politiska ledare och individer på nätet använder det utrymme som yttrandefrihetslagarna ger för att begränsa utrymmet för andra. Flertalet studier med den här inriktningen understryker att demokratiska rättsstater, på ett bättre sätt än vad som görs i nuläget, måste försvara utsatta samhällsgrupper mot de personer och rörelser som begränsar grundläggande rättigheter genom att sprida hot och hat på nätet.

### **Tekniska lösningar för att motverka näthat och desinformation**

– 12 granskade publikationer

Den teknikfokuserade forskningen på området är den inriktning som har tydligast fokus på att söka efter lösningar. Publikationerna på det här området redogör i huvudsak för studier vilka undersöker olika automatiserade funktioner som utifrån maskinlärningstekniker och algoritmer motverkar spridningen av näthat och desinformation. Det rör sig både om studier som presenterar lösningar för att motverka spridningen av näthat och nätmobbning på sociala medier och studier som undersöker tekniska metoder för att skilja mellan faktabaserade och falska nyheter. Det har även gjorts försök att klassificera hatiskt, intolerant och sexistiskt språkbruk på sociala medier i

syfte att minska utsattheten bland kvinnor och HBTQ-personer, samt system som kan identifiera och blockera sociala botar för att minska spridningen av desinformation. Det här forskningsområdet är relativt ungt men är under snabb utveckling.

## 4.2 Kunskapsläget om förhållandet i Sverige

Den existerande kunskapen om förhållandet i Sverige utgörs i hög utsträckning av så kallad grälitteratur. Det kan vara utredningar, rapporter, manuskript och liknande, framtagna av myndigheter, universitet och andra organisationer och företag. Gemensamt är att dessa publikationer inte genomgått en gängse vetenskaplig granskning. Kvaliteten kan likväl hålla mycket hög nivå. Eftersom den här typen av kunskap inte finns samlad i gemensamma databaser, på samma sätt som den vetenskapliga produktionen, är det inte möjligt att söka fram den med samma systematik. Den del av kunskapsöversikten som rör svenska förhållanden kan därför betecknas som ostrukturerad och strategisk i den meningen att vi med olika tekniker har sökt fram kunskap som svarar mot vårt behov.

Publikationerna som beskriver förhållandet i Sverige visar att situationen i hög utsträckning liknar den som beskrivs i den internationella litteraturen. Digitaliseringen har medfört en grogrund för tillväxt av populistiska och antidemokratiska åsikter; samt en normalisering av ett rasistiskt och sexistiskt språkbruk. När det gäller graden av utsatthet råder det delade meningar. Det finns rapporter som pekar på att en förhållandevis låg andel av den svenska befolkningen har utsatts för nätkränkningar som är så allvarliga att de kan bedömas som brottsliga. Andra källor menar att såväl spridning som allvar är betydande.

Det är tydligt att vissa grupper i befolkningen är mer utsatta än andra, det handlar framför allt om HBTQ-personer, personer med funktionsnedsättningar, etniska minoriteter och unga personer. Kvinnor och män utsätt i ungefär lika stor grad – även om formerna och uttrycken förefaller att skilja mellan könen. Kvinnor är i högre grad utsatta för sexistiska och angrepp och angrepp på person. När det gäller unga personer och barn är näthatet inte sällan av en grad att den får bedömas som mobbning. Ungas utsatthet har, precis som vuxnas, ofta en koppling till diskrimineringsgrunderna kön, religion, sexualitet, könsidentitet och funktionsnedsättning. Det finns starka belägg för att ungas utsatthet leder till olika former av psykisk ohälsa, i synnerhet bland unga flickor.

I Sverige utgör hot och hat mot grupper och personer med stor betydelse för demokratin ett betydande och växande problem. Detta gäller särskilt för journalister och förtroendevalda politiker, men även bland andra aktörer som deltar i det demokratiska samtalet. Även i den svenska kontexten är självcensur bland dessa grupper vanligt förekommande.

Slutligen står det klart att tillgänglig kunskap i Sverige pekar på att staten inte förmår att tillhandahålla tillräckligt stöd till brottsoffer som utsätts för hot och hat på nätet. En stor andel av de som har utsatts för trakasserier av en sådan grad att de kan bedömas som olagliga tror inte att en anmälan skulle medföra några åtgärder från rättssamhället. Flera tongivande aktörer menar att det krävs ett omfattande arbete bland myndigheter och det civila samhället för att komma tillrätta med den nuvarande problematiken.

### **4.3 Avslutande reflektioner kring strategier för att motverka problemen**

I en stor andel av de 129 internationella publikationerna finns tankar och idéer som handlar om möjliga lösningar på den problematik som beskrivs. Detsamma gäller för mycket av den grålitteratur som undersökts rörande svenska förhållanden.

Det är möjligt att gruppera de olika lösningarna med utgångspunkt i den modell som Lawrence Lessig, professor vid Harvard, presenterade första gången 1998. Modellen, som gjorts känd genom boken *Code and Other Laws of Cyberspace* utgår från att reglering av nätbeteenden måste ta hänsyn till andra reglerande modaliteter än lagen.

Utifrån den här modellen är det möjligt att gruppera potentiella strategier enligt följande: (a) Juridiska lösningar som inbegriper såväl law in books som law in action. Det vill säga tillskapandet av en adekvat reglering samt tillräckliga resurser för rättsväsendet att arbeta med regelefterlevnad. I detta skall man inte underskatta lagens och åtgärdernas signalvärde. (b) Tekniska lösningar som inbegriper en förståelse för den digitala kodens (algoritmernas) normerande kraft. I en digital kontext utgör den här formen normstrukturer kanske den mest tydligt reglerande modaliteten. (c) Marknadslösningar, vilket i fråga om nathatsproblematiken bland annat kräver de sociala medie-plattformarnas engagemang. Man diskuterar här bland annat hur man kan förändra spelreglerna för marknaden så att en positiv utveckling gynnas. (d) Strategier för att påverka de sociala normerna och den informella sociala kontrollen

människor emellan. Inom ramen för detta ligger bland annat olika typer av bildnings- och påverkanskampanjer.

Den ovan beskrivna modellen understryker vikten av att skapa en optimal regleringsmix som inbegriper flera eller samtliga av de utpekade reglerande modaliteterna. Statsmakternas roll blir med ett sådant synsätt att designa insatser som går i linje med det här sammansatta synsättet.

# Referenslista

## Systematisk litteraturoversikt

### Metod

Tranfield, D., Denyer, D. & Smart, P. (2003). Towards a methodology for developing evidence-informed management knowledge by means of systematic review. **British Journal of Management**, 14, ss. 207–222.

### Det postfaktiska samhället

Barton, C. C. (2019). Critical literacy in the post-truth media landscape. **Policy Futures in Education**, 17(8), ss. 1024–1036. doi: [10.1177/1478210319831569](https://doi.org/10.1177/1478210319831569).

Beam, M. A., Hmielowski, J. D. & Hutchens, M. J. (2018). Democratic digital inequalities: Threat and opportunity in online citizenship from motivation and ability. **American Behavioral Scientist**, 62(8), ss. 1079–1096.  
doi: [10.1177/0002764218764253](https://doi.org/10.1177/0002764218764253).

Bozdag, E. & van den Hoven, J. (2016). Breaking the filter bubble: Democracy and design. **Ethics and Information Technology**, 17(4), ss. 249–265.  
doi: [10.1007/s10676-015-9380-y](https://doi.org/10.1007/s10676-015-9380-y).

Clayton, K. et al. (2020). Real solutions for fake news? Measuring the effectiveness of general warnings and fact-check tags in reducing belief in false stories on social media. **Political Behavior**, 42(4), ss. 1073–1095.  
doi: [10.1007/s11109-019-09533-0](https://doi.org/10.1007/s11109-019-09533-0).

Dahlgren, P. (2018). Media, knowledge, and trust: The deepening epistemic crises of democracy. **Javnost-the Public**, 25(1–2), ss. 20–27.  
doi: [10.1080/13183222.2018.1418819](https://doi.org/10.1080/13183222.2018.1418819).

Duffy, A., Tandoc, E. & Ling, R. (2020). Too good to be true, too good not to share: The social utility of fake news. **Information Communication & Society**, 23(13), ss. 1965–1979. doi: [10.1080/1369118X.2019.1623904](https://doi.org/10.1080/1369118X.2019.1623904).

Estelles, M. & Castellvi, J. (2020). The educational implications of populism, emotions and digital hate speech: A dialogue with scholars from Canada, Chile, Spain, the UK, and the US. **Sustainability**, 12(15), s. 6034.  
doi: [10.3390/su12156034](https://doi.org/10.3390/su12156034).

Frost, C. (2020). The rower of voice: Bots, democracy, and the problem of political ventriloquism. **Journal of Political Power**, 13(1), ss. 6–21.  
doi: [10.1080/2158379X.2019.1701831](https://doi.org/10.1080/2158379X.2019.1701831).

Hannan, J. (2018). Trolling ourselves to death? Social media and post-truth politics. **European Journal of Communication**, 33(2), ss. 214–226.  
doi: [10.1177/0267323118760323](https://doi.org/10.1177/0267323118760323).

Hemsley, J. et al. (2018). Social media for social good or evil: An introduction. **Social Media + Society**, 4(3), s. 2056305118786710.  
doi: [10.1177/2056305118786710](https://doi.org/10.1177/2056305118786710).

Jameel, T., Ali, R. & Malik, K. A. (2019). Social media as an opinion formulator: A study on implications and recent developments. I **2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET)**. Sukkur, Pakistan. doi: [10.1109/iCoMET45996.2019](https://doi.org/10.1109/iCoMET45996.2019)

Kligler-Vilenchik, N., Baden, C. & Yarchi, M. (2020). Interpretative polarization across platforms: How political disagreement develops over time on Facebook, Twitter, and WhatsApp. **Social Media + Society**, 6(3), s. 2056305120944393.  
doi: [10.1177/2056305120944393](https://doi.org/10.1177/2056305120944393).

Lee, T. & Hosam, C. (2020). Fake news is real: The significance and sources of disbelief in mainstream media in Trump's America. **Sociological Forum**, 35, ss. 996–1018. doi: [10.1111/socf.12603](https://doi.org/10.1111/socf.12603).

Lu, J. & Liu, X. (2018). The nation-state in the digital age: A contextual analysis in 33 countries. **International Journal of Communication**, 12, ss. 110–130.

Luhrmann, A. et al. (2018). State of the world 2017: Autocratization and exclusion? **Democratization**, 25(8), ss. 1321–1340.  
doi: [10.1080/13510347.2018.1479693](https://doi.org/10.1080/13510347.2018.1479693).

Marres, N. (2018). Why we can't have our facts back. **Engaging Science Technology and Society**, 4, ss. 423–443. doi: [10.17351/ests2018.188](https://doi.org/10.17351/ests2018.188).

Miletskiy, V. P., Cherezov, D. N. & Strogetskaya, E. V. (2019). Transformations of professional political communications in the digital society (by the example of the fake news communication strategy). I Shaposnhikov, S. & Sharakhina, L. (red.) **Proceedings of the 2019 Ieee Communication Strategies in Digital Society Seminar (2019 Comsds)**, ss. 121–124.

Ross, B. et al. (2019). Are social bots a real threat? An agent-based model of the spiral of silence to analyse the impact of manipulative actors in social networks. **European Journal of Information Systems**, 28(4), ss. 394–412.  
doi: [10.1080/0960085X.2018.1560920](https://doi.org/10.1080/0960085X.2018.1560920).

Schaewitz, L. et al. (2020). When is disinformation (in)credible? Experimental findings on message characteristics and individual differences. **Mass Communication and Society**, 23(4), ss. 484–509.  
doi: [10.1080/15205436.2020.1716983](https://doi.org/10.1080/15205436.2020.1716983).

Sehl, A., Simon, F. M. & Schroeder, R. (2020). The populist campaigns against European public service media: Hot air or existential threat? **International Communication Gazette**, s. 1748048520939868.  
doi: [10.1177/1748048520939868](https://doi.org/10.1177/1748048520939868).

Shao, C. et al. (2018). The spread of low-credibility content by social bots. **Nature Communications**, 9, s. 4787. doi: [10.1038/s41467-018-06930-7](https://doi.org/10.1038/s41467-018-06930-7).

Stetka, V., Mazak, J. & Vochocova, L. (2020). “Nobody tells us what to write about”: The disinformation media ecosystem and its consumers in the Czech Republic. **Javnost-the Public**. doi: [10.1080/13183222.2020.1841381](https://doi.org/10.1080/13183222.2020.1841381).

Sussman, G. (2020). Making enemies: The mainstream media spectacle and US foreign policy. **Perspectives on Global Development and Technology**, 19(1–2), ss. 138–156. doi: [10.1163/15691497-12341546](https://doi.org/10.1163/15691497-12341546).

Valdez, A. C., Burbach, L. & Ziefle, M. (2018). Political opinions of Us and Them and the influence of digital media usage. I Meiselwitz, G. (red.) **Social Computing and Social Media: User Experience and Behavior, Scsm 2018: I**, ss. 189–202. doi: [10.1007/978-3-319-91521-0\\_15](https://doi.org/10.1007/978-3-319-91521-0_15).

Valdez, A. C. & Ziefle, M. (2019). Believability of news: Understanding users perceptions of fake news and the effectiveness of fact-checking badges. I Bagnara, S., Tartaglia, R., Albolino, S., Alexander, T. & Fujita, Y. (red.) **Proceedings of the 20th Congress of the International Ergonomics Association (IEA 2018), Vol V: Human Simulation and Virtual Environments, Work with Computing Systems (WWCS), Process Control**, ss. 469–477.  
doi: [10.1007/978-3-319-96077-7\\_50](https://doi.org/10.1007/978-3-319-96077-7_50).

Yerlikaya, T. (2020). Social media and fake news in the post-truth era: The manipulation of politics in the election process. *Insight Turkey*, 22(2), ss. 177–196. doi: [10.25253/99.2020222.11](https://doi.org/10.25253/99.2020222.11).

## Populism

Alvares, C. & Dahlgren, P. (2016). Populism, extremism and media: Mapping an uncertain terrain. *European Journal of Communication*, 31(1), ss. 46–57. doi: [10.1177/02673231155614485](https://doi.org/10.1177/02673231155614485).

Bonikowski, B. (2017). Ethno-nationalist populism and the mobilization of collective resentment. *British Journal of Sociology*, 68, ss. S181–S213. doi: [10.1111/1468-4446.12325](https://doi.org/10.1111/1468-4446.12325).

Davis, M. (2020). The online anti-public sphere. *European Journal of Cultural Studies*, s. 1367549420902799. doi: [10.1177/1367549420902799](https://doi.org/10.1177/1367549420902799).

Ekman, M. (2015). Online Islamophobia and the politics of fear: Manufacturing the green scare. *Ethnic and Racial Studies*, 38(11), ss. 1986–2002. doi: [10.1080/01419870.2015.1021264](https://doi.org/10.1080/01419870.2015.1021264).

Evolvi, G. (2019). Emotional politics, Islamophobic tweets, the hashtags #Brexit and #chiudiamoiporti. *Partecipazione E Conflitto*, 12(3), ss. 871–897. doi: [10.1285/i20356609vi2i3p871](https://doi.org/10.1285/i20356609vi2i3p871).

Fangen, K. & Holter, C. R. (2020). The battle for truth: How online newspaper commenters defend their censored expressions. *Poetics*, 80, s. 101423. doi: [10.1016/j.poetic.2019.101423](https://doi.org/10.1016/j.poetic.2019.101423).

Jakubowicz, A. (2017). Alt\_Right white lite: Trolling, hate speech and cyber racism on social media. *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, 9(3), ss. 41–60. doi: [10.5130/ccs.v9i3.5655](https://doi.org/10.5130/ccs.v9i3.5655).

Klein, A. (2012). Slipping racism into the mainstream: A theory of information laundering. *Communication Theory*, 22(4), ss. 427–448. doi: [10.1111/j.1468-2885.2012.01415.x](https://doi.org/10.1111/j.1468-2885.2012.01415.x).

Mott, C. & Cockayne, D. (2020). Understanding how hatred persists: Situating digital harassment in the long history of white supremacy. *Gender Place and Culture*. doi: [10.1080/0966369X.2020.1816920](https://doi.org/10.1080/0966369X.2020.1816920).

Schwarzenegger, C. & Wagner, A. (2018). Can it be hate if it is fun? Discursive ensembles of hatred and laughter in extreme right satire on Facebook. *Studies in Communication and Media*, 7(4), ss. 473–498. doi: [10.5771/2192-4007-2018-4-473](https://doi.org/10.5771/2192-4007-2018-4-473).



Splitter, L. J. (2019). Identity and populism begone! The role of philosophy in healing a shattered and divided world. **Childhood and Philosophy**, 15. doi: [10.12957/childphilo.2019.42962](https://doi.org/10.12957/childphilo.2019.42962).

Sponholz, L. & Christofolletti, R. (2019). From preachers to comedians: Ideal types of hate speakers in Brazil. **Global Media and Communication**, 15(1), ss. 67–84. doi: [10.1177/1742766518818870](https://doi.org/10.1177/1742766518818870).

Watson, B. R. & Riffe, D. (2013). Perceived threat, immigration policy support, and media coverage: Hostile media and presumed influence. **International Journal of Public Opinion Research**, 25(4), ss. 459–479. doi: [10.1093/ijpor/eds032](https://doi.org/10.1093/ijpor/eds032).

## Näthat

Byrne, V. L. (2020). Blocking and self-silencing: Undergraduate students' cyberbullying victimization and coping strategies. **Techtrends**. doi: [10.1007/s11528-020-00560-x](https://doi.org/10.1007/s11528-020-00560-x).

Chadha, K. et al. (2020). Women's responses to online harassment. **International Journal of Communication**, 14, ss. 239–257.

Duguay, S., Burgess, J. & Suzor, N. (2020). Queer women's experiences of patchwork platform governance on Tinder, Instagram, and Vine. **Convergence: The International Journal of Research into New Media Technologies**, 26(2), ss. 237–252. doi: [10.1177/1354856518781530](https://doi.org/10.1177/1354856518781530).

Eckert, S. (2018). Fighting for recognition: Online abuse of women bloggers in Germany, Switzerland, the United Kingdom, and the United States. **New Media & Society**, 20(4), ss. 1282–1302. doi: [10.1177/1461444816688457](https://doi.org/10.1177/1461444816688457).

Ekström, M. (2016). Young people's everyday political talk: A social achievement of democratic engagement. **Journal of Youth Studies**, 19(1), ss. 1–19. doi: [10.1080/13676261.2015.1048207](https://doi.org/10.1080/13676261.2015.1048207).

Fernandez Alvarez, O. & Kouadio Alexis, L. (2014). Violence against women and children: A distant and domestic hell. **Procedia – Social and Behavioral Sciences**, 161, ss. 7–11. doi: [10.1016/j.sbspro.2014.12.002](https://doi.org/10.1016/j.sbspro.2014.12.002).

Friess, D., Ziegele, M. & Heinbach, D. (2020). Collective civic moderation for deliberation? Exploring the links between citizens' organized engagement in comment sections and the deliberative quality of online discussions. **Political Communication**. doi: [10.1080/10584609.2020.1830322](https://doi.org/10.1080/10584609.2020.1830322).

- Gearhart, S. & Zhang, W. (2014). Gay bullying and online opinion expression: Testing spiral of silence in the social media environment. **Social Science Computer Review**, 32(1), ss. 18–36. doi: [10.1177/0894439313504261](https://doi.org/10.1177/0894439313504261).
- Jääskeläinen, T. (2020). Countering hate speech through arts and arts education: Addressing intersections and policy implications. **Policy Futures in Education**, 18(3), ss. 344–357. doi: [10.1177/1478210319848953](https://doi.org/10.1177/1478210319848953).
- Kanai, A. & McGrane, C. (2020). Feminist filter bubbles: Ambivalence, vigilance and labour. **Information Communication & Society**. doi: [10.1080/1369118X.2020.1760916](https://doi.org/10.1080/1369118X.2020.1760916).
- Khan, A. U. & Roshan, R. (2020). Fear of cyber communication threat effects on online freedom of expressions. **International Transaction Journal of Engineering Management & Applied Sciences & Technologies**, 11(11), s. 11A11J. doi: [10.14456/ITJEMAST.2020.214](https://doi.org/10.14456/ITJEMAST.2020.214).
- Kokkinos, C. M. et al. (2016). Parenting and internet behavior predictors of cyber-bullying and cyber-victimization among preadolescents. **Deviant Behavior**, 37(4), ss. 439–455. doi: [10.1080/01639625.2015.1060087](https://doi.org/10.1080/01639625.2015.1060087).
- Kruse, L. M., Norris, D. R. & Flinchum, J. R. (2018). Social media as a public sphere? Politics on social media. **Sociological Quarterly**, 59(1), ss. 62–84. doi: [10.1080/00380253.2017.1383143](https://doi.org/10.1080/00380253.2017.1383143).
- Lumsden, K. & Morgan, H. (2017). Media framing of trolling and online abuse: Silencing strategies, symbolic violence, and victim blaming. **Feminist Media Studies**, 17(6), ss. 926–940. doi: [10.1080/14680777.2017.1316755](https://doi.org/10.1080/14680777.2017.1316755).
- McClintock, E. A. (2019). Book review: Sexual harassment online: Shaming and silencing women in the digital age. **Gender & Society**, 33(5), ss. 820–822. doi: [10.1177/0891243219853759](https://doi.org/10.1177/0891243219853759).
- McLaughlin, H. (2019). Sexual harassment online: Shaming and silencing women in the digital age. **Contemporary Sociology: A Journal of Reviews**, 48(4), ss. 440–441. doi: [10.1177/0094306119853809X](https://doi.org/10.1177/0094306119853809X).
- Nadim, M. & Fladmoe, A. (2019). Silencing women? Gender and online harassment. **Social Science Computer Review**, s. 0894439319865518. doi: [10.1177/0894439319865518](https://doi.org/10.1177/0894439319865518).
- Porlezza, C. (2019). From participatory culture to participatory fatigue: The problem with the public. **Social Media + Society**, 5(3), s. 205630519856684. doi: [10.1177/205630519856684](https://doi.org/10.1177/205630519856684).

Porten-Chee, P., Kunst, M. & Emmer, M. (2020). Online civic intervention: A new form of political participation under conditions of a disruptive online discourse. **International Journal of Communication**, 14, ss. 514–534.

Powers, E., Koliska, M. & Guha, P. (2019). “Shouting matches and echo chambers”: Perceived identity threats and political self-censorship on social media. **International Journal of Communication**, 13, ss. 3630–3649.

Rossini, P. (2020). Beyond incivility: Understanding patterns of uncivil and intolerant discourse in online political talk. **Communication Research**, s. 0093650220921314. doi: [10.1177/0093650220921314](https://doi.org/10.1177/0093650220921314).

Sanders, T. (2019). Tania Levey: Sexual harassment online: Shaming and silencing women in the digital age. **Sexualities**, 22(7–8), ss. 1345–1347. doi: [10.1177/1363460718811225](https://doi.org/10.1177/1363460718811225).

Schmuck, D. & Tribastone, M. (2020). Muslims take action: How exposure to anti-Islamic populist political messages affects young Muslims’ support for collective action: A longitudinal experiment. **Political Communication**, 37(5), ss. 635–655. doi: [10.1080/10584609.2020.1736700](https://doi.org/10.1080/10584609.2020.1736700).

Slater, A. (2019). Sexual harassment online: Shaming and silencing women in the digital age: By Tania G. Levey. **Journal of Gender Studies**, 28(6), ss. 747–749. doi: [10.1080/14753634.2019.1630080](https://doi.org/10.1080/14753634.2019.1630080).

Sobieraj, S. (2018). Bitch, slut, skank, cunt: Patterned resistance to women’s visibility in digital publics. **Information Communication & Society**, 21(11), ss. 1700–1714. doi: [10.1080/1369118X.2017.1348535](https://doi.org/10.1080/1369118X.2017.1348535).

Steiner, K. (2015). Images of Muslims and Islam in Swedish Christian and secular news discourse. **Media War and Conflict**, 8(1), ss. 20–45. doi: [10.1177/1750635214531107](https://doi.org/10.1177/1750635214531107).

Syahputra, I. (2019). Expressions of hatred and the formation of spiral of anxiety on social media in Indonesia. **Search: Journal of the Southeast Asia Research Centre for Communications and Humanities**, 11(1), ss. 92–109.

Tontodimamma, A. et al. (2020). Thirty years of research into hate speech: Topics of interest and their evolution. **Scientometrics**. doi: [10.1007/s11192-020-03737-6](https://doi.org/10.1007/s11192-020-03737-6).

Waldman, A. E. (2019). Sexual harassment online: Shaming and silencing women in the digital age: By Tania G. Levey. **American Journal of Sociology**, 124(6), ss. 1946–1948.

Wright, S., Trott, V. & Jones, C. (2020). "The pussy ain't worth it, bro": Assessing the discourse and structure of MGTOW. **Information Communication & Society**, 23(6), ss. 908–925.  
doi: [10.1080/1369118X.2020.1751867](https://doi.org/10.1080/1369118X.2020.1751867).

### Näthat mot yrkesgrupper med viktiga funktioner i demokratin

Barlow, C. & Awan, I. (2016). "You need to be sorted out with a knife": The attempted online silencing of women and people of Muslim faith within academia. **Social Media + Society**, 2(4). doi: [10.1177/2056305116678896](https://doi.org/10.1177/2056305116678896).

Binns, A. (2017). Fair game? Journalists' experiences of online abuse. **Journal of Applied Journalism & Media Studies**, 6(2), ss. 183–206.  
doi: [10.1386/ajms.6.2.183\\_1](https://doi.org/10.1386/ajms.6.2.183_1).

Dorstkar, N. & Preisinger, A. (2017). "Cyber hate" vs. "cyber deliberation": The case of an Austrian newspaper's discussion board from a critical online-discourse analytical perspective. **Journal of Language and Politics**, 16(6), ss. 759–781. doi: [10.1075/jlp.15033.dor](https://doi.org/10.1075/jlp.15033.dor).

Gardiner, B. (2018) "It's a terrible way to go to work": What 70 million readers' comments on the Guardian revealed about hostility to women and minorities online. **Feminist Media Studies**, 18(4), ss. 592–608.  
doi: [10.1080/14680777.2018.1447334](https://doi.org/10.1080/14680777.2018.1447334).

Gorrell, G. et al. (2020). Which politicians receive abuse? Four factors illuminated in the UK general election 2019. **EPJ Data Science**, 9(1), s. 18.  
doi: [10.1140/epjds/s13688-020-00236-9](https://doi.org/10.1140/epjds/s13688-020-00236-9).

Hellmueller, L. et al. (2020). Shaping (non)-discursive social media spaces: Cross-national typologies of news organizations' heavy commenters. **New Media & Society**, ss. 1–19. doi: [10.1177/1461444820946454](https://doi.org/10.1177/1461444820946454).

Hess, K. & Waller, L. (2020). Moral compass: How a small-town newspaper used silence in a hyper-charged controversy. **Journalism**, 21(4), ss. 574–590.  
doi: [10.1177/1464884919886441](https://doi.org/10.1177/1464884919886441).

Koirala, S. (2020). Female journalists' experience of online harassment: A case study of Nepal. **Media and Communication**, 8(1), ss. 47–56.  
doi: [10.17645/mac.v8i1.2541](https://doi.org/10.17645/mac.v8i1.2541).

Marwick, A. E. (2020). Media studies and the pitfalls of publicity. **Television & New Media**, 21(6), ss. 608–615. doi: [10.1177/1527476420919702](https://doi.org/10.1177/1527476420919702).

Meza, R., Vincze, H. O. & Mogos, A. (2018). Targets of online hate speech in context: A comparative digital social science analysis of comments on public facebook pages from Romania and Hungary. **Intersections: East European Journal of Society and Politics**, 4(4), ss. 26–50. doi: [10.17356/ieejsp.v4i4.503](https://doi.org/10.17356/ieejsp.v4i4.503).

Scaramuzzino, G. (2020). Workplace violence: A threat to autonomy and professional discretion. **Sociologisk Forskning**, 57(3–4), ss. 249–270. doi: [10.37062/sf.57.22059](https://doi.org/10.37062/sf.57.22059).

Sobieraj, S. et al. (2020). Politicians, social media, and digital publics: Old rights, new terrain. **American Behavioral Scientist**, 64(11), ss. 1646–1669. doi: [10.1177/00027642200945357](https://doi.org/10.1177/00027642200945357).

Theocharis, Y. et al. (2016). A bad workman blames his tweets: The consequences of citizens' uncivil Twitter use when interacting with party candidates. **Journal of Communication**, 66(6), ss. 1007–1031. doi: [10.1111/jcom.12259](https://doi.org/10.1111/jcom.12259).

Tromble, R. & Koole, K. (2020). She belongs in the kitchen, not in Congress? Political engagement and sexism on Twitter. **Journal of Applied Journalism & Media Studies**, 9(2), ss. 191–214. doi: [10.1386/ajms\\_00022\\_1](https://doi.org/10.1386/ajms_00022_1).

Turner, G. (2018). The media and democracy in the digital era: Is this what we had in mind? **Media International Australia**, 168(1), ss. 3–14. doi: [10.1177/1329878X18782987](https://doi.org/10.1177/1329878X18782987).

Wagner, A. (2020). Tolerating the trolls? Gendered perceptions of online harassment of politicians in Canada. **Feminist Media Studies**. doi: [10.1080/14680777.2020.1749691](https://doi.org/10.1080/14680777.2020.1749691).

Waisbord, S. (2020). Mob censorship: Online harassment of US journalists in times of digital hate and populism. **Digital Journalism**, 8(8), ss. 1030–1046. doi: [10.1080/21670811.2020.1818111](https://doi.org/10.1080/21670811.2020.1818111).

Ziegele, M. et al. (2018). Journalistic counter-voices in comment sections: Patterns, determinants, and potential consequences of interactive moderation of uncivil user comments. **Studies in Communication and Media**, 7(4), ss. 525–554. doi: [10.5771/2192-4007-2018-4-525](https://doi.org/10.5771/2192-4007-2018-4-525).

## Juridik och yttrandefrihet

Bakalis, C. (2018). Rethinking cyberhate laws. **Information & Communications Technology Law**, 27(1), ss. 86–110. doi: [10.1080/13600834.2017.1393934](https://doi.org/10.1080/13600834.2017.1393934).

- Balkin, J. M. (2018). Free speech is a triangle. *Columbia Law Review*, 118(7), ss. 2011–2055.
- Bangstad, S. (2014). The weight of words: The freedom of expression debate in Norway. *Race & Class*, 55(4), ss. 8–28. doi: [10.1177/0306396813519928](https://doi.org/10.1177/0306396813519928).
- Bentham, T. van (2018). Social media actors in the fight against terrorism: Technology and its impact on human rights. *Cambridge International Law Journal*, 7(2), ss. 284–300. doi: [10.4337/cilj.2018.02.06](https://doi.org/10.4337/cilj.2018.02.06).
- Berkey, B. (2017). Business ethics and free speech on the Internet. *Philosophia*, 45(3), ss. 937–945. doi: [10.1007/s11406-016-9785-9](https://doi.org/10.1007/s11406-016-9785-9).
- Cavaliere, P. (2019). Digital platforms and the rise of global regulation of hate speech. *Cambridge International Law Journal*, 8(2), ss. 282–304. doi: [10.4337/cilj.2019.02.06](https://doi.org/10.4337/cilj.2019.02.06).
- Celik, S. (2018). Tertiary-level Internet users' opinions and perceptions of cyberhate. *Information Technology & People*, 31(3), ss. 845–868. doi: [10.1108/ITP-05-2017-0147](https://doi.org/10.1108/ITP-05-2017-0147).
- Cohen-Almagor, R. (2015). Internet architecture, freedom of expression and social responsibility: Critical realism and proposals for a better future. *Innovation: The European Journal of Social Science Research*, 28(2), ss. 147–166. doi: [10.1080/13511610.2015.1025042](https://doi.org/10.1080/13511610.2015.1025042).
- Cohen-Almagor, R. (2017). Balancing freedom of expression and social responsibility on the Internet. *Philosophia*, 45(3), ss. 973–985. doi: [10.1007/s11406-017-9856-6](https://doi.org/10.1007/s11406-017-9856-6).
- Dijck, J. van (2020). Governing digital societies: Private platforms, public values. *Computer Law & Security Review*, 36, s. 105377. doi: [10.1016/j.clsr.2019.105377](https://doi.org/10.1016/j.clsr.2019.105377).
- Edwards, A. et al. (2020). Forecasting the governance of harmful social media communications: Findings from the digital wildfire policy Delphi. *Policing & Society*. doi: [10.1080/10439463.2020.1839073](https://doi.org/10.1080/10439463.2020.1839073).
- Einwiller, S. A. & Kim, S. (2020). How online content providers moderate user-generated content to prevent harmful online communication: An analysis of policies and their implementation. *Policy and Internet*, 12(2), ss. 184–206. doi: [10.1002/poi3.239](https://doi.org/10.1002/poi3.239).

Enarsson, T. & Lindgren, S. (2019). Free speech or hate speech? A legal analysis of the discourse about Roma on Twitter. **Information & Communications Technology Law**, 28(1), ss. 1–18. doi: [10.1080/13600834.2018.1494415](https://doi.org/10.1080/13600834.2018.1494415).

Erdosova, A. (2019). The freedom of political speech in perspective of hate speech. **Vestnik of Saint Petersburg University-Law – Vestnik Sankt-Peterburgskogo Universiteta-Pravo**, 10(2), ss. 353–361. doi: [10.21638/spbu14.2019.210](https://doi.org/10.21638/spbu14.2019.210).

Guiora, A. & Park, E. A. (2017). Hate speech on social media. **Philosophia**, 45(3), ss. 957–971. doi: [10.1007/s11406-017-9858-4](https://doi.org/10.1007/s11406-017-9858-4).

Johnson, B. G. (2018). Tolerating and managing extreme speech on social media. **Internet Research**, 28(5), ss. 1275–1291. doi: [10.1108/IntR-03-2017-0100](https://doi.org/10.1108/IntR-03-2017-0100).

Jones, M. L. (2018). Silencing bad bots: Global, legal and political questions for mean machine communication. **Communication Law and Policy**, 23(2), ss. 159–195. doi: [10.1080/10811680.2018.1430418](https://doi.org/10.1080/10811680.2018.1430418).

Kasakowskij, T. et al. (2020). Network enforcement as denunciation endorsement? A critical study on legal enforcement in social media. **Telematics and Informatics**, 46, s. 101317. doi: [10.1016/j.tele.2019.101317](https://doi.org/10.1016/j.tele.2019.101317).

Liang, L. (2015). Censorship and the politics of micro-fascism. **Television & New Media**, 16(4), ss. 388–393. doi: [10.1177/1527476415573952](https://doi.org/10.1177/1527476415573952).

Limba, T. & Sidlauskas, A. (2018). Peculiarities of anonymous comments' management: A case study of Lithuanian news portals. **Entrepreneurship and Sustainability Issues**, 5(4), ss. 875–889. doi: [10.9770/jesi.2018.5.4\(12\)](https://doi.org/10.9770/jesi.2018.5.4(12)).

Lowe, D. (2020). Far-right extremism: Is it legitimate freedom of expression, hate crime, or terrorism? **Terrorism and Political Violence**. doi: [10.1080/09546553.2020.1789111](https://doi.org/10.1080/09546553.2020.1789111).

Lysaker, O. & Syse, H. (2016). The dignity in free speech: Civility norms in post-terror societies. **Nordic Journal of Human Rights**, 34(2), ss. 104–123. doi: [10.1080/18918131.2016.1212691](https://doi.org/10.1080/18918131.2016.1212691).

Mac Sithigh, D. (2020). The road to responsibilities: New attitudes towards Internet intermediaries. **Information & Communications Technology Law**, 29(1), ss. 1–21. doi: [10.1080/13600834.2020.1677369](https://doi.org/10.1080/13600834.2020.1677369).

Miklaszewska, J. (2019). Freedom of speech in modern political culture. **Avant**, 10(1), ss. 77–88. doi: [10.26913/avant.2019.01.06](https://doi.org/10.26913/avant.2019.01.06).

Milosevic, T. (2016). Social media companies' cyberbullying policies. **International Journal of Communication**, 10, ss. 5164–5185.

Penney, J. W. (2017). Internet surveillance, regulation, and chilling effects online: A comparative case study. **Internet Policy Review**, 6(2), ss. 1–39.  
doi: [10.14763/2017.2.692](https://doi.org/10.14763/2017.2.692).

Poyhtari, R. (2014). Limits of hate speech and freedom of speech on moderated news websites in Finland, Sweden, the Netherlands and the UK. **Annales: Anali Za Istrske in Mediteranske Studije-Series Historia Et Sociologia**, 24(3), ss. 513–524.

Rogers, R. (2020). Deplatforming: Following extreme Internet celebrities to Telegram and alternative social media. **European Journal of Communication**, 35(3), ss. 213–229. doi: [10.1177/0267323120922066](https://doi.org/10.1177/0267323120922066).

Tenove, C. (2020). Protecting democracy from disinformation: Normative threats and policy responses. **International Journal of Press-Politics**, 25(3), ss. 517–537. doi: [10.1177/1940161220918740](https://doi.org/10.1177/1940161220918740).

Wagner, B. (2014). The politics of Internet filtering: The United Kingdom and Germany in a comparative perspective. **Politics**, 34(1), ss. 58–71.  
doi: [10.1111/1467-9256.12031](https://doi.org/10.1111/1467-9256.12031).

### **Tekniska lösningar för att motverka näthat och desinformation**

Artime, O. et al. (2020). Effectiveness of dismantling strategies on moderated vs. unmoderated online social platforms. **Scientific Reports**, 10(1), s. 14392.  
doi: [10.1038/s41598-020-71231-3](https://doi.org/10.1038/s41598-020-71231-3).

Dias Oliva, T., Antonialli, D. M. & Gomes, A. (2020). Fighting hate speech, silencing drag queens? Artificial intelligence in content moderation and risks to LGBTQ voices online. **Sexuality & Culture: An Interdisciplinary Journal**.  
doi: [10.1007/s12119-020-09790-w](https://doi.org/10.1007/s12119-020-09790-w).

Lundberg, J. & Laitinen, M. (2020). Twitter trolls: A linguistic profile of anti-democratic discourse. **Language Sciences**, 79, s. 101268.  
doi: [10.1016/j.langsci.2020.101268](https://doi.org/10.1016/j.langsci.2020.101268).

Monakhov, S. (2020). Early detection of internet trolls: Introducing an algorithm based on word pairs / single words multiple repetition ratio. **Plos One**, 15(8), s. e0236832. doi: [10.1371/journal.pone.0236832](https://doi.org/10.1371/journal.pone.0236832).

Muneer, A. & Fati, S. M. (2020). A comparative analysis of machine learning techniques for cyberbullying detection on Twitter. **Future Internet**, 12(11), s. 187. doi: [10.3390/fi12110187](https://doi.org/10.3390/fi12110187).



- Oliveira, N. R. de, Medeiros, D. S. V. & Mattos, D. M. F. (2020). A sensitive stylistic approach to identify fake news on social networking. **IEEE Signal Processing Letters**, 27, ss. 1250–1254. doi: [10.1109/LSP.2020.3008087](https://doi.org/10.1109/LSP.2020.3008087).
- Pamungkas, E. W., Basile, V. & Patti, V. (2020). Misogyny detection in Twitter: A multilingual and cross-domain study. **Information Processing & Management**, 57(6), s. 102360. doi: [10.1016/j.ipm.2020.102360](https://doi.org/10.1016/j.ipm.2020.102360).
- Pennekamp, J. et al. (2019). Hi Doppelgänger: Towards detecting manipulation in news comments. I **WWW 19: Companion Proceedings of the World Wide Web Conference**, May 2019, ss. 197–205. doi: [10.1145/3308560.3316496](https://doi.org/10.1145/3308560.3316496).
- Rodriguez-Sanchez, F., Carrillo-de-Albornoz, J. & Plaza, L. (2020). Automatic classification of sexism in social networks: An empirical study on Twitter data. **IEEE Access**, 8, ss. 219563–219576. doi: [10.1109/ACCESS.2020.3042604](https://doi.org/10.1109/ACCESS.2020.3042604).
- Singh, V. K., Ghosh, I. & Sonagara, D. (2021). Detecting fake news stories via multimodal analysis. **Journal of the Association for Information Science and Technology**, 72(1), ss. 3–17. doi: [10.1002/asi.24359](https://doi.org/10.1002/asi.24359).
- Ullmann, S. & Tomalin, M. (2020). Quarantining online hate speech: Technical and ethical perspectives. **Ethics and Information Technology**, 22(1), ss. 69–80. doi: [10.1007/s10676-019-09516-z](https://doi.org/10.1007/s10676-019-09516-z).
- Van Royen, K. et al. (2017). “Thinking before posting?”: Reducing cyber harassment on social networking sites through a reflective message. **Computers in Human Behavior**, 66, ss. 345–352. doi: [10.1016/j.chb.2016.09.040](https://doi.org/10.1016/j.chb.2016.09.040).

## Nationell kunskapsöversikt

2018 års tryck- och yttrandefrihetskommitté (2020). **Ett ändamålsenligt skydd för tryck- och yttrandefriheten** (SOU 2020:45). Stockholm: Justitiedepartementet.

Bladini, M. (2017). **Hat och hot på nätet: En kartläggning av den rättsliga regleringen i Norden från ett jämställdhetsperspektiv**. Göteborg: Nordisk information för kunskap om kön, på uppdrag av Nordiska ministerrådet.

Brottsförebyggande rådet (Brå) (2012). **Politikernas trygghetsundersökning 2012: Förtroendevaldas utsatthet och oro för hot, våld och trakasserier**. Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2014). **Politikernas trygghetsundersökning 2013: Utsatthet och oro för trakasserier, hot och våld**. Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2015a). **Hot och våld: Om utsatthet i yrkesgrupper som är viktiga i det demokratiska samhället.** Rapport 2015:12. Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2015b). **Politikernas trygghetsundersökning 2015: Förtroendevaldas utsatthet under valåret 2014.** Rapport 2015:23. Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2016). **It-inslag i brottsligheten och rättsväsendets förmåga att hantera dem.** Rapport 2016:17. Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2017). **Politikernas trygghetsundersökning 2017: Förtroendevaldas utsatthet och oro för trakasserier, hot och våld 2016.** Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2019a). **Hatbrott 2018: Statistik över polisanmälda brott med identifierade hatbrottsmotiv.** Rapport 2019:13. Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2019b). **Hot och kränkningar på internet: En jämförande beskrivning utifrån kön och ålder: Kortanalys.** Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2019c). **Politikernas trygghetsundersökning 2019: Förtroendevaldas utsatthet och oro för trakasserier, hot och våld under valåret 2018.** Rapport 2019:15. Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2019d). **Nationella trygghetsundersökningen 2019: Om utsatthet, otrygghet och förtroende.** Rapport 2019:11. Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2019e). **Kriminalstatistik 2019. Personer lagförda för brott.** Slutlig statistik. Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2020a). **Nationella trygghetsundersökningen 2020: Om utsatthet, otrygghet och förtroende.** Rapport 2020:8. Stockholm: Brå.

Brottsförebyggande rådet (Brå) (2020b). **Skolundersökning om brott 2019.** Rapport 2020:11. Stockholm: Brå.

Carlsson, U. (red) (2018). **Medie- och informationskunnighet i den digitala tidsåldern: En demokratifråga: Kartläggning, analys, reflektioner.** Göteborg: Nordicom.

Childhood (2018). Varannan tjej under femton år sexkontaktad på nätet. <https://childhood.se/nyheter/varannan-tjej-femton-ar-sexkontaktad-pa-natet/>

Civil Rights Defenders (2019) **När samhället tystnar: Eb rapport om hot och hat mot oberoende opinionsbildare i det svenska civilsamhället.** Stockholm: Civil Rights Defenders.

Colliver, C. et al. (2018). **Smearing Sweden: International Influence Campaigns in the 2018 Swedish Election.** London: Institute of Strategic Dialogue (ISD), London School of Economics.

Delegationen mot segregation (DELMOS) (2020). **En demokrati för alla? Deltagandet i demokratin bland boende i områden med socioekonomiska utmaningar jämfört med andra.** Flemingsberg: DELMOS.

Dunkels, E. (2016). **Nätmobbing, näthat och nätkärlek: Kunskap och strategier för en bättre vardag på nätet.** Stockholm: Gothia Fortbildning.

Enarsson, T. (2017). Näthat som rättslig utmaning i nutid och framtid. I Edström, Ö. & Mannelqvist, R. (red.). **Jubileumsskrift till Juridiska institutionen 40 år.** Umeå: Umeå universitet, ss. 103–118.

Fernquist, J. et al. (2018). **Automatiserade konton: En studie av botar på Twitter i samband med det svenska riksdagsvalet 2018.** Karlstad: Myndigheten för samhällsskydd och beredskap.

Fernquist, J. et al. (2019). **Ansiktet bakom hatet: En analys av en grupp Flashback-användare som uttrycker hat mot journalister.** Stockholm: TU-Medier i Sverige.

Fernquist, J. et al. (2020). **Det digitala hatets karaktär: En studie av hat mot kvinnor och män i utsatta yrkesgrupper.** FOI Memo 7429. Stockholm: Totalförsvarets forskningsinstitut (FOI).

Folkhälsomyndigheten (2018). **Skolbarns hälsovanor i Sverige 2017/18: Grundrapport.** Solna: Folkhälsomyndigheten.

Fridh, M. (2018). **Bullying, violence and mental distress among young people: Cross-sectional population-based studies in Scania, Sweden.** Lund: Lunds universitet.

Friends (2017) **Friends nätrapport 2017.** Stockholm: Friends.

Giandomenico, J. & Linderstål, H. (2020). Virtuellt lobbyism: Att påverka makthavare i den digitala eran. I **Det demokratiska samtalet i en digital tid: En antologi om desinformation, propaganda och näthat.** Stockholm: Kommittén Nationell satsning på medie- och informationskunnighet och det demokratiska samtalet.

Hedman, F. et al. (2018). **News and Political Information Consumption in Sweden: Mapping the 2018 Swedish General Election on Twitter**. Data Memo 2018.3. Oxford: Project on Computational Propaganda.

Hoff, E. V., Daukantaitė, D. & Birgerstam, P. (2020). Self-evaluation differences among Swedish children and adolescents over a 30-year period. **Frontiers in Psychology**, 11, artikel 802.

Internetstiftelsen (2019). **Svenskarna och internet 2019: En årlig studie av svenska folkets internetvanor**. Stockholm: Internetstiftelsen.

Justitiedepartementet (2016). **Ett starkt straffrättsligt skydd för den personliga integriteten** (Regeringens proposition 2016/17:222). Stockholm: Regeringskansliet.

Justitiedepartementet (2018). **Brott mot förtroendevalda** (Regeringens proposition 2018/19:154). Stockholm: Regeringskansliet.

Justitiedepartementet (2020). **Ett förstärkt straffrättsligt skydd för vissa samhällsnyttiga funktioner och några andra straffrättsliga frågor** (Dir. 2020:54). Stockholm: Regeringskansliet.

Kaati, L. & Olsson, F. (2018). **Hot och hat mot journalister**. Stockholm: TU-Medier i Sverige.

Kaminski, M. de (2020). Strategier för ett samhälle som inte tystnar: erfarenheter från det civila samhället. I **Det demokratiska samtalet i en digital tid: En antologi om desinformation, propaganda och näthat**. Stockholm: Kommittén Nationell satsning på medie- och informationskunnighet och det demokratiska samtalet.

Kulturdepartementet (2017). **Till det fria ordets försvar: Åtgärder mot utsatthet för hot och hat bland journalister, förtroendevalda och konstnärer**. Ku 17:04. Stockholm: Regeringskansliet.

Landsrådet för Sveriges ungdomsorganisationer (LSU) (2020). **Vi sluter oss inåt: En kartläggning av hat och hot mot Sveriges ungdomsrörelse**. Stockholm: LSU.

Länsförsäkringar (2018). **STUDIE: Unga känner sig misslyckade av sociala medier**. <https://www.lansforsakringar.se/goteborg-och-bohuslan/om-oss/press-media/nyheter/156897/> (hämtad 2021-03-08).

Löfgren Nilsson, M. (2013). **J-panelen: Hotet och hatet mot journalister**. Göteborg: Institutionen för journalistik, medier och kommunikation (JMG), Göteborgs universitet.

Löfgren Nilsson, M. (2017) **Journalisternas trygghetsundersökning: Journalisters utsatthet under 2016**. Arbetsrapport nr 75. Göteborg: Institutionen för journalistik, medier och kommunikation (JMG), Göteborgs universitet.

Medieutredningen (2016). **En gränsöverskridande mediepolitik: För upplysning, engagemang och ansvar** (SOU 2016:80). Stockholm: Kulturdepartementet.

Myndigheten för kulturanalys (2016). **Hotad kultur? En undersökning om hot, trakasserier och våld mot konstnärer och författare i Sverige**. Rapport 2016:3. Göteborg: Myndigheten för kulturanalys.

Myndigheten för samhällsskydd och beredskap (MSB) (2019). **Att möta informationspåverkan: Handbok för kommunikatörer**. Karlstad: MSB.

Myndigheten för samhällsskydd och beredskap (MSB) (2016). **Mediebranschen 2016: Hot, risker och sårbarheter**. Karlstad: MSB.

Myndigheten för ungdoms- och civilsamhällsfrågor (MUCF) (2019). **Unga med attityd 2019: Del 1, Samhällsfrågor och demokrati**. Växjö: MUCF.

Nationell satsning på medie- och informationskunnighet och det demokratiska samtalet (2020). **Det demokratiska samtalet i en digital tid: Så stärker vi motståndskraften mot desinformation, propaganda och näthat** (SOU 2020:56). Stockholm: Kulturdepartementet.

Näthatsgranskaren (2019). **Polismyndighetens arbete med demokrati- och hatbrott: En skuggrapport**. <http://media1.nathatsgranskaren.se/Skuggrapport-2019-09-27-web.pdf>

Polismyndigheten (2019). **Åtgärder för skydd av den fria åsiktsbildningen: Redovisning av regeringsuppdrag**. Stockholm: Polismyndigheten.

Sackemark, A., Schultz, M. & Alderhorn, M. (2015). **Näthat: Rättigheter och möjligheter**. Stockholm: Karnov Group.

Schultz, M. (2020). Demokratiskadliga nätkränkningar: Dags att se över regelverket? I **Det demokratiska samtalet i en digital tid: En antologi om desinformation, propaganda och näthat**. Stockholm: Kommittén Nationell satsning på medie- och informationskunnighet och det demokratiska samtalet.

Statens medieråd (2019). **Ungar & medier 2019**. Stockholm: Statens medieråd.

Statens medieråd (2020). **Utsatt på internet: En internationell forskningsöversikt om nätmobbning bland barn och unga**. Stockholm: Statens medieråd.

Sundell, A. (2018). Fångad i nätet?. I Barrling, K. & Holmberg, S. (red). **Demokratins framtid**. Stockholm: Sveriges riksdag.

Svensson, M. & Dahlstrand, K. (2014). **Nätkränkningar: En studie av svenska ungdomars normer och beteenden**. Växjö: MUCF.

Sveriges Kommuner och Landsting (SKL) (2018a). **Digitalt hat: Hat och hot mot kommunalpolitiker**. Stockholm: SKL.

Sveriges Kommuner och Landsting (SKL) (2018b). **Hot och hat påverkar: Intervjuer med utsatta förtroendevalda**. Stockholm: SKL.

Sveriges Kvinnolobby (2021). **Kvinnor i Sverige 2021. En granskning av hur Sverige lever upp till Kvinnokonventionen**. Stockholm: Sveriges Kvinnolobby.

Tipple, F. (2019). **Det demokratiska samtalet**. SOM-rapport nr 2019:23. Göteborg: SOM-institutet, Göteborgs universitet.

Totalförsvarets forskningsinstitut (FOI) (2020). **Kvinnohat och våldshyllningar i digitala incelmiljöer**. <https://www.foi.se/nyheter-och-press/nyheter/2020-03-04-kvinnohat-och-valdshyllningar-i-digitala-incelmiljoer.html>

TU-Medier i Sverige (2016). **Hot, våld och trakasserier mot medier**. Stockholm: TU-Medier i Sverige.

Utgivarna (2016). **Hot och hat: Så påverkas medierna**. Rapport 2016:1. Stockholm: Utgivarna.

Utredningen om ett modernt och starkt straffrättsligt skydd för den personliga integriteten (2016) **Integritet och straffskydd (SOU 2016:7)**. Stockholm: Justitiedepartementet.

Östensson, I. (2020). Internet är på riktigt. I **Det demokratiska samtalet i en digital tid: En antologi om desinformation, propaganda och näthat**. Stockholm: Kommittén Nationell satsning på medie- och informationskunnighet och det demokratiska samtalet.

## Näthat och demokratiskt deltagande – en kunskapsöversikt

Denna kunskapsöversikt är ett resultat av ett samarbete mellan Brottsoffermyndigheten, Lunds universitet och Högskolan i Halmstad. Den belyser olika frågor om näthat och demokratiskt deltagande, såsom vilka områden som genererar hot och hat, vilka samhälls- och yrkesgrupper som är särskilt utsatta och vilka konsekvenser näthat leder till.

Genom att uppmärksamma frågan om hot och hat på internet vill Brottsoffermyndigheten belysa vilka konsekvenser det kan få för vår demokrati. Myndigheten vill också synliggöra de personer som utsätts för näthat och bidra till en levande samhällsdebatt.

### **Brottsoffermyndigheten**

Box 470, 901 09 Umeå  
Telefon: 090-70 82 00  
[registrator@brottsoffermyndigheten.se](mailto:registrator@brottsoffermyndigheten.se)  
[www.brottsoffermyndigheten.se](http://www.brottsoffermyndigheten.se)



**Brottsoffermyndigheten**