



UPPSALA  
UNIVERSITET

Juridiska institutionen  
Vårterminen 2019

Examensarbete i rättsinformatik  
30 högskolepoäng

# Twitter, Tumblr, troll och porr

Om sociala mediers ansvar för yttrande- och informationsfrihet

## *Twitter, Tumblr, trolls and porn*

*On social media and the responsibility for freedom of speech and  
freedom of information*

Författare: Fredrik Emanuel Thorslund

Handledare: Professor Anna-Sara Lind





*Sätt, o HERRE, en vakt för min mun, bevaka mina läppars dörr.*

Psaltaren 141:3



## Abstract

There has, for the past few decades, been an increasing tendency from states to hold social media and other information intermediaries liable for the illicit or undesirable speech of their users. Much less debated, however, has been the question of intermediary responsibility for the freedom of speech and information of internet users. In the EU, as well as the US, there are no clear legal safeguards for protecting freedom of speech on privately owned networking platforms – albeit the exponential growth of these platforms as spheres for public expression and discourse. This thesis argues that the content moderation practices and policies of some of the largest social media platforms, in several respects, correspond poorly to Swedish and European freedom of speech standards, as established in human rights law. Considering the rationale behind speech rights, as well as the increasing significance of social media for public discourse, the thesis suggests alternative legal measures to establish corporate responsibility and ensuring online freedom of speech.

# Innehåll

<b>Abstract</b> .....	<b>I</b>
<b>Innehåll</b> .....	<b>II</b>
<b>Förkortningslista</b> .....	<b>V</b>
<b>1 Prolog</b> .....	<b>1</b>
<b>2 Om uppsatsen och arbetet</b> .....	<b>3</b>
2.1 Ämne och syfte .....	3
2.2 Avgränsningar .....	3
2.3 Metod, material och disposition .....	5
2.3.1 Allmänt om metod .....	5
2.3.2 Allmänt om material .....	6
2.3.3 Om uppsatsens disposition .....	7
<b>3 Yttrande- och informationsfriheten – en idéhistorisk kontext</b> .....	<b>9</b>
3.1 När ordet blev fritt .....	9
3.1.1 England och tryckfrihetens gryning .....	10
3.1.2 Den svenska tryckfriheten .....	11
3.1.3 Frihet, upplysning, revolution .....	12
3.1.4 Världskriget och totalitarismen .....	13
3.1.5 Rättigheternas renässans .....	15
3.1.6 Yttrandefriheten, här och nu .....	16
3.2 Det fria ordets syfte.....	16
3.2.1 Yttrandefrihet och sanning .....	17
3.2.2 Yttrandefrihet, demokrati och insyn .....	18
3.2.3 Yttrandefrihet och tolerans .....	19
3.3 Nu som då för alltid? .....	20
<b>4 Grindvakter och makthavare – informationslandskap i förändring</b> .....	<b>23</b>
4.1 Grindvakter och mellanhänder.....	24
4.2 Grindvakter och lagen .....	25
4.3 Nya grindvakter i informationssamhället.....	29
4.4 Nya medier, nya remedier? .....	31
<b>5 Villkor och vedergällning – om sociala mediers moderering</b> .....	<b>33</b>
5.1 Sociala medier .....	34

5.2	<i>Användarvillkor, riktlinjer och policy</i> .....	35
5.2.1	Hot, trakasserier, och kränkningar.....	36
5.2.2	”Hate-speech” och hets mot folkgrupp.....	36
5.2.3	Uppmuntran, glorifiering eller skildring av våld och annan brottslighet.....	38
5.2.4	Pornografi och sexuellt innehåll.....	39
5.2.5	Övriga innehållsbegränsningar.....	40
5.3	<i>Personbegränsningar</i> .....	40
5.4	<i>Sorteringsalgoritmer</i> .....	41
5.5	<i>Regelefterlevnad och sanktioner</i> .....	43
<b>6</b>	<b>Kongruens och kompetens – problemet med modereringen</b> .....	<b>46</b>
6.1	<i>Om yttrandefrihetens ratione materiae</i> .....	47
6.2	<i>Kongruensproblemet</i> .....	49
6.2.1	Tillgänglighet och precision.....	49
6.2.2	Subjektiva begränsningar.....	50
6.2.3	Innehållsbegränsningar.....	51
6.3	<i>Kompetensproblemet</i> .....	55
6.3.1	Intressen och incitament.....	57
6.3.2	Rättssäkerhet och rättvisa.....	57
6.3.3	Transparens och insyn.....	58
6.4	<i>Staten och kapitalet: det utkontrakterade ansvaret</i> .....	59
<b>7</b>	<b>Ansvar och skyldigheter – nya medier, nya remedier</b> .....	<b>62</b>
7.1	<i>Om yttrande- och informationsfrihetens ratione personae</i> .....	63
7.2	<i>Ansvar vid innehållsmoderering</i> .....	65
7.3	<i>En hållbar lösningsmodell</i> .....	67
7.3.1	Indirekt direkt ansvar.....	68
7.3.2	Preciserat ansvar.....	70
7.3.3	Effektiva rättsmedel.....	71
7.4	<i>Sammanfattning</i> .....	73
<b>8</b>	<b>Epilog</b> .....	<b>74</b>
<b>9</b>	<b>Källförteckning</b> .....	<b>76</b>
9.1	<i>Svenskt offentligt tryck</i> .....	76
9.2	<i>Utländskt offentligt tryck</i> .....	76
9.2.1	Förenta nationerna.....	76
9.2.2	Europeiska unionen.....	76
9.2.3	Förenta staterna.....	77

9.3	<i>Litteratur och tidskrifter</i> .....	78
9.4	<i>Domar från svenska och utländska domstolar</i> .....	84
9.4.1	Svenska domstolar .....	84
9.4.2	EU-domstolen.....	84
9.4.3	Europadomstolen .....	84
9.4.4	Amerikanska domstolar .....	86
9.5	<i>Internetkällor</i> .....	87
9.5.1	Facebook .....	87
9.5.2	Twitter.....	87
9.5.3	Tumblr.....	88
9.5.4	YouTube.....	88
9.5.5	Instagram.....	90
9.5.6	Övrigt .....	90



## Förkortningslista

CSR	Corporate Social Responsibility
Dataskyddsförordningen	Europaparlamentets och rådets förordning (EU) 2016/679 av den 27 april 2016 om skydd för fysiska personer med avseende på behandling av personuppgifter och om det fria flödet av personuppgifter och om upphävande av direktiv 95/46/EG (allmän dataskyddsförordning) (Text av betydelse för EES)
Dataskyddslagen	Lag (2018:218) med kompletterande bestämmelser till EU:s dataskyddsförordning.
EDRi	European Digital Rights
EFF	Electronic Frontier Foundation
E-handelsdirektivet	Europaparlamentets och rådets direktiv 2000/31/EG av den 8 juni 2000 om vissa rättsliga aspekter på informationssamhällets tjänster, särskilt ekonomisk handel, på den inre marknaden ("Direktiv om ekonomisk handel")
E-handelslagen	Lag (2002:562) om elektronisk handel och andra informationssamhällets tjänster
EKMR	Europeiska konventionen om skydd för de mänskliga rättigheterna och de grundläggande friheterna (Europakonventionen)

EU-stadgan	Europeiska unionens stadga om de grundläggande rättigheterna
FCC	Federal Communications Commission
HD	Högsta domstolen
HFD	Högsta förvaltningsdomstolen
OHCHR	Office of the United Nations High Commissioner for Human Rights
RF	Kungörelse (1974:152) om beslutad ny regeringsform (Regeringsformen)
FOSTA-SESTA	Allow States and Victims to Fight Online Sex Trafficking Act of 2017
TF	Tryckfrihetsförordning (1949:105)
TSM-förordningen	Europaparlamentets och rådets förordning (EU) 2015/2120 av den 25 november 2015 om åtgärder rörande en öppen internetanslutning och o ändring av direktiv 2002/22/EG om samhällsomfattande tjänster och användares rättigheter avseende elektroniska kommunikationsnät och kommunikationstjänster och förordning (EU) nr 531/2012 om roaming i allmänna mobilnät i unionen (Text av betydelse för EES)
YGL	Yttrandefrihetsgrundlag (1991:1469)

# 1 Prolog

*If liberty means anything at all, it means the right  
to tell people what they do not want to hear.*

George Orwell

I George Orwells kanoniska *1984* (1949) arbetar Winston Smith på Sanningsministeriet, en myndighet under ”Storebror” med uppgift att förvanska språket och informationen så att de bättre passar den totalitära statens syften.<sup>1</sup> Karin Boyes *Kallocain* (1940) skildrar kemisten Leo Kall, uppfinnare av ett sanningsserum som ska tvinga Världsstatens medborgare att ge till känna sina hemliga, statsfientliga tankar.<sup>2</sup> I Aldous Huxleys *Du sköna nya värld* (1932) distribueras istället drogen soma för att beröva de kuvade medborgarna deras tankefrihet, och i Jevgenij Zamjatin *Vi* (1924) blir Den Enda Statens opponenter gripna av Beskyddarbyrå och avrättade på Kubtorget.<sup>3</sup>

I 1900-talets stora dystopiska romaner ställs det fria ordet mot sin självklara nemeses – den totalitära staten. De är ekon från sin samtid, skrivna med färsk minnen från Stalins terrorregim, Hitlers skräckvälde och Mussolinis fascistdiktatur. De är tidsdokument från ett Europa där frispråkighet var livsfarligt.

På många håll i världen, även i Europa, är det fria ordet fortfarande hårt ansatt från maktgiriga stater. Men den samtid som skildras av Orwell, Boye, Huxley och Zamjatin framstår närmast som överspelad. Efter årtionden av fred, internationalisering, rättighetsdiskurs och överstatligt samarbete lever vi i ett samhälle där yttrandefrihet och informationsfrihet är konstitutionella grundbultar.

Men har vi lärt oss någonting av vår historia borde det vara att våra rättigheter aldrig kan tas för givna, och att hoten mot den fria åsiktsbildningen kan alstras lika plötsligt som oväntat. Likaså kan vi anta att hoten inte alltid kommer att uppstå i samma skepnad som förr – eller, för den delen, att de inte alltid kommer kunna tillskrivas ”Storebror” eller ”Världsstaten”.

Denna uppsats behandlar de nya utmaningar som yttrande- och informationsfriheten står inför i ett digitaliserat samhälle där privata aktörer fått en allt starkare ställning som ”grundvakter” i informationslandskapet.

---

<sup>1</sup> Orwell, *1984*.

<sup>2</sup> Boye, *Kallocain*.

<sup>3</sup> Huxley, *Du sköna nya värld* och Zamjatin, *Vi*.

Dessa aktörer har revolutionerat kommunikationstekniken och möjliggjort gränslös masskommunikation för varje individ med en internetuppkoppling och en dator. De har etablerat fora för de som tidigare saknade möjlighet att göra sig hörda och demokratiserat kommunikationsformerna på sätt som saknar motstycke i vår moderna historia. De har, kort sagt, åstadkommit ett av de största genombrotten för yttrande- och informationsfriheten på århundraden.

Samtidigt kan vi se oroande tendenser i hur dessa företag hanterar vår information och kommunikation. Från att tidigare ha agerat mer som passiva möjliggörare av mellanmänsklig kommunikation har flera av företagen, i allt högre utsträckning, anammat en mer aktiv hållning som kuratorer i informationslandskapet. I en intervju från 2019 konstaterar Twitters verkställande direktör, Jack Dorsey:

I don't believe that we can afford to take a neutral stance anymore. I don't believe that we should optimize for neutrality.<sup>4</sup>

Dorsey talar om neutralitet – neutralitet i förhållande till vad som får sägas, spridas och visas på Twitter. Neutralitet som en typ av innehållsmässig agnosticism; en laissez faire-ideologi till förmån för yttrande- och informationsfriheten. Denna neutralitet har nu dödförklarats. Den har ersatts av en ny typ av interventionism, där plattformar som Twitter, Facebook och Instagram istället fått axla rollen som yttrandefrihetens skiljedomare. De faktiska avvägningarna kring det fria ordet har hamnat i Facebooks och Twitters arbetsordningar, vare sig de vill eller inte.

Även om staten historiskt sett varit det fria ordets nemesis, är det också staten som haft ansvaret för att garantera medborgarna deras yttrande- och informationsfrihet och göra de nödvändiga avvägningarna mot konkurrerande intressen. De juridiska ramverken kring yttrande- och informationsfrihet som samhället lutar sig mot är utformade kring stater som den relevanta, ansvariga aktören. Så, om makten över det fria ordet nu har förskjutits från staten till privata aktörer, kan detsamma sägas om ansvaret? Eller, har det uppstått en form av juridiskt ansvarsvakuum på de fora där yttrande- och informationsfriheten idag får sin praktiska betydelse?

---

<sup>4</sup> Dorsey & Harris, *Making Sense with Sam Harris*, #148 – Jack Dorsey, Making Sense Podcast 5/2 2019.

## 2 Om uppsatsen och arbetet

### 2.1 Ämne och syfte

Denna uppsats behandlar sociala mediers ansvar för yttrande- och informationsfriheten. Eftersom de sociala medierna i högre utsträckning har fått kontrollen över informationen och den mellanmänniska kommunikationen, är det angeläget att undersöka vilka juridiska ramverk som aktualiseras på området. Syftet med uppsatsen är tudelat. Dels vill jag reda ut och klargöra hur de sociala mediernas informationshantering kan utgöra ett yttrande- och informationsfrihetsproblem. Dels vill jag diskutera om de gällande ansvarskonstruktionerna kring yttrande- och informationsfrihet är tillfredställande utformade för att bemöta problemet, och, annars, vilka nya lösningar som hade kunnat skapa ett mer tillfredställande resultat utifrån ett yttrande- och informationsfrihetsperspektiv.

### 2.2 Avgränsningar

Några inledande avgränsningar av ämnet kan etableras genom att precisera ämnesbeskrivningen. Uppsatsen tar sikte på det ansvar som åligger eller bör åligger sociala medier såsom *enskilda aktörer*. Åtgärder som vidtas av statliga eller offentlig-rättsligt reglerade aktörer ligger utanför ämnet. Brottsbekämpande myndigheters insatser mot brott som begås på medieplattformar, exempelvis väckande av åtal på grund av förtal på en blogg, faller således utanför uppsatsens ämnesområde, även om sådana insatser givetvis kan inkräkta på yttrande- och informationsfriheten.

Gränsdragningen kan förstås nyanseras: inte sällan vidtar enskilda aktörer på eget bevåg åtgärder som direkt eller indirekt kan härledas till offentlig-rättslig reglering eller myndigheters ingripanden. Som exempel kan en social medieplattform välja att blockera visst innehåll som skildrar barnpornografi, i syfte att undgå straffrättsligt ansvar eller skadeståndsansvar.<sup>5</sup> Eftersom åtgärden i sig vidtas av den enskilde aktören kommer emellertid sådana åtgärder att omfattas av framställningen.

Uppsatsen behandlar *sociala medier*, men fokus ligger framförallt på de så kallade sociala nätverkstjänsterna – särskilt de ledande aktörerna på marknaden, som Facebook, Twitter, Youtube, Tumblr och Instagram.

---

<sup>5</sup> Jfr ansvarskonstruktionen i 5 § och 7 § lagen (1998:112) om ansvar för elektroniska anslagstavlor.

Uppsatsen tar vidare sikte på de sociala mediernas ansvar för *yttrande- och informationsfriheten*. Dessa friheter – rätten att yttra sig och rätten att ta del av yttranden – är oupplösligen sammanlänkade och behandlas därför som en enhet. Det finns ingen enda auktoritativ definition av dessa friheter; de formuleras på olika sätt och har olika stor omfattning beroende på vilken rättskälla som konsulteras. Även om yttrande- och informationsfriheten är universella som begrepp och koncept, är de i juridisk bemärkelse avhängiga den reglering som är tillämplig inom en given jurisdiktion. De yttrande- och informationsfrihetsnormer som avhandlas inom ramen för denna uppsats är därför avgränsade till sådana som är direkt tillämpliga och relevanta ur ett svenskt perspektiv. Därmed aktualiseras regleringen i regeringsformens (RF) andra kapitel, tryckfrihetsförordningen (TF) och yttrandefrihetsgrundlagen (YGL), Europakonventionen om mänskliga rättigheter (EKMR) och Europeiska unionens stadga om de grundläggande rättigheterna (EU-stadgan). Att föra argumentationen utifrån alla dessa rättskällor innebär en mer heltäckande och rättvisande bild av gällande rätt. Samtidigt gör det framställningen något mer otymplig. Därför har ytterligare regelverk lämnats därhän. FN:s allmänna förklaring om de mänskliga rättigheterna får anses principiellt betydelsefull i sammanhanget – men eftersom dess rättsliga betydelse framstår som begränsad, kommer den inte att behandlas närmare i uppsatsen.<sup>6</sup>

I utvalda delar frångår jag dock det svenska perspektivet. Som utförligare kommenteras nedan, innefattar delar av uppsatsen jämförelser med amerikansk rätt. Jämförelsen framstår som motiverad, inte minst på grund av den relativa regleringslikhet som råder mellan Sverige, EU och USA i fråga om mellanhandsansvar i digitala miljöer. Eftersom de komparativa elementen är av underordnad betydelse i arbetet, finns det dock skäl att avgränsa förekommande jämförelser till amerikansk rätt.

Som framgår av syftet kommer en del av diskussionen att ägnas åt att utvärdera regleringslösningar ur ett yttrande- och informationsfrihetsperspektiv. Med detta inte sagt att framställningen är avgränsad till ett visst intressentperspektiv. Intresset för yttrande- och informationsfrihet kommer istället att vägas mot motstående intressen då normativa argument framförs – bland annat de sociala mediernas rätt till näringsfrihet och det allmännas intressen i relevanta avseenden. Därmed har jag som mål att presentera ett arbete som inte är präglad av särskild partshänsyn.

---

<sup>6</sup> Den allmänna förklaringen utgör inte en konvention, och anses heller inte – till skillnad från EKMR eller EU-stadgan – utgöra ett juridiskt bindande instrument. Se Shaw, *International Law*, s. 203–204.

Slutligen, i fråga om alternativa regleringslösningar, har jag valt att inte lägga fokus på frågor kring jurisdiktion, regleringskompetens och praktisk regelefterlevnad. Fokus ligger istället på principfrågorna kring yttrande- och informationsfrihet och mellanhandsansvar.

## 2.3 Metod, material och disposition

### 2.3.1 Allmänt om metod

Den metod som ligger till grund för arbetet får primärt betraktas som rättsdogmatisk. Arbetet präglas av en ”rekonstruktion” av rätten; en beskrivning, systematisering och analys av gällande rätt (yttrande- och informationsfriheten) utifrån de vedertagna rättskällorna, varefter den ”rekonstruerade” rätten appliceras på ett visst rättsligt problem (de sociala mediernas informationshantering).<sup>7</sup> Argumentationen i detta hänseende grundas främst i lag och andra bindande folkrättsliga normer, domstolsavgöranden, förarbeten och doktrin. I den mån diskussionen övergår i de lege ferenda-resonemang (vilken ansvarsreglering som *bör* gälla) stödjer sig framställningen dock på friare ändamålsargumentation. Även detta får emellertid anses rymmas väl inom den rättsdogmatiska metoden.<sup>8,9</sup>

Historiska, komparativa och empiriska inslag förekommer i arbetet. Sådana inslag avser berika och komplettera den rättsdogmatiska framställningen – främst av pedagogiska skäl. För tydlighets skull redovisas detta närmare i dispositionen nedan.

---

<sup>7</sup> Se Sandgren, *Rättsvetenskap för uppsatsförfattare*, s. 49, Jareborg, SvJT 2004 s. 1–10, Peczenik, SvJT 2005 s. 249–250 samt Kleineman, *Juridisk metodlära* s. 21–29.

<sup>8</sup> ”Rättsdogmatik förknippas primärt med rekonstruktion av gällande rätt och därmed också med begränsningar som uppställs genom maktutövning och beslutsfattande av lagstiftare och domstolar. Detta hindrar dock på intet sätt att man vid rättsdogmatisk argumentation vidgar perspektivet och går utanför gällande rätt. All vetenskaplig verksamhet går ut på att finna nya svar och bättre lösningar. Det är helt legitimt för rättsdogmatiker att söka efter ideala lösningar”. Jareborg, SvJT 2004 s. 4. Jfr även Kleineman, *Juridisk metodlära* s. 36–41, som skulle karaktärisera detta som en ”kritisk rättsdogmatisk metod”. Se dock Sandgren, *Rättsvetenskap för uppsatsförfattare*, s. 49.

<sup>9</sup> Jag är medveten om att synen på den rättsdogmatiska metoden inom svensk juridik är splittrad. Jag anser emellertid att det varken är intressant eller betydelsefullt att, inom ramen för detta arbete, gå närmare in på distinktionerna mellan den rättsdogmatiska metoden och den ”rättsanalytiska” eller ”rättspolitiska” metoden, såsom de ibland beskrivs i litteraturen (jfr Sandgren, *Rättsvetenskap för uppsatsförfattare*, s. 48–54). Inte heller anser jag det betydelsefullt att göra skillnad mellan en rättsdogmatisk metod och vad som ibland benämns den EU-rättsliga metoden (jfr Reichel, *Juridisk metodlära*, s. 109–142). I den mån jag har analyserat EU-rätten, eller andra utländska rättssystem, har jag givetvis beaktat de institutionella skillnaderna som präglar de olika systemen. Det väsentliga, ur en metodologisk synvinkel, får anses vara vilken metod som *faktiskt* tillämpats, snarare än hur metodiken ska betecknas.

### 2.3.2 *Allmänt om material*

Som konstaterat utgår den juridiska argumentationen i arbetet från de vedertagna rättskällorna. Av tillgänglighetsskäl följer en närmare redogörelse för det material som använts i dispositionen nedan, i förhållande till relevant avsnitt. Dessförinnan är dock ett par klargöranden på sin plats angående det material som ligger till grund för framställningen.

I avsnitt fyra, fem och sju förekommer enstaka tabeller och figurer. Tabellerna i avsnitt fyra och sju (fig. 1, 2 och 5) är illustrationer som förtydligar de i löptexten förda resonemangen och slutsatserna, främst i pedagogiskt syfte. Tabellerna i avsnitt fem (fig. 3 och 4) utgör till löptexten kompletterande information, avsedda att demonstrera vissa konkreta skillnader i de sociala mediernas innehållsmoderering. De grundar sig på plattformarnas egna villkor och riktlinjer, men är självständigt systematiserade och sammansatta.

I samband med arbetet har även ett antal intervjuer genomförts med yrkesverksamma jurister, journalister och forskare inom medie- och kommunikationsvetenskap. Detta redovisas här enbart som en transparent redogörelse för arbetssättet: i övrigt sker inga hänvisningar till intervjuer eller muntliga källor i uppsatsen. Intervjuerna har uteslutande syftat till att bidra med inspiration, erfarenhet och perspektiv på ämnet. Varje rättslig bedömning och varje ställningstagande bör tillskrivas författaren, om inte annat framgår, inbegripet eventuella felbedömningar, feltolkningar eller misstag.<sup>10</sup>

Slutligen en kommentar angående de rättskällor som åberopas. Yttrande- och informationsfriheten, såsom den kodifieras i svensk rätt, i EU-rätten och i EKMR, är förhållandevis vagt utformad. Dess verkliga innebörd har fastställts successivt av domstolar, som tolkat och preciserat rättigheternas utsträckning i olika sammanhang. Då gällande rätt presenteras och systematiseras har jag i så stor utsträckning som möjligt valt att förankra min argumentation direkt i domstolarnas avgöranden. Jag har inte bortsett från den åtskilliga och ambitiösa doktrin som utverkats av rättsvetare på området, inte heller har jag avsett uppfinna hjulet på nytt. Att jag hellre åberopat domstolsavgöranden för att fastställa yttrande- och informationsfrihetens innebörd är av strikt rättskällehierarkiska skäl. Det framstår som lämpligare, för ett arbete av denna typ, att hänvisa direkt till primärkällan.

---

<sup>10</sup> Med det sagt, är jag verkligen tacksam till dem som generöst ställt sin tid till förfogande för att bidra med värdefulla synpunkter i ämnet och på arbetet.



### 2.3.3 *Om uppsatsens disposition*

Den följande framställningen inleds med ett avsnitt om yttrande- och informationsfriheten ur ett historiskt-filosofiskt perspektiv (avsnitt tre). Detta syftar till att sätta ämnet i en bredare kontext, men också att betona yttrande- och informationsfrihetens betydelse för samhället. Avsnittet utgår från en rättshistorisk metodik – även om de källor som konsulterats, med hänsyn till uppsatsens fokus, till stor del utgörs av sekundära historiska källor.

I det fjärde avsnittet följer en genomgång av de stora förändringar som skett i informationssamhället under modern tid, och vilka problem som dessa förändringar fört med sig ur ett regleringsperspektiv. Syftet är att ge en regulatorisk bakgrund till de nya ansvarsproblem som uppstått i informationssamhället, med utgångspunkt i gällande lagar, prejudicerande domar och doktrin. Här introduceras olika typer av mellanhänder i informationssamhället – varav de sociala medierna utgör en typ – och de regleringsmodeller som har tillämpats för att etablera ansvar för de olika mellanhänderna.

Svenska och EU-rättsliga normer jämförs här med amerikansk reglering, vilket ger avsnittet en komparativ karaktär. Den amerikanska rätten blir ett intressant jämförelseobjekt av flera anledningar. Dels är europeiska och amerikanska förhållanden jämförbara sett till ekonomi, demografi och utveckling. Dels har yttrandefrihetsintresset en stark roll i det amerikanska rättslivet. Både i lagstiftningsammanhang och praxis blottläggs ofta förhållandet mellan yttrandefriheten och ansvarsregleringen i amerikansk rätt på ett sätt som är mer utförligt och explicit än i svensk rätt eller EU-rätt. De många beröringspunkterna mellan europeisk och amerikansk reglering tyder också på ett samband mellan rättsordningarna – en samsyn i fråga om yttrandefrihet och mellanhandsansvar i informationssamhället.

Det femte avsnittet handlar specifikt om de sociala medierna och deras roll som kuratorer och moderatorer i informationssamhället. Här beskrivs hur de sociala medierna på olika sätt formar och sätter gränser för informationsdistributionen på deras plattformar: deras innehållsmoderering i form av regler, sanktioner och innehålls-algoritmer. Som grund för framställningen ligger huvudsakligen plattformarnas egna användaravtal och riktlinjer, som granskats, jämförts, sammanställts och systematiserats i en tillgänglig form. Därtill har befintlig forskning använts för att komplettera beskrivningen.

I det sjätte avsnittet diskuteras de sociala mediernas innehållsmoderering i förhållande till yttrande- och informationsfriheten, såsom den utvecklats i svensk och europeisk rätt. Syftet är framför allt att analysera hur väl plattformarnas modereringspraxis stämmer överens med gällande yttrande- och informationsfrihetsstandarder. Som grund för analysen ligger regleringen i RF 2 kap., TF, YGL, EKMR och EU-stadgan; förarbeten, där sådana finns; domstolspraxis från HD, HFD, och Europadomstolen som berör yttrande- och informationsfrihetens villkor och gränser; samt doktrin och forskningsrapporter.

Det sjunde avsnittet gäller de sociala medierna och deras ansvar för yttrande- och informationsfriheten, de lege lata såväl som de lege ferenda. Här ligger i princip samma källmaterial till grund för diskussionen som i avsnitt sex, men istället för yttrande- och informationsfrihetens materiella innehåll – dess *ratione materiae* – diskuteras normernas subjektiva omfång: dess *ratione personae*. Därefter förs en diskussion gällande den nuvarande regleringens lämplighet, och behovet av nya regler, utifrån ändamålsresonemang.

Det sista avsnittet utgörs av en kort epilög, med ett par iakttagelser och reflektioner som får avsluta uppsatsen.

### 3 Yttrande- och informationsfriheten – en idéhistorisk kontext

*There is freedom of speech,  
but I cannot guarantee freedom after speech.*

Idi Amin, Ugandas president 1971–1979

Detta första avsnitt syftar till att sätta yttrande- och informationsfriheten i en bredare historisk-filosofisk kontext. Vi kommer att följa det fria ordets utveckling från tiden för Gutenbergs tryckpress, genom den engelska censurens avskaffande, den svenska tryckfrihetsförordningens stiftande, den franska och amerikanska revolutionen och den europeiska krigs- och efterkrigstiden under 1900-talet.<sup>11</sup> Vi kommer också att ta del av några av de centrala argument som förts fram genom historien, mot censur och för yttrandefrihet. Med förståelse för yttrande- och informationsfrihetens själva existensberättiganden, och de hinder som lagts i frihetens väg under historien, står vi sedan bättre rustade inför en diskussion kring de utmaningar som friheterna möter idag.

#### 3.1 När ordet blev fritt

Med boktryckarkonstens uppkomst i Europa, kring mitten av 1400-talet, skedde sannolikt det största jordskredet i informationsteknikens historia. Från en muntlig tradition av informationsförmedling; från manuell handskrift och brevväxling inom en begränsad krets av skriv- och läskunniga, skulle det nu bli möjligt att massdistribuera information inom vida kretsar. Även om det skulle ta ytterligare ett par århundraden till innan ett fungerande pressväsende växte fram och kunde sprida information till allmänheten i praktiken, hade nyhetsförmedling, opinionsbildning och idéutbyten aldrig kunnat ske i samma skala utan tryckpressens frammarsch.<sup>12</sup>

Denna insikt var emellertid inte glädjande för alla. Den katolska kyrkan förstod snart att den egna makthegemonin skulle störas om åsikter och idéer kunde distribueras fritt utan kyrkans tillsyn, och prästerskapet reagerade genom att inrätta egna censurdekret och svartlista sådan tryckt litteratur som ansågs stå i strid med ”den sanna läran”.<sup>13</sup>

---

<sup>11</sup> Det sägs att den amerikanska journalisten I F Stones stora opus om rättegången mot Sokrates (*The Trial of Sokrates*, 1988) ska ha börjat som en historisk översikt om yttrandefriheten. Stone började utforska yttrandefriheten i det antika Grekland, lärde sig klassisk grekiska – och kom aldrig riktigt vidare. Resultatet blev en bästsäljare, och en påminnelse om riskerna med att gräva alltför djupt i den rättshistoriska myllan. Så trots att kampen mellan makten och det fria ordet antagligen kan spåras hur långt tillbaka som helst kommer avsnittet ta vid en bit in, ungefär två millenier efter det att Sokrates avrättats för sin frispråkighet. Se Von Vegesack, *När ordet blev fritt*, s. 13.

<sup>12</sup> Eek, *Om tryckfriheten*, s. 27.

<sup>13</sup> Nordin, *Fritt ord 250 år*, s. 28–33.

På egen hand saknade kyrkan dock medel att åstadkomma en verkligt effektiv censur. Det var först i samband med reformationen, då den kyrkliga och världsliga makten började förenas, som kontrollen över det fria ordet blev verklighet.<sup>14</sup> I Sverige lät Gustav Vasa förstatliga tryckerierna, förbjuda kättersk litteratur och helt stoppa importen av tryckta skrifter. I Frankrike föreskrevs dödsstraff för den som skrev, tryckte, sålde eller innehade förbjuden litteratur, och i England förbjöds ömsom katolsk, ömsom anti-katolsk litteratur, beroende på den sittande regentens trosuppfattning.<sup>15</sup>

### 3.1.1 *England och tryckfrihetens gryning*

Det var naturligtvis inte bara kyrkliga intressen som styrde yttrandefrihetsinskränkningarna – snart insåg också den politiska makten vikten av att kontrollera informationsflödena. Den litteratur som cirkulerade skulle inte bara överensstämma med lagen, utan också statens politiska intressen. I England infördes förhandsgranskning av alla tryckta skrifter under 1600-talet. Författaranonymitet gjordes omöjligt, tryckerier avkrävdes statliga tillstånd och anländande resenärer fick sina bagage rannsaka efter förbjudna skrifter. Under en kort period i slutet av seklet förbjöds till och med all nyhetsrapportering. Det var en tid av kraftig repression, men en tid som också skulle bli början till kampen för tryckfrihet i västvärlden.<sup>16</sup>

1644 formulerade författaren John Milton det första brett utformade försvaret för pressfrihet. I skriften ”Areopagitica” liknade Milton de engelska censurreglerna vid ”fiendeskepp” som blockerade landets hamnar och inlopp från den mest värdefulla av alla handelsvaror – sanningen.<sup>17</sup> Miltons skrift blev inflytelserik, och kom att bana väg för de engelska författare, jurister och filosofer vars idéer skulle bli bärande för det fria ordets frammarsch i Europa – däribland Daniel Defoe, Matthew Tindal, John Locke, och senare James och John Stuart Mill.<sup>18</sup>

Efter en reformbegäran till det brittiska parlamentet, undertecknad John Locke själv, kom England snart att bli det första landet att avskaffa censuren år 1695. Censurens avskaffande innebar emellertid bara en avreglering, och ingen substantiell tryckfrihet.

---

<sup>14</sup> Eek, *Om tryckfriheten*, s. 28–29, Von Vegesack, *När ordet blev fritt*, s. 21–22 och Strömberg & Lundell, *Grundlagsskyddad yttrandefrihet*, s. 19.

<sup>15</sup> Eek, *Om tryckfriheten*, s. 43, Nygren, *Fritt ord 250 år*, s. 167–168, Von Vegesack, *När ordet blev fritt*, s. 22–24 och Von Vegesack, *Smak för frihet. Opinionsbildningen i Sverige 1755–1830*, s. 314–317.

<sup>16</sup> Eek, *Om tryckfriheten*, s. 29 och Von Vegesack, *När ordet blev fritt*, s. 22–24.

<sup>17</sup> Milton, *Areopagitica; A Speech of Mr John Milton for the Liberty of Vnlicenc'd Printing, to the Parliament of England*, s. 29.

<sup>18</sup> Von Vegesack, *När ordet blev fritt*, s. 26–30 och 109–111.

Lagstridiga publiceringar kunde fortfarande leda till långa fängelsestraff, straffarbete och förverkande av de otillåtna skrifterna genom bokbål. Därtill infördes ett slags ekonomiskt censursystem, där stämpelskatter, annonsskatter, borgenskrav och importskatt på papper användes för att stävja spridningen av tidningar. Den engelska censurens avskaffande var en historisk reform, om än bara en delseger i kampen för yttrandefrihet.<sup>19</sup>

### 3.1.2 *Den svenska tryckfriheten*

Nästa stora framsteg skedde i Sverige. De inskränkningar i tryckfriheten som införts under Gustav Vasa hade förvärrats kraftigt under den karolinska eran och kulminerat i ett helt och hållet despotiskt åsiktsförtryck.<sup>20</sup> När enväldet avbröts och frihetstiden inträdde i början av 1700-talet var tiden mogen för reformer. Den riksdag som övertagit lagstiftningsmakten från kungen var delad mellan det regerande hattpartiet och det opponerande mösspartiet. Hattpartiet hade under sin långvariga mandatperiod gjort sig ökända för korruption, och när det slutligen vankades maktskifte gick mösspartiet till val på att öppna upp den slutna byråkratin och öka insynen i maktapparaten. Det var i denna kontext som offentlighetsprincipen, med rätten för envar att ta del av allmänna handlingar, utkristalliserades. Därtill föreslogs en rätt för envar att fritt trycka och sprida skrifter, i syfte att ge offentlighetsprincipen så stort praktiskt genomslag som möjligt. Dessa två principer lagstodgades i 1766 års tryckfrihetsförordning, och gavs därmed grundlagsställning.<sup>21</sup> Den unika svenska tryckfrihetsförordningen innebar, till skillnad från den engelska regleringen, en *positiv rätt* att trycka och sprida skrifter utan allmänhetens inblandning.<sup>22</sup>

Helt fritt var ordet förstås inte för det. Det statliga censursämbetet avskaffades med förordningen, men den teologiska censuren kvarstod och utövades av kyrkliga konsortium. Därtill kvarstod importrestriktionerna för utländska skrifter.<sup>23</sup> Men tryckfrihetsförordningen, som gjorde frihet till huvudregel, var banbrytande.

---

<sup>19</sup> Eek, *Om tryckfriheten*, s. 32–38, Nordin, *Fritt ord 250 år*, s. 40–49 och Von Vegesack, *När ordet blev fritt*, s. 26–30 och 109–111.

<sup>20</sup> Nygren, *Fritt ord 250 år*, s. 168–173 och Von Vegesack, *Smak för frihet. Opinionsbildningen i Sverige 1755–1830*, s. 314–317.

<sup>21</sup> Hirschfeldt, *Fritt ord 250 år*, s. 601 och Nygren, *Sex brytpunkter i svensk konstitutionell rätt*, s. 15–49.

<sup>22</sup> Nygren, *Fritt ord 250 år*, s. 197–200.

<sup>23</sup> Von Vegesack, *Smak för frihet. Opinionsbildningen i Sverige 1755–1830*, s. 314–317. Se även, för utvecklingen av tryckfriheten efter 1766, Funke, *Offentlighetsprincipen – praktik och teori*, s. 181–184 och Axberger, *Fritt ord 250 år*, s. 241–282.

### 3.1.3 *Frihet, upplysning, revolution*

Den engelska "förupplysningens" idéer som lett fram till censurens avskaffande 1695 visade sig vara början på någonting mycket större. Genom filosofer som Montesquieu och Voltaire kom Lockes ideologi kring maktindelning, empirism, tolerans och politisk frihet att importeras till Frankrike, för att sedan spridas bland franska encyklopedister och revolutionärer.<sup>24</sup>

Synen på tryck- och yttrandefrihet bland de franska filosoferna var inte strömlinjeformad. Encyklopedister som Voltaire, Diderot, d'Alembert och Rousseau var kanske flitigare med att själva kringgå den ofta korrupta och ineffektiva franska censurregimen än att formulera slagkraftiga argument för bred och allmän yttrandefrihet.<sup>25</sup>

Men andra, som de Condorcet och Sieyès, var ihärdigare ideologer. De konstaterade att yttrandefriheten var av naturen given, och inte behövde förkunnas genom lag för att vara giltig. Dessutom lanserades idén om att yttrandefriheten också innebar en garantiplikt för staten: en plikt att utforma lagarna så att myndigheter och andra enskilda inte inskränker medborgarens naturgivna rättigheter.<sup>26</sup>

Upplysningsidealen fick också fäste på andra sidan Atlanten. Under frihetskriget från England hade framstående amerikanska revolutionärer som Thomas Jefferson, Benjamin Franklin och James Madison inspirerats av de lockesianska rättighetsidealen, många i tätt samröre med sina franska gelikar.<sup>27</sup> När de första amerikanska kolonierna nedtecknade sin självständighetsförklaring 1776 genomsyrades den av de naturrättsliga idéerna om att *alla* människor var innehavare av vissa okränkbara rättigheter. Vissa av dessa rättigheter hade tidigare kodifierats i delstaternas egna konstitutioner – varav Virginias rättighetskatalog från 1776 var den första – och när det var dags för de nya förenta staterna att anta en gemensam konstitution 1789 lades dessa rättigheter till som orubbliga tillägg till konstitutionen.<sup>28</sup> Det första tillägget försäkrar var och en rätten till yttrandefrihet, pressfrihet och religionsfrihet:

Congress shall make no law [...] abridging the freedom of speech, or of the press [...]

---

<sup>24</sup> Bring, *De mänskliga rättigheternas väg*, s. 173–195.

<sup>25</sup> Von Vegesack, *När ordet blev fritt*, s. 35–40. Citatet "Jag ogillar vad du säger, men jag är beredd att gå i döden för din rättighet att säga det" brukar tillskrivas Voltaire, men härstammar i själva verket från Evelyn Beatrice Hall, i boken *Voltaires vänner* från 1906. Se Bring, *De mänskliga rättigheternas väg*, s. 191.

<sup>26</sup> Eek, *Om tryckfriheten*, s. 42–44.

<sup>27</sup> Bring, *De mänskliga rättigheternas väg*, s. 206–210 och Von Vegesack, *När ordet blev fritt*, s. 57.

<sup>28</sup> Bring, *De mänskliga rättigheternas väg*, s. 208–212.

Den amerikanska yttrandefrihetsförklaringen innebar för första gången, till skillnad från den engelska ordningen, ett förbud mot *varje* lag som inskränker yttrandefriheten – vare sig ren censur eller de facto inskränkande åtgärder som stämpelskatter, tillståndsplikter och dylikt.<sup>29</sup>

Utvecklingen mot mänskliga rättigheter och yttrandefrihet i USA följdes intensivt i Frankrike, där revolutionen nu var ett faktum. När nationalförsamlingen sammanträtt och övertagit makten var yttrandefriheten hett omdebatterad. Förslag framfördes om allt mellan en oinskränkt yttrandefrihet till ingen yttrandefrihetslagstiftning alls. Till slut antogs en rättighetsdeklaration, inspirerad av den amerikanska, vars artikel XI anger:

Varje medborgare äger rätt att fritt yttra sig i tal, skrift och tryck, men får stå till svars för missbruk av denna frihet i de fall som lagen stadgar

Den andra satsen återspeglar bland annat Rousseaus idé om att rättigheter, om än av naturen givna, inte är absoluta. Att utnyttja sina rättigheter till skada för andra är inte försvarbart, och sådant missbruk kan beivras med stöd i lag.<sup>30</sup>

Den franska rättighetsförklaringen från 1792 blev dock inte särskilt långlivad. Redan året därpå suspenderades tryckfriheten, och under den napoleonska eran som följde återinfördes censuren, och åsiktsförtrycket blev till och med värre än under *l'ancien regime*.<sup>31</sup> Men de idéer som gett upphov till revolutionen och rättighetsdeklarationen levde vidare.

### 3.1.4 *Världskriget och totalitarismen*

Under seklet som följde den franska och amerikanska revolutionen svepte en lavin av liknande yttrandefrihetsförklaringar över Europa. Den tyska revolutionen 1848 gav varje tysk rätt att i ”ord, skrift, tryck eller framställning i bild fritt uttala sin mening”. Samma år fick Italien, Schweiz och Nederländerna liknande yttrandefrihetsbestämmelser. De övriga Beneluxländerna hade fått yttrandefrihetslagar redan 1831 respektive 1815. I Grekland infördes pressfrihet 1864, i Bulgarien 1879, i Rumänien 1884, och i Serbien 1901.<sup>32</sup>

I breda penseldrag kan det alltså konstateras att ordet blev allt friare i Europa. Vägen mot yttrandefrihet i västvärlden var dock, som framgått, inte helt linjär. Den kantades av många och svåra återgångar till censur och repression, såväl i Frankrike som i Sverige.

---

<sup>29</sup> Von Vegesack, *När ordet blev fritt*, s. 110.

<sup>30</sup> Bring, *De mänskliga rättigheternas väg*, s. 222–223.

<sup>31</sup> Von Vegesack, *När ordet blev fritt*, s. 67–75.

<sup>32</sup> Von Vegesack, *När ordet blev fritt*, s. 102–107 och Eek, *Om tryckfriheten*, s. 55–56.

Ett av de mest förödande bakslagen i yttrandefrihetens historia skedde i samband med världskriget under första hälften av 1900-talet.

I samband med första världskrigets utbrott blev de europeiska nationsstaterna ivrigare att hindra sådan nyhetsrapportering och opinionsbildning som kunde kompromettera den egna statens ställning i kriget. I Tyskland återinfördes censuren omedelbart vid krigsutbrottet, och Frankrike, England och Italien infördes undantagslagar som i praktiken gjorde den fria pressen till en chimär. Särskilda militära pressbyråer försågs med mandat att förhandsgranska pressmeddelanden, tidningar, och böcker. Budskap som kunde inverka menligt på den allmänna moralen, folkstämningen eller statsintresset kriminaliserades.<sup>33</sup>

Ännu värre blev läget med det andra kriget. De auktoritära diktaturer som då tagit form i bland annat Sovjet, Tyskland och Italien förde inte längre en presspolitik med syfte att gjuta olja på vågorna och upprätthålla nationell säkerhet. Istället blev pressen ett instrument för statens agenda.

I Sovjet förstatligades pressväsendet, som en förvisso naturlig del i den kommunistiska ideologin om offentligt ägda industrier. Pressen var underställd partiet, och tidningsmännen ersattes av statliga tjänstemän. Snarare än för maktgranskning användes medierna i propagandistiskt och pedagogiskt syfte – för att sprida partiets budskap och fostra fogliga medborgare. I Tyskland och Italien skedde ingen formell monopolisering av tidningsväsendet; istället skedde ett indirekt förstatligande genom uppställandet av strikta yrkeskrav för journalister, genom perverterade omtolkningar av yttrandefrihetslagstiftningen och genom hårdare straff-, civil- och yrkesrättsliga sanktioner för den som vågade utöva sitt yrke i strid med diktaturens politik.<sup>34</sup>

Krigseran innebar en fullkomlig lemlästning av den frihetstradition som hade börjat utvecklas i Europa, som i vissa fall skulle fortgå långt efter krigens slut. Men den brutala nonchalansen inför människans rättigheter som demonstrerats under världskriget kom också att bli en katalysator för de än starkare rättighetsideal som följde i krigens kölvatten.

---

<sup>33</sup> Eek, *Om tryckfriheten*, s. 58–62.

<sup>34</sup> Eek, *Om tryckfriheten*, s. 126–128.



### 3.1.5 Rättigheternas renässans

Tre år efter krigsslutet, i december 1948, nedtecknades FN:s allmänna förklaring om de mänskliga rättigheterna, med explicit hänvisning till den ”ringaktning och förakt för de mänskliga rättigheterna” som hade lett till ”barbariska gärningar” under krigen. Dess nittonde artikel garanterar var och en yttrandefrihet i form av:

[F]rihet att utan ingripande hysa åsikter samt söka, ta emot och sprida information och idéer med hjälp av alla uttrycksmedel och oberoende av gränser

FN:s deklARATION följdes snart av en rad liknande, regionala konventioner. Knappt fem år senare trädde den *Europeiska konventionen om skydd för de mänskliga rättigheterna och de grundläggande friheterna* i kraft mellan medlemmarna i Europarådet, inspirerad av FN-deklARATIONEN. 1969 tillkom en interamerikansk rättighetskonvention, och 1981 en afrikansk motsvarighet.<sup>35</sup> Kring millennieskiftet påbörjades därtill ett arbete med en egen rättighetsstadga för EU, som trädde i kraft i samband med Lissabonfördraget 2009.<sup>36</sup>

Rättighetsrenässansen fick uttryck också i nationell rätt. I samband med grundlagsreformen 1974 intogs vissa medborgerliga rättigheter i den nya regeringsformen, däribland yttrande- och informationsfriheten, förbudet mot retroaktiv strafflagstiftning och landsförvisning av svenska medborgare.<sup>37</sup> Frågan om fri- och rättigheternas konstitutionella ställning blev dock föremål för fortsatt utredning, och redan 1975 föreslogs att rättighetsskyddet skulle utvidgas ytterligare. De medborgerliga rättigheterna, däribland opinionsfriheterna, fick ett eget kapitel och möjligheten att inskränka rättigheterna genom lag underkastades hårdare restriktioner.<sup>38</sup>

Samma år föreslogs en svensk massmediegrundlag, som skulle behandla ”yttrandefrihet i tryckt skrift, radio, television och film”, och därmed ersätta 1949 års tryckfrihetsförordning.<sup>39</sup> Förslaget antogs inte – istället kom tryckfrihetsförordningen att kompletteras med en fjärde grundlag med bestämmelser om yttrandefrihet i radio, TV, film och liknande medier – 1991 års yttrandefrihetsgrundlag.<sup>40</sup>

---

<sup>35</sup> Sevastik et al., *En bok i folkrätt*, s. 215–218 och Cameron, *An Introduction to the European Convention on Human Rights*, s. 38.

<sup>36</sup> Spaventa, *European Union Law*, s. 236–237.

<sup>37</sup> Prop. 1973:90 s. 95–96 och SOU 1972:15 s. 106–108.

<sup>38</sup> Holmberg & Stjernquist, *Vår författning*, s. 35–36 och SOU 1975:75 s. 19–25.

<sup>39</sup> SOU 1975:49 s. 11.

<sup>40</sup> Regner, SvJT 1991 s. 428–444.

### 3.1.6 Yttrandefriheten, här och nu

De individuella fri- och rättigheternas frammarsch, från efterkrigstiden till idag, har resulterat i ett omfattande och överlappande rättighetskomplex i svensk rätt – inte minst för yttrande- och informationsfriheten. I RF 2 kap. 1 § tillförsäkras var och en:

[F]rihet att i tal, skrift eller bild eller på annat sätt meddela upplysningar samt uttrycka tankar, åsikter och känslor, [och] frihet att inhämta och ta emot upplysningar samt att i övrigt ta del av andras yttranden

Därtill finns de mediebaserade grundlagarna, varav TF behandlar särskilda yttrandefrihetsregler för tryckta skrifter och YGL omfattar yttrandefrihetsregler för radio, tv, ljudupptagningar, databaser och vissa andra massmedier. För yttranden som omfattas av TF och YGL gäller grundlagarna exklusivt.<sup>41</sup>

Dessutom är Sverige sedan 1952 folkrättsligt bundet av Europakonventionen, vars artikel 10 garanterar varje enskild ”åsiktsfrihet samt frihet att ta emot och sprida uppgifter och tankar”. Konventionen gäller sedan 1994 som svensk lag, och genom RF 2 kap. 19 § gäller att svensk lag inte får stiftas i strid med konventionsbestämmelserna.<sup>42</sup> På EU-rättslig nivå skyddas yttrandefriheten dessutom av artikel 11.1 i EU-stadgan, som har samma lydelse som EKMR artikel 10 och är juridiskt bindande för EU:s institutioner och medlemsstaterna vid tillämpning av unionsrätten.<sup>43</sup>

## 3.2 Det fria ordets syfte

Om ingenting annat, visar historien att det fria ordets väg sällan varit rak och följsam. Varje steg från censur och åsiktsförtryck har kantats av bakslag och uppoffringar, och viljan att tysta det oppositionella, det hädiska, banbrytande och okonventionella verkar lika gammal som tiden. Men i varje era, parallellt med varje bokbål och censurdekret, har det också funnits människor som placerat sig i yttrandefrihetens ringhörna.

Somliga har försvarat yttrande- och informationsfriheten som en av naturen given rättighet, ett uttryck för individens självklara suveränitet, vars rättfärdigande inte behöver förankras i termer av samhällliga intressen. Friheten har ett intrinsiskt värde och är således ett självändamål, snarare än ett medel för kollektiv nytta (en *deontologisk* hållning).

---

<sup>41</sup> Strömberg & Lundell, *Grundlagsskyddad yttrandefrihet*, s. 21–43.

<sup>42</sup> Jfr Cameron, *An Introduction to the European Convention on Human Rights*, s. 178–184.

<sup>43</sup> Jfr artikel 51(1) EU-stadgan. Se även Spaventa, *European Union Law*, s. 236–252. Angående förhållandet mellan RF, EKMR och EU-stadgan i svensk rättstillämpning, se Cameron, *An Introduction to the European Convention on Human Rights*, s. 184–186.

John Stuart Mill, och senare libertarianer som Robert Nozick, ansåg att friheten inte med rätta kunde begränsas annat än i syfte att skydda andra från skada:

[T]he sole end for which mankind are warranted, individually or collectively, in interfering with the liberty of action of any of their number, is self-protection. [T]he only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others [...] Over himself, his own body and mind, the individual is sovereign.<sup>44</sup>

Den deontologiska hållningen utesluter emellertid inte en diskussion om yttrandefrihetens praktiska fördelar i ett samhälle – dess instrumentella värde (en *konsekventialistisk* hållning).<sup>45</sup> Det framgår inte minst av Mills försvarstal för yttrandefriheten i *On Liberty*.

Yttrandefrihetens *raison d'être*, ur samhällelig synpunkt, har också varit bland de mest omskrivna frågorna i debatten kring det fria ordets vara eller icke-vara. Och argumenten har varit vittskilda. Under tryckfrihetens formativa år kunde vissa filosofer bland annat åberopa yttrande- och informationsfriheten som ett medel för despoter och tyranner att säkerställa inblick i – och således kontroll över – medborgarnas tankeliv.<sup>46</sup> Ett sådant rättfärdigande kan förstås tyckas bakvänt och främmande i en modern kontext. Mer livskraftiga är istället de argument som formulerats ur allmänhetens perspektiv. Dessa argument belyser yttrande- och informationsfrihetens vitala funktion ur medborgarkollektivets och mänsklighetens synpunkt.

### 3.2.1 Yttrandefrihet och sanning

Ett ofta åberopat argument för yttrande- och informationsfriheten är dess nytta i jakten på sanning. Argumentet har formulerats av Locke i *An Essay Concerning Human Understanding*:

While the parties of men cram their beliefs down the throats of everyone they can get into their power, without allowing them to examine their truth or falsehood, and won't let truth have a fair run for its money in the world or allow men the freedom to search for it, what improvements of this kind can be expected?<sup>47</sup>

Locke skriver inte bara om ”sanning” i strikt empirisk bemärkelse, utan även om sanningen i moraliska, filosofiska, metafysiska och till och med estetiska frågor.

---

<sup>44</sup> Mill, *On Liberty*, s. 17. Se även Nozick, *Anarki, Stat och Utopi*.

<sup>45</sup> Även autonomiargumentet har klassats som konsekventialistiskt, då själslig autonomi kan fostra mer mogna och reflekterande medborgare till samhällets nytta. Se Barendt, *Freedom of Speech*, s. 13.

<sup>46</sup> Von Vegesack, *När ordet blev fritt*, s. 43–44.

<sup>47</sup> Locke, *An Essay Concerning Human Understanding*, IV.iii.20.

Om makten tillåts begränsa den information som får komma till uttryck, lyder resonemanget, hur vore det då möjligt för folket att jämföra motstridiga uppfattningar – att stöta och blöta olika argument i offentlighetens ljus för att närma sig det mest riktiga, lämpliga och sanna?

Resonemanget utvecklades ytterligare av Mill, som menar att argumentet är lika applicerbart vare sig de förtryckta och förbjudna åsikterna skulle visa sig sanna eller falska. Om den förtryckta opinionen skulle vara sann, menar Mill, berövas människorna sanningen. I alla tidsåldrar har människor hållit fast vid uppfattningar som senare visat sig falska, om inte absurda – och det vore inte orimligt att anta att också våra moderna uppfattningar skulle kunna bli föremål för samma retrospektiva förakt. Det enda som berättigar oss att hålla ett faktum för sant är att faktumet är öppet för kritik och klander.

Om de förtryckta uppfattningarna med säkerhet kunde antas vara falska, menar Mill, vore åsiktsförtryck emellertid lika olämpligt. De vedertagna sanningarna skulle då snart reduceras till ”döda dogmer”, sårbara inför de mest banala och uppenbart falska motargumenten, eftersom medborgarna aldrig hade fått lära sig försvara sanningen. Likaså vore innehavarna av ”falska” uppfattningar oförbätterliga, eftersom deras föreställningar aldrig skulle få lov att utmanas i det öppna.<sup>48</sup> Den senare utgångspunkten har liknats vid ”homeopatisk masochism”: om vi tillåts möta det falska i små, vardagliga doser, kommer vi att vara bättre rustade inför mötet med de stora lögnerna.<sup>49</sup>

### 3.2.2 *Yttrandefrihet, demokrati och insyn*

Relaterat till sanningsargumentet är demokratiargumentet. En reell demokrati – i bemärkelsen folkstyre genom fria, allmänna val – kan naturligtvis inte utövas utan fritt åsiktsutbyte. Detta fastslås redan i regeringsformens portalparagraf: ”Den svenska folkstyrelsen *bygger på fri åsiktsbildning* och på allmän och lika rösträtt” (egen kursivering).

I en bredare bemärkelse kan demokrati också avse allas lika rätt att delta i samhället och den offentliga diskursen, vilket förutsätter en fundamental yttrandefrihet för varje individ, vare sig i majoritet eller minoritet. Den senare definitionen är vad Ronald Dworkin kallat för det *konstitutionella demokratibegreppet*.<sup>50</sup>

---

<sup>48</sup> Mill, *On Liberty*, s. 25.

<sup>49</sup> Peters, *Courting the Abyss: Free Speech and the Liberal Tradition*, s. 6.

<sup>50</sup> Barendt, *Freedom of Speech*, s. 19–20.

Men beröringspunkterna mellan demokrati och yttrande- och informationsfriheten är fler än de delaktighetsresonemang som formulerats ovan. En av de tidiga förkämparna för tryckfriheten i Sverige, Peter Forsskål, erbjöd följande resonemang i sin skrift ”Tankar om borgerliga friheten” från 1759:

Skriffrihet uppdrifwer wetenskaperna till sin höjd, röjer alla skadeliga författningar, tyglar alla ämbetsmäns orättwisa, och är Regeringens tryggaste försvar i ett fritt rike.<sup>51</sup>

Den första satsen anknyter tydligt till sanningsargumentet: tryckfriheten bidrar till sökandet efter sanning och vetenskap. De följande två argumenten – att tryckfriheten kan rensa ut skadliga lagar och tygla ämbetsmäns orättvisa – verkar snarare handla om yttrandefrihetens instrumentella värde inom ramen för folkstyret. Yttrandefrihet kan tjäna till att sprida kunskap om, och bilda opinion kring, orättvis och korrupt maktutövning.

I det hänseendet är yttrandefriheten en nödvändig förutsättning för maktgranskning. Det var, som konstaterat, också denna uppfattning som låg till grund för den svenska tryckfrihetsförordningens tillkomst, bara tre år efter Forsskåls död.<sup>52</sup> Argumentet om yttrandefriheten som demokratins väktare har således haft en stark genomslagskraft genom historien, och utgör en av de centrala försvarspunkterna för yttrandefriheten än idag.<sup>53</sup>

### 3.2.3 Yttrandefrihet och tolerans

Det tredje argumentet ligger nära sannings- och demokratiargumentet, men har en annan utgångspunkt. Argumentet innebär att yttranden bör tolereras, även om de varken bidrar till sanningen, vetenskapen, samhällsutvecklingen eller den politiska debatten i någon betydande mån. Dessa yttranden kan tänkas innefatta de mest subversiva, kränkande, skadliga eller farliga åsikter – som likväl bör tolereras. Till grund för detta kan läggas två skilda anledningar.

När Forsskål skriver att skriffriheten är ”Regeringens tryggaste försvar” är det inte långsökt att associera till ett annat av hans uttalanden:

Tvenne utvägar äro att avböja skadliga följder av missnöje. Den ena fordrar bläck, den andra blod.<sup>54</sup>

---

<sup>51</sup> Forsskål, *Tankar om borgerliga friheten*, s. 10.

<sup>52</sup> Funcke, *Offentlighetsprincipen – praktik och teori*, s. 181–182.

<sup>53</sup> Barendt, *Freedom of Speech*, s. 18.

<sup>54</sup> Forsskål, *Tankar om borgerliga friheten*, s. 27.

Tanken är att förbud, förpassning och stigmatisering av åsikter, oavsett hur avskyvärda de är, kan vålla större skada än nytta. Att tolerera åsiktsyttringar, att låta dem luftas i offentligheten, är bättre än att fördriva dem till samhällets dunklare vrår där de riskerar att radikaliserar och i förlängningen ta sig till uttryck i våldsamma missnöjesyttringar.

Ett annorlunda toleransargument, mindre illavarslande, har framförts av Tomas Bull inom ramen för den senaste tidens debatt kring hetsbrottet. Bull menar att det finns en pedagogisk funktion i att uppvisa tolerans även mot det förkastliga. Toleransen är ett grundläggande värde i det demokratiska samhället, men toleransen är också självuppfyllande. Då samhället uppvisar tolerans också mot det förkastliga – och därmed en motvilja att ta till intoleransens egna medel – förstärks känslan av empati och tålamod hos medborgarna.<sup>55</sup> Att ge utrymme också åt de yttranden som varken är nödvändiga eller önskvärda har således en viktig fördel. Tolerans föder tolerans.

### 3.3 Nu som då för alltid?

Den mest övergripande och centrala slutsatsen som kan formuleras, med hänsyn till den historisk-filosofiska exkursionen ovan, kan verka ganska banal. Men den är värd att betona i sammanhanget: yttrande- och informationsfriheten är ett pågående arbete. Det finns ingen rimlig anledning att anta att den nuvarande juridiska konstruktionen som skyddar yttrandefriheten är vare sig den rätta, slutgiltiga eller absoluta. För att parafrasera Mill: det vore naivt att anta att vi inte en dag riskerar att se tillbaka på dagens yttrandefrihetslagstiftning som lika otillfredsställande och primitiv som vi nu betraktar den engelska yttrandefrihetsregimen i slutet av 1600-talet.

Med det sagt kan vi dra ytterligare ett antal intressanta slutsatser av historien. De första gäller de inskränkningar i yttrandefriheten som manifesterats genom tiden. Inledningsvis kan konstateras att de mest påtagliga inskränkningarna i yttrandefriheten har härstammat från nationalstaten. Detta gäller dock inte utan reservationer. Den första maktenheten att motarbeta spridningen av tryckta skrifter var inte någon stat, utan den katolska kyrkan, i syfte att hindra kätteri och moraliskt förfall. Därtill finns många exempel på fall där privata aktörer med makt över informationsdistributionen, som medel för statsintresset, mer eller mindre frivilligt har låtit begränsa yttrande- och informationsfriheten.

---

<sup>55</sup> Bull, *På yttrandefrihetens yttersta gren*, JT 2005/06 s. 534–535.

I forna Östtyskland utövades till exempel statlig censur av privata bokförlag, maskerat som sedvanlig förlagsgranskning.<sup>56</sup> Denna form av ”outsourcad” censur har ofta varit än mer inskränkande och förrädisk än den som utövats öppet av staten själv. Privata aktörer har alltså spelat en både direkt och indirekt roll i censuren och yttrandefrihetsinskränkningarna.

Vidare kan konstateras att det inte alltid har varit direkta censurlagar eller föreskrifter om yttrandefrihetsbrott som inverkat på den praktiska yttrandefriheten. Makthavare har ofta varit mycket kreativa vad gäller att begränsa det fria ordet genom mer subtila styrmedel. Här kan den ekonomiska censuren i England lyftas, där stämpelskatter, importskatter och borgenskrav på förläggare nyttjades flitigt för att begränsa den faktiska pressfriheten. Och bland de många halvdolda censuråtgärder som genomfördes i Frankrike under Napoleons tid kan nämnas förbudet mot författaranonymitet, som verkade avkylande på eventuella oppositionella krafter.<sup>57</sup> Indirekt reglering har alltså spelat en stor roll för att begränsa yttrandefriheten.

Vad gäller lagen kan nämnas att yttrandefrihetsregleringens formativa period utspelade sig över en tid då tryckpressen var det huvudsakliga medlet för att sprida uppgifter. Detta förklarar att en stor del av regleringen, som ofta lever kvar än idag, är utformad med hänsyn till ett visst medium: den tryckta skriften. Men den motivering som uppbär yttrandefrihetslagstiftningen, som formulerats av Milton, Locke, Mill, Forsskål och andra jurister, filosofer och författare, inte är tids- eller mediebunden. Här kan det första tillägget i den amerikanska konstitutionen, som garanterar var och en yttrandefrihet, kontrasteras mot det tredje tillägget, som garanterar var och en frihet från påtvingad inkvartering av soldater i fredstid. Det senare intresset kan betraktas som obsolet i en modern kontext. Men intresset av individuell autonomi, sanning, progression, demokrati, insyn, maktgranskning och tolerans är konstanta värden. Därmed är också yttrandefriheten ett konstant värde, oaktat tid, kultur eller samhälle.

Slutligen kan betonas att det aldrig verkar ha rått konsensus kring yttrande- och informationsfrihetens verkliga innebörd. Även om det går att urskönja en ”hård kärna” i friheten, har dess omfång och gränser alltid varit föremål för diskussion. Medan vissa har ställt yttrandefriheten i relation till staten har andra, som Mill och Forsskål, påpekat att yttrandefriheten också gäller i förhållande till andra enskilda.

---

<sup>56</sup> Von Vegesack, *När ordet blev fritt*, s. 14.

<sup>57</sup> Von Vegesack, *När ordet blev fritt*, s. 69.

Medan vissa har betraktat yttrandefriheten som en negativ rättighet har andra, som Sieyès, hävdad att yttrandefriheten också ger upphov till positiva förpliktelser hos staten. Ytterligare diskussioner har förts kring yttrandefrihetens innebörd – huruvida yttrandefrihet enbart innebär frihet från censur, eller också en rätt att inte straffas för det som uttrycks. Vidare har yttrandefrihetens gränser diskuterats i det oändliga. Om yttrandefrihet antas innebära en rätt att uttrycka vilka uppfattningar som helst så länge inte andra skadas – hur ska skaderekvisitet kvalificeras? Räknas kränkningar och obehag som skada? Är fördomsfulla, subversiva och smaklösa yttranden skadliga? Dessa frågor debatteras än idag, och för att anknyta till den inledande poängen, får frågan om yttrandefrihetens natur och gränser anses utgöra ett pågående arbete.



## 4 Grindvakter och makthavare – informationslandskap i förändring

*Governments of the Industrial World, you weary giants of flesh and steel, I come from Cyberspace, the new home of Mind. On behalf of the future, I ask you of the past to leave us alone. You are not welcome among us. You have no sovereignty where we gather.*

John Perry Barlow, *A Declaration of the Independence of Cyberspace*

När den katolska kyrkan försökte hejda den spridning av avvikande uppfattningar som möjliggjorts genom tryckpressens frammarsch, visade det sig snart förgäves. Det var inte förrän reformationen som kyrkans anspråk kunde få stöd av den världsliga maktapparaten och därmed uppnå verklig effektivitet.<sup>58</sup> Detta har illustrerat hur yttrandefrihetsrestriktioner inte nödvändigtvis behöver härstamma från statliga anspråk. Men det kan också få illustrera ett annat faktum: att maktanspråk och kontroll inte alltid löper samman. Det kan framstå som ett banalt påstående, men det är ett bärande tema i yttrande- och informationsfrihetens narrativ. Makt förutsätter effektiv kontroll, och kontrollen över vad som tryckts, sagts, skrivits, förmedlats, och distribuerats har ofta tillkommit icke-statliga aktörer. Det kan förklara varför staten ofta har valt att ingripa mot *mellanhänderna* i en informationskedja – boktryckare, bokhandlare, redaktioner, och distributörer – istället för de enskilda avsändarna och mottagarna av informationen.

De aktörer som haft faktisk kontroll över informationsdistributionen i samhället har skiftat i takt med den tekniska utvecklingen. Tidigt i historien var skriv- och läskunniga klostermunkar, statstjänstemän och universitetslärda de som hade förmågan att varaktigt förmedla information genom handtecknad skrift. Vid tryckpressens tillkomst möjliggjordes för första gången massdistribution av information. Men distributionen i sig kontrollerades ännu av det relativa fåtal som kunde betala för den kostsamma tryckpressen: förutom staten rörde det sig väsentligen om universiteten, nyhetspressen och den rikare klassen. Med den elektroniska kommunikationens framväxt skedde ytterligare en landvinning. Information kunde nu distribueras blixtnabbt över långa avstånd. Men kontrollen över informationsdistributionen låg fortfarande hos de aktörer som kunde bekosta tv- och radionätverken och samtidigt beviljas licenser för sina sändningar, på marknader som ofta var helt eller delvis monopoliserade.<sup>59</sup>

---

<sup>58</sup> Jfr ovan avsnitt 3.1.

<sup>59</sup> Weaver, *Information and Law in Transition* s. 24–33.

Utöver staten har det alltid funnits vissa mellanhänder i informationskedjan som, på grund av sin ställning, kunskap eller förmögenhet, kunnat kontrollera kommunikationen och informationsdistributionen i samhället. Gemensamt har dessa mellanhänder kallats för informationssamhällets ”grindvakter”.

I det följande avsnittet, som syftar till att ge en regulatorisk bakgrund till det problem som snart kommer att presenteras, riktar vi blicken mot dessa grindvakter. Vi undersöker olika typer av grindvakter i informationssamhället, och förhållandet mellan grindvakterna och lagen. I fokus ligger det ansvar och de skyldigheter som ålagts olika mellanhänder i informationssamhället, i förhållande till deras användare. Här jämförs regleringen i Sverige och EU med den i USA, där liknande överväganden har gjorts under de senaste 30 åren. Slutligen kompliceras bilden, då vi introducerar en ny typ av grindvakt i informationssamhället – de sociala medierna.

#### 4.1 Grindvakter och mellanhänder

Termen ”grindvakt”, eller *gatekeeper*, härstammar från socialpsykologin, där det myntades av Kurt Lewin 1943 för att illustrera hur amerikanska hemmafruar kontrollerade familjens födointag. Mat, liksom information, rör sig inte av egen kraft. Istället slussas den genom olika kanaler på sin väg från producenten till slutmottagaren – från åkern till fabriken, från fabriken till affären, från affären till skafferiet, och från skafferiet till middagsbordet. Den som råder över de olika ”kanalerna” har kontrollen över vad som slutligen når fram till konsumenten. I Lewins fall var det hemmafrun som utövade denna kontroll. Hon handlade maten, lagrade den, tillredde den och serverade den på middagsbordet. Hon var en ”grindvakt” i konsumtionskedjan.<sup>60</sup>

Lewins grindvaksallegori fick snart fäste inom andra discipliner. Inom journalistik och kommunikation är grindvakter individer i redaktörsställning som kurerar nyhetsutbudet till morgontidningen, och i finansindustrin kan kreditinstitut och investmentbanker, på grund av sitt marknadsanseende, utgöra grindvakter i diverse kapitalflöden.<sup>61</sup> Inom informationsvetenskapen är grindvakterna de aktörer som besitter kontrollen över informationsflödena i en informationskedja. De kan generellt beskrivas som ”those who carry, host, or index others' content”.<sup>62</sup>

---

<sup>60</sup> Lewin, *The Problem of Changing Food Habits: Report of the Committee on Food Habits* s. 35–37.

<sup>61</sup> Laidlaw, *International Review of Law, Computers & Technology* 2010 s. 264–265.

<sup>62</sup> Zittrain, *Harvard Journal of Law & Technology* 2006 s. 253.

Allegoriskt sett utför de alltså samma typ av arbete som Lewins amerikanska hemmafru när hon bär maten, sorterar den, lagrar den, behandlar den och presenterar den för familjen.

Traditionellt har det emellertid gjorts en distinktion mellan två typer av grindvakter i informationslandskapet: sådana redaktionsliknande aktörer som förmedlar och kontrollerar tillgången till information, och aktörer som råder över de resurser och tjänster som används för att transportera information och koppla ihop avsändare och mottagare i ett informationsnätverk.<sup>63</sup>

Ett exempel på den tidigare typen vore redaktören på en nyhetswebb, eller utgivaren på en streamingtjänst, som väljer ut och ordnar det tillgängliga informationsutbudet. För tydlighetens skull kan vi kalla dessa grindvakter för ”redaktionella mellanhänder”. Ett exempel på den senare kategorin vore istället internetleverantören, som tillhandahåller den nödvändiga infrastrukturen för kommunikation över internet. Dessa aktörer kan vi kalla för ”infrastrukturella mellanhänder”.

## 4.2 Grindvakter och lagen

Distinktionen mellan infrastrukturella mellanhänder och redaktionella mellanhänder kan anas också i lagstiftningssammanhang. I common law-länder som USA och Storbritannien har tillhandahållare av allmänna kommunikationsnät och infrastruktur ofta klassificerats som ”common carriers”. Tidigare har beteckningen avsett bland annat järnvägsbolag och telegrafoperatörer, men på senare tid har också internetleverantörer kommit att omfattas av klassificeringen.<sup>64</sup> Karaktäristiskt för ”common carriers” är att de lyder under en kvasi-offentligrättslig skyldighet att opartiskt och utan diskriminering leverera sådana varor eller meddelanden som lagligen anförtrotts dem för transport.<sup>65</sup> Internetleverantörer har således ansetts lyda under en ”neutralitetsprincip” – en skyldighet att inte intervensera i informationsflödena som de kontrollerar.<sup>66</sup>

---

<sup>63</sup> Helberger et al., *Info: The Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media* 2015 s. 52–53.

<sup>64</sup> Under senare år har flera omklassificeringar skett avseende internetleverantörers ställning som ”common carriers”. Diskussionerna har varit turbulenta och präglade av starka politiska motsättningar. Sedan 2017 anses internetleverantörer emellertid inte längre skyldiga att iakttä särskilt ”common carrier”-ansvar. Se vidare *Verizon v. FCC*, FCC Open Internet Order (2015) p. 39 och 59 och FCC Restoring Internet Freedom Order (2017) p. 1–4.

<sup>65</sup> Marsden, *Network neutrality: from policy to law to regulation*, s. 17.

<sup>66</sup> För ytterligare diskussion kring konceptet ”nätneutralitet”, se Wu, *Journal of Telecommunications and High Technology Law* 2003, Ubena, *How to Regulate Information and Communications*

Ett bärande skäl bakom denna neutralitetsprincip är att främja det öppna internet, och därigenom medborgarnas yttrande- och informationsfrihet.<sup>67</sup> Det vore yttrandefrihetsvidrigt om de stora internetleverantörerna tilläts blockera information som de av olika skäl ogillade. En motsvarande neutralitetsprincip finns också inom EU-rätten, som likt sin amerikanska förlaga kräver att internetleverantörer "[...] ska behandla all trafik likvärdigt, utan diskriminering, begränsningar eller störanden [*sic*], och oberoende av sändare och mottagare, det innehåll användarna tar del av eller distribuerar [...]". Även här var yttrande- och informationsfriheten ett tungt vägande skäl bakom regleringen.<sup>68</sup>

I svensk nationell rätt finns en dylik princip. Vid stiftandet av 1949 års TF infördes en bestämmelse i grundlagen som innebar att tillståndspliktiga postbefordringsföretag och andra allmänna kommunikationstjänster, såväl statliga som privata, inte fick förhindra spridningen av en viss skrift på grund av skriftens innehåll. En motsvarande "distributionsplikt" infördes senare i YGL avseende tekniska upptagningar. Distributionsplikten infördes för att förhindra att även icke-statliga aktörer, på den mycket konkurrenssvaga post- och telemarknaden, skulle få för stora möjligheter att stävja informationsdistributionen i samhället.<sup>69</sup>

Neutralitetsprincipen har ofta förenats med en idé om att de infrastrukturella mellanhänderna ska ha ett begränsat ansvar för det innehåll som tredjeparter publicerar genom deras kanaler.<sup>70</sup> Detta kan illustreras med det amerikanska CompuServe-fallet:

CompuServe var en internet- och tjänsteleverantör som gav sina kunder internetaccess till hundratals elektroniska bibliotek, forum och databaser. På ett av dessa forum, ett intresseforum för journalistik, publicerades ett nyhetsbrev som innehöll nedsättande uppgifter om en viss Mr. Blanchard. Varken forumet eller nyhetsbrevet administrerades av CompuServe, utan av externa företag. Likväl valde Blanchard att stämma CompuServe för ärekränkningen (*libel*). Domstolen konstaterade att CompuServe var att betrakta som en distributör i förhållande till innehållet på forumet, på samma sätt som ett bibliotek eller en tidningskiosk. Till skillnad från en publicist skulle internetleverantören därför hållas ansvarig för innehållet på forumet endast om de haft vetskap om, eller anledning att känna till, de kränkande uppgifterna. Om en distributör ålades ett lika strikt ansvar som en publicist – resonerade domstolen – skulle det inverka avkylande på yttrandefriheten.<sup>71</sup>

---

*Technology? A Jurisprudential Inquiry into Legislative and Regulatory Techniques* och Marsden, *Network neutrality: from policy to law to regulation*.

<sup>67</sup> FCC, Open Internet Order (2010) p. 4, 23, 50, 53, 75–78, 94–95 samt FCC, Open Internet Order (2015) p. 1, 77, 545, 554–555.

<sup>68</sup> Artikel 3.3 och skäl 1, 3, 6 och 33 TSM-förordningen.

<sup>69</sup> TF 6 kap. 4 §, YGL 3 kap. 15 § och SOU 2012:55 s. 339–353. Distributionsplikten omfattar dock inte internetleverantörer. Notera lagrummens reviderade lydelse från den 1 januari 2019.

<sup>70</sup> Jfr Thompson, *Vanderbilt Journal of Entertainment and Technology Law* 2016 s. 784–785.

<sup>71</sup> *Cubby v. CompuServe*.

Den rättsliga ansvarsdistinktionen mellan distributörer ("passive conduites" eller "deliverers") och publicister i CompuServe har upprätthållits i senare amerikansk praxis.<sup>72</sup> Den har sedermera blivit känd som "the content/conduit distinction".<sup>73</sup>

En liknande ansvarskonstruktion finns inom EU-rätten, där mellanhänders ansvar regleras genom e-handelsdirektivet (2000/31/EC). Enligt direktivets artikel 12 kan mellanhänder inte åläggas ansvar för information som tredjeparter publicerar genom deras kanaler, så länge de enbart sysslar med passiv överföring av innehållet ("mere conduit"). Bestämmelsen omfattar alltså internetleverantörer, som typiskt sett inte ändrar eller intervenerar i den information som de vidarebefordrar genom sina nät. Liknande ansvarsbegränsningar finns i artikel 13 och 14 för caching- och hosting-tjänster. Den kritiska förutsättningen för ansvarsfrihet är att mellanhanden förblir passiv och inte utövar någon form av redaktionell kontroll över informationsflödena,<sup>74</sup> genom att exempelvis redigera, arrangera eller optimera information från tredjeparter.<sup>75</sup>

E-handelsdirektivets ansvarsbegränsningar har implementerats i svensk rätt genom 16–18 §§ e-handelslagen. I likhet med direktivets föreskrifter innebär ansvarskonstruktionen att internetleverantörer, såväl som caching- och hosting-tjänster, inte kan bli skyldiga att betala skadestånd eller sanktionsavgifter på grund av innehållet i den information som de befattar sig med – åtminstone i den mån de inte känner till att informationen är lagstridig. Denna ansvarskonstruktion skiljer sig väsentligt från den som gäller för publicister enligt svensk rätt. Enligt både TF och YGL har den ansvarige utgivaren formalansvar för allt innehåll som publiceras i exempelvis en periodisk skrift, webbtidning eller i ett radio- eller tv-program.<sup>76</sup> Utgivaren anses ha kännedom om all information som publiceras under dennes kontroll, vare sig så är fallet eller inte, och kan bara undgå ansvar för olagligt innehåll under "extrema omständigheter".<sup>77</sup>

---

<sup>72</sup> *Auvil v. CBS 60 Minutes* (gällande lokala tv-stationer) och *Stratton Oakmont v. Prodigy* (gällande en elektronisk anslagstavla). Ansvarsbegränsningen för mellanhänder enligt amerikansk lag är förvisso inte helt konsekvent. Enligt Communications Decency Act, som reglerar mellanhandsansvar i allmänhet, råder i princip immunitet för mellanhänder i avsaknad av vållande (*fault*). Enligt Digital Millennium Copyright Act, som reglerar mellanhandsansvar specifikt vid immaterialrättsliga intrång, gäller ansvarsimmunitet under förutsättning att internetleverantören svarar på klagomål från rättighetsinnehavaren. Se Savin, *EU Internet Law* s. 108–110.

<sup>73</sup> Gillespie, *Custodians of the internet* s. 31.

<sup>74</sup> Savin, *EU Internet Law* s. 112–117. Förutsättningen för ansvarsfrihet enligt direktivets preambel (skäl 42) är att mellanhanden "varken har kännedom om eller kontroll över vidarebefordrad eller lagrad information" (egen kursivering). Begreppet "kontroll" får i detta sammanhang förstås i snävare bemärkelse än ovan (avsnitt 4.1).

<sup>75</sup> C-236/08–C-238/08 *Google France mot Louis Vuitton* och C-324/09 *L'Oréal mot eBay*.

<sup>76</sup> Jfr TF 8 kap. 1 § samt YGL 6 kap. 1 §.

<sup>77</sup> Jfr TF 8 kap. 14 § och YGL 6 kap. 9 §. Se även NJA 1991 s. 155.

Så här långt går det alltså att uppfatta en principiell skillnad i hur lagstiftare, både nationellt och internationellt, har betraktat infrastrukturella respektive redaktionella mellanhänder i informationssamhället. De infrastrukturella mellanhänderna har, till skillnad från redaktionella mellanhänder, i vissa sammanhang ålagts en *skyldighet* att förhålla sig neutralt till information från användare och tredjeparter som överförs genom deras kanaler. De infrastrukturella mellanhänderna har också åtnjutit ett begränsat *ansvar* i förhållande till den information som användare och tredjeparter publicerar genom deras kanaler, vilket står i stark kontrast mot den strikta ansvarskonstruktion som gäller för redaktionella mellanhänder. Båda regleringslösningar har motiverats av intresset för yttrande- och informationsfriheten. Schematiskt kan regleringslösningarna ställas upp på följande vis:

	Infrastrukturella mellanhänder (internetleverantörer, distributörer)	Redaktionella mellanhänder (redaktörer, utgivare)
Skyldighet att iaktta neutralitet i förhållande till information från användare/tredjeparter	✓	—
Begränsat ansvar för information som tillhandahålls av användare/tredjeparter	✓	—

*Fig. 1. Regleringen av infrastrukturella och redaktionella mellanhänder i Europa och USA, schablon.*

Tabellen ovan är ingen heltäckande beskrivning av gällande rätt, och kan givetvis nyanseras och ifrågasättas på olika grunder.<sup>78</sup> Den avser endast utgöra en skissartad illustration över hur lagstiftare i EU och USA har betraktat vissa typer av mellanhänder i informationssamhället, som ovan beskrivits som ”infrastrukturella” och ”redaktionella” grindvakter. Av de potentiella invändningar som kan riktas mot uppställningen ovan skulle de starkaste dock röra giltigheten i den skarpa distinktionen. Flera av de rättsliga källor som skapat dikotomier mellan ”infrastrukturella” och ”redaktionella” mellanhänder, ”aktiva” och ”passiva” överförare och ”distributörer” och ”publicister” är snart ett eller flera decennier gamla – och teknologiska dikotomier tenderar att bli övermogna i allt snabbare takt.

<sup>78</sup> Se exempelvis not 64 och 72 ovan.

### 4.3 Nya grindvakter i informationssamhället

Under slutet av 1900-talet och början av 2000-talet, samtidigt som lagstiftare runt om i världen mobiliserade för att åstadkomma en hållbar regleringslösning för de infrastrukturella och redaktionella mellanhänderna, började en ny typ av mellanhänder växa fram på internet. 2004 lanserade Harvardstudenten Mark Zuckerberg en virtuell bildmatrikel för universitetsstudenter under domännamnet "thefacebook.com".<sup>79</sup> Samma år introducerade Jack Dorsey mikrobloggen "twtr",<sup>80</sup> och året innan, den 23 april 2005, hade datavetaren Jawed Karim publicerat ett 19 sekunder långt videoklipp där han kommenterar elefanterna på San Diego Zoo. Videoklippen "Me at the Zoo" blev det första i YouTubes historia.<sup>81</sup>

Under de år som följde upplevde de sociala medierna ett explosionsartat uppsving. En mer prosaisk beskrivning av mediernas utveckling ligger utanför ramen för den här framställningen, men resultaten talar för sig själva. Under tredje kvartalet 2018 hade Facebook 2,27 miljarder aktiva användare per månad.<sup>82</sup> Siffran överträffar antalet invånare i världens största stat, Kina, med nästan en miljard. På Twitter skickas mer än en halv miljard "tweets" per dag, motsvarande över 5 500 per sekund.<sup>83</sup> Och varje dag strömmas mer än en miljard timmar video genom YouTube, i 91 olika länder runtom i världen.<sup>84</sup> Varje superlativ blir nästintill banalt för att beskriva de sociala mediernas betydelse för kommunikation, nyheter, nöje och opinionsbildning världen över.

Utvecklingen inom digital informationsdistribution har också medfört en komplicerad i den traditionella teoribildningen kring mellanhänder och grindvakter i informationskedjan. Dels är onlinekontexten generativ. Medieanvändare förväntas inte längre bara vara passiva konsumenterna av information, utan också producera information, vilket grumlar den typiska distinktionen mellan avsändare-mottagare i en informationskedja. Dessutom har relationen mellan olika avsändare, mellanhänder och mottagare blivit mer dynamisk och mångfacetterad. Den som driver en blogg med ett modererat kommentarsfält kan till exempel anses utgöra en sorts redaktionell mellanhand i förhållande till sina följare, likt en traditionell tidningsredaktör.

---

<sup>79</sup> Burns, *Social Media – A Reference Handbook*, s. 17–19 och Houston, *The Facebook story: from inception to IPO*, The Verge 18/5 2012.

<sup>80</sup> Burns, *Social Media – A Reference Handbook*, s. 25–27 och Abell, *March 21, 2006: Twitter Takes Flight*, Wired 21/3 2011.

<sup>81</sup> Snickars & Vonderau, *The Youtube Reader*, s. 9.

<sup>82</sup> Facebook, *Company Info*.

<sup>83</sup> Twitter, *Business*.

<sup>84</sup> YouTube, *YouTube för pressen*.

Samtidigt kan bloggaren vara bunden av föreskrifter, användaravtal och restriktioner från innehavaren av bloggplattformen, som i sin tur agerar mellanhand gentemot bloggaren. Dessutom är både bloggaren och bloggplattformen beroende av den infrastruktur som tillhandahålls av deras internetleverantörer, som på så sätt agerar infrastrukturell mellanhand i förhållande till bägge aktörer.<sup>85</sup>

Slutligen finns det skäl att frågå den traditionella dikotomin mellan ”infrastrukturella” och ”redaktionella” mellanhänder i fråga om de sociala medierna. En social medieplattform, som Facebook, består av mjukvara och tjänster som kanaliserar information genom olika nätverksprotokoll. Plattformarna utgör alltså en infrastruktur för informationsöverföring mellan olika tredjeparter, eller slutanvändare. Kanaliseringen, som vi snart ska undersöka närmare, är emellertid inte neutral. De tillämpade protokollen formar informationsutbudet, och utformningen av användargränssnitt och funktioner jämkar dessutom användarnas handlingsutrymme och beteende.<sup>86</sup>

Informationsutbudet på sociala medier påverkas också på mer direkta sätt. En användares innehållsexponering styrs ofta av algoritmer som skräddarsyr användarupplevelsen. Det kan ske genom personaliseringsfilter, rekommendationsfunktioner eller sortering av innehåll så att viss information premieras framför annan. Informationsutbudet styrs också av de restriktioner, villkor och policyer som plattformen tillämpar för användargenererat material. Sorteringen och begränsningen av användargenererat innehåll innebär att de sociala medierna, i olika stor utsträckning, vidtar aktiva åtgärder för att kurera informationsutbudet gentemot användarna.

Som mellanhand i informationssamhället får sociala medier anses ha en roll som varken överensstämmer med den passiva, infrastrukturella mellanhanden eller den aktiva, redaktionella mellanhanden. De är både och, eller mitt emellan.

---

<sup>85</sup> Laidlaw, *International Review of Law, Computers & Technology* 2010 s. 265–266.

<sup>86</sup> Van Dijk, *Convergence: The International Journal of Research into New Media Technologies* 2012 s. 148.



	Infrastrukturella mellanhänder	Sociala medier (Facebook, Twitter, etc.)	Redaktionella mellanhänder
Skyldighet att iaktta neutralitet i förhållande till information från användare/tredjeparter	✓		–
Begränsat ansvar för information som tillhandahålls av användare/tredjeparter	✓		–

Fig. 2. Regleringen av infrastrukturella och redaktionella mellanhänder jämfört med sociala medier.

#### 4.4 Nya medier, nya remedier?

Den traditionella skiljelinjen mellan ”infrastrukturella” och ”redaktionella” mellanhänder verkar alltså ha förlorat sin udd i ljuset av den generativa informationsteknologins framväxt. Det har uppstått nya ”grindvakter” i informationssamhället, som både erbjuder en kommunikativ infrastruktur för slutanvändare i form av mjukvara och tjänster, och samtidigt kurerar användarnas tillgång till information. Dessa fakta inbjuder till ny diskussion kring regleringen av informationssamhällets mellanhänder. Om de tidigare regleringsmodellerna i princip korresponderade med den klassiska indelningen av mellanhänder, är det förstås angeläget att överväga lämpliga regleringsmodeller också för de ”hybridmellanhänder” som uppstått under de senaste femton åren.

Så långt har också den rättsliga diskursen nått. Under senare tid har det rättsliga ansvaret för sociala medier som Facebook, Twitter och YouTube diskuterats flitigt bland lagstiftare runtom i världen. I Sverige har det gjorts anspråk på att skärpa sociala mediers ansvar för att hindra förekomsten av ”näthat” och andra allvarliga integritetskränkningar.<sup>87</sup> Sedan 2016 har det också pågått intensiva diskussioner inom EU kring antagandet av ett nytt upphovsrättsdirektiv, som skulle ålägga sociala medier ett mycket större ansvar för att identifiera och förekomma användares upphovsrätts-intrång.<sup>88</sup> I USA har ny antitrafickinglagstiftning gett upphov till inskränkningar i den ansvarsbegränsning som annars gäller för mellanhänder på internet.<sup>89</sup>

<sup>87</sup> Se exempelvis SOU 2016:7 s. 498–523.

<sup>88</sup> Artikel 13 och skäl 37–38, *Förslag till Europaparlamentets och rådets direktiv om upphovsrätt på den digitala inre marknaden* (COM(2016) 593 final).

<sup>89</sup> FOSTA-SESTA. Jfr ovan avsnitt 4.2.

Den rättsliga diskursen kring regleringen av sociala medier har dock tenderat att fastna vid denna ena aspekt av regleringsproblemet: de sociala mediernas ansvar för användarnas förehavanden.<sup>90</sup> Mindre vikt har lagts vid den andra änden av problemet: de sociala mediernas skyldigheter gentemot användarna, då de nyttjar plattformarna för att distribuera och ta del av information.<sup>91</sup> De intressen som burit upp mycket av den reglering som gäller för infrastrukturella mellanhänder – intresset av en bred yttrande- och informationsfrihet – har därigenom framstått som underprioriterat i lagstiftnings-sammanhang. Det har saknats ett användarcentriskt perspektiv i regleringsdebatten kring sociala medier,<sup>92</sup> och sällan framförs invändningar mot den redaktionella verksamhet som plattformarna bedriver.

Frågan om lämpligheten i detta förhållningssätt, och frågan om de sociala mediernas skyldigheter i yttrande- och informationsfrihetshänseende, återkommer vi till senare i framställningen. Dessförinnan ska undersökas vilken typ av kontroll som de sociala medierna faktiskt utövar i ”den nya offentligheten”, och hur den tar sig i uttryck.

---

<sup>90</sup> Representerat ovan som den mellersta undre raden i Fig. 2.

<sup>91</sup> Representerat ovan som den mellersta övre raden i Fig. 2.

<sup>92</sup> Jfr Barzilai-Nahon, *Journal of the American Society for Information Science and Technology* 2008 s. 1494–1495, Laidlaw, *International Review of Law, Computers & Technology* 2010 s. 267, Helberger et al., *Info: The Journal of Policy, Regulation and Strategy for Telecommunications* 2015 s. 63–64.

## 5 Villkor och vedergällning – om sociala mediers moderering

*[A] lot of the most sensitive issues that we faced today are conflicts between our real values, right? Freedom of speech and hate speech and offensive content. Where is the line, right? [...] I feel fundamentally uncomfortable sitting here in California at an office, making content policy decisions for people around the world. I have to, because [I lead Facebook], but I'd rather not.*

Mark Zuckerberg, grundare av Facebook<sup>93</sup>

Hittills har vi kunnat bevittna hur såväl statliga som icke-statliga entiteter har genererat inskränkningar i yttrande- och informationsfriheten. Hur dessa ingrepp kommit till uttryck och hur effektiva de har varit har berott på de bakomliggande intressena, men också den bakomliggande aktörens kontrollsfär. Förhållandet mellan regleringsintresse och kontroll har inneburit att stater inte sällan reglerat yttrande- och informationsfriheten genom inflytelserika mellanhänder i informationskedjorna – informations-samhällets grindvakter.

Vilka skyldigheter och rättigheter som åvilat dessa aktörer har typiskt sett berott på deras kontrollställning, roll och funktion i informationssamhället. ”Infrastrukturella mellanhänder”, som internetleverantörer, har behandlats ur ett mer användarcentriskt perspektiv. De har anförtrotts ansvarsfrihet för tredjeparters lagöverträdelser, men också ålagts vissa skyldigheter gentemot användarna. De ”redaktionella mellanhänderna”, som tidningsredaktioner, har istället betraktats som mer autonoma och mindre ansvarsfria. Någonstans mellan dessa traditionella stereotyper ligger de sociala medierna, som bär drag av bådadera.

I det följande avsnittet kommer vi att granska de sociala medierna och den informationskontroll som de utövar på sina plattformar: de sociala mediernas innehållsmoderering. Vi kommer att se närmare på plattformarnas innehållspolicyer, de sorteringsalgoritmer som formar informationsutbudet på plattformarna, och de sanktioner som plattformarna tillämpar vid regelöverträdelser. Beskrivningen konkretiserar de sociala mediernas grindväktarfunktion, men utgör också grunden till nästföljande avsnitt, där innehållsmodereringen ställs i relation till yttrande- och informationsfriheten.

---

<sup>93</sup> Swisher & Wager, *Here's the transcript of Recode's interview with Facebook CEO Mark Zuckerberg about the Cambridge Analytica controversy and more*, Recode 22/3 2018.

## 5.1 Sociala medier

Liksom internetleverantörer och tidningar är de sociala medierna värddjur för mellanmänsklig kommunikation. De erbjuder datormedierad kommunikation i form av en interaktiv, digital plattform där användarna förväntas representera sig själva, knyta kontakter med andra och generera innehåll.<sup>94</sup> Till skillnad från internetleverantören erbjuder Facebook således en plattform, och till skillnad från dagstidningen kan användarna producera och distribuera eget innehåll i sina nätverk. Sociala medier är det nya offentliga rummet i det digitala samhället:

Over the last 15 years, Facebook and Instagram have helped people connect with friends, communities, and interests in the digital equivalent of a town square.<sup>95</sup>

Tjänsterna tjänar flera funktioner. De sociala medierna möjliggör för användaren att söka och ta del av information från andra, och samtidigt dela med sig av sina egna åsikter, värderingar och erfarenheter. De katalyserar identitetsbyggande genom att göra användaren medveten om och låta manifesteras sina egna preferenser och reaktioner. De underbygger mellanmänskliga relationer genom att påbjuda interaktion mellan gamla och nya bekantskaper. Vissa tjänster syftar också till att underlätta arbete (som arbetsnätverket LinkedIn och nätverksapplikationen Slack) eller tillhandahålla underhållning (som videoplattformen YouTube eller bloggportalen Tumblr).<sup>96</sup> Därtill utgör varje socialt medium ett eget funktionellt mikrokosmos. De har med tiden och med användarkollektivens insatser utvecklat en särpräglad diskurs och förväntade användarbeteenden. På så sätt har Twitter blivit en nyhets- och opinionsbildningsportal, medan Instagram har fått en livsstils- och underhållningstygdpunkt.

I samtliga fall tillämpar plattformarna någon form av innehållsmoderering. Alla sociala medier sållar, kurerar, arrangerar, raderar, ordnar och sammanställer det användargenererade innehållet utefter olika principer och intressen. Modereringen kan bero på en genuin vilja att skapa en etiskt eller estetiskt hållbar tjänst. Den styrs av ekonomiska imperativ – ett intresse av att inte försaka användarnas trivsel och förtroende, att inte skada det egna varumärket eller mota bort annonsörer på grund av olämpligt eller otrevligt innehåll.

---

<sup>94</sup> Manning, *Encyclopedia of Social Media and Politics* s. 1158–1161, Lin & Lu, *Computers in Human Behavior* 2011 s. 1152, Boyd & Ellison, *Journal of Computer-Mediated Communication* 2001 s. 211–214.

<sup>95</sup> Zuckerberg, *A Privacy-Focused Vision for Social Networking*.

<sup>96</sup> Manning, *Encyclopedia of Social Media and Politics* s. 1158–1161.

Den kan också botten i juridiska incitament – som ovan diskuterat, en vilja att undgå legalt ansvar.<sup>97</sup> Det är modereringen som gör de sociala medierna till grindvakter i informationssamhället. Långt ifrån perifert, har det till och med hävdats att modereringen utgör en del av de sociala mediernas kärnverksamhet:

[M]oderation is, in many ways, the commodity that platforms offer. Though part of the web, social media platforms promise to rise above it, by offering a better experience of all this information and sociality: curated, organized, archived, and moderated.<sup>98</sup>

I praktiken sker denna moderering på olika sätt. Synligast kommer den till uttryck genom plattformens användaravtal, riktlinjer eller policyer. Men en stor portion sker på sätt som är mindre uttalade: genom innehålls algoritmer, automatiserade rapporterings-system och dunkla compliance-mekanismer.<sup>99</sup> All denna moderering sätter gränser för informationsdistributionen på de tjänster som nu kommit att bli några av världens största kommunikationsredskap.

## 5.2 Användarvillkor, riktlinjer och policy

Den sociala nätverkstjänsten utgår typiskt sett från att en användare skapar en profil för att publicera innehåll interagera med andra.<sup>100</sup> Antingen representerar profilen användarens verkliga identitet (som på Facebook) eller så bygger profilen på ett alias (som på internetforumet Reddit). Användarprofilen skapas med förbehåll om att användaren förbinder sig att följa tjänstens användarvillkor. Användarvillkoren reglerar ofta förutsättningarna för att få använda tjänsten, förutsättningarna för att publicera innehåll genom tjänsten, plattformens rättigheter till användargenererat material, information om datahantering och ansvarsfriskrivningar.<sup>101</sup> Förutsättningarna för att publicera innehåll genom tjänsten kompletteras vanligtvis genom ett separat, bilagt avtal innehållandes ”community guidelines”, plattformens regler och policyer.<sup>102</sup> Dessa regler och riktlinjer reglerar vilket innehåll och vilken information som får distribueras på plattformen. De utgör således en utgångspunkt för innehållsmodereringen.

---

<sup>97</sup> Gillespie, *Custodians of the internet* s. 10.

<sup>98</sup> Gillespie, *Custodians of the internet* s. 13.

<sup>99</sup> Everett, *Kansas Journal of Law & Public Policy* 2018 s. 119–121.

<sup>100</sup> Manning, *Encyclopedia of Social Media and Politics* s. 1158–1161, Boyd & Ellison, *Journal of Computer-Mediated Communication* 2001 s. 211–214.

<sup>101</sup> Facebook, *Terms of Service*, Twitter, *Twitter Terms of Service*, YouTube, *Användarvillkor för Tjänsten*, Tumblr, *Terms of Service*, Instagram, *Användarvillkor*.

<sup>102</sup> Facebook, *Community Standards*, Twitter, *The Twitter Rules*, YouTube, *Riktlinjer för Communityn*, Tumblr, *Community Guidelines*, Instagram, *Communityregler*.

Riktlinjernas innehåll och utformning varierar mellan olika plattformar, men många villkor är gemensamma för de flesta större plattformarna.

### 5.2.1 *Hot, trakasserier, och kränkningar*

Alla användarriktlinjer förbjuder riktade, trovärdiga hot mot andra individer. Såväl direkta som indirekta hot om våld omfattas.<sup>103</sup> I Instagrams och Tumblrs riktlinjer nämns, utöver fysiskt våld, även ”hot om stöld, skadegörelse eller annan ekonomisk skada” som stridande mot användarriktlinjerna. Instagram nämner därtill ”hot om att publicera intima bilder på andra”.

Därtill förbjuds riktade trakasserier mot andra.<sup>104</sup> Facebook definierar trakasserier som ”oönskade och illvilliga kontakter”, men i andra riktlinjer framgår inte närmare hur begreppet ska tolkas. Oftare förekommer begreppet i samband med specifika former av trakasserier. Twitter och Tumblr nämner särskilt oönskade sexuella närmanden, ”sexuell misskötsamhet” och sexuell objektifiering av andra. Instagram tar upp delandet av ”personlig information som delas för att utöva utpressning eller trakassera någon” (så kallad ”doxing”). Både Facebook och YouTube har omfattande regelverk kring ”cybermobbing”.<sup>105</sup>

Riktlinjerna för otillåtna trakasserier är stundom mycket brett utformade, och går mot vad som snarare skulle kunna uppfattas som kränkande eller förolämpande innehåll. YouTube tolererar inte ”innehåll med sårande eller negativa kommentarer/videor om en annan person”, Instagrams riktlinjer omfattar ”innehåll riktat mot privatpersoner i syfte att nedvärdera eller kränka dem” och Facebook förbjuder ”riktade svordomar” mot individer eller kollektiv.

### 5.2.2 *”Hate-speech” och hets mot folkgrupp*

Utöver riktade hot, trakasserier och kränkningar mot enskilda användare innehåller användarvillkoren också förbud mot missaktande eller hatiskt tal mot vissa kollektiv. De skyddade kollektiven omfattar dels de folkgrupper som kommer till uttryck i hetslagstiftningen i 16 kap. 8 § brottsbalken, innefattandes ras eller etnicitet, nationellt ursprung, tro och religiös tillhörighet, sexuell läggning, könsidentitet, men även kön och funktionshinder eller annan allvarlig sjukdom.

---

<sup>103</sup> Facebook, *Community Standards*, II. 10. “Harassment”, Twitter, *The Twitter Rules*, “Abusive and hateful conduct”, YouTube, *Policy om trakasserier och cybermobbing*, Tumblr, *Community Guidelines*, “Violent Content and Threats, Gore and Mutilation” och “Harassment”, Instagram, *Communityregler*, “Respektera andra medlemmar i Instagram-communityn” och “Följ lagen”.

<sup>104</sup> Se not 103.

<sup>105</sup> Jfr Facebook, *Community Standards*, II. 9 “Bullying”.

Twitter, YouTube och Tumblr nämner även ålder, Facebook och YouTube listar ”immigrantstatus”, Tumblr och YouTube ”veteranstatus” och Facebook är ensamma om att också omnämna ”kast” som skyddat kollektiv.<sup>106</sup>

Vilka typer av yttranden som betraktas som ”hatiska” skiljer sig vida åt mellan plattformarna. Medan Instagram lakoniskt nämner ”hets mot folkgrupp” och Tumblr ”innehåll som syftar till att stödja eller framkalla hat eller avhumanisering”, har Facebook, Twitter och YouTube separata, detaljerade och exemplifierande policyer om hatiskt tal. Vanligen innefattas uttryck för önskan eller förhoppning om att medlemmarna ur ett skyddat kollektiv utsätts för våld, sjukdom eller annat missöde, reproducerandet av negativa stereotyper, avhumanisering och förmedlandet av hatfulla symboler – exempelvis svastikor.

Facebooks riktlinjer är de mest utarbetade. Där klassificeras hatiskt tal i tre klasser, från mest till minst allvarligt. Mest allvarligt anses stödjande av våld mot ett skyddat kollektiv eller avhumanisering, exempelvis genom att jämföra medlemmar av ett skyddat kollektiv med djur, bakterier, eller kriminella. I mellanklassen nämns mer subjektiva former av missaktning, som att medlemmar av ett skyddat kollektiv har negativa fysiska eller psykiska egenskaper, värdeomdömen som ”jag tycker inte om [medlemmar ur ett skyddat kollektiv]”, eller riktade svordomar och missaktande kommentarer. I den tredje klassen nämns ”anspråk på att exkludera eller segregera” medlemmar ur ett skyddat kollektiv (undantaget åsikter i immigrationspolitik).

Kollektiv	Facebook	Twitter	YouTube	Tumblr	Instagram
Ras/etnicitet	x	x	x	x	x
Nationalitet	x	x	x	x	x
Religion	x	x	x	x	x
Sexuell läggning	x	x	x	x	x
Könsidentitet	x	x		x	x
Kön	x	x	x	x	x
Funktionshinder och sjukdom	x	x	x	x	x
Ålder		x	x	x	
Immigrantstatus	x		x		
Veteranstatus			x	x	
Kast	x				

Fig. 3. Skyddade kollektiv enligt plattformarnas riktlinjer för hatiskt tal och hets mot folkgrupp.

<sup>106</sup> Facebook, *Community Standards*, III. 12 “Hate Speech”, Twitter, *Hateful conduct policy*, Youtube, *Policy om hatretorik*, Tumblr, *Community Guidelines*, “Hate Speech”, Instagram, *Communityregler*, “Respektera andra medlemmar i Instagram-communityn”.

### 5.2.3 Uppmuntran, glorifiering eller skildring av våld och annan brottslighet

Förbud mot att uppmuntra till våld och andra brott mot enskilda eller skyddade kollektiv framgår typiskt sett av plattformens riktlinjer mot hat, trakasserier eller hatiskt tal. Därtill finns vanligen riktlinjer som förbjuder uppvigling, uppmuntran, anstiftan eller erbjudande om att utföra brott.<sup>107</sup> Det innefattar ofta narkotikabrott, vapenbrott, prostitution och människohandel, samt andra reglerade varor och tjänster.

Därtill finns vanligen regler som mer generellt förbjuder glorifiering eller lovordande av visst beteende eller vissa organisationer. Alla plattformar förbjuder yttranden som stödjer, glorifierar, hyllar eller lovordar terrororganisationer, brottsyndikat, ”hatgrupper” eller andra ”farliga organisationer”.<sup>108</sup> Förbudet omfattar typiskt sett också lovordande av enstaka medlemmar ur en sådan organisation, och deras gärningar. Återigen har Facebook den mest rigorösa begreppsbildningen kring sådana grupper och individer, och definierar bland annat en kriminell organisation som ”en organisation med tre eller flera medlemmar som samlas under ett namn, en färg, handgest eller erkända indicium, som medverkar i eller hotar att medverka i kriminell aktivitet, såsom (men inte uteslutande) mord, narkotikasmuggling, vapensmuggling, identitetsstöld, pengatvätt, utpressning [...]”.

Plattformarna förbjuder därutöver skildringar av våld och vissa brott. Instagrams modereringspolicy omfattar enbart förbud mot skildring av grovt våld, men i övrigt omfattas typiskt sett skildringar av våld, tortyr och övergrepp samt döda, stympade eller svårt skadade människor och djur.<sup>109</sup> I vissa fall utsträcks förbuden till obscen, grafiskt eller chockerande material. Tumblr tillåter inte blodigt material (”gore”) och YouTube förbjuder ”video eller bilder som avbildar kroppsvätskor, som blod eller kräkningar”. I många fall är tröskeln för otillåtet material lägre när materialet skildrar underåriga.

---

<sup>107</sup> Jfr Facebook, *Community Standards*, I. 3 “Promoting or publishing crime”, I. 4 “Coordinating harm” och I. 5 “Regulated goods”, Twitter, *The Twitter Rules*, “Unlawful use”, Youtube, *Policyer om skadligt eller farligt innehåll, Försäljning av olagliga eller reglerade varor, Policyer för innehåll med skjutvapen*. Tumblr, *Community Guidelines*, ”Unlawful Uses or Content” och “Human Trafficking and Prostitution”, Instagram, *Communityregler*, “Följ lagen”.

<sup>108</sup> Facebook, *Community Standards*, I. 2 “Dangerous individuals and organizations”, Twitter, *Violent threats and glorification of violence, Violent extremist groups*, YouTube, *Brottsliga organisationer*, Tumblr, *Community Guidelines*, “Terrorism” och “Violent Content and Threats, Gore and Mutilation”, Instagram, *Communityregler*, “Följ lagen”.

<sup>109</sup> Facebook, *Community Standards*, III. 13 “Violence and graphic content”, Twitter, *Twitter media policy*, YouTube, *Policyer om våldsamt eller chockerande innehåll*, Tumblr, *Community Guidelines*, “Violent Content and Threats, Gore and Mutilation”, Instagram, *Communityregler*, “Tänk igenom vad du gör när du publicerar händelser med nyhetsvärde”.



I nästan samtliga fall konstateras eller antyds att förvisso otillåtna grafiska skildringar kan tillåtas under vissa omständigheter. Facebook, Twitter och YouTube kan godta våldsskildringar eller annat grafiskt material under förutsättning att det markeras som ”känsligt innehåll” eller görs otillgängligt för underåriga. I flera fall konstateras att grafiskt eller våldsamt innehåll som syftar till att chocka eller äckla är mindre tolererbart än innehåll som sprids med avsikt att informera, fördöma eller diskutera, eller om informationen i fråga har ”nyhetsvärde”.

#### 5.2.4 Pornografi och sexuellt innehåll

Utöver våldsamt, grafiskt, eller obscen material tillåter ingen av plattformarna pornografi eller innehåll som annars är ”ägnat att väcka sexuell tillfredsställelse”.<sup>110</sup>

I detta hänseende har det nyligen skett stora förändringar i plattformarnas innehållsmoderering. Tumblr tillät tidigare nakenhet och erotiskt innehåll på sin plattform, och upplät till användare att dölja sådant material genom att aktivera ett ”säkerhetsläge”. I december 2018 ändrade Tumblr sin hållning, och valde att helt förbjuda naket och erotiskt innehåll. Beslutet hade effekt både framgent och retroaktivt. Genom innehållsfilter och mänsklig moderering raderades material som lades upp, och som redan hade publicerats på plattformen, efter att ha flaggats som ”vuxeninnehåll”.<sup>111</sup>

Även Facebook har skärpt sin syn på sexuellt innehåll. Under 2018 kompletterade Facebook sina plattformsregler med nya riktlinjer kring ”sexuellt kontaktsökande” (”sexual solicitation”). De nya riktlinjerna innebär att också innehåll som ”förenklar, uppmuntrar eller koordinerar sexuell kontakt mellan vuxna” är otillåtet på plattformen. Förbudet innefattar sexuellt explicit språk som kan *leda till* sexuellt kontaktsökande. Riktlinjerna omfattar ”sexualiserad slang”, omnämnande av sexuella preferenser eller partnerpreferenser, omnämnande av ”vanligen sexualiserade kroppsdelar”, och ”vaga, suggestiva uttalanden” som kan innebära implicit sexuell kontaktsökande.

---

<sup>110</sup> Facebook, *Community Standards*, III. 14 “Adult nudity and sexual activity”, Twitter, *Twitter media policy*, YouTube, *Policy om nakenhet och sexuellt innehåll*, Tumblr, *Community Guidelines*, “Adult content”, Instagram, *Communityregler*, “Publicera foton och videoklipp som är lämpliga för en spridd målgrupp”.

<sup>111</sup> Tumblr, *Update to Tumblr’s Community Guidelines*. I ett missiv förklarar Tumblrs vd Jeff D’Onofrio den ändrade hållningen: ”We recognize Tumblr is also a place to speak freely about topics like art, sex positivity, your relationships, your sexuality, and your personal journey. As Tumblr continues to grow and evolve, and our understanding of our impact on our world becomes clearer, we have a responsibility to consider that impact across different age groups, demographics, cultures, and mindsets. We spent considerable time weighing the pros and cons of expression in the community that includes adult content. In doing so, it became clear that without this content we have the opportunity to create a place where more people feel comfortable expressing themselves”. D’Onofrio, *A better, more positive Tumblr*.

### 5.2.5 Övriga innehållsbegränsningar

De ovan nämnda innehållsrestriktionerna utgör ett axplock av de som återkommer i plattformarnas riktlinjer och policyer. Därtill kommer en mängd regler och riktlinjer av varierande karaktär och innehåll. Typiskt sett tolererar plattformarna inte spam, identitetsförfalskning och bedrägeri,<sup>112</sup> upphovsrättsintrång,<sup>113</sup> eller publicerandet av integritetskränkande eller känsliga uppgifter om andra,<sup>114</sup> för att nämna några.

Facebook utmärker sig på en punkt. Plattformens riktlinjer tar nämligen också sikte på ”falsa nyheter” och desinformation.<sup>115</sup> ”Falsa nyheter” ges ingen närmare definition, men får antas avse medvetna, empiriskt oriktiga sakomdömen i frågor av allmänintresse (vilket kan beskriva vitt skilda företeelser, från rent bedrägeri till propaganda, ”trollande” och satir).<sup>116</sup> Facebook konstaterar att de inte raderar innehåll som betraktas som ”falsa nyheter”, men att de ingriper för att ge sådant innehåll lägre exponering och cirkulation. I ett annat avsnitt anges emellertid att ”desinformation som bidrar till nära förestående våld eller fysisk skada” är otillåtet, och kan raderas.<sup>117</sup>

## 5.3 Personbegränsningar

Frånsett de innehållsrelaterade restriktionerna innefattar plattformarnas riktlinjer ofta också begränsningar kring *vem* som får nyttja tjänsten. I regel innefattar villkoren en åldersgräns, samt förutsättningen att användaren inte är förhindrad att använda tjänsten på grund av tillämplig lag.<sup>118</sup> Åldersgränsen är typiskt sett 13 eller 16 år, med förbehåll för nationella normer om rättshandlingskapacitet och minimumålder för samtycke till personuppgiftsbehandling enligt dataskyddsförordningen.<sup>119</sup>

---

<sup>112</sup> Facebook, *Community Standards*, IV. 16 “Spam and Security”, IV. 17 “Misrepresentation”, Twitter, *The Twitter Rules*, “Spam”, YouTube, *Policyer för spam, bedrägligt beteende och bedrägerier*, Tumblr, *Community Guidelines*, “Spam” och “Confusion or Impersonation”, Instagram, *Communityregler*, “Främja meningsfulla och äkta interaktioner”.

<sup>113</sup> Facebook, *Community Standards*, V. 20 “Intellectual Property”, Twitter, *The Twitter Rules*, “Intellectual property”, YouTube, *Användarvillkor för Tjänsten*, p. 7.7, Tumblr, *Community Guidelines*, “Copyright or Trademark Infringement”, Instagram, *Communityregler*, “Dela enbart foton och videoklipp som du har tagit eller har rätt att dela”.

<sup>114</sup> Facebook, *Community Standards*, II. 11 “Privacy breaches and image privacy rights”, Twitter, *About private information on Twitter*, Tumblr, *Community Guidelines*, “Privacy Violations”.

<sup>115</sup> Facebook, *Community Standards*, IV. 18.

<sup>116</sup> Jfr Baade, *The European Journal of International Law* 2018, s. 1358–1359 och Verstraete et al., *Arizona Legal Studies*, Discussion Paper 17-15, s. 5–9.

<sup>117</sup> Facebook, *Community Standards*, I. 1.

<sup>118</sup> Facebook, *Terms of Service*, 3.1 “Who can use Facebook”, Twitter, *Twitter Terms of Service*, “1. Who May Use the Services”, YouTube, *Användarvillkor för Tjänsten*, 2.3, Tumblr, *Terms of Service*, “3. Use of the Services”, Instagram, *Användarvillkor*, “Dina åtaganden”.

<sup>119</sup> I Sverige är gränsen 13 år, se 4 § dataskyddslagen.

Även i detta hänseende går Facebook, och Facebook-ägda Instagram, längre i sina villkor. Båda plattformarna förbehåller sig rätten att neka tillgång till tjänsten om en användare tidigare blivit avstängd från tjänsten. Dessutom konstaterar båda att dömda sexualförbrytare är förbjudna från att använda tjänsten.

#### 5.4 Sorteringsalgoritmer

Utöver användarprofilen innefattar en social medieplattform typiskt sett ett eller flera informationsflöden som synliggör innehåll och interaktioner som har publicerats på plattformen eller i användarens nätverk (en så kallad ”feed”). Facebook, Instagram och Twitter utgår till exempel från en nyhetsfeed som synliggör uppdateringar från och interaktioner mellan användarens kontakter. Informationsflödena är inte strikt kronologiskt ordnade – istället kureras och arrangeras innehållet i flödena genom algoritmiska urvalmekanismer inbyggda i plattformens mjukvara.<sup>120</sup> Att det sker ett urval, snarare än en neutral och kronologisk återgivning av allt innehåll som distribueras i en användares nätverk, lär bero på användarvänligheten – en enskild Facebook-användare skulle exempelvis mötas av ett tusental nya inlägg och interaktioner i sitt flöde varje dag, om inget urval skedde.<sup>121</sup>

Urvalsalgoritmerna utgör ytterligare en aspekt av de sociala mediernas innehållsmoderering. De avgör, i allt högre utsträckning, vilken information som blir tillgänglig för den enskilde användaren, och i förlängningen, den enskilde användarens insyn och delaktighet i den samhälleliga diskursen. De avgör vad som ska uppfattas som omdebatterat, trendigt, eller kontroversiellt, vilka kontakter inom ett användarnätverk som synliggörs, och vilka som osynliggörs.<sup>122</sup>

Det är svårt att redogöra för exakt hur dessa urvalmekanismer opererar. Algoritmerna har med tiden blivit extremt komplexa, och utgör dessutom konkurrenskänslig information som företagen ogärna offentliggör. Generellt får algoritmerna dock antas premiera sådan information som har hög *relevans* för den enskilde användaren. Produktchefen för Facebooks ”newsfeed”, Greg Marra, formulerar urvalsprinciperna som följande:

---

<sup>120</sup> Bucher, *New Media & Society* 2012, s. 1167.

<sup>121</sup> Jfr Mosseri, *News Feed Ranking in Three Minutes Flat*.

<sup>122</sup> Gillespie, *Media Technologies: Essays on Communication, Materiality, and Society*, s. 167.

Frågan vi utgår från när vi rangordnar inläggen i newsfeed är alltid: hur förser vi användarna med det innehåll som är mest relevant för just dem?<sup>123</sup>

Om Facebook är det känt att deras innehållsmoderering tidigare utgick från en algoritm som kallades för Edgerank. Edgerank rangordnade varje inlägg och interaktion inom en användares nätverk (varje så kallad "edge") utefter tre huvudkriterier: relationen mellan användaren och innehållets upphovsman ("affinity"), innehållets vikt baserat på dess popularitet och form ("weight") och tidsåtgången från det att innehållet publicerats ("time decay"). Utifrån de tre kriterierna tillmättes varje "edge" ett eget värde i förhållande till den enskilda användaren, vilket avgjorde om innehållet skulle synliggöras för användaren, och i så fall, hur högt det skulle rangordnas i användarens flöde. Innehåll fick således större exponering om det härstammade från nära vänner, om det utgjordes av bild eller video snarare än text, om många andra användare redan hade interagerat med innehållet genom att kommentera eller "gilla" det, om det var nytt, och så vidare.<sup>124</sup>

---

$$\sum_{edge\ e} u_e w_e d_e$$

**Där:**

$u_e$  = värde, "affinity"

$w_e$  = värde, "weight"

$d_e$  = värde, "time decay"

---

*Fig. 4. Approximerad formel för rangordning av användargenererat innehåll på Facebook enligt algoritmen Edgerank. De tre variablerna kvantifieras och multipliceras till ett samlat rankingvärde som avgör innehållets placering i en given användares flöde.<sup>125</sup>*

EdgeRank-algoritmen tillämpades fram till 2011, varefter den ersattes av en mer komplex maskininlärningsalgoritm med förmåga att utveckla sig själv i takt med att den exponeras för ny data.<sup>126</sup> Genom maskininläring är de nya urvalsmekanismerna i princip höljda i dunkel – de är "svarta lådor", vars beslutsskäl knappt är urskiljbara ens för algoritmens upphovsmän.<sup>127</sup>

---

<sup>123</sup> Gelin & Pettersson, *Internet är trasigt*, s. 31.

<sup>124</sup> Bucher, *New Media & Society* 2012, s. 1167–1169.

<sup>125</sup> Bucher, *New Media & Society* 2012, s. 1167–1169.

<sup>126</sup> Chandler & Munday, *A Dictionary of Social Media*, "Edgerank" och Bucher, *Information, Communication & Society* 2017 s. 41.

<sup>127</sup> Introna, *Science, Technology & Human Values* 2016 s. 41.

Den nya ”newsfeed”-algoritmen uppdateras varje vecka, men några av de grundläggande relevansparametrarna för Edgerank – ursprung, form och popularitet – uppges dock fortfarande vara centrala för innehållssorteringen.<sup>128</sup>

## 5.5 Regelefterlevnad och sanktioner

Med komplexa innehållsriktlinjer och astronomiska mängder användargenererad data varje dag möter de sociala medierna ett svårt problem med regelefterlevnad.<sup>129</sup> Vanligen tillämpar de olika plattformarna egna system med användarbaserad innehållsflaggning och rapportering, tillsammans med algoritmisk och mänsklig granskning av rapporterat innehåll för att komma till rätta med avvikelser från plattformsriktlinjerna.<sup>130</sup> Vissa plattformar tillämpar system med ”betrodna flaggare” – enskilda användare eller organisationer som tillmäts särskild trovärdighet och prioritet när de rapporterar otillåtet innehåll.<sup>131</sup>

Vid bedömningen av om ett visst yttrande innebär ett brott mot användarriktlinjerna, betonar plattformarna ofta vikten av sammanhang och omständigheter.<sup>132</sup> Som ovan nämnt bedöms våldsamt och grafiskt innehåll hårdare om det anses avsett att chockera, eller om det skildrar underåriga.<sup>133</sup> En bedömning kan samtidigt påverkas av att innehållet är pedagogiskt eller vetenskapligt betingat, eller om det anses vara av ”nyhetsvärde” eller ”allmänintresse”. I fråga om hatiskt tal eller hets mot folkgrupp tas hänsyn till humoristiska eller satiriska ambitioner.<sup>134</sup> Likaså kan upphovsmannens egenskaper påverka bedömningen. En person eller organisation som ”existerar i det allmänna medvetandet” kan åtnjuta större bedömningsmarginaler. Förekomsten av tidigare överträdelser mot plattformsriktlinjerna kan dock försvåra bedömningen.

I användarvillkoren förbehåller sig plattformarna rätt att vidta sanktioner mot användare som bryter mot användarriktlinjerna.<sup>135</sup> Urvalet av möjliga sanktioner varierar kraftigt från plattform till plattform.

---

<sup>128</sup> Gelin & Pettersson, *Internet är trasigt*, s. 32.

<sup>129</sup> Se ovan avsnitt 4.3.

<sup>130</sup> Crawford & Gillespie, *New Media & Society* 2016 s. 410–413. Jfr Tumblr, *Content moderation on Tumblr*.

<sup>131</sup> Youtube, *YouTube-programmet Betrodda flaggare*.

<sup>132</sup> Jfr Twitter, *Our approach to policy development and enforcement philosophy* och YouTube, *Vikten av sammanhang*.

<sup>133</sup> Se ovan avsnitt 5.2.3.

<sup>134</sup> Facebook, *Community Guidelines*, III. 12 “Hate speech”.

<sup>135</sup> Se not 101.

Twitters sanktionskatalog innefattar exempelvis innehållsrelaterade sanktioner som innehållsrestriktioner (som åldersbegränsningar, kommentarsblockering eller avkommersialisering), interimistisk blockering av innehåll och radering av innehåll, samt användarrelaterade sanktioner som ändrings- eller raderingsförelägganden mot användare, restriktioner på användarkonton (som förbud mot innehållsdistribution), interimistisk blockering av användarkonto och permanent avstängning av användarkonto. Valet av sanktion avses i princip stå i proportion till överträdelsens allvar. Direkta hot mot andra användare medför direkt och permanent avstängning av användarkontot, medan mindre allvarliga överträdelser kan innebära anmodan till rättelse genom föreläggande.<sup>136</sup> YouTube tillämpar istället ett system med eskalerande varningar. En första överträdelse leder till anmärkning. Därefter sanktioneras ytterligare överträdelser med varningar, kombinerat med restriktioner på användarkontot. Den tredje varningen innebär att användarkontot stängs av permanent.<sup>137</sup>

I vissa fall utfärdas inga direkta, officiella sanktioner vid överträdelser av användarriktlinjerna. Istället nedprioriteras innehållet genom plattformens sorteringsalgoritmer, och får således begränsad spridning och synlighet. Så hanterar Facebook exempelvis innehåll klassat som ”falska nyheter”.<sup>138</sup> YouTube och Twitter tillämpar också nedprioritering av innehåll som klassats som ”suggestivt” eller ”profant”.<sup>139</sup>

Sanktion	Facebook	Twitter	YouTube	Tumblr	Instagram
Nedprioritering av innehåll	×	×	×		
Restriktioner på innehåll	×	×	×		
Interimistisk blockering av innehåll		×			
Radering av innehåll	×	×	×	×	×
Anmärkning			×		
Varning	×		×		
Föreläggande mot användarkonto		×		×	
Restriktioner på användarkonto	×	×	×		
Interimistisk blockering av konto		×			
Temporär avstängning av användarkonto	×				×
Avstängning av användarkonto	×	×	×	×	×

Fig. 5. Sanktionsmöjligheter, som framgår av plattformarnas användarvillkor, riktlinjer och policyer.

<sup>136</sup> Se Twitter, *Hateful conduct policy*.

<sup>137</sup> YouTube, *Grundläggande information om varningar om brott mot communityns riktlinjer*.

<sup>138</sup> Facebook, *Community Guidelines*, IV. 18 ”False news”.

<sup>139</sup> Gillespie, *Media Technologies: Essays on Communication, Materiality, and Society*, s. 172.

Så mycket kan konstateras om de sociala mediernas innehållsmoderering: sociala medier är inte neutrala i förhållande till det innehåll som användarna producerar och distribuerar på plattformen. De begränsar, formar och arrangerar informationsflödena genom gränssnittskonstruktioner, genom bindande användarvillkor och riktlinjer, genom algoritmiskt kurerande av nyhetsflöden och genom olika rapporteringsverktyg och sanktionsmekanismer. De avgör vem som får yttra sig på deras plattformar, vad som får yttras, hur stor exponering ett yttrande ska få, och vilka konsekvenser som ska följa av ett otillåtet yttrande. Att detta sker är knappast ägnat att förvåna. Mer uppseendeväckande, om något, lär vara modereringens omfattning och de grunder på vilka modereringen sker.

## 6 Kongruens och kompetens – problemet med modereringen

*You have the right to free speech  
As long as  
You're not dumb enough to actually try it*

Joe Strummer, The Clash (“Know Your Rights”, *Combat Rock*, 1982)

Innehållsmodereringen på sociala medier kan på det stora hela tyckas förnuftig. Det verkar krävas en viss grad av informationssortering för att upprätthålla en användarvänlig tjänst. Det är heller inte svårt att rättfärdiga en nolltolerans mot mordhot, hämndporr eller vapenhandel på kommersiella, sociala nätverksplattformar. Men därom råder inte heller tvistighet. Frågan om yttrandefrihet på sociala medier är knappast begränsad till om nynazister ska få angripa journalister på Twitter eller om det bör anses godtagbart att sprida terrorpropaganda på YouTube. Frågan är istället hur stort godtycke som plattformarna bör åtnjuta, i egenskap av informationssamhällets nya grindvakter.

I det följande diskuteras några av de huvudsakliga problem som aktualiseras i samband med de sociala mediernas innehållsmoderering, ur ett yttrande- och informationsfrihetsperspektiv. Inledningsvis sker en kort genomgång av gällande rätt på yttrande- och informationsfrihetens område: de rättsliga normernas materiella tillämpningsområde, samt förutsättningarna för att genomföra inskränkningar i yttrande- och informationsfriheten. Därefter analyseras de sociala mediernas innehållsmoderering i ljuset av gällande rätt. Det som nedan kallas kongruensproblemet kommer att innefatta en abstrakt bedömning av överensstämmelsen – *kongruensen* – mellan plattformarnas riktlinjer och de yttrande- och informationsfrihetsstandarder som uppställts genom prejudicerande domstolsavgöranden. Därefter diskuteras vad som här benämns kompetensproblemet: de problem som uppstår i den konkreta tillämpningen av plattformarnas riktlinjer. Båda problembeskrivningar utgår från en fingerad tillämpning av yttrande- och informationsfrihetsnormerna – en förutsättningslös jämförelse mellan praxis och ideal.



## 6.1 Om yttrandefrihetens *ratione materiae*

I svensk rätt skyddas yttranden i tryckta skrifter alltså genom tryckfrihetsförordningen, medan yttranden i andra traditionella massmedier, som radio och tv, omfattas av yttrandefrihetsgrundlagen. TF och YGL är således teknikberoende – de medför ett särskilt starkt skydd för yttranden som manifesteras i form av exempelvis skrifter framställda i tryckpress, genom stencilering och fotokopiering samt ljudradio, videogram och databaser.<sup>140</sup> Teknikberoendet, tillsammans med grundlagarnas högdetaljerade utformning, innebär att de är svårtillämpade och illa anpassade för digital kommunikation.<sup>141</sup>

Skyddet för yttrandefriheten enligt RF 2 kap. 1 § är emellertid medieneutralt. Detsamma gäller enligt artikel 10 i EKMR och artikel 11 i EU-stadgan, som garanterar var och en ”åsiktsfrihet samt frihet att ta emot och sprida uppgifter och tankar”. Yttrande- och informationsfriheten, enligt dessa rättighetskataloger, omfattar också informationsdistribution som äger rum på internetplattformar.<sup>142</sup> Den omfattar varje typ av yttrande, vare sig politiska meningsyttringar eller kommersiellt betingade yttranden.<sup>143</sup> Den omfattar yttranden i tal och text såväl som i visuella skildringar.<sup>144</sup>

Yttrande- och informationsfriheten enligt RF, EKMR, och EU-stadgan är emellertid inte absolut. Under vissa förutsättningar kan friheterna inskränkas. Ett första krav härvid är att inskränkningen har stöd i lag.<sup>145</sup> Legalitetsprincipen innebär ett krav på viss kvalitet i den lagstiftning som ligger till grund för inskränkningen: lagen måste vara allmänt tillgänglig och tillräckligt precist utformad för att en enskild ska kunna förutse tillåtligheten i och konsekvenserna av sitt agerande.<sup>146</sup>

---

<sup>140</sup> SOU 2012:55 s. 264–267 samt Strömberg & Lundell, *Grundlagsskyddad yttrandefrihet*, s. 27–43.

<sup>141</sup> Se SOU 2012:55 s. 226–232 samt Axberger, *Yttrandefrihetsgrundlagarna*, s. 15.

<sup>142</sup> Ahmet Yıldırım v. Turkey p. 48–50, Delfi AS v. Estonia p. 11, Times Newspapers Ltd (nos. 1 and 2) v. the United Kingdom. Enligt Europadomstolen kan skillnaden mellan traditionell press och internetkommunikation, särskilt med avseende på den högre risken för integritetskränkningar i det senare fallet, i vissa fall motivera en skillnad i tillämplig policy för de olika medierna. Se Editorial Board of Pravoye Delo and Shtekel v. Ukraine p. 64.

<sup>143</sup> Ashby Donald and others v. France p. 34. Politiska yttranden och yttranden av allmänintresse åtnjuter dock typiskt sett ett vidare skydd än yttranden av marknadsföringskaraktär. Jfr Axel Springer AG v. Germany p. 90 och Mouvement Raëlien Suisse v. Switzerland p. 51–52.

<sup>144</sup> Ashby Donald and others v. France p. 34.

<sup>145</sup> Se RF 2 kap. 20 §, EKMR artikel 10.2 samt EU-stadgan artikel 51.1.

<sup>146</sup> Se Editorial Board of Pravoye Delo and Shtekel v. Ukraine p. 51–52, NJA 2012 s. 400 p. 16.

Därtill måste inskränkningen tjäna ett legitimt ändamål. I fråga om inskränkningar i yttrande- och informationsfriheten finns en uttömmande flora av legitima ändamål kodifierad i rättighetskatalogerna.<sup>147</sup> Dessa omfattar skydd för statens säkerhet och allmän säkerhet, upprätthållandet av allmän ordning och förebyggandet av brott, samt skydd för andras anseende och privatliv. EKMR nämner också skyddet för territoriell integritet, hälsa och moral, andras rättigheter, förtroliga underrättelser och domstolars auktoritet och opartiskhet. RF, i sin tur, lägger till ”folkförsörjningen” och andra ”särskilt viktiga skäl”.

Dessutom krävs, för att en inskränkning i yttrande- och informationsfriheten ska vara godtagbar, att inskränkningen är nödvändig och proportionerlig i förhållande till dess syften.<sup>148</sup> Enligt Europadomstolen innebär detta kriterium att inskränkningen måste svara mot ett angeläget samhällsligt behov (”pressing social need”), att inskränkningen måste stå i proportion till det legitima ändamålet i fråga, och att skälen för inskränkningen måste vara ”relevanta och tillräckliga”.<sup>149</sup> I varje fall måste beaktas vikten av en bred yttrande- och informationsfrihet. Europadomstolen har flera gånger manat till försiktighet vid inskränkningar i yttrande- och informationsfriheten:

Freedom of expression constitutes one of the essential foundations of a democratic society and one of the basic conditions for its progress and for each individual's self-fulfilment. Subject to paragraph 2 of Article 10, it is applicable not only to 'information' or 'ideas' that are favourably received or regarded as inoffensive or as a matter of indifference, but also to those that offend, shock or disturb. Such are the demands of pluralism, tolerance and broadmindedness without which there is no 'democratic society'. As set forth in Article 10, this freedom is subject to exceptions, which – as the Court has already said above – must, however, be construed strictly, and the need for any restrictions must be established convincingly.<sup>150</sup>

---

<sup>147</sup> RF 2 kap. 23 § samt EKMR artikel 10.2. EU-stadgan innehåller ingen uttömmande uppräkningslista av legitima ändamål, utan stadgar att inskränkningar måste svara mot ”mål av allmänt samhällsintresse som erkänns av unionen eller behovet av skydd för andra människors rättigheter och friheter”. Dock stadgas, i artikel 52.3, att rättigheternas innehåll och räckvidd åtminstone måste motsvara det som gäller enligt EKMR.

<sup>148</sup> Se RF 2 kap. 21 §, som stadgar att en inskränkning inte får ”gå utöver vad som är nödvändigt med hänsyn till det ändamål som har föranlett den” och inte heller ”sträcka sig så långt att den utgör ett hot mot den fria åsiktsbildningen”, och EKMR artikel 10.2, som endast erkänner inskränkningar ”som i ett demokratiskt samhälle är nödvändiga”.

<sup>149</sup> Zana v. Turkey p. 51, Hertel v. Switzerland p. 46, Steel and Morris v. the United Kingdom p. 87. För tillämpningen av proportionalitetsprincipen enligt RF 2 kap. 21 §, jfr NJA 2012 s. 400 p. 19–20, NJA 2015 s. 45 p. 15–17 och HFD 2015 ref. 80.

<sup>150</sup> Se Hertel v. Switzerland p. 46 och Handyside v. the United Kingdom p. 49. Jfr även, för dylika uttalanden, RF 2 kap. 23 § andra stycket samt NJA 2012 s. 400 p. 20.

Yttrande- och informationsfriheten omfattar inte bara yttranden som framstår som smakfulla, positiva eller neutrala, utan också yttranden som är “stötande, störande och chockerande” och, vad avser pressfriheten, likaså yttranden som innefattar en viss mån av “överdrift och provokation”.<sup>151</sup>

## 6.2 Kongruensproblemet

Den moderering som utövas av de olika sociala medieplattformarna innebär en markant begränsning av den enskilde användarens möjlighet att fritt yttra sig på plattformen. Modereringen utgör, de facto, en inskränkning i yttrandefriheten på internet. De policyer, sanktioner och sorteringsåtgärder som plattformarna vidtar har naturligtvis inte grund i lag – plattformarna saknar hitintills lagstiftningskompetens – men kan i övrigt bedömas i förhållande till de kriterier och krav som annars gäller för yttrandefrihetsinskränkningar.

I det följande poneras alltså, tills vidare, att plattformarna vore underställda gällande yttrande- och informationsfrihetsstandards, för att undersöka om det råder kongruens mellan dessa standards och plattformarnas modereringspolicyer.

### 6.2.1 Tillgänglighet och precision

En första förutsättning för godtagbara ingrepp i yttrande- och informationsfriheten är att den regel som ligger till grund för ingreppet är tillgängligt för den tilltänkta mottagarkretsen och tillräckligt preciserad för att skapa viss förutsägbarhet för mottagarna.

I de flesta fall är riktlinjer och policyer enkelt tillgängliga för användarna. Användaravtalen stadgar normalt sett att ändrade riktlinjer kommer att delges användarna, och inte tillämpas retroaktivt. Det finns emellertid exempel som visar på att så inte alltid är fallet. Då Tumblr inrättade nya, strängare regler för sexuellt explicit innehåll kom reglerna att tillämpas också retroaktivt. De nya riktlinjerna innebar att innehåll, som vid publiceringstillfället harmonierade med användarvilkoren, i efterhand kom att bli föremål för sanktioner i form av blockering och radering. En sådan modereringspraxis kan svårligen anses ha varit tillgänglig för användarna, och skulle antagligen inte tillfredsställa en legalitetsprincip.

---

<sup>151</sup> Jfr Pedersen and Baadsgaard v. Denmark p. 65, Handyside v. the United Kingdom p. 49 och Lindon, Otchakovsky-Laurens and July v. France p. 62.

Också vad gäller precision lämnar riktlinjerna ofta mer att önska. Många av de förbud och restriktioner som stadgas i användaravtalen är mycket vaga och svårtolkade. Det gäller exempelvis förbuden mot "trakasserier", "hets mot folkgrupp", stödjande av "hatgrupper" och "farliga organisationer", och "falska nyheter" – begrepp som ofta inte definieras eller kopplas till en befintlig rättslig definition.<sup>152</sup>

Europadomstolen har uppmärksammat att lagar till sin natur är mer eller mindre vaga och öppna för tolkning, och att lagens omfång ofta får fastställas närmare genom praxis.<sup>153</sup> Detta stämmer förstås också för andra typer av normer tänkta att appliceras i en dynamisk, pluralistisk kontext. Ett problem är emellertid att plattformarna inte, på samma sätt som det nationella rättsväsendet, prövar sina normer på ett sätt som möjliggör prejudikatbildning och precisering av reglernas innehåll. Även då plattformarna erbjuder överprövning av sanktionsbeslut, sker tillämpningen av reglerna sällan på ett sätt som är öppet och transparent och som på sikt kan leda till ökad klarhet för användarna.<sup>154</sup> Den ofrånkomliga vagheten i användarriktlinjerna kan därför inte avhjälpas på samma sätt som den förutsätts göra inom ramen för en ordinarie rättsordning.

### 6.2.2 *Subjektiva begränsningar*

Även i de fall då riktlinjerna är mycket precist utformade, med rigorösa begreppsbildningar och exempelfall, kvarstår en del bekymmer ur yttrande- och informationsfrihetsperspektiv. De kanske mest inskränkande villkoren, ur yttrande- och informationsfrihetsperspektiv, är de subjektiva begränsningarna. Det är riktlinjer som a priori inskränker en individs möjlighet att delta i den offentliga diskursen på sociala medieplattformar.

Enligt Facebook och Instagrams användarvillkor omfattas dömda sexualförbrytare av de subjektiva användarbegränsningarna. I brist på utförligare begreppsförklaring får någon som "dömts för sexualbrott" antas omfatta den som lagförts för sexualbrott enligt gällande nationell lagstiftning<sup>155</sup> (i svenskt hänseende skulle detta motsvara brott enligt 6 kap. brottsbalken, innefattandes bland annat våldtäkt, men också samlag med syskon, sexuellt ofredande och köp av sexuell tjänst).

---

<sup>152</sup> Jfr FN:s generalförsamling, A/HRC/38/35 s. 10.

<sup>153</sup> Se Hertel v. Switzerland p. 35.

<sup>154</sup> Jfr FN:s generalförsamling, A/HRC/38/35 s. 13 och 19.

<sup>155</sup> Jfr Facebook, *Terms of Service*, p. 4.4.

Förbudet är utan tvivel avsett att skydda bland annat minderåriga användare från oönskade sexuella närmanden genom plattformen. Skyddet för minderårigas fysiska och moraliska välfärd har också ansetts utgöra ett tungt vägande syfte enligt konventionspraxis – så tungt att stater har ålagts en positiv förpliktelse att verka för att underåriga skyddas i internetmiljöer.<sup>156</sup> Europadomstolens praxis har emellertid gällt situationen då stater underlåtit att beivra redan begångna brott mot underåriga, och inte situationer där det anses föreligga en förhöjd risk att underåriga utsätts för brott, grundat i generaliserade uppfattningar om hög recidivrisk bland sexualbrottslingar.

Frågan är om inte ett generellt förbud för dömda sexualbrottslingar att använda sociala medieplattformar hade ansetts utgöra en oproportionerlig inskränkning i deras yttrande- och informationsfrihet, även med hänsyn till det stadgade syftet. Denna hållning har anammats av amerikanska domstolar, som ansett att ett dylikt förbud (då inrättat genom strafflag) inte var förenligt med yttrande- och informationsfriheten enligt det första konstitutionella tillägget. Trots en förhållandevis låg lagprövningsstandard ansågs förbudet varken nödvändigt, lämpligt eller proportionerligt i förhållande till ändamålet.<sup>157</sup>

### 6.2.3 *Innehållsbegränsningar*

Likaså framstår de innehållsbaserade begränsningarna stundom som långtgående i förhållande till gängse rättighetspraxis. Plattformarnas förbud mot hot, kränkningar och förolämpningar är ett behjärtansvärt verktyg för att komma till rätta med cybermobbing och trakasserier, som annars riskerar att inverka avkylande på de utsattas yttrandefrihet. Europadomstolen har konstaterat att det finns utrymme att inskränka yttrande- och informationsfriheten vid orättfärdiga attacker mot andra, särskilt vid uttalanden som uppmanar till våld mot enskilda, men också vid förtalsartade sakpåståenden om andra.<sup>158</sup>

Annorlunda förhåller det sig dock med kränkande värdepåståenden. Sådana påståenden kan inte påvisas som sanna eller falska. De ger också uttryck för talarens åsikter, vilket tangerar yttrandefrihetens hårda kärna. Därför har domstolen varit försiktig med att inskränka yttrandefriheten vid kränkningar som blott ger uttryck för ett värdeomdöme.<sup>159</sup>

---

<sup>156</sup> K.U. v. Finland p. 46.

<sup>157</sup> Doe v. Marion County. Se även Hitz, *Indiana Law Journal* 2014 s. 1336–1358.

<sup>158</sup> Sürek v. Turkey p. 61.

<sup>159</sup> Se *Lingens v. Austria* p. 46 och *Jerusalem v. Austria* p. 42.

Att offentligt kalla en politiker för ”idiot” (*trottel*) har till exempel ansetts omfattas av den materiella yttrandefriheten, även om politikern i fråga tog illa vid sig.<sup>160</sup> Härvid bör alltså påminnas om att yttrandefriheten gäller också för yttranden som är stötande, störande eller chockerande.<sup>161</sup>

Plattformarnas riktlinjer mot hot och uppvigling till våld får, med bakgrund av detta, anses vara berättigade och väl förankrade i europeisk rättspraxis. Men då riktlinjerna också förbjuder förolämpande, negativa eller sårande yttranden, eller, i Facebooks fall, ”riktade svordomar”, framstår kongruensen med gällande praxis som haltande. Sannolikt ligger en del av de stötande yttranden, som prekluderas enligt riktlinjerna, väl inom yttrandefrihetens ramar.

Detsamma gäller för plattformarnas regler kring ”hate-speech” och hets mot folkgrupp. Även i detta hänseende har Europadomstolen förklarat att yttrande- och informationsfriheten kan inskränkas för att förhindra yttranden som sprider, uppviglar till, stödjer eller rättfärdigar hat grundat i intolerans.<sup>162</sup> Också yttranden som förolämpar, förlöjligar eller förtalar medlemmar ur en viss folkgrupp har ansetts kunna beivras, som en godtagbar begränsning av yttrandefriheten.<sup>163</sup> Därmed inte sagt att varje missaktande eller kränkande yttrande om en folkgrupp eller dess medlemmar kvalificerar som hatiskt tal. Återigen lär rent subjektiva värdeomdömen, som förvisso uttrycker missaktning mot en folkgrupp men som inte uppmuntrar till hat, omfattas av yttrandefriheten.<sup>164</sup>

Också i detta hänseende framstår de sociala mediernas policyer kring ”hate-speech” och hets mot folkgrupp som dels harmonierande med, dels avvikande från rättighetspraxis. Förbud mot våldsbejakande och avhumaniserande yttranden mot folkgrupper utgör utan tvivel en godtagbar inskränkning i yttrandefriheten. Men vissa riktlinjer, som Facebooks och YouTubes, ligger utanför självklarhetens sfär. Yttranden som ”jag tycker inte om män”, ”de som stred i Koreakriget är avskum”, eller enbart svordomar riktade mot en folkgrupp, framstår knappast som hatiskt tal i konventionens mening. Dylika yttranden vore, med andra ord, omfattade av yttrandefriheten.

---

<sup>160</sup> Oberschlick v. Austria p. 33–35.

<sup>161</sup> Jfr ovan not 149-150.

<sup>162</sup> Gündüz v. Turkey p. 40.

<sup>163</sup> Féret v. Belgium p. 73 och Vejdeland and Others v. Sweden p. 55.

<sup>164</sup> Jfr Gündüz v. Turkey p. 48-49, NJA 2005 s. 805 (där yttranden om homosexualitet som en ”cancersvulst på hela samhällskroppen” ansågs omfattas av yttrandefriheten enligt artikel 10) och Österdahl, SvJT 2006 s. 222–223. Se även NJA 2006 s. 467 och Vejdeland and Others v. Sweden, med motsatt utgång.

Avvägningen mellan politisk yttrandefrihet och stävandet av terrorism är en högkvalificerad balansakt. I målet *Leroy v. France* hade en baskisk serietecknare, knappt två dagar efter 11 september-attackerna, publicerat en teckning föreställande de brinnande tvillingtornen med texten ”NOUS EN AVIONS TOUS RÊVÉ ... LE HAMAS L’A FAIT” (ungefär: ”Vi har alla drömt om det ... Hamas gjorde det”). Europadomstolen kunde konstatera att teckningen utgjorde lovordande av terrorism, vilket ansågs berättiga en inskränkning av tecknarens yttrandefrihet.<sup>165</sup>

Härifrån skulle det till synes gå att finna täckning för de sociala medieplattformarnas förbud mot hyllande och glorifierande av terrorism, brottslingar och farliga organisationer. Men liksom för ”hate speech” och kränkningar går det dock att urskilja nyanser i Europadomstolens hållning, som saknas i plattformarnas rigorösa riktlinjer. I *Leroy v. France* var domstolen uppenbart tveksam mot att låta inskränka yttrandefriheten, särskilt då det rörde sig om ett politiskt budskap, dessutom i form av satir. Av avgörande betydelse för utgången var att teckningen publicerats så snart efter en svår tragedi, utan taktkänsla i uttrycket. Dessutom lades vikt vid de känsliga relationerna mellan Frankrike och Baskien, som tidigare präglats av terror och våld. Att härifrån utläsa ett generellt förbud mot lovordande av extremistgrupper – än mindre brottsliga eller ”farliga” organisationer – vore olämpligt.<sup>166</sup>

Också där plattformarna söker precisera vad som avses med en ”farlig organisation” tycks definitionen mycket omfattande. Facebooks definition av en kriminell organisation omfattar Hamas och Ku Klux Klan, men verkar också inrymma hackernätverket ”Anonymous” och andra aktiviströrelser. Den tunna gränsen mot politisk yttrandefrihet får plattformarnas riktlinjer att framstå som mycket tvivelaktiga.

Ett sista exempel är plattformarnas regler kring osedligt och explicit innehåll. Vad som anses osedligt eller moraliskt tvivelaktigt skiftar kraftigt mellan kulturer och nationer. Därför har det ansetts finnas ett visst svängrum för beslutsfattare vad gäller inskränkningar i yttrandefriheten som grundar sig i skyddet för moral och sedlighet.<sup>167</sup> Detta svängrum innebär dock inte ett fritt skön. Återigen bör betonas att yttrandefriheten inte bara inrymmer sådant som kommer att mottas positivt av samhället, utan också information som stöter, stör eller chockerar.

---

<sup>165</sup> *Leroy v. France*.

<sup>166</sup> Att notera är att den franska lagstiftning som tillämpats i *Leroy v. France* förbjöd bland annat anstiftan, stödande och ”ursäktande” av vissa brott, och inte lovordande av vissa organisationer. Se artikel 23–25 i den franska presslagen (*Loi du 29 juillet 1881 sur la presse*).

<sup>167</sup> *Handyside v. the United Kingdom* p. 48 och *Müller v. Switzerland* p. 35.

När det gäller sexuellt explicit, subversivt eller osedligt innehåll har Europadomstolen ofta tagit i beaktande om informationen har varit riktad mot unga människor, och om den uppmanat till brott eller skadligt beteende.<sup>168</sup> Sådana omständigheter kan antas inverka negativt vid fastställandet av yttrandefrihetens gränser.

På en social medieplattform som samlar medlemmar ur vitt skilda kulturer och samhällen, kan det tas för givet att hårdpornografi, skildringar av bestialitet, barnpornografi, ”grooming” och sexuella trakasserier begränsas, utan att det uppstår någon diskrepans med gällande yttrandefrihetsstandards. Men som visat sträcker sig plattformarnas riktlinjer längre än så. Även ”sexuellt kontaktsökande”, vaga sexuella anspelningar och nakenhet förbjuds enligt plattformarnas policyer. Om sådana förbud hade ansetts förenliga med artikel 10, tillika RF 2 kap. 1 § och EU-stadgans artikel 11, kan ifrågasättas. Här får beaktas att plattformarna ofta har separata regler mot barnpornografi, sexuella trakasserier och annat brottsligt beteende. Facebooks vidsträckta förbud kring ”sexual solicitation” omfattar dessutom uttryckligen sexuellt innehåll som delas mellan vuxna människor, och tar således inte sikte på ”grooming” eller sexuellt explicit material som riktas mot unga. Därtill är plattformarna åldersbegränsade. Minderåriga – typiskt sett personer under 13 år – har inte tillgång till innehåll på plattformarna. Skyddet för unga, och intresset av att motverka skadligt eller brottsligt beteende kopplat till osedligt innehåll, gör sig knappast gällande i hög grad för förbuden mot nakenhet, sexuella referenser eller sexuellt kontaktsökande. En betydande del av de yttranden som träffas av plattformarnas förbud, vore sannolikt att anse som legitima förfoganden av yttrande- och informationsfriheten.

Ovan har diskuterats ett par exempel där plattformarnas policyer och gällande yttrandefrihetsstandards står i bristande kongruens. Listan hade kunnat göras längre. Exempelen bör visa på att det inte heller handlar om perifera eller banala avvikelser. Det är diskrepanser som kan få stor betydelse för den praktiska yttrandefriheten, och som vållat oro och kritik bland fri- och rättighetsorganisationer världen över.<sup>169</sup>

---

<sup>168</sup> Handyside v. the United Kingdom p. 52 och Müller v. Switzerland p. 36.

<sup>169</sup> Se exempelvis Article 19, *Twitter Rules: Analysis against international standards on freedom of expression*, Article 19 5/9 2018, Article 19, *ARTICLE 19 and 70+ civil society groups write to Mark Zuckerberg over content removal*, Article 19 13/11 2018, Austin, *We can no longer talk about sex on Facebook in Europe*, EDRi 16/1 2019, Maebe, *ENDitorial: The fake fight against fake news*, EDRi 25/7 2018, Trendacosta & York, *Dear Tumblr: Banning "Adult Content" Won't Make Your Site Better But It Will Harm Sex-Positive Communities*, EFF 5/12 2018, Harmon, *Facebook's Sexual Solicitation Policy is a Honeypot for Trolls*, EFF 7/12 2018, York, *Blunt Policies and Secretive Enforcement Mechanisms: LGBTQ+ and Sexual Health on the Corporate Web*, EFF 24/10 2018, Human Rights Watch, *An Open Letter to Mark Zuckerberg*, *The World's*



Kongruensproblemet, som häri benämnt, är i sig värt att begrunda. Att riktlinjerna möjliggör olämpliga ingrepp i den faktiska yttrande- och informationsfriheten innebär emellertid inte att de tillämpas på ett sätt som faktiskt ger upphov till illegitima inskränkningar. Detta faktum riktar ljuset mot hur användarvillkoren tillämpas i praktiken.

### 6.3 Kompetensproblemet

Bedömningen av godtagbara yttrandefrihetsinskränkningar utgår från omständigheterna i det enskilda fallet. Ett yttrandes form, innehåll, kontext, avsändare, mottagare och syfte, samt de sanktioner som yttrandet resulterat i, utgör grunden för en samlad bedömning i varje enskilt fall.<sup>170</sup>

En sådan kontextuell bedömning är, som ovan visat, inte heller främmande för de sociala medieplattformarna då de tillämpar sina riktlinjer. Plattformarna betonar vikten av kontext vid bedömningen av tillåtligt innehåll: om innehållet är avsett att chockera eller engagera; om innehållet är av nyhetsvärde eller allmänintresse; om det riktar sig mot en enskild individ eller en politiker; om yttrandet är en del av en bredare tendens mot bristande regelefterlevnad; om innehållet skildrar eller riktar sig mot barn; om innehållet, i den mån det är stötande eller obscen, innefattar varningar eller förklaringar – och så vidare. Flera av dessa bedömningsgrunder har täckning i Europadomstolens praxis. Dessutom tillämpas olika slags sanktioner vid olika allvarliga överträdelser, vilket möjliggör en mån av proportionalitet i ingreppen. Sanktionernas allvar i förhållande till överträdelserna är en fundamental bedömningsgrund enligt EKMR.

Men, plattformarnas dynamiska tillämpning av reglerna är långt ifrån en försäkran för yttrande- och informationsfriheten. Inom ramen för en vanlig rättsordning bedöms yttrandefrihetsbrott typiskt sett av kompetenta domstolar eller nämnder, genom en kontradiktorisk process där parter och allmänhet har insyn i processen och där beslut kan överprövas och ställas i relation till konstitutionell eller internationell rätt. Ett sådant system är ägnat att iaktta legitimitet, transparens, förutsägbarhet, proportionalitet och rättvisa i tillämpningen.

---

*Freedom of Expression is in Your Hands*, Human Rights Watch 14/11 2018, samt FN:s generalförsamling, A/HRC/38/35 s. 9–14.

<sup>170</sup> Jfr Jerusalem v. Austria p. 34–35, Leroy v. France p. 37–47.

Systemen präglas av rättssäkerhetsstandards som naturligtvis inte kan matchas vid den extremt omfattande modereringen som äger rum på sociala medier. Där sker beslutsfattande istället genom summariska eller automatiserade processer, som ovan visat, genom enskilda moderatorers bedömningar, genom flaggning av innehåll och ibland genom filtrering och nedprioritering av innehåll. Regeltillämpningen, om än i enlighet med plattformens riktlinjer, kan därför ge upphov till mycket otillfredsställande resultat.

#### Fallet Aftenposten

I september 2016 publicerade den norske författaren Tom Egeland, genom sin Facebook-profil, ett antal fotografier som skildrade krigsföring genom historien. Ett av fotografierna var Nick Uts Pulitzerpris-vinnande bild föreställande en napalmskadad ung flicka under Vietnamkriget. Kort därpå lät Facebook meddela att de hade raderat bilden på grund av att den ansågs skildra grafiskt våld och barnpornografi. Egeland försökte publicera bilden ytterligare två gånger, men blev följaktligen avstängd från plattformen under ett respektive tre dygn. Även Aftenposten, en norsk kvällstidning, försökte genom sin Facebooksida publicera samma fotografi tillsammans med ett öppet brev till Mark Zuckerberg, kritiskt mot plattformens modereringspolicy. Aftenpostens publicering raderades kort därpå. Norges premiärminister Erna Sohlberg, liksom andra norska ministrar, lät publicera samma fotografi i protest. Även dessa publiceringar raderades, med hänvisning till Facebooks användarvillkor. Efter omfattande kritik och censuranklagelser från plattformens medlemmar gick Facebook senare ut och ursäktade hanteringen, som påstods vara en följd av felaktig regeltillämpning. Det har därefter framkommit att samma bild nyttjats internt i Facebooks modereringsutbildning, som ett exempel på otillåtet innehåll.<sup>171</sup>

#### Fallet EastEnders

Sommaren 2011 publicerade bloggaren Richard Metzger, på sin Facebook-sida, en artikel om en kommande demonstration i London. Artikeln delades tillsammans på en bild på två män som kysstes – en skärmdump från såpoperan EastEnders – för att protestera mot att ett homosexuellt par avhysts från en bar i London. Inom ett dygn hade Metzgers inlägg ”flaggats” ett antal gånger av användare som ogillade bilden, och snart meddelade Facebook att de raderat bilden på grund av att den stred mot plattformspolicyn. Bilden bedömdes som ”sexuellt explicit”. Efter stark kritik tvingades Facebook till att återpublicera bilden och be om ursäkt.<sup>172</sup>

Exemplen ovan är medvetet valda för att konkretisera ett antal problem med de sociala mediernas informationshantering, och bör inte betraktas som representativa för plattformarnas moderering. Det finns ingen anledning att anta att Facebook eller deras anställda är vare sig historierevisionister eller homofober. Men likväl belyser exemplen en del oegentligheter i sättet på vilket Facebook tillämpar sina policier.

---

<sup>171</sup> Gillespie, *Custodians of the internet* s. 1–5 och Hansen, *Dear Mark. I am writing this to inform you that I shall not comply with your requirement to remove this picture*, Aftenposten 8/9 2016.

<sup>172</sup> Crawford & Gillespie, *New Media & Society* 2016 s. 410–411.

### 6.3.1 *Intressen och incitament*

I en ordinarie rättsprövning bör en domstol förhålla sig till konflikten mellan yttrande- och informationsfriheten å ena sidan, och de angelägna samhällsliga intressen som aktualiseras å andra sidan, då de prövar huruvida ett yttrande ska kunna bli föremål för inskränkningar eller inte. De sociala medieplattformarna kan förvisso antas göra en i princip motsvarande prövning vid tillämpningen av riktlinjerna. Men, som ovan konstateras, drivs plattformarna också av andra intressen. De måste förhålla sig till användarkollektivets önskemål och inställningar, för att undvika att användarna lämnar plattformen. De måste förhålla sig till annonsörernas intressen, för att säkra annonsintäkter. De måste dessutom förhålla sig till en mängd olika rättsordningar och kulturer, för att säkerställa att tjänsten inte blir föremål för sanktioner eller repressalier i ett annat land med en annan standard för yttrandefrihet, censur och mellanhandsansvar. I exemplet Aftenposten kan tyckas självklart att en bild som skildrar krigets illdåd, trots att den innefattar ett naket barn, skulle ligga väl inom ramen för yttrandefriheten. I fallet framstår det inte heller som att ingreppen mot bilden grundades i enskilda felbedömningar, utan att det tvärt om rörde sig om en medveten modereringspolicy avseende bilden i fråga. Varför Facebook tillämpat denna modereringspolicy är svårt att veta, men det kan röra sig om ett intresse av att inte exponera användare och annonsörer för stötande och förargelseväckande material – det vill säga, mer ekonomiskt betingade avvägningar.

### 6.3.2 *Rättssäkerhet och rättvisa*

En ordinarie rättsprövning kring yttrandefrihetens gränser präglas också av vissa minimikrav på kompetens, utredning och rättssäkerhet. I fallet EastEnders, till skillnad från Aftenposten, var det uppenbart att modereringsbeslutet fattats på ett sätt som inte var förenligt med plattformens riktlinjer. Flaggningsystemet, tillsammans med en knapphändig granskning och bedömning i ärendet, gav ändå vid handen att bilden raderades. Detta är särskilt bekymmersamt, då systemet med användarinitierat beivrande möjliggör riktade angrepp mot förvisso tillåtet innehåll, som dock inte väcker andra användares sympati. Det har medfört att särskilt minoritetsgrupper, regimkritiker och politiska oppositioner blivit föremål för oproportionerligt mycket innehållsmoderering.<sup>173</sup>

---

<sup>173</sup> FN:s generalförsamling, A/HRC/38/35 s. 10-13.

Särskilt problematiskt är fallet också då innehållsfilter tillämpas för att upptäcka brott mot plattformsriktlinjerna. I enklare fall kan algoritmisk moderering ge goda resultat, men tillämpat för mer intrikata avvägningsfrågor – såsom upptäckandet av innehåll som glorifierar terror – är automatiserade bedömningar mycket olämpliga.<sup>174</sup> Men även mänsklig moderering kan vara ett bristande remedium, då moderatorerna sällan besitter samma kompetens eller utredningspotential som en domstol. Facebooks ”klickarbetare” – innehållsmoderatorer på lägsta nivå, som bedömer flaggat innehåll och dess förenlighet med användarvillkoren – arbetar sig exempelvis igenom flaggat innehåll inom intervaller om tio sekunder.<sup>175</sup> En sådan ultrasummarisk bedömning kan knappast anses adekvat ens för de mer flagranta fallen av överträdelse, än mindre för den presumerade lejonpart av innehåll på plattformen som fordrar mer kvalificerade bedömningar av riktlinjernas tillämplighet. Skillnaden mellan exempelvis falska nyheter, som enligt Facebooks riktlinjer ska nedprioriteras, och satir, som inte bör bli föremål för ingrepp, består inte sällan i avsändarens syfte och motiv. Vad som ska anses utgöra satir respektive falska nyheter blir därmed mycket svårt att avgöra utifrån en summarisk bedömning utan god kontextuell och kulturell kännedom.

### 6.3.3 *Transparens och insyn*

I en ordinarie rättsprövning ställs typiskt sett höga krav på transparens i rättstillämpningen. Det kan förklaras av rättssäkerhetsskäl, men transparensen bidrar också till att utveckla en konsekvent rättstillämpning och tydliggöra rättsläget i fall där normerna är vaga eller tvetydiga. Som ovan nämnt brister plattformarnas regeltillämpning kraftigt vad gäller transparens. I båda fallen Aftenposten och East-Enders tillställdes upphovsmännen en kortfattad beslutsformulering, som konstaterade att innehållet raderas på grund av att det stred mot Facebooks riktlinjer. Inte heller vid överprövning av modereringsbeslut är prövningen särdeles transparent. Bristen på insyn och öppenhet i plattformarnas modereringspraxis innebär en risk för bristande förutsebarhet och godtycklighet i regeltillämpningen.<sup>176</sup>

---

<sup>174</sup> FN:s generalförsamling, A/HRC/38/35 s. 10-13.

<sup>175</sup> Langvardt, Georgetown Law Journal 2018 s. 1362.

<sup>176</sup> FN:s generalförsamling, A/HRC/38/35 s. 13-14.

#### 6.4 Staten och kapitalet: det utkontrakterade ansvaret

Istället för att avhjälpa en eventuell diskrepans mellan plattformarnas policyer och fri- och rättighetsstandards, verkar plattformarnas modereringspraxis snarare förvärra saken. Den mänskliga modereringen ger förvisso plattformarna möjlighet att göra mer nyanserade bedömningar, men samtidigt innebär praxisen att riktlinjerna tillämpas olika, oförutsägbart, och i värsta fall, orättvist.

Kompetensproblemet rymmer också en annan bekymrande aspekt. De sociala medieplattformarna modererar med hänsyn till sina egna etiska övertygelser, med hänsyn till användarkollektivets presumerade uppfattningar, och i viss mån till annonsörernas tycke. Men till stor del drivs plattformarna också, i sin moderering, av ansvarsincitament. Mycket av plattformarnas riktlinjer kan nämligen härledas till statlig lagstiftning som ålägger plattformarna skyldigheter att ingripa mot visst innehåll, under hot om mellanhandsansvar. Som nämnt ovan tenderar lagstiftare i ökad utsträckning att ålägga plattformarna att komma till rätta med oönskat användarbeteende, såsom nätkränkningar eller upphovsrättsintrång. Som exempel verkar Facebooks och Tumblrs nya hållning kring sexuellt innehåll – särskilt Facebooks ingripande riktlinjer kring ”sexual solicitation” – härstamma från ny amerikansk lagstiftning mot trafficking och prostitution.<sup>177</sup> Men även i fråga om terror- och extremistrelaterat innehåll, trakasserier, hets mot folkgrupp och annan cyberbrottslighet har stater över hela världen ålagt plattformarna mellanhandsansvar som tvingat dem att utvidga sina innehållspolicyer och ingripa mot visst innehåll.<sup>178</sup>

I början av 2018 utfärdade EU-kommissionen en rekommendation till medlemsstaterna, som uppmuntrade medlemsstaterna att anta verktyg för att effektivt komma tillrätta med olagligt material online. Rekommendationen vände sig huvudsakligen mot hosting-tjänster och digitala tjänstetillhandahållare, med precisa rekommendationer kring vilket ansvar sådana aktörer bör åläggas för tredjeparters publiceringar. Bland annat föreslogs att sådana tjänstetillhandahållare bör inrätta flaggningssystem för misstänkt olagligt innehåll, och helst system med ”betrodda flaggare” vars innehållsflaggningar skulle äga prioritet framför andra flaggningar. Det föreslogs också att tjänstetillhandahållarna skulle inrätta ”automatiserade metoder” för att upptäcka och beivra brottsligt innehåll på deras plattformar.

---

<sup>177</sup> Jfr Austin, *We can no longer talk about sex on Facebook in Europe*, EDRi 16/1 2019.

<sup>178</sup> Gillespie, *Custodians of the Internet* s. 33–40, Laidlaw, *Regulating Speech in Cyberspace* s. 4–6 och FN:s generalförsamling, A/HCR/38/35 s. 7–8.

Vidare ansåg man att tjänstetillhandahållarna skulle åläggas att samarbeta med varandra, och med medlemsstaternas myndigheter och Europol, genom att bland annat rapportera misstänkta brott begångna på deras plattformar. Särskilt riktade rekommendationen in sig på terror-relaterat innehåll. Rekommendationen föreskrev uttryckligen att plattformarna, i sina innehållsriktlinjer, skulle förbjuda förekomsten av sådant innehåll på plattformen, och inrätta automatiserade system för att upptäcka och radera sådant innehåll.<sup>179</sup>

Denna typ av kvasi-statlig innehållsmoderering har också tagit sig i uttryck i mindre formella former. I maj 2016 beslutade EU-kommissionen, tillsammans med Microsoft, Facebook, Twitter och YouTube om en ”uppförandekodex” som innebar att plattformarna, genom sina användarvillkor och riktlinjer, skulle ”prohibit the promotion of incitement to violence and hateful conduct [and] promoting independent counter-narratives”. Även detta regelverk föreskrev att plattformarna skulle inrätta system med innehållsflaggning och ”betrodna flaggare”, samarbeten med myndigheter och andra privata aktörer, och beivningsmekanismer som möjliggjorde att otillåtet material kunde raderas inom 24 timmar.<sup>180</sup> Senare anslöt sig också Instagram, Google+, SnapChat och flera andra sociala medieplattformar till uppförandekodexen.<sup>181</sup>

Även om stävjandet av terrorism och hetsbrott kan tyckas välkommet i sig, är det svårt att betona precis hur bekymrande denna tendens är. I en FN-rapport från 2017 skriven av David Kaye, FN:s särskilde rapportör för stöd och skydd för åsikts- och yttrandefriheten, konstateras bland slutsatserna:

[S]tates increasingly exercise censorship through the private sector.<sup>182</sup>

Samma bedömning gör Joe McNamee, verkställande direktör vid EDRI:

The European Commission is pushing ‘voluntary’ censorship to internet giants to avoid legislation that would be subject to democratic scrutiny and judicial challenge.<sup>183</sup>

Det är uppenbart att de sociala medieplattformarna inte längre utövar innehållsmoderering enbart utifrån egna intressen, utan också som en förlängning av staters anspråk på informationskontroll.

---

<sup>179</sup> EU-kommissionen, *COMMISSION RECOMMENDATION of 1.3.2018 on measures to effectively tackle illegal content online*, C(2018) 1177 final.

<sup>180</sup> EU-kommissionen, *Code of Conduct on illegal online hate speech*.

<sup>181</sup> EU-kommissionen, *The EU Code of conduct on countering illegal hate speech online*.

<sup>182</sup> FN:s generalförsamling, A/HCR/35/22 s. 20.

<sup>183</sup> EDRI, *EU Commission’s Recommendation: Let’s put internet giants in charge of censoring Europe*, EDRI 1/3 2018.

Det rör sig om en slags "outsourcing" av rättsskipning och reglering, där privata aktörer ges både mandat och ansvar att bevaka yttrandefrihetens gränser å statens vägnar. De sociala medierna blir en sorts institutionaliserade grindvakter. Särskilt oroväckande blir detta, då grindvakterna inte tycks omfattas av de krav och skyldigheter som annars åvilar stater vid inskränkningar i yttrandefriheten.

## 7 Ansvar och skyldigheter – nya medier, nya remedier

*[Protection] against the tyranny of the magistrate is not enough: there needs protection also against the tyranny of the prevailing opinion and feeling; against the tendency of society to impose, by other means than civil penalties, its own ideas and practices as rules of conduct on those who dissent from them; to fetter the development, and, if possible, prevent the formation, of any individuality not in harmony with its ways, and compel all characters to fashion themselves upon the model of its own.*

John Stuart Mill, *On Liberty*<sup>184</sup>

Inom loppet av ett knappt kvartssekel har yttrande- och informationsfrihetens spelplan ritats om i grunden. Förändringen kan knappast ha undgått någon. Internet, och i synnerhet de sociala medierna, har åstadkommit en kommunikativ demokratiseringsprocess som få lär ha förväntat sig då världskrigen blåst över och 1900-talets rättighetskataloger sattes i pränt.

Men på den nya spelplanen råder nya regler och nya ledare. En oöverskådlig del av all mellanmänsklig kommunikation lyder idag under villkor som författats inte av grundlagsberedningar eller konstitutionsdomstolar, utan privata medieföretag. Villkoren fyller egna syften och funktioner, och bottnar inte alltid i de rättighetsavvägningar som vuxit fram genom försiktiga domstolsförfaranden. Villkoren begränsar möjligheten att distribuera information och idéer på några av världens största fora, och dess omfattning verkar öka stadigt under press från användarkollektiv, annonsörer och regeringar. Samtidigt har yttrande- och informationsfriheten på samma fora kommit att betraktas snarast som ett självregleringsproblem.

När Facebook, Twitter och YouTube bereds större makt över den praktiska yttrandefrihetens gränser, vare sig av egen vilja eller som instrument för statliga anspråk, uppstår ett sorts vakuum i den traditionella allokeringen av ansvar för yttrande- och informationsfrihetens förverkligande. Denna fråga kommer att beröras i detta sista avsnitt. Om det förra avsnittet utgick från en fingerad tillämpning av gällande yttrandefrihetsstandards i relation till plattformarnas modereringspraxis, kommer vi alltså nu att bemöta elefanten i rummet. Har de sociala medierna något ansvar för att upprätthålla yttrande- och informationsfriheten? Inledningsvis redogörs för yttrande- och informationsfrihetsnormernas subjektiva tillämpningsområde – vilket ansvar normerna ger upphov till, och för vem.

---

<sup>184</sup> Mill, *On Liberty* s. 10.



Därefter lämnar vi frågan om gällande rätt, för att istället rikta fokus mot *lex ferenda*. Om de befintliga ansvarskonstruktionerna skulle visa sig otillräckliga – hur skulle då en mer lämplig reglering se ut, i ljuset av de ovan beskrivna problemen?

### 7.1 Om yttrande- och informationsfrihetens *ratione personae*

Som framgått av ovan finns det en uppsjö av olika inskränkningar och restriktioner i yttrande- och informationsfrihetens materiella tillämpningsområde – regler som förbjuder bland annat hets mot folkgrupp, lovordande av terrorbrott, förtal, ärekränkningar och obsceniteter. En i sammanhanget mer avgörande begränsning av yttrande- och informationsfriheten ligger dock inte i dess *ratione materiae*, utan dess *ratione personae*.

Enligt RF 2 kap. 1 § är var och en tillförsäkrad yttrande- och informationsfrihet ”gentemot det allmänna”. Enligt EKMR och EU-stadgan innebär yttrande- och informationsfriheten en rätt att uttrycka, sprida och ta del av information ”utan offentlig myndighets inblandning”. Enligt ordalydelsen gäller de individuella fri- och rättigheterna, yttrande- och informationsfriheten, i förhållande till staten eller sådana aktörer vars aktiviteter kan tillskrivas staten. Som utgångspunkt kan de inte åberopas gentemot en annan enskild (så kallad direkt horisontell effekt eller *Drittwirkung*).<sup>185</sup>

Samma hållning har anammats i praxis, där domstolar varit mycket skeptiska mot att låta rättighetskatalogerna få verkan direkt emellan enskilda:

The Court does not consider it desirable, let alone necessary, to elaborate a general theory concerning the extent to which the Convention guarantees should be extended to relations between private individuals *inter se*.<sup>186</sup>

Facebook, Twitter, YouTube, Tumblr och Instagram är privata företag. Utifrån gällande rätt vore det mycket osannolikt att de skulle kunna förpliktigas upprätthålla yttrande- och informationsfrihet i förhållande till sina användare, genom att exempelvis tillåta eller tolerera vissa yttranden på sina plattformar, med direkt grund i EKMR, RF eller EU-stadgan.

---

<sup>185</sup> Se Eka et al., *Regeringsformen – med kommentarer* s. 80, Cameron, *An Introduction to the European Convention on Human Rights* s. 48–49 samt Lebeck, *EU-stadgan om grundläggande rättigheter* s. 136 och 337.

<sup>186</sup> Verein gegen Tierfabriken (VgT) v. Switzerland p. 46. Se även NJA 2007 s. 747 (om artikel 8) och NJA 2015 s. 899 (om artikel 11 och artikel 1 tilläggsprotokoll 1).

Här bör nämnas att FN har utvecklat riktlinjer för företagsverksamhet och mänskliga rättigheter, som föreskriver att privata företag ska iaktta de mänskliga rättigheterna – däribland yttrande- och informationsfriheten – i sin verksamhet.<sup>187</sup> Till skillnad från de rättighetskataloger som diskuterats ovan tar riktlinjerna således direkt sikte mot privata företag i deras förhållande till andra enskilda. Det rör sig emellertid inte om ett juridiskt bindande dokument, utan soft-law, som saknar effektiva mekanismer för tillsyn och genomdrivande. Även om riktlinjerna skulle kunna tänkas få en viss normerande effekt innebär de, utifrån ett juridiskt perspektiv, inte särskilt mycket mer än en uppmaning till självreglering och ”corporate social responsibility”.

Avsaknaden av direkta horisontella verkningar innebär emellertid inte att yttrande- och informationsfriheten är utan relevans i förhållandet mellan sociala medier och deras användare. Artikel 1 i Europakonventionen stadgar att staten är skyldig att säkerställa att var och en kommer i åtnjutande av konventionsrättigheterna. Trots att yttrande- och informationsfriheten typiskt sett får uppfattas som en negativ rättighet – en rätt att inte bli utsatt för sanktioner på grund av det egna agerandet – kan friheterna därför, under vissa omständigheter, innebära en positiv förpliktelse för staten att skydda eller främja yttrande- och informationsfriheten, också i förhållandet mellan enskilda.<sup>188</sup> Detta kan betraktas som en indirekt horisontell effekt av fri- och rättigheterna. Till skillnad från Drittwirkung innebär den indirekta horisontella effekten att staten alltjämt är föremål för ansvaret, även om ansvaret kan ge upphov till lagstiftning eller beslut som i sin tur lägger ansvar på andra enskilda.

När statens positiva förpliktelser kan ge upphov till ansvar varierar beroende på de relevanta faktorer och omständigheter som gör sig gällande i situationen.<sup>189</sup> I fallet *Appleby and Others v. the United Kingdom* från 2003 hade de klagande parterna nekats tillstånd att demonstrera och dela ut politiska flygblad utanför stadens stora köpcentrum. Köpcentrumet var privat ägt, men fungerade som stadens centrum. Klagandena ansåg att den brittiska staten hade brustit i sin plikt att försvara deras yttrandefrihet, genom att inte ålägga köpcentrumets ägare att låta dem demonstrera där. Europadomstolen bedömde saken annorlunda och friade Storbritannien, men påpekade, apropå yttrandefriheten i förhållandet mellan enskilda medborgare och privata företag:

---

<sup>187</sup> OHCHR, *Guiding Principles on Business and Human Rights* artikel 11–24. Jfr FN:s allmänna förklaring om de mänskliga rättigheterna artikel 19.

<sup>188</sup> *Özgür Gündem v. Turkey* p. 42–43, *X and Y v. the Netherlands* p. 23 (om artikel 8) och *NJA* 2007 s. 747 (om artikel 8).

<sup>189</sup> *Özgür Gündem v. Turkey* p. 42–43.

Where [...] the bar on access to property has the effect of preventing any effective exercise of freedom of expression or it can be said that the essence of the right has been destroyed, the Court would not exclude that a positive obligation could arise for the State to protect the enjoyment of the Convention rights by regulating property rights. A corporate town where the entire municipality is controlled by a private body might be an example.<sup>190</sup>

Fem år senare, I fallet *Khurshid Mustafa and Tarzibachi v. Sweden*, fälldes Sverige i Europadomstolen för att ha brutit i sina positiva förpliktelser att skydda och främja yttrande- och informationsfriheten enligt artikel 10 EKMR. I fallet hade två boende i ett hyresgästhus i Rinkeby nyttjat en fasadmonterad parabolantenn för att ta del av utländska kanaler som sände på deras modersmål. Hyresvärderna ansåg att parabolantennen stred mot hyresavtalet och beslutade att vräka hyresgästerna. Efter överklagande fastställdes hyresvärdens beslut av Svea hovrätt.

Europadomstolen konstaterade att klagandena hade ett starkt och berättigat intresse av att få ta del av information på deras modersmål, särskilt då denna information innefattade politiska, sociala och kulturella nyheter av stor vikt för deras känsla av samhörighet och delaktighet. De skäl som hyresvärderna och staten åberopat till stöd för förbudet, bland annat säkerhetshänsyn och estetiska hänsyn, ansågs inte väga tungt i förhållande till klagandenas yttrande- och informationsfrihet. Därför hade staten, genom att upprätthålla vräkningsbeslutet, brutit mot klagandenas rätt enligt artikel 10.<sup>191</sup>

## 7.2 Ansvar vid innehållsmoderering

Det står alltså klart att de sociala medieplattformarna inte bär något ansvar för sina användares yttrande- och informationsfrihet, åtminstone inte direkt grundat på rättighetskatalogernas bestämmelser. Lika klart står det emellertid att samma förhållande, mellan plattformarna och användarna, skulle kunna ge upphov till en indirekt horisontell verkan av rättighetsbestämmelserna på grund av statens positiva skyldigheter. Oaktat Facebooks icke-statliga ställning skulle det alltså vara möjligt att extrahera ett visst ansvar för yttrande- och informationsfriheten ur rättighetskatalogerna.

---

<sup>190</sup> *Appleby and Others v. the United Kingdom* p. 47. Intressant att notera är att domstolen härvid hänvisar till det amerikanska fallet *Marsh v. Alabama*, där konstitutionens första tillägg om yttrandefrihet ansågs tillämpligt i en tvist mellan två enskilda: en fysisk person och en privatägd stad.

<sup>191</sup> *Khurshid Mustafa and Tarzibachi v. Sweden* p. 30–50.

Det första som talar för gångbarheten i en sådan ansvarslösning är att synen på yttrande- och informationsfriheten är under utveckling. Inte bara har synen på yttrandefriheten vidgats från att omfatta enbart förhandscensur, till att också omfatta sanktioner och indirekta ingrepp mot vissa yttranden – även synen på yttrandefriheten som en primärt negativ rättighet verkar ha skiftat under åren. I *Appleby and Others v. the United Kingdom* konstaterar den estniske domaren Rait Maruste i en skiljaktig mening:

The old traditional rule that the private owner has an unfettered right to eject people from his land and premises without giving any justification and without any test of reasonableness being applied is no longer fully adapted to contemporary conditions and society.

I takt med att samhället och tekniken förändras finns det anledning att omvärdera den ursprungliga räckvidden av rättigheterna, så att deras funktion kan upprätthållas i praktiken.

Den andra anledningen att tillkänna yttrande- och informationsfriheten indirekt horisontell verkan för sociala medier är deras funktion. De sociala medierna är både medier för informationsdistribution, som i fallet *Khurshid Mustafa and Tarzibachi v. Sweden*, och plattformar för utövandet av samma friheter, som i *Appleby and Others v. the United Kingdom*. De är både publicister, likt tidningsredaktioner, och distributörer, på samma sätt som internetleverantörer. De dubbla funktionerna, och den vertikala integrationen mellan infrastrukturell och redaktionell verksamhet, innebär att de sociala medierna har en än mer utpräglad grindväktarposition i informationssamhället än traditionella medier. Det borde, om något, innebära ökade hänsyn till användarnas yttrande- och informationsfrihet.<sup>192</sup>

En tredje anledning är omfattningen i de sociala mediernas verksamhet. Plattformarna brukas av flera miljarder människor världen över, och informationen som distribueras på plattformarna innefattar allt från underhållning till nyheter, opinionsbildning, och sociokulturellt identitetsskapande. Via plattformarna kommer användarna i kontakt med familj och vänner, men också arbetsmarknaden, föreningslivet, det politiska livet och samhällets institutioner och myndigheter. Inte sällan är plattformarna bärare av sådan information som utgör yttrande- och informationsfrihetens hårda kärna.

---

<sup>192</sup> Jfr Pasquale, *Theoretical Inquiries in Law* 2016 s. 487–513.

En fjärde anledning är sambandet mellan privaträttslig och offentligrättslig reglering på sociala medier. I många fall är det inte det privatägda sociala medieföretaget i sig som ligger bakom de villkor som inverkar på informationsdistributionen på plattformen – utan stater, regeringar och myndigheter. Den statliga delaktigheten i innehållsmodereringen borde rimligen öka statens positiva förpliktelser i förhållande till modereringens konsekvenser. Om stater i ökad utsträckning ”utkontrakterar” till sociala medier att avgränsa yttrandefriheten, bör de inte undgå ansvar med på grund av att utkontrakteringen sker till icke-statliga entiteter.

Alla dessa argument talar för att yttrande- och informationsfrihetens indirekta horisontella verkningar skulle och borde kunna ge upphov till ett ansvar för de sociala medieplattformarna. Men eftersom det rör sig om ett ansvar som härstammar från statens positiva förpliktelser, och inte en direkttillämpning av rättighetsnormer mellan enskilda, måste ansvaret utstakas genom offentligrättslig reglering. Hur skulle då en sådan regleringsmodell kunna utformas?

### 7.3 En hållbar lösningsmodell

Att ålägga enskilda särskilda förpliktelser gentemot andra enskilda kräver alltid en varsam balansgång mellan olika intressen och hänsyn. Den enes rätt kan inte ge upphov till oproportionerligt tunga eller rent av omöjliga förpliktelser för den andre, även om den andre råkar vara ett multinationellt företag med hög omsättning. Om nu en ansvarsreglering för sociala medier och yttrandefrihet skulle etableras, vore det därför angeläget att överväga vilka grundläggande krav som bör ställas på regleringen i fråga.

I denna fråga har Europadomstolen konstaterat att statens positiva förpliktelser inte får ge upphov till omöjliga eller oproportionerliga bördor för myndigheterna, och att hänsyn också måste tas till den reglerades rättigheter.<sup>193</sup> I förhållande till problemet som här ligger för handen borde följande minimikrav kunna formuleras:

1. *En ansvarsreglering bör vara effektiv och förutsägbar.* En regleringsmodell som inte åtgärdar problemet på ett tillfredställande sätt, eller som är alltför svårtillämpad i praktiken, vore naturligtvis inte lämplig.

---

<sup>193</sup> Appleby and Others v. the United Kingdom p. 40–43.

2. *En ansvarsreglering bör inte gå utöver vad som är nödvändigt för att komma tillrätta med problemet.* En reglering borde därför inte omfatta alla sociala medier, utan enbart sådana som är att betrakta som en typ av ”offentlig plats”<sup>194</sup> där skyddet för yttrande- och informationsfriheten gör sig gällande särskilt starkt.
3. *En ansvarsreglering bör ta rimlig hänsyn till de sociala medieplattformarnas rättigheter.* Här gör sig främst de sociala medieplattformarnas näringsfrihet och äganderätt gällande.<sup>195</sup>
4. *En ansvarsreglering bör ta rimlig hänsyn också till allmänna intressen och andra enskildas rättigheter.* Här gör sig många olika intressen gällande, varav flera utgör grunden till plattformarnas nuvarande modereringspraxis.

Det låter sig nästan gå osagt att det rör sig om en mycket komplex avvägning mellan motstående viljor och intressen. Ett antal lösningsmodeller har dock lagts fram i doktrin, varav några kommer att diskuteras i det följande.

### 7.3.1 Indirekt direkt ansvar

En regleringsmodell som förts fram bygger i princip på att låta sociala medier, som Facebook, tillskrivas ansvar för yttrande- och informationsfriheten genom offentligrättslig lagstiftning. Det skulle alltså inte formellt sett röra sig om en direkttillämpning av exempelvis artikel 10 EKMR i förhållande till Facebook, utan snarare om att implementera ansvarsnormer som *hänvisar* till yttrande- och informationsfriheten såsom den är definierad i någon av rättighetskatalogerna. Resultatet skulle sannolikt bli att artikel 10 blev indirekt tillämplig för Facebook, till sitt fulla materiella tillämpningsområde. Denna indirekta direkta tillämpning av yttrande- och informationsfrihetsnormer kan uppfattas som ett sätt att kringgå yttrandefrihetens *ratione personae*, men idéer i den riktningen har ändå lanserats.

---

<sup>194</sup> Ordvalet kan diskuteras. Med ”offentlig plats” avses inte enbart en plats som är öppen och tillgänglig för envar, utan en *offentlighet* – en bärare av den allmänna opinionen och ett medium för fri debatt, åsiktsbildning och demokratiskt självbestämmande. Jämför härvid det engelskspråkiga konceptet *public sphere*: ”the intermediary realm that supports the communication of opinions, the discovery of problems that need to be dealt with collectively, the channeling of these problems through the filter of the media and political institutions, and the realization of the collective will of the people in the act of democratic self-government”. Trenz, *International Encyclopedia of Political Communication*, ”Public Sphere”.

<sup>195</sup> Jfr RF 2 kap. 15 § och 2 kap. 17 §, EKMR artikel 11 och artikel 1 tilläggsprotokoll 1 samt EU-stadgan artikel 16–17.

Till dessa hör bland annat ett förslag om att göra FN:s riktlinjer för företagsverksamhet och mänskliga rättigheter juridiskt bindande.<sup>196</sup> Detta skulle i princip innebära att företag som Twitter förpliktigades att respektera yttrande- och informationsfriheten, såsom den definieras enligt FN:s deklaration om mänskliga rättigheter, i sin verksamhet. Också på andra håll har det framställts förslag om tvingande lagstiftning som underställer de sociala medierna konstitutionella eller folkrättsliga yttrandefrihetsnormer.<sup>197</sup>

I samma anda har David Kaye, FN:s ovan nämnde särskilde rapportör för stöd och skydd för åsikts- och yttrandefriheten, föreslagit att företagen ska ”erkänna att den auktoritativa globala standarden för att säkerställa yttrandefrihet på deras plattformar är mänskliga rättighetsnormer”. Företagen borde följaktligen omdefiniera sina innehållsriktlinjer och modereringspraxis i överensstämmelse med yttrande- och informationsfrihetsnormer.<sup>198</sup>

Samtliga dessa förslag bygger på att ”artificiellt” utsträcka de befintliga yttrande- och informationsfrihetsnormerna – såsom RF 2 kap. 1 § eller EKMR artikel 10 – så att också YouTube eller Twitter i praktiken omfattas av dem. En sådan ansvarsmodell har både fördelar och nackdelar. Fördelen är att de befintliga yttrande- och informationsfrihetsnormerna, såsom de definierats och utvecklats genom domstolspraxis, speglar en mycket eftertänksam avvägning mellan yttrande- och informationsfriheten å ena sidan, och de motstående samhällliga intressena samt andras rättigheter å andra. Vad gäller intresseavvägningen mellan olika hänsyn, rättigheter och skyldigheter hade de sociala medierna alltså inte behövt uppfinna hjulet på nytt.

Lösningen har också en del slagsidor. De konstitutionella rättighetsnormerna är förhållandevis öppna och vaga till sin utformning, och avpassade för försiktiga och eftertänksamma domstolsprocesser snarare än de ögonblicksbeslut som präglar de sociala mediernas faktiska modereringspraxis. Det framstår som osannolikt att en rimlig avvägning mellan den konstitutionella yttrandefriheten och rätten till privatliv eller skyddet för upphovsrätten hade låtit sig göras av en Facebook-anställd inom ett tiosekundersintervall.

---

<sup>196</sup> Se *Statement on behalf of a Group of Countries at the 24rd [sic] Session of the Human Rights Council*, presenterat inför FN:s råd för mänskliga rättigheter i september 2013. Tillgänglig via <https://www.business-humanrights.org/en/pdf-statement-on-behalf-of-a-group-of-countries-at-the-24rd-session-of-the-human-rights-council>.

<sup>197</sup> Jfr Everett, *Kansas Journal of Law & Public Policy* 2018 s. 130–131.

<sup>198</sup> FN:s generalförsamling, A/HCR/38/35 s. 20.

En sådan lösning hade också inneburit – om implementerad på ett seriöst och ändamålsenligt sätt av plattformarna – en extremt stor börda i deras verksamhet. Om de förväntas bedöma rättighetstvister på samma sätt som en konstitutionsdomstol hade det mycket sannolikt resulterat i oförsvarliga kostnader och arbetsbördor för företagen. Detta hade med andra ord utgjort ett ganska besvärande ingrepp i plattformarnas näringsfrihet.

### 7.3.2 *Preciserat ansvar*

En ytterligare ansvarsmodell vore att introducera särskilda skyldigheter för de sociala medierna, *ägnade* att främja yttrande- och informationsfriheten snarare än att återspegla dess materiella innehåll. Som exempel på en sådan regleringslösning kan nämnas EU:s dataskyddsförordning, som ålägger företag vissa specifika skyldigheter som i slutändan tjänar till att förbättra fysiska personers skydd för personuppgifter och privatliv.<sup>199</sup> Som mer närliggande exempel kan också nämnas de nätneutralitetsregler som gäller för internetleverantörer i EU, vars syfte delvis varit att skydda och främja yttrande- och informationsfriheten på internet.<sup>200</sup>

En rimlig utgångspunkt vore att undersöka möjligheten att tillämpa samma typ av neutralitetsprincip som gäller för internetleverantörer också för sociala medieplattformar. En sådan ansvarsmodell skulle ta sig i uttryck genom bindande så kallade *must carry*-regler – regler som förpliktigar plattformarna att neutralt förmedla den data som användarna publicerar och söker ta del av på plattformen, utan att diskriminera viss data med hänsyn till dess innehåll. En sådan modell får, liksom för internetleverantörer, antas vara ägnad att främja en fri åsiktsbildning och en användarcentrisk informationsmarknad.

Fördelen med en sådan lösningsmodell är dess precision. Det rör sig om klara skyldigheter och snävt formulerade undantag som bör gå lätt att applicera i praktiken.<sup>201</sup> Dessutom kan en sådan lagstiftningsmodell anpassas så att den inte går längre än nödvändigt får att komma till rätta med det faktiska problemet. En neutralitetsprincip skulle kunna formuleras så att den enbart är gällande för sociala medieplattformar som har en viss omfattning och ett visst inflytande.

---

<sup>199</sup> Jfr dataskyddsförordningen, skäl 1–13.

<sup>200</sup> Se ovan avsnitt 4.2.

<sup>201</sup> Jfr TSM-förordningen artikel 3.



Utan att peka ut vissa företag skulle principen kunna formuleras så att den blir tillämplig för företag vars omsättning eller marknadsandelar överstiger ett visst tröskelvärde.<sup>202</sup> På så sätt avgränsas skyldigheten så att den inte blir tillämplig för små eller nischade plattformar, vars innehållsmoderering inte har samma praktiska effekter för yttrande- och informationsfriheten.

Nackdelen med samma modell ligger i avvägningen mellan motstående hänsyn och intressen. Modellen innebär alltså en inskränkning i plattformarnas näringsfrihet. Därtill är avvägningen mellan användarnas yttrande- och informationsfrihet och andra individuella rättigheter och samhällliga intressen inte lika balanserad som vid en direkt tillämpning av yttrande- och informationsfrihetsnormer. Must carry-regler, strikt utformade, skulle innebära att material som kan uppfattas som mycket kränkande eller subversivt skulle kunna publiceras på plattformarna. I detta hänseende finns dock utrymme för anpassning, så att avvägningen blir mer tillfredställande. Must carry-regler skulle exempelvis kunna villkoras gentemot användarna med identifieringskrav och autentiseringsmekanismer, som säkerställer att användare kan hållas ansvariga för det som de själva publicerar på plattformen.<sup>203</sup> En sådan lösning kan tänkas minska förekomsten av brottsligt innehåll, kränkningar och trakasserier i sig, och därigenom jämka de negativa effekterna av en neutralitetsprincip. En regleringsmodell som föreskriver specifika skyldigheter för plattformarna, som exempelvis must carry-regler, har således också fördelen av att vara anpassningsbara till de särskilda problem som aktualiseras vid kommunikation på sociala medieplattformar.

### 7.3.3 *Effektiva rättsmedel*

En tredje lösning kan övervägas, något väsensskild från de ovan diskuterade. Denna ansvarsmodell hade inneburit att yttrande- och informationsfriheten på sociala medieplattformar fortsatt reglerades genom soft law och CSR, men att användarna erbjöds möjlighet att i efterhand överpröva modereringsbeslut i enlighet med konstitutionella yttrande- och informationsfrihetsnormer.

---

<sup>202</sup> Jfr härmed det ungefärliga tröskelvärde som gäller enligt FEUF artikel 102 och 2 kap. 7 § konkurrenslagen för företag med dominerande marknadsställning: omkring 50 % marknadsandelar på den relevanta marknaden. Se C-62/86 AKZO mot Kommissionen och C-85/76 Hoffman-La Roche mot Kommissionen.

<sup>203</sup> En sådan ordning är på sätt och vis implicit föreställd vad gäller internetleverantörernas must carry-skyldigheter, eftersom att internetleverantörerna typiskt sett har tillförlitlig information om deras användares faktiska identitet.

Modellen skulle delvis likna den som gäller inom det svenska pressetiska systemet, där redaktioner lyder under självreglering i form av de pressetiska reglerna, och publiceringsbeslut kan prövas i Pressens Opinionsnämnd i enlighet med dessa regler.<sup>204</sup>

I praktiken hade en sådan modell kunnat införlivas genom en offentligt finansierad överprövningsnämnd som särskilt sysslade med att pröva modereringsbeslut på sociala medieplattformar. Nämnden hade med fördel utrustats med mandat att utfärda bindande förelägganden mot plattformarna med syfte att avhjälpa modereringsbeslut som tagits i strid med användarens yttrande- och informationsfrihet. En raderad Twitter-post eller ett blockerat YouTube-konto kunde således bli återpublicerat eller återställt.

Fördelen med en sådan lösning vore främst att plattformarna inte belastades med en alltför stor börda i sin moderering. Twitter hade kunnat moderera enligt bästa förmåga och välvilja, och i efterhand kunde särskilt svåra eller avvikande fall bli föremål för utomstående granskning i en opartisk nämnd. En sådan ansvarslösning hade sannolikt uppfattats som en betydligt mindre inskränkning i plattformarnas näringsfrihet, då den i praktiken verkställs och bekostas av ett utomstående organ. Därtill kan en efterhandsprövning ge upphov till mer rimliga och kompetenta avvägningar mellan motstående hänsyn, intressen, rättigheter och skyldigheter, som knappast kan åstadkommas vid plattformarnas vardagliga, summariska modereringsprocesser. De stora konflikterna mellan yttrandefrihet, privatliv, allmän säkerhet och moral hade på så sätt förts tillbaka till det offentliga sfären. Detta faktum gör också lösningen principiellt attraktiv, eftersom ansvaret i grunden bygger på statens positiva förpliktelser att främja och skydda yttrande- och informationsfriheten.

Baksidan med en sådan regleringsmodell vore dess effektivitet. Den samlade modereringen på alla sociala medieplattformar uppgår till sådana volymer att enbart en bråkdel av alla modereringsbeslut hade kunnat bli föremål för efterhandsprövning. Kanske skulle detta volymproblem till viss del kunna avhjälpas genom att överprövningarna får en prejudicerande verkan, som plattformarna kan anpassa sin fortsatta modereringspraxis utefter. Men en efterhandsprövning innebär också en risk för att skadan redan är definitiv, då beslutet kan prövas och korrigeras. Både nyheter och opinionsbildning är till stor del en färskvara. En senare återpublicering av raderat innehåll, i sådana fall, skulle antagligen inte effektivt avhjälpa de negativa effekterna av den ursprungliga raderingen.

---

<sup>204</sup> Strömberg & Lundell, *Grundlagsskyddad yttrandefrihet*, s. 137–139 och Strömberg & Axberger, *Yttrandefrihetsrätt*, s. 101–108.

## 7.4 Sammanfattning

Innehållsmodereringen på sociala medier utgör ett verkligt problem ur yttrande- och informationsfrihetssynpunkt. Modereringens form och omfattning stämmer illa överens med de yttrande- och informationsfrihetsstandarder som utvecklats inom rätts-tillämpningen. Varken yttrande- och informationsfrihetsnormerna i sig, eller någon annan mellanhandsreglering, etablerar några substantiella skyldigheter för de sociala medierna i förhållande till användarnas yttrande- och informationsfrihet. I takt med ett mer omfattande mellanhandsansvar tenderar modereringen snarare att tillta, på ytterligare bekostnad av den fria informationsdistributionen.

	Infrastrukturella mellanhänder	Sociala medier (Facebook, Twitter, etc.)	Redaktionella mellanhänder
Skyldighet att iaktta neutralitet i förhållande till information från användare/tredjeparter	✓	–	–
Begränsat ansvar för information som tillhandahålls av användare/tredjeparter	✓	–	–

*Fig. 5. Regleringen av infrastrukturella och redaktionella mellanhänder jämfört sociala medier. Jfr ovan avsnitt 4.2–4.4.*

Det finns goda skäl att adressera problemet genom någon form av ansvarsregim. De ansvarsmodeller som presenterats här ovan är ett fåtal, som vardera präglas av egna för- och nackdelar. Därtill kvarstår ett stort antal problem som häri gått förbi okommenterade: frågor om jurisdiktion och territorialitet, regleringskompetens, effektiv offentlig-privat samverkan, mekanismer för regelefterlevnad, och så vidare. Sådana problem kan förhoppningsvis lösas inom ramen för någon av modellerna. Eller också kan en verkligt hållbar lösning uppstå först genom en variation av endera lösningarna, en kombination av de diskuterade lösningarna, eller, för den delen, en helt annan regleringsmodell.

## 8 Epilog

*Huru aflöper riksdagen och skriffriheten i Sverige?*

Peter Forsskål, brev till Carl von Linné (Alexandria, oktober 1761)

På en konferens i Davos 2018 talar finansmannen George Soros inför företagsledare, politiker och ekonomer från hela världen. Soros är knappast känd som en teknikpessimist eller bakåtsträvare, men under sitt anförande är han märkbart upprörd. Han talar om it-jättarna och de sociala medierna:

Vid horisonten lurar en skrämmande möjlighet. Det kan skapas en ohelig allians mellan de nya auktoritära regimerna och it-monopolens tillgångar på personlig data. Storföretagens övervakningsapparat kan förenas med den redan väl utvecklade statliga övervakningsapparaten. Det kan mycket väl sluta med ett nätverk av totalitär kontroll som inte ens Aldous Huxley eller George Orwell kunde ha föreställt sig.<sup>205</sup>

Det Soros talar om – övervakning, kartläggning och personuppgiftsinsamling – har varit på agendan under ett par år. Först med Snowden-läckorna,<sup>206</sup> och senast Cambridge Analytica-skandalen.<sup>207</sup> Den oheliga alliansen mellan företagens informationskontroll och statens övervakningsanspråk har knappast kunnat avfärdas som en paranoid dystopi.

Nu ser vi samma tendenser i fråga om yttrande- och informationsfriheten. Informationskontrollen är koncentrerad till en liten mängd mellanhänder. I brist på verkligt inflytande börjar stater runtom i världen utöva sin kontrollmakt genom mellanhänderna. Med de godaste av intentioner överlämnas ansvaret för personlig integritet, upphovsrätt och upprätthållandet av allmän moral och säkerhet till Facebook, YouTube och Twitter. Yttrande- och informationsfriheten, som ska utgöra den mest fundamentala bland de individuella fri- och rättigheterna, lämnas åt CSR-avdelningen.

Häri ligger pudelns kärna. Den innehållsmoderering som de sociala medieplattformarna vidtar på eget bevåg är värd att diskutera – likaså rätten att kunna nyttja sin yttrande- och informationsfrihet i förhållande till andra enskilda. Men den är harmlös i förhållande till den statliga moderering som kan komma att utövas genom medierna: genom nya regler om strikt mellanhandsansvar eller överenskommelser med myndigheter.

---

<sup>205</sup> Gelin & Pettersson, *Internet är trasigt* s. 92.

<sup>206</sup> Se Gellman, Tate och Soltani, *In NSA-intercepted data, those not targeted far outnumber the foreigners who are*, The Washington Post 5/7 2014.

<sup>207</sup> Se Cadwalladr & Graham-Harrison, *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach*, The Guardian 17/3 2018.

I bästa fall rör vi oss mot en ny makthierarki, där it-företagen absorberar en del av den traditionella myndighetsutövningen och rättsskipningen. I värsta fall ser vi återigen en "outsourcing" av statlig censur.

## 9 Källförteckning

### 9.1 Svenskt offentligt tryck

Prop 1973:90 med förslag till ny regeringsform och ny riksdagsordning m. m.

SOU 1972:15. Ny regeringsform, ny riksdagsordning.

SOU 1975:49. Massmediegrundlag.

SOU 1975:75. Medborgerliga fri- och rättigheter.

SOU 2012:55. En översyn av tryck- och yttrandefriheten.

SOU 2016:7. Integritet och straffskydd.

### 9.2 Utländskt offentligt tryck

#### 9.2.1 *Förenta nationerna*

FN:s generalförsamling, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (A/HRC/38/35)*

FN:s generalförsamling, *Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression (A/HRC/35/22)*

OHCHR, *Guiding Principles on Business and Human Rights (HR/PUB/11/04)*

#### 9.2.2 *Europeiska unionen*

EU-kommissionen, *Förslag till Europaparlamentets och rådets direktiv om upphovsrätt på den digitala inre marknaden COM(2016) 593 final*

EU-kommissionen, *COMMISSION RECOMMENDATION of 1.3.2018 on measures to effectively tackle illegal content online C(2018) 1177 final*

EU-kommissionen, *Code of Conduct on illegal online hate speech*  
[https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/countering-illegal-hate-speech-online\\_en](https://ec.europa.eu/info/policies/justice-and-fundamental-rights/combating-discrimination/racism-and-xenophobia/countering-illegal-hate-speech-online_en).

### 9.2.3 *Förenta staterna*

United States Congress, *Allow States and Victims to Fight Online Sex Trafficking Act of 2017*, Public Law No: 115–164, 11 april 2018 (cit: FOSTA-SESTA)

Federal Communications Committee, *In the Matter of Preserving the Open Internet Broadband Industry Practices - REPORT AND ORDER*, GN Docket No. 09–191, 2010, 21 december 2010 (cit: FCC, Open Internet Order (2010))

Federal Communications Committee, *In the Matter of Protecting and Promoting the Open Internet – REPORT AND ORDER ON REMAND, DECLARATORY RULING, AND ORDER*, GN Docket No. 14–28, 2015, 26 februari 2015 (cit: FCC, Open Internet Order (2015)).

Federal Communications Committee, *In the Matter of Restoring Internet Freedom - DECLARATORY RULING, REPORT AND ORDER, AND ORDER*, WC Docket No. 17–108, 2017, 14 december 2017 (cit: FCC, Restoring Internet Freedom Order (2017))

### 9.3 Litteratur och tidskrifter

Axberger, H-G, *Yttrandefrihetsgrundlagarna*, 3 uppl., Wolters Kluwer 2017

Axberger, H-G, *Den svenska tryckfrihetens rättsliga utveckling*, i Wennberg, B & Örtén, K (red) *Fritt ord 250 år, Tryckfrihet och offentlighet i Sverige och Finland – ett levande arv från 1766* (s. 241–289), Sveriges Riksdag 2016

Baade, B, *Fake News and International Law*, *The European Journal of International Law* 2018 (29:4) s. 1358–1376

Barendt, E, *Freedom of Speech*, 2 uppl., Oxford University Press 2005

Barzilai-Nahon, *Toward a Theory of Network Gatekeeping: A Framework for Exploring Information Control*, *Journal of the American Society for Information Science and Technology* 07/2008 (59:9) s. 1493–1512

Boye, K, *Kallocain*, Albert Bonniers Förlag 2014

Boyd, D, Ellison, N, *Social Network Sites: Definition, History, and Scholarship*, *Journal of Computer-Mediated Communication* 2007 (13:1) s. 210–230

Bring, O, *De mänskliga rättigheternas väg*, Atlantis 2011

Bucher, T, *Want to be on the top? Algorithmic power and the threat of invisibility on Facebook*, *New Media & Society* 2012 (14:7) s. 1164–1180

Bucher, T, *The algorithmic imaginary: exploring the ordinary affects of Facebook algorithms*, *Information, Communication & Society* 2017 (20:1) s. 30–44

Bull, T, *På yttrandefrihetens yttersta gren*, *JT* 2005/06 s. 514–535

Burns, K, *Social Media – A Reference Handbook*, ABC-CLIO 2017

Cameron, *An Introduction to the European Convention on Human Rights*, 7 uppl., Iustus 2014

Chandler, D, & Munday, R (red), *A Dictionary of Social Media*, Oxford University Press 2016



Crawford, K & Gillespie, T, *What is a flag for? Social media reporting tools and the vocabulary of complaint*, *New Media & Society* 2016 (18:3) s. 410–428

Eek, H, *Om tryckfriheten*, Geber 1942

Eka, A, Hirschfeldt, J, Jermsten, H & Svahn Starrsjö, K, *Regeringsformen – med kommentarer*, 2 uppl., Karnov Group 2018

Everett, C, *Free Speech on Privately-Owned Fora: A Discussion on Speech Freedoms and Policy for Social Media*, *Kansas Journal of Law & Public Policy* 2018 (28:1) s. 113–145

Forsskål, P, *Tankar om borgerliga friheten*, Bokförlaget Rediviva 1984

Funcke, N, *Offentlighetsprincipen – praktik och teori*, Studentlitteratur 2014

Gelin, M & Pettersson, K, *Internet är trasigt – Silicon Valley och demokratins kris*, Natur & Kultur 2019

Gillespie, T, *Custodians of the internet: platforms, content moderation, and the hidden decisions that shape social media*, Yale University Press 2018

Gillespie, T, *The Relevance of Algorithms* i Gillespie, T, Boczkowski, P & Foot, K (red) *Media Technologies: Essays on Communication, Materiality, and Society* (s 167–194), MIT Press 2014

Helberger, N, Kleinen-von Königslöw, K & van der Noll, R, *Regulating the new information intermediaries as gatekeepers of information diversity*, *Info: The Journal of Policy, Regulation and Strategy for Telecommunications, Information and Media* 2015 (17:6) s. 50–71

Hirschfeldt, J, *Konstitutionellt yttrandefrihetsskydd i Sverige* i Wennberg, B & Örtengren, K (red) *Fritt ord 250 år, Tryckfrihet och offentlighet i Sverige och Finland – ett levande arv från 1766* (s. 591–639), Sveriges Riksdag 2016

Hitz, J, *Removing Disfavored Faces from Facebook: The Freedom of Speech Implications of Banning Sex Offenders from Social Media*, *Indiana Law Journal* 2014 (89:3) s. 1327–1363

- Holmberg, E, & Stjernquist, N, *Vår författning*, 13 uppl., Norstedts juridik 2003
- Huxley, A, *Du sköna nya värld*, Wahlström & Widstrand 1984
- Introna, L, *Algorithms, Governance, and Governmentality: On Governing Academic Writing*, Science, Technology, & Human Values 2016 (41:1) s. 17–49
- Jareborg, N, *Rättsdogmatik som vetenskap*, SvJT 2004 s. 1–10
- Kleineman, J, *Rättsdogmatisk metod* i Nääv, M & Zamboni, M (red) *Juridisk metodlära* (s. 21–77), Studentlitteratur 2018
- Laidlaw, E, *Regulating Speech in Cyberspace – Gatekeepers, Human Rights and Corporate Responsibility*, Cambridge University Press 2015
- Laidlaw, E, *A framework for identifying Internet information gatekeepers*, International Review of Law, Computers & Technology 2010 (24:3) s. 263–276
- Langvardt, K, *Regulating Online Content Moderation*, Georgetown Law Journal 2018 (106:5) s. 1353–1358
- Lebeck, C, *EU-stadgan om grundläggande rättigheter*, 2 uppl., Studentlitteratur 2016
- Lewin, K, *Forces Behind Food Habits and Methods of Change* i *The Problem of Changing Food Habits: Report of the Committee on Food Habits 1941–1943* (s. 35–65), National Research Council 1943
- Lin, K, Lu, H, *Why people use social networking sites: An empirical study integrating network externalities and motivation theory*, Computers in Human Behavior 2011 (17:3) s. 1152–1161
- Locke, J, *An Essay Concerning Human Understanding*, Clarendon Press 1975
- Manning, J, *Social Media, Definition and Classes of* i Harvey, K (red) *Encyclopedia of Social Media and Politics* (s. 1158–1161), SAGE Publications 2014
- Marsden C, *Network neutrality: from policy to law to regulation*, Manchester University Press 2017

- Mill, J S, *On Liberty*, Penguin Books 2010
- Milton, J, *Areopagitica; A Speech of Mr John Milton for the Liberty of Vnlicenc'd Printing, to the Parliament of England*, 1644
- Nozick, R, *Anarki, Stat och Utopi*, Timbro 2001
- Nygren, R, *Sex brytpunkter i svensk konstitutionell rätt* i Karlsson, N, Berggren, N, & Nergelius, J, (red) *Demokratins konstitutionella val* (s. 15–49), City University Press 2001
- Nygren, R, *TF 1766 i sin historiska och rättsliga kontext* i Wennberg, B & Örtenhed, K (red) *Fritt ord 250 år, Tryckfrihet och offentlighet i Sverige och Finland – ett levande arv från 1766* (s. 167–205), Sveriges Riksdag 2016
- Orwell, G, *1984*, Bokförlaget Atlantis 2010
- Pasquale, F, *Platform Neutrality: Enhancing Freedom of Expression in Spheres of Private Power*, *Theoretical Inquiries in Law* 2016 (17:2) s. 487–513
- Peczenic, A, *Juridikens allmänna läror*, SvJT 2005 s. 249–272
- Peters, J D, *Courting the Abyss: Free Speech and the Liberal Tradition*, University of Chicago Press 2005
- Regner, G, *En utdragen grundlagshistoria*, SvJT 1991 s. 428–444
- Reichel, J, *EU-rättslig metod* i Nääv, M & Zamboni, M (red) *Juridisk metodlära* (s. 109–142), Studentlitteratur 2018
- Sandgren, C, *Rättsvetenskap för uppsatsförfattare*, 4 uppl., Norstedts juridik 2018
- Savin, A, *EU Internet Law*, Edward Elgar Publishing Limited 2013
- Sevastik, P (red), Nyman-Metcalf, K, Spiliopoulou Åkermark, S & Mårsäter, O, *En bok i folkrätt*, Norstedts juridik 2009
- Shaw, M, *International Law*, 7 uppl., Cambridge University Press 2014

Snickars, P, & Vonderau, P, *Introduction* i Snickars, P, & Vonderau, P (red), *The Youtube Reader* (s. 9–21), National Library of Sweden 2009

Spaventa, E, *Fundamental Rights in the European Union* i Barnard, C & Peers, S (red), *European Union Law* (s. 226–254), Oxford University Press 2014

Strömberg, H & Lundell, B, *Grundlagsskyddad yttrandefrihet*, 15 uppl., Studentlitteratur 2013

Strömberg, H & Axberger, H-G, *Yttrandefrihetsrätt*, Studentlitteratur 2004

Thompson, M, *Beyond Gatekeeping: The Normative Responsibility of Internet Intermediaries*, *Vanderbilt Journal of Entertainment and Technology Law* 2016 (18:4) s. 783–848

Trenz, H-J, *Public Sphere* i Gianpietro Mazzoleni (red) *International Encyclopedia of Political Communication*, 1 uppl., Wiley 2015

Ubena, J, *How to Regulate Information and Communications Technology? A Jurisprudential Inquiry into Legislative and Regulatory Techniques*, Jure Förlag 2015

van Dijk, J, *Facebook and the engineering of connectivity: A multi-layered approach to social media platforms*, *Convergence: The International Journal of Research into New Media Technologies* 2012 (19:2) s. 141–155

Verstraete, M, Bambauer, D & Bambauer, J, *Identifying and Countering Fake News*, *Arizona Legal Studies Discussion Paper No. 17-15* 2007

Von Vegesack, T, *När ordet blev fritt*, Natur och Kultur 2003

Von Vegesack, T, *Smak För Frihet: Opinionsbildningen i Sverige 1755–1830*, 2 uppl., Natur och Kultur 2001

Weaver, R, *Free Speech and Democracy in an Internet Era* i Lind, A-S, Reichel, J & Österdahl, I (red) *Information and Law in Transition* (s. 25–32), Liber 2015

Wu, T, *Network Neutrality, Broadband Discrimination*, *Journal of Telecommunications and High Technology Law* 2003 (2) s. 141–176

Zamjatin, J, *Vi*, Modernista 2016

Zittrain, J, *A History of Online Gatekeeping*, Harvard Journal of Law & Technology 2006 (19:2) s. 253–298

Österdahl, I, *Åke Green och missaktande men inte hatiskt tal*, SvJT 2006 s. 213–226

## 9.4 Domar från svenska och utländska domstolar

### 9.4.1 Svenska domstolar

NJA 1991 s. 155

NJA 2005 s. 805

NJA 2006 s. 467

NJA 2007 s. 747

NJA 2012 s. 400

NJA 2015 s. 45

NJA 2015 s. 899

HFD 2015 ref. 80

### 9.4.2 EU-domstolen

Dom av den 13 februari 1979 i målet C-85/76, *Hoffman-La Roche mot Kommissionen*

Dom av den 3 juli 1991 i målet C-62/86, *AKZO mot Kommissionen*

Dom av den 23 mars 2010 i förenade målen C-236/08–C-238/08, *Google France mot Louis Vuitton*

Dom av den 12 juli 2011 i målet C-324/09, *L'Oréal mot eBay*

### 9.4.3 Europadomstolen

*Ahmet Yildirim v. Turkey*, application no. 3111/10, dom av den 18 december 2012

*Appleby and Others v. the United Kingdom*, application no. 44306/98, dom av den 6 maj 2003

*Ashby Donald and others v. France*, application no. 36769/08, dom av den 10 januari 2013

*Axel Springer AG v. Germany*, application no. 39954/08, dom av den 7 februari 2012

*Delfi AS v. Estonia*, application no. 64569/09, dom av den 10 oktober 2013

*Editorial Board of Pravoye Delo and Shtekel v. Ukraine*, application no. 33014/05, dom av den 5 maj 2011

*Féret v. Belgium*, application no. 15615/07, dom av den 16 juli 2009

*Gündüz v. Turkey*, application no. 35071/97, dom av den 4 december 2003

*Handyside v. the United Kingdom*, application no. 5493/72, dom av den 7 december 1976

*Hertel v. Switzerland*, application no. 25181/94, dom av den 25 augusti 1998

*Jerusalem v. Austria*, application no. 26958/95, dom av den 27 februari 2007

*Khurshid Mustafa and Tarzibachi v. Sweden*, application no. 23883/06, dom av den 16 december 2008

*K.U. v. Finland*, application no. 2872/02, dom av den 2 december 2008

*Leroy v. France*, application no. 36109/03, dom av den 2 oktober 2008

*Lindon, Otchakovsky-Laurens and July v. France*

*Lingens v. Austria*, application no. 9815/82, dom av den 8 juli 1986

*Mouvement Raëlien Suisse v. Switzerland*, application no. 16354/06, dom av den 13 juli 2012

*Müller v Switzerland*, application no. 10737/84, dom av den 24 maj 1988

*Oberschlick v. Austria*, application no. 11662/85, dom av den 23 maj 1991

*Pedersen and Baadsgaard v. Denmark*, application no. 49017/99, dom av den 16 juni 2003

*Steel and Morris v. the United Kingdom*, application no. 68416/01, dom av den 15 februari 2005

*Sürek v Turkey*, application no. 26682/95, dom av den 8 juli 1999

*Times Newspapers Ltd (nos. 1 and 2) v. the United Kingdom*, applications 3002/03 och 23676/03, dom av den 10 mars 2009

*Vejdeland and Others v. Sweden*, application no. 1813/07, dom av den 9 februari 2012

*Verein gegen Tierfabriken (VgT) v. Switzerland*, application no. 32772/02, dom av den 30 juni 2009

*X and Y v. the Netherlands*, application no. 8978/80, dom av den 26 mars 1985

*Zana v. Turkey*, application no. 18954/91, dom av den 25 november 1997

*Özgür Gündem v. Turkey*, application no. 23144/93, dom av den 16 mars 2000

#### 9.4.4 Amerikanska domstolar

*Cubby, Inc. v. Compuserve, Inc.*, 776 F.Supp. 135 (S.D.N.Y 1991) (cit: Cubby v. CompuServe)

*Doe v. Prosecutor, Marion County, Indiana*, 705 F.3d 694 (7<sup>th</sup> Cir. 2013) (cit: Doe v. Marion County)

*Grady Auvil et ux; et al. v CBS "60 Minutes" et al.*, 800 F. Supp. 928 (E.D. Wash. 1992) (cit: Auvil v. CBS 60 Minutes)

*Marsh v. State of Alabama*, 326 U.S. 501 (cit: Marsh v. Alabama)

*Stratton Oakmont, Inc. v. Prodigy Services Co.*, 1995 WL 323710 (N.Y. Sup. Ct. 1995) (cit: Stratton Oakmont v. Prodigy)

*Verizon Communications Inc. v. Federal Communications Commission*, 740 F.3d 623 (D.C. Cir. 2014) (cit: Verizon v. FCC)



## 9.5 Internetkällor

### 9.5.1 Facebook

Facebook, *Terms of Service*, <https://www.facebook.com/terms.php> (9/4 2019)

Facebook, *Community Standards*, <https://www.facebook.com/communitystandards/> (12/3 2019)

Facebook, *Company Info*, <https://newsroom.fb.com/company-info/> (12/3 2019)

Mosseri, A, *News Feed Ranking in Three Minutes Flat*, Facebook 22/5 2018 (9/4 2019)  
<https://newsroom.fb.com/news/2018/05/inside-feed-news-feed-ranking/>

Zuckerberg, M, *A Privacy-Focused Vision for Social Networking*, Facebook 6/3 2019  
(7/4 2019) <https://www.facebook.com/notes/mark-zuckerberg/a-privacy-focused-vision-for-social-networking/10156700570096634/>

### 9.5.2 Twitter

Twitter, *Twitter Terms of Service*, <https://twitter.com/en/tos> (9/4 2019)

Twitter, *The Twitter Rules*, <https://help.twitter.com/en/rules-and-policies/twitter-rules> (9/4 2019)

Twitter, *Business*, <https://business.twitter.com/>, (7/4 2019)

Twitter, *Hateful conduct policy*, <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy> (9/4 2019)

Twitter, *Violent threats and glorification of violence*, <https://help.twitter.com/en/rules-and-policies/violent-threats-glorification> (9/4 2019)

Twitter, *Violent extremist groups*, <https://help.twitter.com/en/rules-and-policies/violent-groups> (9/4 2019)

Twitter, *Twitter media policy*, <https://help.twitter.com/en/rules-and-policies/media-policy> (9/4 2019)

Twitter, *About private information on Twitter*, <https://help.twitter.com/en/rules-and-policies/personal-information> (9/4 2019)

Twitter, *Our approach to policy development and enforcement philosophy*, <https://help.twitter.com/en/rules-and-policies/enforcement-philosophy> (10/4 2019)

### 9.5.3 *Tumblr*

Tumblr, *Terms of Service*, <https://www.tumblr.com/policy/en/terms-of-service> (9/4 2019)

Tumblr, *Community Guidelines*, <https://www.tumblr.com/policy/en/community> (9/4 2019)

Tumblr, *Updates to Tumblr's Community Guidelines*, <https://support.tumblr.com/post/180758979032/updates-to-tumblrs-community-guidelines> 3/2 2018 (9/4 2019)

Tumblr, *Content moderation on Tumblr*, <https://tumblr.zendesk.com/hc/en-us/articles/360011799473-Content-moderation-on-Tumblr> (10/4 2019)

D'Onofrio, J, *A better, more positive Tumblr*, <https://staff.tumblr.com/post/180758987165/a-better-more-positive-tumblr> 3/12 2018 (8/1 2019)

### 9.5.4 *YouTube*

YouTube, *Användarvillkor för Tjänsten*, <https://www.youtube.com/t/terms> (9/4 2019)

YouTube, *Youtube för pressen*, <https://www.youtube.com/intl/sv/yt/about/press/>, (7/4 2019)

YouTube, *Riktlinjer för Communityn*, <https://www.youtube.com/intl/sv/yt/about/policies/#community-guidelines> (9/4 2019)

YouTube, *Policy om trakasserier och cybermobbing*, [https://support.google.com/youtube/answer/2802268?hl=sv&ref\\_topic=2803176](https://support.google.com/youtube/answer/2802268?hl=sv&ref_topic=2803176) (9/4 2019)

YouTube, *Policy om hatretorik*,  
[https://support.google.com/youtube/answer/2801939?hl=sv&ref\\_topic=2803176](https://support.google.com/youtube/answer/2801939?hl=sv&ref_topic=2803176) (9/4  
2019)

YouTube, *Policyer om skadligt eller farligt innehåll*,  
[https://support.google.com/youtube/answer/2801964?hl=sv&ref\\_topic=2803176](https://support.google.com/youtube/answer/2801964?hl=sv&ref_topic=2803176) (9/4  
2019)

YouTube, *Försäljning av olagliga eller reglerade varor*,  
[https://support.google.com/youtube/answer/9229611?hl=sv&ref\\_topic=2803176](https://support.google.com/youtube/answer/9229611?hl=sv&ref_topic=2803176) (9/4  
2019)

YouTube, *Policyer för innehåll med skjutvapen*,  
[https://support.google.com/youtube/answer/7667605?hl=sv&ref\\_topic=2803176](https://support.google.com/youtube/answer/7667605?hl=sv&ref_topic=2803176) (9/4  
2019)

YouTube, *Brottsliga organisationer*,  
[https://support.google.com/youtube/answer/9229472?hl=sv&ref\\_topic=2803176](https://support.google.com/youtube/answer/9229472?hl=sv&ref_topic=2803176) (9/4  
2019)

YouTube, *Policyer om våldsamt eller chockerande innehåll*,  
[https://support.google.com/youtube/answer/2802008?hl=sv&ref\\_topic=2803176](https://support.google.com/youtube/answer/2802008?hl=sv&ref_topic=2803176) (9/4  
2019)

YouTube, *Policy om nakenhet och sexuellt innehåll*,  
[https://support.google.com/youtube/answer/2802002?hl=sv&ref\\_topic=2803176](https://support.google.com/youtube/answer/2802002?hl=sv&ref_topic=2803176) (9/4  
2019)

YouTube, *Policyer för spam, bedrägligt beteende och bedrägerier*,  
[https://support.google.com/youtube/answer/2801973?hl=sv&ref\\_topic=2803176](https://support.google.com/youtube/answer/2801973?hl=sv&ref_topic=2803176) (9/4  
2019)

YouTube, *YouTube-programmet Betrodda flaggare*,  
[https://support.google.com/youtube/answer/7554338?hl=sv&ref\\_topic=2803138](https://support.google.com/youtube/answer/7554338?hl=sv&ref_topic=2803138) (10/4  
2019)

YouTube, *Vikten av sammanhang*,

<https://support.google.com/youtube/answer/6345162?hl=sv> (10/4 2019)

YouTube, *Grundläggande information om varningar om brott mot communityns*

*riktlinjer*, <https://support.google.com/youtube/answer/2802032?hl=sv> (10/4 2019)

#### 9.5.5 Instagram

Instagram, *Användarvillkor*, <https://www.tumblr.com/policy/en/terms-of-service> (9/4

2019)

Instagram, *Communityregler*, <https://help.instagram.com/477434105621119> (9/4 2019)

#### 9.5.6 Övrigt

Abell, J, *March 21, 2006: Twitter Takes Flight*, Wired 21/3 2011 (3/6 2019),

<https://www.wired.com/>

Article 19, *Twitter Rules: Analysis against international standards on freedom of expression*, Article 19 5/9 2018 (29/4 2019), <https://www.article19.org/>

Article 19, *ARTICLE 19 and 70+ civil society groups write to Mark Zuckerberg over content removal*, Article 19 13/11 2018 (29/4 2019), <https://www.article19.org/>

Austin, E, *We can no longer talk about sex on Facebook in Europe*, EDRI 16/1 2019 (29/4 2019), <https://edri.org/>

Cadwalladr, C & Graham-Harrison, E, *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach*, The Guardian 17/3 2018 (3/6 2019), <https://www.theguardian.com/>

Dorsey, J & Harris, S, *Making Sense with Sam Harris, #148 – Jack Dorsey*, Making Sense Podcast 5/2 2019 (4/6 2019), <https://samharris.org/podcast/>

EDRI, *EU Commission's Recommendation: Let's put internet giants in charge of censoring Europe*, EDRI 1/3 2018 (29/4 2019), <https://edri.org/>

Gellman, B, Tate, J och Soltani, A, *In NSA-intercepted data, those not targeted far outnumber the foreigners who are*, The Washington Post 5/7 2014 (3/6 2019), <https://www.washingtonpost.com/>

Hansen, E E, *Dear Mark. I am writing this to inform you that I shall not comply with your requirement to remove this picture*, Aftenposten 8/9 2016 (29/4 2019), <https://www.aftenposten.no/>

Harmon, E, *Facebook's Sexual Solicitation Policy is a Honeypot for Trolls*, EFF 7/12 2018 (29/4 2019), <https://www.eff.org/>

Houston, T, *The Facebook story: from inception to IPO*, The Verge 18/5 2012 (3/6 2019), <https://www.theverge.com/>

Human Rights Watch, *An Open Letter to Mark Zuckerberg, The World's Freedom of Expression is in Your Hands*, Human Rights Watch 14/11 2018 (29/4 2019), <https://www.hrw.org/>

Maebe, J, *ENDitorial: The fake fight against fake news*, EDRi 25/7 2018 (29/4 2019), <https://edri.org/>

Swisher, K & Wagner, K, *Here's the transcript of Recode's interview with Facebook CEO Mark Zuckerberg about the Cambridge Analytica controversy and more*, Recode 22/3 2018 (12/3 2019), <https://www.recode.net/>

*Statement on behalf of a Group of Countries at the 24rd [sic] Session of the Human Rights Council*, Business & Human Rights Resource Centre 1/9 2013 (3/6 2019) <https://www.business-humanrights.org/>

Trendacosta, K & York, J, *Dear Tumblr: Banning "Adult Content" Won't Make Your Site Better But It Will Harm Sex-Positive Communities*, EFF 5/12 2018 (29/4 2019), <https://www.eff.org/>

York, J, *Blunt Policies and Secretive Enforcement Mechanisms: LGBTQ+ and Sexual Health on the Corporate Web*, EFF 24/10 2018 (29/4 2019), <https://www.eff.org/>