



LUNDS
UNIVERSITET

Despoti online: En tematisk analys av Twitters regelverk utifrån Fuller och Bourdieu.

Av: Joakim Marklund

LUNDS UNIVERSITET
Rättssociologiska institutionen

Kandidatuppsats (RÄSK02)
Vårterminen 2021



Handledare: Peter Bergwall

Examinator: Per Wickenberg

Abstract

As our lives ever increasingly move online, the influence of social media is constantly growing. It has however proven difficult for legislators to keep up with rapid development of online spaces. Therefore, legislative control has largely been left to the social media platforms. To see if this is a good solution, this paper aims to analyze Twitters terms of use, rules, guidelines, and policies, in order to see how they measure up to established legal principles such as due process, equality under the law, and the rule of law as well as desiderata taken from the writings of Lon Fuller. As a result of a thematic analysis with theories from Lon Fuller and Pierre Bourdieu as well as themes based in legal principles, some glaring problems within both Twitters terms of use and the rules, guidelines and policies that regulate behavior could be uncovered. The most common problem being a lack of clarity that risks alienating large portions of the userbase from a clear understanding of the rules that govern them. Most troubling however is the despotic nature of the twitter system of rules. The paper concludes with some suggestions on how the legislation of online spaces could improve.

Key words: Desiderata, Legal principles, Lon Fuller, Pierre Bourdieu, Social media

Innehållsförteckning

1. Inledning	1
1.1. Syfte och frågeställning.....	2
1.2. Bakgrund	2
1.3. Avgränsningar	4
1.4. Rättssociologisk relevans	5
2. Disposition.....	5
3. Tidigare forskning.....	6
3.1. Litteratursökning	6
3.2. Sammanfattning av tidigare forskning	7
4. Teori.....	11
4.1 Fuller: Desiderata	11
4.2. Bourdieu: Fält, doxa och praktiskt förnuft	12
5. Metod.....	13
5.1. Val av teman.....	13
5.2. Urval.....	14
5.3. Övrigt material	14
5.4. Tillvägagångssätt.....	15
5.5. Presentation och analys av material	15
5.6. Generaliserbarhet och etiska överväganden	16
6. Resultat.....	16
6.1. Genomgång av teman.....	16
6.2. Problematiska policys	23
6.2.1. Hateful conduct policy.....	23
6.2.2. Coordinated harmful activity	25
7. Diskussion	27

7.1. I förhållande till rättsliga principer.....	27
7.2. I förhållande till samhällskonsekvenser	29
8. Slutsats.....	32
8.1. Förslag till lagstiftare	34
9. Framtida forskning	35
Referenser	36
Bilaga 1: Litteratursökningar	42
Bilaga 2: Teman.....	45

Begrepp och förkortningar

Desiderata

Betyder önskemål, används av Lon Fuller för att beskriva vad ett rättssystem behöver uppfylla. Används även i singular: Desideratum

Despoti

Människostyre. Styre som kommer från den eller de som agerar som suverän där suveränen står över de regler som finns. I kontrast till nomokrati (regelstyre).

Tweet

Primära meddelandeformen på Twitter. Ett kort meddelande på max 280 bokstäver som läggs ut offentligt.

ToS

Terms of use, terms of service, terms and agreements. Olika plattformar använder olika namn på deras användaravtal. Har därför valt att förkorta alla med *ToS* för att undvika förvirring.

1. Inledning

Under det senaste decenniet har sociala medier blivit mer och mer centralt i vårt samhälle. Med teknologisk utveckling har samhällsdebatten alltmer flyttats över online, där den i stor utsträckning sker på sociala medieplattformar. Plattformarna har varit ett fantastiskt verktyg för kommunikation samt för att ge en möjlighet att göra sin röst hörd till människor som tidigare inte haft tillgång till liknande megafoner. Samtidigt har det gett utrymme för aktörer med mindre goda intentioner att utnyttja plattformarna för desinformation, hot och andra destruktiva beteenden. Detta har lett till en evigt ökande påtryckning på företagen, från olika håll, att reglera vad som sägs på deras plattformar. Svaret har blivit att utveckla regelverk genom *ToS*, regler och riktlinjer som agerar som lag på plattformarna. Eftersom utvecklingen har gått så pass fort, har det varit svårt för traditionella rättssystem att hänga med i regleringen av internet. Därför har plattformarnas regelverk till stor del fått hålla i taktipinnen för att reglera beteenden online. Men hur väl fungerar det egentligen? Den här uppsatsen tacklar denna fråga från ett rättssociologiskt perspektiv genom att använda en tematisk analys för att undersöka Twitters regelverk. Med stöd i teorier från Lon Fuller och Pierre Bourdieu, undersöks hur väl regelverket förhåller sig till rättsliga principer samt vilka samhällskonsekvenser som kan komma av att fortsätta lämna över kontrollen till plattformarna.

1.1. Syfte och frågeställning

Syftet med denna uppsatts är att, med hjälp av en tematisk analys av Twitters *ToS*, *Rules and policies* och *General guidelines and policies*, undersöka i vilken utsträckning regelverket förhåller sig till rättsliga principer och därmed lämpar sig som den primära regleringen online. Detta för att få en insikt i hur regleringen på plattformarna fungerar och vilka potentiella konsekvenser ett fortsatt överlåtande av regleringen som sker online kan få i samhället. För att undersöka detta utgår uppsatsen från följande forskningsfrågor:

- I vilken utsträckning följer Twitters *ToS*, *Rules and Policies* och *General guidelines and policies* rättsliga principer?
- Vad kan potentiella samhällskonsekvenser bli av att tillåta sociala medieföretagen fortsätta styra regleringen av nätet?

1.2. Bakgrund

Det förekommer en konstant diskurs angående sociala medier och hur de ska modereras. Från ena sidan kommer ett tryck för mer reglering där man anser att sociala medieföretagen måste göra bättre jobb med att ta bort innehåll med hatisk, hotfull, missledande eller olaglig information. Medan det från den andra sidan lyfts fram att plattformarna censurerar för mycket samt att de har en bias i hur de censurerar. Efter upploppet vid Kapitolium i Washington DC, den sjätte januari i år (Levinson, 2021) ställdes båda dessa perspektiv emot varandra. Många ansåg att sociala medieplattformarna bar ett ansvar, då stora delar av protesterna var organiserade bland annat via Facebook (Sharma, 2021). Detta resulterade i en stor rensning på plattformarna av individer som följde den så kallade QAnon (Wendling, 2021). Samtidigt var det många som reagerade över Twitters avstängning av Amerikas dåvarande president Donald J. Trump (Twitter Inc., 2021). Bland annat kom Tysklands Förbundskansler Angela Merkel med kritik där hon fördömde avstängningen och ansåg det problematiskt att privata företag kan förtrycka yttrandefriheten, även Frankrikes finansminister Bruno Le Maire gick ut hårt mot Twitter och kallade plattformarna för "ett av demokratins största hot" (TT, 2021). Avstängningen gick även emot ett utlåtande som Twitter tidigare gett där de sagt att de inte skulle stänga

av världsledare annat än vid grova övertramp (Twitter Inc., 2019). En annan händelse som blev väldigt uppmärksammas var när New York Post, inför det amerikanska presidentvalet 2020, släppte en artikel angående en laptop tillhörande dåvarande presidentkandidat Joe Bidens son Hunter. Laptopsen skulle tydligen innehålla en del information som potentiellt kunde vara skadligt för Bidens kampanj (Morris & Fonrouge, 2020). Twitter var snabba att plocka bort artikeln och stänga av New York Post med argumentet att det var potentiellt hackat material (Manskar, 2020). De gav dock inget stöd för argumentet och blev snabbt anklagade för att förtrycka journalistisk frihet och för att försöka hjälpa Bidens kampanj genom att förtrycka negativ information (Manskar, 2020). Twitter gjorde sedan ändringar i sin policy om hackat material (Gadde, 2020) och tog tillbaka avstängningen av New York Post som respons på kritiken (McMillan, 2020).

När det kommer till rättslig reglering av plattformarna diskuteras det även mycket, på flera håll. I USA har problematiken kring sociala medieplattformarnas inflytande diskuteras frekvent. Bland de mest noterbara är kommentarer från Clarence Thomas, domare i USAs högsta domstol. I en dom angående huruvida valda tjänstemän ska ha rätt att blockera väljare på Twitter skrev Thomas, att stoppa tillvalda från att blockera väljare förvisso faller i linje med yttrandefrihetslagstiftningen. Men det skulle skapa en problematik genom att ge privata företag mer makt över yttrandefriheten än folkvalda politiker har, både när det gäller att yttra sig till politikerna och politikernas egen förmåga att yttra sig. Detta genom att en dom som stoppar politikerna från att blockera folk, inte skulle stoppa Twitter från att stänga av folket från plattformen och därmed uppfylla samma effekt som man vill stoppa (BIDEN v. KNIGHT , 2021). Thomas fortsatte med att kommentera att en diskussion behövde föras kring huruvida plattformarna skulle ha lika skyldigheter som exempelvis brevbärare när det kommer till informationen de bär (BIDEN v. KNIGHT , 2021).

Även inom EU har en del diskussioner förts. I en debatt i Europaparlamentet i februari tidigare i år fördes argument från olika perspektiv fram. Dels argumenterades för ett större ansvar för plattformarna i att kontrollera hat och desinformation medan andra lyfte fram problematiken kring plattformarnas kontroll över yttrandefrihet och vad som anses vara acceptabel diskurs. Kontentan var att det behövs lagar, inte riktlinjer för plattformarna (Europaparlamentet, 2021).

Europeiska kommissionen har även lagt fram två lagförslag ämnade att reglera plattformarna *rättsakten om digitala marknader* (EUROPEISKA KOMMISSIONEN, 2020) och *rättsakten om digitala tjänster* (EUROPEISKA KOMMISSIONEN, 2020). *Rättsakten om digitala marknader* inriktar sig främst på att göra marknaden mer tillgänglig. Reducera så kallade "gate-keepers" möjligheter att kontrollera marknaden och på så sätt främja innovation och ge konsumenter större valmöjligheter (EUROPEISKA KOMMISSIONEN, 2020). *Rättsakten om digitala tjänster* är något bredare. Men, för användare, som är fokus för denna uppsats, ämnar den bland annat att öka transparens i plattformarnas *ToS*, tydliggöra varför innehåll rekommenderas till användare, ge användare rätt att välja bort annonsering som är baserat på profilering. Den skulle även lägga ett större ansvar på de största plattformarna att skydda användare från illegalt eller skadligt innehåll och ge bättre tillgång till auktoriteter och forskare, så att vi kan få en bättre förståelse för plattformarnas påverkan på samhället (EUROPEISKA KOMMISSIONEN, 2020).

1.3. Avgränsningar

Sociala medier har blivit en del av vardagen för de flesta och är integrerade i väldigt stora delar av människors liv. Det går därför att analysera dem från ett nästan oändligt antal vinklar. Det har därför varit viktigt att göra en del avgränsningar. Uppsatsen fokuserar främst på sociala medieföretagens sätt att styra sina plattformar och vilka konsekvenser det kan få. På grund av tidsbegränsning har tematiska analysen begränsats till att analysera Twitter specifikt. Upplägget med *ToS* och algoritmer används dock av de flesta plattformarna såsom Facebook eller Instagram, med stöd i tidigare forskning gjord även på andra plattformar finns det ett visst utrymme för generaliserbarhet. Även Twitters regelverk är mycket brett, därför kommer den tematiska analysen begränsas till de delar av regelverket som riktar sig mot Twitters användare i form av *ToS*, *Rules and policies* och *General guidelines and policies*.

En annan avgränsning har gjorts när det kommer till andra parter beteende på plattformarna. Detta involverar individuella användares beteende såsom hot, hat, förtal osv. men även tredjepartsföretags agerande såsom Cambridge Analytics involvering i

amerikanska presidentvalet 2016 (Detrow, 2018) samt företag som manipulerar algoritmer.

Avgränsningar har även gjorts när det kommer till Twitters kompetens i sin reglering, exempelvis huruvida algoritmerna har förmåga att fånga alla olika felstavelser av hot som förekommer eller ej.

Dessa avgränsningar kan till viss del ses som mjuka avgränsningar eftersom uppsatsen tar ett brett perspektiv på Twitters regelverk som till stor del påverkas av exempelvis användares beteende och algoritmernas förmåga. Därför kan man inte helt frånga att åtminstone hålla dem i åtanke.

1.4. Rättssociologisk relevans

Rättssociologin befinner sig i en ganska unik position i att kunna undersöka ett regelverk som Twitters som om det vore lag samtidigt som de samhälleliga konsekvenser ett sådant regelsystem kan få. Den rättssociologiska relevansen för denna uppsats kommer därför av det faktum att den befinner sig väldigt centralt i det rättssociologiska fältet där samhälle möter rätt.

2. Disposition

Del ett i denna uppsats börjar med att förklara uppsatsens utgångspunkt samt dess syfte och frågeställningar. Efter det kommer ett bakgrundsavsnitt där den samhällskontext som uppsatsen grundar sig i, samt en del av den rättsliga diskursen runt ämnet förklaras. Detta följs av en beskrivning av de begränsningar som gjorts så väl som den rättssociologiska relevansen. I del tre presenteras forskningsläget runt uppsatsens ämne samt hur den litteratursökning som tagit fram den tidigare forskningen har utförts. Del fyra består av ett teoriavsnitt där uppsatsens teoretiska ansats beskrivs. Denna uppsats grundar sig i teorier från två forskare, Lon Fuller och Pierre Bourdieu. Sedan i del fem beskrivs den metod som har använts för att utföra forskningen samt hur resultat presenteras. Del sex är

en resultatdel där presenteras resultaten. På grund av det metodval som har gjorts, presenteras även en stor del av analysen tillsammans med resultaten för att bidra med nödvändig kontext. I del sju förs en diskussion kring resultaten med hjälp av den teoretiska ansatsen och tidigare forskningen. Del åtta består av de slutsatser som gjorts utifrån resultat, analys och diskussion. Del åtta har även en sektion där en del förslag för framtida lagstiftning presenteras grundat i slutsatserna. Uppsatsen avslutas sedan med några förslag för framtida forskning.

3. Tidigare forskning

3.1. Litteratursökning

Litteratursökningen för denna uppsats gjordes via sökmotorn LUBSearch i syfte att hitta hur forskningsläget ser ut angående sociala medier och deras regelverk samt deras inflytande på yttrandefrihet och diskussioner kring lagstiftning. För att uppnå detta gjordes totalt sex sökningar med sökord som "Terms of use" och "algorithm*" tillsammans med sökord relaterade till yttrandefrihet samt synonymmer till dessa ord. "Social media" inkluderades även i varje sökning för att säkra relevansen till området (Full redovisning över alla sökningar finns i Bilaga 1). Sökningarna specificerades till träffar med sökorden i abstract samt att alla skulle vara peer-reviewed. Sökresultaten behandlades sedan i referenshanteringsprogrammet Zotero. Efter alla sökningar samt rensning av dubletter och ett konferenspapper återstod 127 artiklar. Alla artiklar som ej var på engelska uteslöts och lämnade kvar 120 artiklar (fick inga träffar på Svenska). Efter det lästes alla abstract och alla artiklar som visade någon potential att vara relevanta behölls, medan allt som saknade relevans till uppsatsen rensades bort. 67 artiklar återstod. En till genomgång av de återstående abstract gjordes med exkluderingskriterier, främst rensades artiklar som antingen behandlade användares beteende på plattformarna, snarare än plattformarna själva och artiklar ämnade åt att förbättra algoritmer. Efter denna rensning återstod 34 träffar. Av dessa var fem topciterade (över 50 citeringar). Efter kursivläsning av dessa bedömdes endast två vara relevanta. Därför gjordes en till genomgång av de 34 där

citationsgränsen sänktes till 20 citeringar och artiklar skrivna under senaste året inkluderades. Dessa lästes kursivt och ytterligare två bedömdes relevanta. Sedan lades det till en relevant artikel som citerats i en av artiklarna samt en bok efter tips från handledare.

3.2. Sammanfattning av tidigare forskning

Lawrence Lessig har länge studerat internet och lagen. I hans bok Code 2.0 (2006) skriver han om hur kodningen av internet kan användas som lag (Lessig, 2006). Lessig tittade på det tidiga internet där många ansåg att reglering var mer eller mindre omöjligt, men kom fram till en annan synpunkt. Han ansåg att internet inte bara gick att reglera utan att det hade kapaciteten att bli det mest reglerbara området vi har (Lessig, 2006, s. 32). I Lessigs mening handlar reglering av internet om den arkitektur som nätet har. Där beroende på hur man kodar nätet och hur algoritmerna som används ser ut, kan nätet vara mer eller mindre reglerbart (Lessig, 2006, ss. 33-36). Lessig menade att svårigheterna med att reglera nätet var den anonymitet som användare åtnjöt och att verkliga världen skulle vara lika svår att reglera om människor var lika anonyma där (Lessig, 2006, s. 36). Han påpekade dock att det redan när han skrev boken 2006 fanns olika mekanismer som spårar individer på nätet, såsom TCP/IP och cookies (Lessig, 2006, ss. 43-49). Enlig Lessig var det fria internet bara en tillfällighet och internet skulle sträva till att bli mer och mer reglerat. Men detta skulle inte ske genom suveräner, utan främst genom kommersiella intressen och påtryckningar från användare (Lessig, 2006, s. 38). Lessig skrev Code 2.0 2006, detta var innan sociala medier riktigt hade tagit i kraft, vilket gjort hans syn på ett internet reglerat av kommersiella intressen närmast profetisk.

Nicolas Suzor (2018) analyserade, från ett amerikanskt perspektiv, 14 internetplattformars användarkontrakt, för att se hur de förhöll sig till tre rättsliga principer baserade på regelstyre. Principerna var *samtycke från de styrda, likhet inför lagen och förutsägbarhet* samt *vederbörlig process* (Suzor, 2018). Suzor hittade att *ToS*, likt en konstitution gav makt till de styrande. Men till skillnad från en konstitution anger de inga regler eller begränsningar i hur denna makt kan användas (Suzor, 2018, s. 6). Användare var också mycket begränsade i sin förmåga att förhandla de avtal som band dem. Detta

eftersom kontrakten var av en "take it or leave it" karaktär, vilket kommer i kontrast med principen av samtyckligt styre (Suzor, 2018, s. 6). De flesta plattformarna (inklusive Twitter) gav sig själva även rätten att ta bort innehåll på plattformarna av *alla och inga anledningar* samt att reglerna ofta var arbiträra och beslut fattades i hemlighet, vilket ger en låg jämställdhet och förutsägbarhet (Suzor, 2018, ss. 6-7). Suzor kom även fram till att användare egentligen saknade någon verklig förmåga att utmana ett felaktigt beslut, där många avtal innebär att man säger upp rätten till vissa rättsliga verktyg och man tvingas in i en omständlig intern process (Suzor, 2018, s. 8). Suzor menade att *ToS* är designade för att öka plattformarnas makt och handlingsmöjligheter samtidigt som de minimerar sitt eget ansvar (Suzor, 2018, s. 8). Han anser därför att det behövs ett regelstyre hos plattformarna i stället för den despotiska karaktär de har idag (Suzor, 2018, ss. 8-9).

Ett annat problem som studerats när det kommer till *ToS* är huruvida användare faktiskt läser och förstår dem. Studier har visat att användare ofta inte läser *ToS*. En av anledningarna kan tänkas vara att de ofta är skrivna med ett ganska svårt språk (Robinson & Zhu, 2020, ss. 1-2). 2012 valde Google därför att prova uppdatera sina *ToS* med ett enklare språk i hopp om att fler skulle läsa dem, samt att det skulle vara lättare för användare att förstå sig på dem (Robinson & Zhu, 2020, s. 1). Eric P. Robinson och Yiecheng Zhu utnyttjade denna förändring för att studera huruvida individer hade en bättre förståelse av innehållet i de nya *ToS* än i de gamla, genom att låta användare läsa antingen de nya eller de gamla *ToS*. De hade även en kontrollgrupp som inte fick läsa alls. De tre grupperna testades sedan i deras förståelse av olika innehåll som *ToS* behandlade (Robinson & Zhu, 2020, ss. 4-5). Resultatet visade att de som läst antingen de nya *ToS* från 2012 eller de gamla svårare *ToS* från 2007 hade båda en bättre förståelse än de som inte hade läst alls (Robinson & Zhu, 2020, ss. 5-7). Något som dock är intressant är att de inte kunde se någon förbättring i förståelse hos de som läst de lättare *ToS* än de som läst de svårare. I ett av testen hade till och med den svårare gruppen en marginellt större förståelse av innehållet (Robinson & Zhu, 2020, ss. 6-7). Detta påvisar att en enklare text i *ToS* inte nödvändigtvis skulle öka användares förståelse över vad de binder sig till genom att använda dessa plattformar.

När det kommer till sociala medieplattformar styrs de dock enbart delvis av *ToS* och *community guidelines*. Grovarbetet av den kontroll som bedrivs på plattformarna kommer från de algoritmer som automatiskt styr innehållet på plattformarna. Tarleton Gillespie (2015) påpekar att plattformarna genom dessa algoritmer både vinklar, väljer, tar bort och omprioriterar information som delas (Gillespie, 2015, s. 1). De flesta användarna märker dock inte det inflytande som utförs över dem, på grund av att de sällan kolliderar med eller läser igenom de regler som finns (Gillespie, 2015, s. 1). Algoritmerna används alltså inte enbart för att omdirigera information utan även för att prioritera viss information över annan. Plattformarna tenderar att främja eller tona ned den kontroll de har beroende på vad som gynnar dem själva (Gillespie, 2015, s. 1). Gillespie menar också att den konstanta intervention plattformarna utför studeras sällan, utan man studerar i stället resultat som produceras på plattformarna (Gillespie, 2015, ss. 1-2). Gillespie påpekar även att kontrollen inte sker neutralt utan påverkas av flera olika faktorer som får plattformarna att styra om deras prioriteringar såsom explicita och implicita normer, kulturella antaganden angående smak och etikett på bevåg av kränkta användare eller oroliga lagstiftare samt för att främja deras egna ekonomiska mål (Gillespie, 2015, s. 2).

Algoritmerna är designade för att ge varje användare personliga rekommendationer där informationen vi får till oss på plattformarna filtreras baserat på våra personliga preferenser, våra tidigare val och våra vänners preferenser (Helberger, et al., 2018, s. 191). Denna information påverkas även av kommersiella strategiska beslut angående vilken information vi får tillgång till och vilken information som exkluderas eller inkluderas (Helberger, et al., 2018, s. 191; 202). Detta leder till att vi väldigt enkelt hamnar i så kallade filterbubblor där vi bara blir utsatta för information som faller i linje med det vi redan tror på (Helberger, et al., 2018, s. 192). Helberger et al. (2018) har undersökt denna problematik utifrån vad de kallar för exponeringsmångfald (Helberger, et al., 2018). De menar att exponeringsmångfald är önskvärd ur flera olika perspektiv. Det är önskvärt ur ett liberal-individualistiskt perspektiv bland annat för att det utökar individuella valmöjligheter (Helberger, et al., 2018, ss. 194-195), ur ett reflekterande perspektiv bland annat för att det öppnar upp för en mer inkluderande samhällsdebatt (Helberger, et al., 2018, s. 195) samt ur ett kontradiktoriskt perspektiv bland annat för att det kan tillåta att nya perspektiv och ny

kritik kommer in (Helberger, et al., 2018, ss. 195-196). De argumenterar också för att en åsiktsmångfald kan uppnås antingen genom mjukvarudesign eller genom rättslig reglering (Helberger, et al., 2018, ss. 196-200).

Trots de problem som existerar med både *ToS* och algoritmerna går det tyvärr inte komma ifrån att någon form av innehållsmoderering krävs för att kunna ha fungerande plattformar (Langvardt, 2018, s. 1357). Kyle Langvardt (2018) menar att kategorisering av chockerande innehåll såsom groteskt våld, trakasserier eller obscen innehåll länge varit en stor utmaning för domstolar vilket innebär en mycket större utmaning för plattformarna med tanke på den hastighet och den volym av innehåll de måste moderera (Langvardt, 2018, s. 1362). Det är därför förståeligt varför man mer och mer förlitar sig på automatiserade algoritmer för att moderera plattformarna. Langvardt menar att de kraven som finns på att moderera plattformarna i kontrast till yttrande- och deltaganderättigheter gör att vi har en riskfylld väg framför oss och föreslår fyra val vi kan göra när det kommer till dessa plattformar: vi kan (1) göra svepande begränsningar av plattformarnas förmåga att moderera för att skydda rättigheter till yttranden, (2) kräva att plattformarna ger användare möjligheten att stänga av och på innehållsbegränsningar, (3) kräva att plattformarna är öppna med deras policy och procedurer eller (4) lämna plattformarnas rättigheter att moderera som de gör idag (Langvardt, 2018, s. 1357). Enligt Langvardt har alla stora för och nackdelar, men det fjärde alternativet är klart det farligaste då det leder till en väldigt despotisk framtid (Langvardt, 2018, ss. 1385-1386). Langvardt menar att under dessa plattformar blir användare utsatta för den närmaste och snabbaste censurregim världen någonsin har skådat (Langvardt, 2018, s. 1353). Samtidigt är det också en fråga om alla plattformar ska behandlas lika, om exempelvis ett hobbyforum inte tycker att politiska diskussioner hör hemma på deras forum bör de då behandlas lika som öppna plattformar som Facebook eller Twitter? (Langvardt, 2018, s. 1371).

4. Teori

Den här uppsatsens analys kommer främst bygga på två teorier. Först när den tematiska innehållsanalysen genomförs, kommer Lon Fullers desiderata för hur ett rättsliknande regelsystem kan misslyckas, användas (Fuller, 1969, s. 39). Sedan för att diskutera vilka samhällskonsekvenser resultaten från innehållsanalysen kan få, kommer jag utgå ifrån Pierre Bourdieus teorier kring fält, doxa och praktiskt förnuft (Bourdieu, 1999).

4.1. Fuller: Desiderata

I sin bok *The Morality of Law* (1969) försöker Lon Fuller hitta ett mellanting mellan traditionell naturrätt och rättspositivism (Fuller, 1969). Fuller kom därigenom fram till en lista med åtta desiderata eller krav på ett rättsligt regelsystem presenterade som åtta fundamentala misslyckanden (Fuller, 1969, s. 39). Dessa åtta misslyckanden är: *(1) att misslyckas med att skapa regler alls, (2) att misslyckas med att publicera eller åtminstone att misslyckas informera den som förväntas följa regeln, (3) missbruk av retroaktiv lagstiftning, (4) misslyckande att göra regler som går att förstå, (5) användande av motsägande regler, (6) regler som kräver agerande som ligger utanför förmågan hos den som förväntas följa reglerna, (7) ändra reglerna så ofta att det inte går att förhålla sig till dem och slutligen (8) ett misslyckande i att överensstämna angivna regler med hur de utförs* (Fuller, 1969, s. 39). Fuller menade att ett totalt misslyckande i en av dessa punkter resulterar i något som inte ens kan kallas för ett rättsligt system (Fuller, 1969, ss. 39-40). Fuller var dock medveten att det ofta finns brister på punkterna i de flesta rättssystemen och att en utopisk lag är omöjlig (Fuller, 1969, s. 41). Därför menade han att lagens inre moral drivs av en aspirationsmoral snarare än den pliktmoral den yttre moralen drivs av (Fuller, 1969, s. 43). Med det menas att, medan lagar är skrivna med en plikt i fokus, ett förbud eller påbud, drivs formandet av lagarna och rättssystemet med en aspiration i fokus, med andra ord en moral för att uppnå ett ideal (Fuller, 1969, s. 5). Fuller argumenterade också för att synen på lag inte skulle vara bunden till regler som kommer från staten, utan ansåg lag vara verksamheten att underkasta människors agerande under ett styre av regler. Därmed argumenterade Fuller

att även regelverk utanför staten skulle ses som lag (Fuller, 1969, ss. 124-127). Därför är Fullers teori väldigt användbar när Twitters regelverk ska analyseras.

4.2. Bourdieu: Fält, doxa och praktiskt förnuft

De flesta av Bourdieus idéer går att förstå genom vad han kallar för de sociala positionernas rum (Bourdieu, 1999, ss. 16-18). Bourdieu menade att våra positioner i rummet utgörs av en blandning av olika kapital där de främsta är kulturellt och ekonomiskt kapital. Tillsammans utgör alla olika kapital vårt totala kapital (Bourdieu, 1999, ss. 16-18). Beroende på objektiva relationer mellan individer som innehar liknande positioner i rummet uppstår vad Bourdieu kallar ett fält (Bourdieu, 1999, s. 55). Exempel på fält kan vara konstens fält, akademins fält eller politikens fält. Varje fält har sin egen doxa, vilken utgör vad, inom fältet anses vara sunt förnuft, rätt och fel, normalt och konstigt samt vad som kvalificerar inkludering (Andersen & Kaspersen, 2007, s. 273). Bourdieu menade att de fält vi tillhör agerar som en lins genom vilken vi ser och dömer världen (Bourdieu, 1999, ss. 184-195). Det som produceras i fältet kan även förstås till fullo enbart med en full förståelse av fältet, dess medlemmar och dess historia. Exempelvis tar Bourdieu upp konst ur konstfältet eller litteratur ur litteraturfältet (Bourdieu, 1999, ss. 55-56).

Argumenten kring fälten ger oss även en insikt i hur Bourdieu tänkte kring det praktiska förnuftet (Bourdieu, 1999). Det praktiska förnuftet går bäst att förstå i kontrast till det skolastiska synsättet, applicerat av forskare (Bourdieu, 1999, ss. 181-182). Bourdieu menade att, när forskare studerar människor med hjälp av det skolastiska synsättet, applicerar de det vetenskapliga fältets logik på de som studeras, och använder den logiken i sin tolkning (Bourdieu, 1999, ss. 184-195). I kontrast till den objektiva logik som utgör det skolastiska synsättet, innebär det praktiska förnuftet, det förnuft man får genom att förstå den icke-logiska logik människor agerar efter, som de fått genom historien och som existerar inom det fält i vilket individerna befinner sig (Andersen & Kaspersen, 2007, ss. 264-266). Praktiska förnuftet kan förstås som att man ser det man studerar på en marknivå, i verkligheten, i kontrast till att se en karta som beskriver det man studerar (Andersen & Kaspersen, 2007, s. 265). Med grund i det praktiska förnuftet kritiserade

Bourdieu vetenskapliga modeller, han menade att modellerna enbart fungerar så länge de ses som just modeller (Bourdieu, 1999, s. 266). Man kan förstå det genom att modellerna är simplifierade versioner av verkligheten där kontexten de är studerade i är tvungen att minimeras. Vilket innebär att kontexten är försvunnen när modellen appliceras på verkligheten (Bourdieu, 1999).

5. Metod

Metoden för denna uppsats är en kvalitativ tematisk innehållsanalys (Bryman, 2018, ss. 702-707) där teman valts ut i form av rättsliga principer. Tematisk analys är en ganska öppen metod utan en tydlig bakgrund (Bryman, 2018, s. 702). Därför har metoden anpassats för att bäst passa syftet med uppsatsen. Eftersom det forskats utförligt kring rätten, bedömdes det bäst att förankra analysen i några på förhand utvalda, etablerade rättsliga principer (Fuller, 1969, s. 39). Fördelen med detta är att de etablerade principerna ger en ökad reliabilitet och reducerar påverkan av subjektiva värderingar. Valet att förankra analysen i teman valda i förväg underlättar också för en mer strukturerad analys vilket bidrar till att bibehålla en hög validitet.

5.1. Val av teman

I syfte att hitta etablerade rättsliga principer bedömdes Lon Fullers åtta desiderata (se teoriavsnitt) högst lämpliga. Eftersom Fuller tog fram dessa desiderata i ett försök att slå ihop perspektiv från naturrätten och rättspositivismen (Fuller, 1969), har de redan inslag av en tematisk syntetisering (Bryman, 2018, ss. 715-716). Fuller menade också att de inte enbart var applicerbara på statsrättsliga system, utan även på andra system med rättslig karaktär (Fuller, 1969, ss. 124-125), vilket lämpar sig väl för syftet med denna uppsats. Eftersom Fullers desiderata delvis riktar sig in på processen i ett rättssystem och kan vara svår att bedöma i ett regelverk, övervägdes huruvida vissa desiderata skulle utestlutas. Eftersom Fuller presenterar alla desiderata tillsammans, togs ett beslut att ändå behålla alla, även om vissa riskerade att inte bli applicerbara.

Efter litteraturgenomgången kompletterades de teman tagna från Fuller med tre teman inspirerade av Nicolas Suzors (2018) undersökning av online plattformars kontrakt (se tidigare forskning) (Suzor, 2018). Dessa teman var *vederbörlig process, likhet inför lagen* och *nomokrati (regelstyre)*.

För att bibehålla en viss flexibilitet och för att se till att inte missa någonting potentiellt av vikt, bibehölls även en öppenhet för att nya teman skulle dyka upp i analysen av regelverket eftersom detta är en fördel med tematisk analys (Bryman, 2018, ss. 702-707). De elva teman som valts ut i förväg visades dock uppfylla en tematisk mättnad och kunde täcka allt av intresse i regelverket.

5.2. Urval

Urvalet till denna uppsatts har varit Twitters regelverk. Medan det hade varit av intresse att undersöka fler plattformar, var en begränsning till Twitter nödvändigt under den tidsram som funnits till förfogande. Anledningen att Twitter valdes i stället för andra, större plattformar såsom Facebook, Youtube eller Instagram, är för att Twitter används i stor utsträckning av offentliga personer exempelvis politiker, journalister och aktivister för öppen samhällsdebatt.

De delar av Twitters regelverk som undersökts, är de som specifikt riktar sig till Twitters användare, dessa är Twitters *ToS* (Twitter Inc., 2020), *Rules and Policies* (Twitter Inc., 2021) samt *Generals Guidelines and Policies* (Twitter Inc., 2021). Specifikt valdes *ToS* för användare utanför Europa och resten av världen (utanför USA). Tillsammans uppgick regelverket till ca 64 000 ord.

5.3. Övrigt material

Utöver Twitters regelverk används en del andra källor och utlåtande från Twitter samt en del nyhetsrapporteringar. Dessa är dock inte systematiskt utvalda och används därför enbart för att bidra med förståelse och kontext.

5.4. Tillvägagångssätt

För att underlätta genomgången av materialet, kopierades hela Twitters *Rules and policies* och *General guidelines and policies* över från Twitters hemsida (Twitter Inc., 2021) till Word-dokument och *ToS* bearbetades som PDF. Eftersom repetition är viktigt i en tematisk analys (Bryman, 2018, s. 705) gjordes flera genomgångar av materialet. Först lästes allt material igenom, samtidigt markerades allt av potentiellt intresse med hjälp av färgkodning baserat på utvalda teman och potentiella ytterligare teman. Detta lade grunden till en andra genomgång av materialet, då texten lästes igenom med huvudfokus på de markerade delarna av texten, där en striktare bedömning över vad som borde inkluderas gjordes samt en undersökning om några ytterligare teman dykt upp. En sammanfattning skrevs sedan för varje tema, i ett nytt dokument. Med hjälp av dessa sammanfattningar skrevs sedan resultatet samtidigt som allt dubbelkollades och verifierades i den ursprungliga texten.

5.5. Presentation och analys av material

De teman som använts kräver en viss tolkning och kontextualisering för att på ett korrekt och begripligt sätt kunna presenteras. Därför är det nödvändigt att inkludera motivationen till varför de teman som presenteras är applicerbara. Således bedömdes det lämpligast att använda en presentationsmetod där presentationen av resultaten vävs ihop med analysen (Lindstedt, 2019, s. 214).

Eftersom en del av analysen vävs ihop med presentationen av resultatet används en diskussionsdel i stället för den analysdel som normalt används. Där diskuteras resultaten utifrån den tidigare forskning och det teoretiska ramverket (Lindstedt, 2019, s. 214). Detta underlättar även i att undvika förvirring mellan den tematiska analysen och analysen av resultatet. Diskussionsdelen är uppdelad i två delar: Först en analys och diskussion ur ett rättsligt perspektiv med stöd i Fullers teori. Andra delen tar ett bredare perspektiv för att analysera samhällsliga implikationer delvis med stöd i Bourdieus Praktiska förnuft.

5.6. Generaliserbarhet och etiska överväganden

Medan Twitters regler förvisso inte är applicerbara utanför Twitter, är de flesta sociala medieplattformarnas regler utformade enligt en liknande struktur. Därför, med hjälp av tidigare forskning, lämnas ett visst utrymme för generalisering.

Eftersom det enda som hanterats är offentliga dokument, har inga etiska övervägande behövt göras, utöver att se till att allt görs korrekt och ärligt.

6. Resultat

Här presenteras resultatet från den tematiska analysen av Twitters *ToS, Rules and policies* samt *General guidelines and policies*. Varje tema presenteras i ordning. Eftersom det krävs tolkning och kontextualisering för att kunna förmedla varför varje tema är applicerbart, presenteras resultatet tillsammans med motivation för hur varje tema är applicerbart, vilket utgör en del av analysen för uppsatsen.

I delar av Twitters regler och policys kunde väldigt stora problem identifieras. I Twitters *Hateful Conduct policy* (Twitter Inc., 2021) samt deras policy mot *Coordinated harmful activity* (Twitter Inc., 2021) finns problem som relaterar till flera teman samtidigt. För att informationen inte ska repeteras om och om igen under varje applicerbart tema, adresseras dessa policys enskilt, i stället för under varje tema. Hur dessa förhåller sig till applicerbara teman hittas under rubriken "Problematiska policys" med hänvisning till de teman som är applicerbara.

6.1. Genomgång av teman

(1) att misslyckas med att skapa regler alls

Det är det ganska uppenbart att Twitter lyckats skapa regler, eftersom de har ett regelverk. Det går däremot att ifrågasätta huruvida man faktiskt kan kalla det ett regelverk när de ger sig själva rätten att stänga av användare för *alla och inga anledningar* (Twitter Inc., 2020) och om reglerna därför bör ses som riktlinjer snarare än regler.

(2) att misslyckas med att publicera eller åtminstone informera den som förväntas följa regeln

Detta tema är svårt att applicera utan tillgång till relativt mycket information utanför regelverket eftersom ett misslyckande att publicera troligtvis inte står publicerade i regelverket. Några exempel gick dock att hitta. När Twitter ska redogöra för vad som involveras i en regel använder de sig ofta av *...but is not limited to* (Twitter Inc., 2021) utan att redogöra eller ge någon indikation för vad regeln omfattar utöver det som räknas upp. I sin policy *About specific instances when a Tweet's reach may be limited* säger de bland annat att det är möjligt att de kan begränsa *tweets* baserat på om de bedömer att de är av låg kvalitet (Twitter Inc., 2021), men de redogör aldrig för vad låg kvalitet innebär.

(3) missbruk av retroaktiv lagstiftning

För att fullt undersöka detta tema behövs studier av information utanför regelverket för att göra definitiva bedömningar. Twitter skriver dock i *ToS* att dess ändringar ej kommer att ske retroaktivt (Twitter Inc., 2020). Rätten att stänga av användare för *alla och inga anledningar* (Twitter Inc., 2020) gör dock det åtagandet ganska meningslöst. Något sådant åtagande om retroaktiv applicering görs dock inte i *Rules and policies* eller i *General guidelines and policies* (Twitter Inc., 2021). Ett problem kunde också identifieras i uppdateringen av *Coordinated harmful policy* och presenteras därför längre ned under resultatet specifikt för den policyn.

(4) Misslyckande att göra regler som går att förstå.

Detta tema kunde identifiera flest problem. Medan det inte nödvändigtvis finns någonting som är omöjligt att förstå i de regler och policys som presenteras, är språket som används ett stort problem för den som vill skaffa sig en förståelse för vad som gäller.

Först, när det kommer till språkval som Twitters *Rules and Policies* och *General Policies* erbjuds på. De erbjuds på 21 olika språk (Twitter Inc., 2021), medan Twitters tjänst erbjuds med 47 språkval (Twitter Inc., 2021). Detta innebär att det finns 26 språkval som inte erbjuds i Twitters regelverk. För någon som använder ett av de 26 språken och inte kan,

eller har bristande kunskaper i, något av de 21 vilka reglerna erbjuds i, blir det därför väldigt svårt att skaffa sig en förståelse för regelverket.

Språk är även ett problem i valet av termer. Genomgående, framför allt i *Rules and Policies* och *General Policies*, används ord och fraser vilka är öppna för tolkning, utan att ge förklaringar för hur dessa ska tolkas, och därför gör det svårt att veta exakt vad som gäller. Dessa kan delas in i fyra primära kategorier: Ord och fraser som indikerar en subjektiv gräns som kan variera från person till person såsom "*excessively*", "*apparent*", "*significant*", "*primary*", "*particularly egregious*", "*demonstrably*", "*widely*" eller "*credibly*". Andra kategorin är ord som indikerar intentioner och kräver en subjektiv tolkning av någons avsikt från Twitter såsom "*With intent*", "*for the purpose of*", "*intended to*", "*whose purpose is*" eller "*dedicated to*". Tredje kategorin är ord och fraser som direkt hänvisar till Twitters subjektiva tolkningar såsom "*that we believe*", "*that appears to have*" eller "*with apparent intent*". Till sist övriga ord och fraser som kräver en subjektiv tolkning och vars bedömning kan variera från person till person såsom "*misleading*", "*newsworthy events*", "*significant public interest*", "*glorify*", "*widely debunked*", "*tropes*", "*content that is relevant*", "*legitimate public interest*", "*extremist groups*". (Twitter Inc., 2021) (Twitter Inc., 2020). Det är inte ovanligt att en del av dessa typer av ord och fraser används i lagstiftning. Tillgång till förarbeten, rättsfall och doktriner ger dock ofta en indikation för hur dessa ska tolkas så att man kan ta reda på hur man ska förhålla sig till lagarna. Twitter håller dock stora delar av sin process dold och offentliggör oftast inte sina beslut, vilket leder till att det blir väldigt svårt att veta hur de gör sina tolkningar. Lagar definierar också ofta sina ord i själva lagarna, det gör även Twitter på ett bra sätt på många håll exempelvis i sin *Child sexual exploitation policy*, *violent threats policy* samt definitionerna i *Counterfeit policy* och *self-harm policy* (Twitter Inc., 2021). Det saknas dock att det görs konsekvent.

Twitters regler, terms och policys är i stor del skrivet med ett juridiskt språk, vilket kan förväntas. Medan det i sig ofta leder till att många får svårt att förstå, är det inte det enda områdesspecifika språket som används. Twitter använder även termer som är väldigt IT-specifika eller specifika till Twitter som plattform, utan att ge någon definition för att förtydliga. I *ToS* används IT-termer som "*crawling*", "*scraping*", "*TCP/IP packet header*", "*source-identifying information*", "*overloading*", "*flooding*" och "*spamming*" samt Twitter

specifika termer som "*super hearts*", "*coins*" och "*stars*" (Twitter Inc., 2020). I *Rules and Policys* används termer som "*Synthetic media*", "*pre-roll video ads*" och "*creepshots*" samt Twitter specifika termer "*bio*", "*fleet*", "*Twitter space*" och "*Username squatting*". I general policies används termer och förkortningar som "*APIs*", "*widgets*", "*reCAPTCHA*", "*synthetic or manipulated media*", "*spliced*" och "*isolative editing*". Användning av denna mix av ord, fraser och förkortningar med ett juridiskt språk innebär att man behöver ha en förståelse både för IT- och Twitterspråk samt juridiskt språk för att fullt kunna förstå regelverket.

(5) Användande av motsägande regler

I Twitters regelverk görs inga direkta distinktioner mellan regler och policys. Twitters policys involverar både förbud, påbud, instruktioner och Twitters åtaganden. Några förbud eller påbud som motsäger varandra har inte identifierats. Eftersom ingen direkt distinktion för regler och åtaganden görs, finns det däremot ett antal direkta motsägelser i regelverket. Detta främst när det kommer till yttrandefrihet. Twitter repeterar vid flera tillfällen att yttrandefrihet, åsiktsfrihet och friheten att uttrycka sig ligger i grunden för plattformen samt att dessa är mänskliga rättigheter (Twitter Inc., 2021). De hänvisar även till bland annat amerikanska *Bill of Rights* och *Europakonventionen för mänskliga rättigheter* när de anger yttrandefrihet som deras grundläggande åtagande (Twitter Inc., 2021). De skriver även att de inte blockerar, begränsar eller tar bort innehåll baserat på individers åsikter eller synsätt (Twitter Inc., 2021).

Twitters åtaganden upprätthålls dock inte i regelverket. Eftersom sociala medier främst är sociala plattformar och består nästan enbart av just yttranden, innebär det att nästan alla regler är skrivna för att kontrollera eller begränsa användares förmåga till yttranden. Att applicera en absolut standard är dock inte nödvändigtvis rättvist. Applicerar vi dock en mer realistisk standard på yttrandefrihet, där vi accepterar begränsningar som typiskt görs även i fria länder såsom förtal, bedrägerier, hatbrott, olaga hot etc. finns det fortfarande stora problem. Exempelvis begränsar Twitter ofta vad de kallar falsk eller missledande information bland annat när det kommer till Covid-19 där de säger sig utgå från auktoritativa källor, dock utan att nämna vilka dessa källor är (Twitter Inc., 2021). Med andra ord kan du bli begränsad eller straffad för att ha fel, eller i de fall där de auktoritativa källorna potentiellt har fel kan du även bli begränsad eller straffad när du säger något som

är sant. En begränsning i att kunna ifrågasätta auktoritet står i strid med de mest fundamentala yttrandefrihetsprinciperna. Även när det kommer till åtagandet att de inte begränsar användare baserat på åsikt eller synsätt finns det problem. Exempelvis begränsar Twitter ofta att man önskar eller hoppas på att något ska hända, oftast skada på individer. Även under Twitters *Glorification of violence policy* förbjuds åsikter som uttrycker stöd för våld som har hänt där exempel som anges i policyn är att säga: "man är glad detta hände", "den här personen är min hjälte", "jag önskar att fler gjorde såhär" eller "jag hoppas detta inspirerar fler att agera" (Twitter Inc., 2019). Medan detta förståeligt inte är önskvärt och ofta direkt destruktivt, är det direkta förbud på åsikter och därför står de i strid mot åtagandet om att inte begränsa åsikter.

Även när det kommer till Twitters agerande visar de sig stå i strid med principerna kring yttrandefrihet. Exempelvis står avstängningen av ett nyhetsbolag som New York Post på grund av nyhetsrapportering i direkt strid med både amerikanska *Bill of Rights* och Europakonventionen för mänskliga rättigheter som Twitter säger sig grundas i (Twitter Inc., 2021). Detta eftersom censur av journalister brukar ses väldigt allvarligt på både från USA och EU.

(6) regler som kräver agerande som ligger utanför förmågan hos den som förväntas följa reglerna

För det här temat identifierades ett problem enbart på ett ställe. I Twitters *Ban evasion policy* står det att, om någon som blivit permanent avstängd försöker kringgå avstängningen erhåller Twitter rätten att stänga av konton de tror tillhör samma entitet som det ägde avstängda kontot (Twitter Inc., 2020). Eftersom det räcker med att Twitter tror att kontot tillhör den eller de som blivit avstängda tidigare, kan någon bli avstängd helt utom deras egen kontroll enbart för att Twitter trodde att någon annan kontrollerade kontot. Det är därmed en regel som potentiellt ligger utom personens kontroll.

(7) ändra reglerna så ofta att det inte går att förhålla sig till dem

ToS ändras inte särskilt ofta, i genomsnitt ungefär en gång per år och det finns ett arkiv som visar när de har ändrats och hur de tidigare *ToS* ser ut (Twitter Inc., 2021). Twitter skriver även i *ToS* att de ger åtminstone 30 dagars förvarning när *ToS* kommer att ändras,

undantaget påtvingande av lag (Twitter Inc., 2020). Ändringar en till två gånger om året är knappast att anse som för ofta. När det kommer till *Rules and policies* och *General guidelines and policies* däremot, finns det ett antal problem. Det finns inte någon logg som redovisar när ändringar sker. Vissa ändringar annonseras genom Twitters företagsblogg (Twitter Inc., 2021). Det är dock inte tydligt om alla förändringar annonseras där eller om det bara är vissa. Det är också långt ifrån det enda som skrivs i bloggen vilket gör det svårt att navigera igenom för att hitta ändringar. Av totalt 57 regler, policys och guidelines, har 16 ett årtal och en månad i toppen, två har ett helt datum (Twitter Inc., 2021). Det redogörs dock inte om detta antyder när policyn var skriven eller om det är när den senast var uppdaterad. En kraftig brist på transparens innebär därför att reglerna kan ha ändrats utan att det meddelats och kan ändras när som helst. En stor del av Twitter styrs också genom algoritmerna som arbetar i bakgrunden, i Twitters mjukvara görs flera ändringar i månaden (Twitter Inc., 2021). Eftersom Twitter listar omprioritering av Tweets bland sina sanktioner (Twitter Inc., 2021), kan delar av regelverket förändras väldigt frekvent.

(8) Misslyckande i att överensstämna angivna regler med hur de utförs

Under detta tema hittades enbart problem i policyn för *Coordinated harmful activity* som presenteras nedan.

(9) Vederbörlig process

Under detta tema är främst brist på transparens ett problem. Twitter ger rätten att överklaga beslut under de flesta policys (Twitter Inc., 2021) undantaget om Twitter tvingar en markör för känsligt innehåll på ett konto (Twitter Inc., 2021). De förklarar hur man överklagar men ej hur processen går till, ingen garanti ges heller på att överklagan kommer att tas i akt. Processen för hur beslut fattas kring sanktioner är också med vissa undantag dold. I *Synthetic and manipulated media policy* anger Twitter även att de i gränsfall kommer att överväga att ta bort innehåll som möjligtvis inte bryter mot policyn (Twitter Inc., 2021), vilket innebär en skuldpresumption och därför inte kompatibelt med *vederbörlig process*.

(10) Lika inför lagen

Problem med detta tema hittades främst i *Hateful conduct policy* och policyn för *Coordinated harmful activity* och presenteras därför längre ned.

(11) Nomokrati (regelstyre)

Här har Twitter väldigt allvarliga brister. Som nämnts i sektionen för motsägande regler, begränsar Twitter vad de kallar falsk eller missledande information i förhållande till Covid-19 och säger sig använda *auktoritativa källor* för att bestämma vad som är falskt eller missledande (Twitter Inc., 2021). Problemet är att de inte ger några kriterier för vad som ska anses som *auktoritativa källor*. Auktoriteter kring Covid-19 har också skiftat åsikter ofta och olika auktoriteter har haft olika ståndpunkter, vissa har varit bevisligen felaktiga. Exempelvis sade WHO tidigt att Covid-19 inte spreds från människa till människa baserat på rapporter från Kina (World Health Organization (WHO), 2021), trots att Taiwan ska ha rapporterat att spridning mellan människor faktiskt skedde (Watt, 2020). Donald Trump blev faktakontrollerad ett flertal gånger på Twitter för felaktig information kring Covid-19 (Scola, 2020). Medan Trump knappast är en säker källa och var ofta missledande kring Covid-19, går argument göras för att den amerikanska presidenten, med tillgång till landets främsta experter, kan ses som en auktoritet. Med andra ord finns det kriterier som Twitter använder för att göra ett aktivt val angående vem eller vilka som ska anses vara auktoriteter. Problemet ligger i att, utan en redogörelse för vilka källor som gäller samt på vilka grunder dessa källor är valda, blir i slutändan Twitter den enda auktoriteten som räknas. Detta då de har möjligheten att välja auktoriteter utifrån vad de vill få sagt.

Ett annat problem är det som nämnts under sektionen för regler som är svåra att förstå, där termer som är mycket öppna för tolkning ofta används genom regelsystemet. Detta innebär en risk att genomdrivningen av reglerna till stor del lämnar mycket makt till den som ska tolka reglerna på Twitters sida och inte så mycket till reglerna själva. Otydligheter angående exakt vem eller vilka det är som gör tolkningen bidrar till problemen.

Största problemet ligger dock i en mening från *ToS* där Twitter ger sig rätten att stänga av eller avsluta ett konto samt avsluta service för alla eller delar av tjänsten när som helst

för alla och inga anledningar (Twitter Inc., 2020). Detta innebär att Twitters regler och policys enbart existerar som en riktlinje för hur de kan förväntas att agera, men bär egentligen ingen vikt. I slutändan styrs plattformen helt despotiskt.

6.2. Problematiska policys

När Twitters regelverk analyserades var det två policys som stack ut, dessa var *Hateful conduct policy* (Twitter Inc., 2021) och policyn för *Coordinated harmful activity* (Twitter Inc., 2021). Båda dessa har problem med flera teman, ibland samtidigt. Resultatet för dem blir därför bättre att presentera utifrån varje policy i stället för att dela upp dem under de teman som är applicerbara. De siffror inom parentes hänvisar till varje temas numrering i syfte att förtydliga vilket tema som adresseras. Eftersom båda dessa policys hanterar ganska kontroversiella ämnen är det viktigt att betona att det som kritiseras i reglerna inte är substansen i det de försöker adressera, utan enbart hur de är utformade och hur språket används för att förklara dem.

6.2.1. Hateful conduct policy

Problematiken med Twitters *Hateful conduct policy* ligger främst i dess otydligheter. Policyn inleds med att redogöra ett förbud mot att uppmana till våld mot, direkt attackera eller hota andra baserat på *ras, etnicitet, nationellt ursprung, kast, sexuell läggning, kön, könsidentitet, religiös tillhörighet, ålder, handikapp* eller *allvarlig sjukdom* och benämner dessa "kategorier" (Twitter Inc., 2021). Bortsett från att det inte är tydligt vad exakt "attackera" innebär, är detta ganska likt hur lagar av detta slag är utformade exempelvis i svensk lagstiftning (SFS 2018: 1744). Senare i policyn används ofta termen "skyddade kategorier" (Twitter Inc., 2021), huruvida det innebär de kategorier som räknats upp är dock inte tydligt. De nämnda kategorierna är de enda som beskrivs som "kategorier", så det hade kunnat vara fallet. Problemet är dock hur *skyddade kategorier* används. Exempelvis skriver Twitter:

“We prohibit content that wishes, hopes, promotes, incites, or expresses a desire for death, serious bodily harm, or serious disease against an entire protected category and/or individuals who may be members of that category.” (Twitter Inc., 2021)

Den här meningen kan framstå som ganska självklar vid första anblick, men när man tittar närmare uppstår en del otydligheter. Förbjudna ageranden mot en hel kategori är relativt tydligt. Exempelvis om kategorin är kön skulle det antyda att man inte får gå efter alla kvinnor. Problemet kommer i den andra delen. Där är individer skyddade. Skyddet gäller dock inte enbart förbjudna ageranden med anspelning på tillhörigheten i kategorin, exempelvis att man inte får gå efter en kvinna för att hon är kvinna likt vi ser i svensk lagstiftning. Skyddet går ett steg längre och gäller i stället förbjudna ageranden oavsett motiv. För kategorin kön innebär det att du inte får gå efter någon som har ett kön. Detta skulle dock betyda att alla är skyddade eftersom alla tillhör någon av de nämnda kategorierna, på något sätt. Exempelvis har alla en ålder. Om alla är skyddade blir dock indelningen i kategorier meningslös. Därför, för att regel ska fungera, måste *”skyddade kategorier”* betyda enbart vissa grupper under kategorierna. Det är dock inte tydligt vem som då räknas. Sen kan man fråga sig om det i så fall är okej under regeln att göra de beskrivna agerandena mot individer som inte tillhör dessa kategorier?

I en sektion av policyn beskriver Twitter sitt resonemang för varför policyn existerar, vilket kan ge en indikation för vad de möjligtvis kan mena med skyddade kategorier:

“We recognize that if people experience abuse on Twitter, it can jeopardize their ability to express themselves. Research has shown that some groups of people are disproportionately targeted with abuse online. This includes; women, people of color, lesbian, gay, bisexual, transgender, queer, intersex, asexual individuals, marginalized and historically underrepresented communities.” (Twitter Inc., 2021)

Där nämner de även att de vill skydda de som varit *”historiskt marginaliserade”* från att få sina röster tystade (Twitter Inc., 2021). Dessa beskrivs dock ej som kategorier, utan enbart vad de använder som grund för policyn. Är detta vad som används som skyddade kategorier uppstår även vissa problem. Exempelvis definieras aldrig vad *historiskt underrepresenterade ”communities”* eller *historiskt marginaliserade* innebär. Med tanke på att Twitter är en global plattform, spelar både geografisk och temporal kontext en stor roll för båda dessa termer. Det är heller inte helt klart hur man ska tolka kategorier som ålder

eller religiös tillhörighet utifrån de grupper som nämns om de nu är de skyddade kategorierna.

För att förtydliga gäller problemet inte det faktum att vissa beteenden begränsas eller att man vill ge skydd baserat på grupptillhörighet. Problemet ligger i de otydligheter som existerar kring hur kategorierna ska tolkas samt vem som tillhör dem.

En annan problematik är att Twitter säger att den som blir utsatt inte behöver vara del av en skyddad grupp, för att de ska använda sanktioner och att de aldrig kommer att försöka verifiera huruvida den som blir utsatt tillhör en skyddad grupp (Twitter Inc., 2021).

I förhållande till de teman som undersökts skapar naturligtvis allt detta stora problem. Först och främst de olika otydligheterna innebär en del *(4) misslyckanden att göra regler som går att förstå*, eftersom de kategorier som bedöms är potentiellt helt odefinierade och historiskt underrepresenterade eller marginaliserade saknar tolkningsriktlinjer. Att inte definiera vad skyddade kategorier innebär är även *(2) ett misslyckande med att publicera*. Eftersom det inte krävs att den som blir utsatt, blir det med anspråk på tillhörande i en skyddad kategori, utan regeln appliceras oavsett motiv, har man även skapat en odefinierad skyddad klass där det är möjligt reglerna applicerar annorlunda beroende på vem man är. Vilket innebär att alla inte är *(10) lika inför lagen*. Samt att inte verifiera att den som blir utsatt faktiskt tillhör en skyddad kategori kan innebära ett *(8) ett misslyckande att överensstämna angivna regler med hur de utförs* då man riskerar att bli straffad för beteenden mot någon som inte tillhör en skyddad kategori, om de anmäler och påstår att de tillhör kategorin.

6.2.2. Coordinated harmful activity

Coordinated harmful activity, är en policy som uppdaterades i respons till kravallerna vid Kapitolium i Washington D.C. den sju januari tidigare i år (Twitter Safety, 2021). Exakt vilka ändringar som gjordes i uppdateringen annonseras inte. Men med hjälp av en arkiverad version av policyn från december gick det, efter jämförelse med nuvarande policy, verifiera att ändringen som gjordes var att de lade till förmågan att stänga av konton under policyn (Virginia Tech Libraries, 2020). Samtidigt som Twitter annonserade att

policyn uppdaterats, annonserade de även att 70 000 konton med anknytning till konspirationsgruppen QAnon stängdes av från Twitter med antydning till policyn (Twitter Safety, 2021). Detta indikerar ett *(3) missbruk av retroaktiv lagstiftning*. Då det mesta tyder på att många av de konton som stängdes av blev avstängda på grund av ändringen i policyn.

Policyns innehåll är också problematiskt i förhållande till de teman som använts i analysen. Problemen grundar sig främst i att policyn är extremt bred. Policyn riktar in sig på kampanjer, grupper och rörelser med koordinerad aktivitet som leder till skada på och utanför plattformen (Twitter Inc., 2021). Problemet kommer i att det räcker med att individer som visat stöd för gruppen, rörelsen eller kampanjen orsakar skada för att räknas in i policy, samt att alla andra individer som visat stöd, på eller utanför plattformen, sedan kan sanktioneras (Twitter Inc., 2021). Med andra ord, riskerar en grupp, rörelse eller kampanj och alla som visat stöd för dem, bli sanktionerade för agerande som ligger helt utom deras kontroll, alltså ett problem med *(6) regler som kräver agerande som ligger utanför förmågan hos den som förväntas följa reglerna*. Exempelvis riskerar alla som någonsin visat stöd för *Black Lives Matter (BLM)*, vare sig stödet är uttryckt på Twitter eller någon annan stans, bli sanktionerade på Twitter. Detta på grund av delar av protesterna under fjolåret blev våldsamma och ledde till skador (Kingson, 2020). Så länge de som utförde våldet också uttryckt stöd för rörelsen, spelar det ingen roll om våldet är accepterat inom rörelsen eller ej, policyn gör ingen sådan distinktion. Att sanktionera alla som påvisat stöd för BLM skulle naturligtvis vara ett väldigt extremt tilltag och knappast önskvärt för vare sig Twitter eller dess användare. Men det innebär också att utformningen av policyn är ett problem då den är så brett utformat att den oundvikligen leder till *(8) ett misslyckande i att överstämna angivna regler med hur de utförs* och att alla inte blir *(10) lika inför lagen*.

7. Diskussion

7.1. I förhållande till rättsliga principer

En stor del av analysen av regelverket som lag är redan gjord i resultatdelen. Här kommer det därför främst utvecklas på några viktiga punkter samt göras en återkoppling till frågeställningen.

Ett av de största problemen som kom fram ur den tematiska analysen är den mängd otydligheter som existerar i regelverket. Detta blev speciellt tydligt i hur ofta termer öppna för subjektiv tolkning användes. Fuller menade att problemet med att förstå regler är ofta på grund av att de är trubbiga och inte specifika nog (Fuller, 1969, s. 84). De subjektiva termerna gör det därför väldigt svårt att förstå hur exakt man ska tolka regelverket. Vilket, om Googles *ToS* är av liknande karaktär, kan hjälpa att förklara varför en förenkling av det juridiska språket inte var till hjälp för förståelsen av *ToS* i Robinson och Zhus studie (Robinson & Zhu, 2020). Risken är att reglerna är så otydliga att det inte går att veta huruvida de till och med motsäger varandra (Fuller, 1969, s. 111) eftersom tolkningsutrymmet är så pass brett. Fuller var mycket medveten om att regelverk alltid kommer att ha brister när det kommer till hans desiderata, han menade dock att brister i ett desiderata ibland kan kompenseras för, med hjälp av ett annat (Fuller, 1969, s. 104). Exempelvis kan man kanske undra varför termer som *"excessively"* och *"particularly egregious"* tas upp som otydliga när vi ofta i svensk lagstiftning använder termer som *"grov"* eller *"synnerligen grov"* vilka är av en liknande karaktär. Skulle man gå runt på gatan och fråga människor när de anser att en misshandel övergår till grov misshandel och sedan till synnerligen grov misshandel, skulle man troligtvis få en relativt stor spridning mellan svaren. Man löser det dock i svensk rätt genom att publicera förarbeten, doktriner och rättsfall som går att läsa och skaffa sig en bild av var gränsen går. Därmed kan man utnyttja vad Fuller kallar generaliseringens makt (Fuller, 1969, s. 48) och skriva generella lagar utan att skapa stora brister i förståelse. I Twitters regelverk saknas dock vidspridd publicering av beslut samt av den processen som leder fram till beslutet. Därför blir det desideratum för svårigheter att förstå reglerna extra påtagligt på grund av bristerna i det desideratum

angående misslyckande att publicera (Fuller, 1969, s. 39). När Twitter visar brister i så pass många desiderata som de gör, riskerar de därmed att få en eskalerande effekt.

Viktigt att komma ihåg är dock att *ToS, Rules and Policies* och *General Guidelines and policies* endast är en del av regelverket som används för att utföra kontroll över Twitters användare. Twitter har ett andra dolt regelverk som konstant arbetar något mer osynligt. Detta är de algoritmer som används för att styra innehållet på Twitter. Som vi lärt från Lawrence Lessig (2006), kan även kod ses som lag (Lessig, 2006). Twitter använder, enligt regelverket, vissa kontrollmekanismer som sanktioner via algoritmerna, såsom att reducera synlighet eller sökresultat (Twitter Inc., 2021). Men den kraftiga bristen på transparens, som även Nicolas Suzor påpekade (Suzor, 2018, s. 5), gör att man aldrig kan vara säker på i vilken utsträckning algoritmerna används som kontrollmekanismer vilket naturligtvis bryter starkt mot det desideratum angående misslyckande att publicera (Fuller, 1969, s. 39). Utifrån det som Gillespie påpekat angående hur plattformarna styrs (Gillespie, 2015) är det troligt att algoritmernas kontroll sträcker sig något längre än Twitter är villiga att medge.

Ett intressant mönster som noterades är att, i många av de regler som tar ställningstagande som är väldigt långt från kontroversiella, i exempelvis *Child exploitation policy* och *Violent threats policy* (Twitter Inc., 2020) är det väldigt tydligt i vad policyn innebär. Däremot när det kommer till de två policys som lyfts fram som mest problematiska kan vi se att de adresserar områden där det funnits påtryckningar för att Twitter ska agera. Det diskuteras mycket över hat på nätet vilket *Hateful conduct policy* är menat att adressera samt att det blev en stor storm kring upplöpan vid Kapitolium som ledde till uppdateringen av *Coordinated harmful activity* policyn (Twitter Inc., 2021) ger en del stöd till Gillespie angående att regleringen styrs till stor del av externa påtryckningar (Gillespie, 2015). Det är möjligt att dessa påtryckningar leder till förhastade förändringar som även reflekteras i reglernas kvalitet.

7.2. I förhållande till samhällskonsekvenser

De flesta sociala medieföretagen är centrerade i Silicon Valley i San Francisco. Detta reflekteras till viss del i Twitters regelverk där delar uppstår väldigt amerikacentriskt, exempelvis i deras riktlinjer för Twitter moments nämns hur man bör benämna olika folkgrupper, där alla nämnda är del av anglosfären (Twitter Inc., 2021). De säger även att de grundar sin syn på yttrandefrihet i amerikanska *Bill of Rights* (Twitter Inc., 2021). Silicon Valley är ett ganska unikt område, som domineras av IT-företag. San Francisco är också en politiskt väldigt homogen stad (Department of Elections, 2020). Detta leder till att det *fält* som plattformarna befinner sig i är relativt isolerat, där de som styr befinner sig i ett område där de flesta arbetar inom samma bransch och har liknande värderingar. Mot bakgrunden av detta är det kanske inte så underligt att stora delar av Twitters regelverk är väldigt öppen för tolkning. Isolering från kontakt med andra fält och åsikter riskerar att skapa en miljö där den doxa som existerar i fältet börjar ses som allmänvett. Ta exempelvis termen "*historiskt marginaliserade*" som används i Twitters *Hateful conduct policy*. Med en insikt i det fält Twitter befinner sig i, är det kanske inte så svårt att förstå vad denna term hänvisar till. Bor man i Sverige eller USA är det sannolikt att man har en relativt bra bild av hur Twitter resonerar. Men Twitter existerar i en global kontext. Vem som är marginaliserad kan variera väldigt mycket beroende på var man befinner sig. Samma sak gäller för den historiska aspekten. När ska man ha varit marginaliserad? Hur länge? Var i världen räknas? Exempelvis var Fransmän marginaliserade i sitt eget land när de fann sig under tysk ockupation under andra världskriget. Räknas då fransmän som historiskt marginaliserade i Frankrike? Dessa frågor kan låta löjliga. Men, likt de konstverk och litteratur, skapade i sina egna fält där Bourdieu menade att man behövde en full förståelse av fälten för att fullt förstå sig på (Bourdieu, 1999, ss. 55-56), produceras Twitters regelverk inom ett fält. Utan insikt i fältet finns risker för att det blir svårt att få en full förståelse för hur reglerna ska tolkas. Samma sak går att säga om det innehåll användare lägger upp på Twitter. Innehållet produceras i ett nästan oändligt antal fält världen över. Tolkningen kommer alltid att ske genom en lins som utgörs av det fält den individ eller de individer som fattar beslutet om det som sägs ska tillåtas på Twitter eller ej. Frågan är om Twitter har, eller ens har en möjlighet att skaffa sig, ett praktiskt förnuft för ens en bråkdel

av de fält som täcks. En indikator angående Twitters förmåga att se utanför sin egen doxa kan vara de IT-specifika begrepp som togs upp i resultatdelen. De nämns i regelverket utan förklaring för vad de betyder vilket indikerar ett antagande att de ska förstås av den som läser. En annan möjlig indikator är det faktum att regelverket inte existerar på lika många språk som Twitters plattform erbjuds på.

En stor fara med isolerade åsiktshomogena fält, är något som man inom socialpsykologin kallar för gruppolarisering (Gilovich, et al., 2019, ss. 422-425). Gruppolarisering är ett fenomen som sker när en grupp är väldigt åsiktshomogen och det är en avsaknad av konträra åsikter. Gruppen tenderar då att drivas längre ut i den riktning gruppen är benägen och blir mer extrem (Gilovich, et al., 2019, ss. 422-425). Exempelvis om man har en isolerad grupp där alla tillhör den politiska högern som nästan aldrig blir utsatt för åsikter från politiska vänstern, kommer gruppen att sakta men säkert drivas längre ut till höger. Tar vi sedan hänsyn i den tendens som betonas av Helberger et.al (2018), där den algoritmiska sortering av information på plattformarna, enbart visar oss vad vi vill se till den grad att vi isoleras från åsikter vi inte håller med om (Helberger, et al., 2018), riskerar vi att få ett stort problem. Detta innebär att vi först riskerar ett regelverk som har bias inskrivet i sig baserat på den doxa som existerar i ett relativt isolerat fält. När sedan de som styr plattformen själva använder den blir de konstant rekommenderade mer information de redan håller med om, där algoritmerna sorterar bort åsikter de inte håller med om. Samtidigt rensar regelverket åsikter som ligger längre utanför doxan och gruppen därmed fortsätter att polariseras i brist på utmanande åsikter. Därmed finns det en risk att det har skapats en ond cirkel där fältets isolering orsakar en doxa som verkar allmängiltig och blir inbyggd i regelverket och algoritmerna. Åsiktsisoleringen riskerar därmed att öka som i sin tur driver fältet i en riktning där regelverket måste uppdateras för att passa en doxa i konstant rörelse mot en mer och mer extrem position. Om fältet flyttas för långt kan det också bli svårare för individer i andra fält att skaffa sig en förståelse.

Twitter har en del mekanismer som arbetar mot dessa effekter. Bland annat görs ett försök att ta kontext i åkt när beslut fattas och de anlitar föreningar och organisationer runt om i världen för att hjälpa till med att få en mer global förståelse genom så kallade *Trust & Safety Council* (Twitter Inc, 2021). Problemet är dock att valet av organisationer alltid

kommer att ske i enlighet med Twitters doxa. Kritik utifrån från politiker, journalister, aktivister och kommentatorer kan också hjälpa en del med att trycka tillbaka, så att åtminstone företagets toppskikt blir utsatta för en del motsatta åsikter. Den stora problematiken ligger dock i den brist på transparens som existerar inom algoritmerna samt angående vem som fattar beslut i hur individer ska sanktioneras vilket gör att vi inte kan få en riktigt bra insikt i hur de styrs av de fält de tillhör och den doxa som styr fältet. Ännu större är problemet med den despotiska karaktär Twitters regelverk har, där vi ej egentligen kan veta hur sanktionerna sker och om de sker i enlighet med regelverket eller om det finns andra värderingar som styr. Samtidigt har algoritmerna skapats i det fältet. Eftersom algoritmer inte är förmögna att göra val, har de också väldigt svårt för att se nyanser vilket gör dem oförmögen att ens kunna skaffa sig ett praktiskt förnuft för de fält som regleras, då alla potentiella kontexter måste programmeras in i algoritmerna för att tas i akt.

Detta är något att tänka på när man pressar plattformarna att ta bort oönskat material såsom hat eller missinformation. Hat är en väldigt bred term och är sällan väldefinierad när det ställs krav på att hat ska tas bort från nätet. Att bedöma vad som är sanning är också långt ifrån enkelt. Filosofer är inte ens överens om vad sanning är (Baggini, 2018). Frågan man måste ställa sig är då, huruvida man vill ge odemokratiska sociala medieföretag tolkningsföreträde för vad som ska anses vara hat och därför plockas bort från samhällsdebatten? Vill man tillåta dem vara grindvakter för vad som ska anses vara sanning på nätet och kan man lita på att de ska hålla vid sanningen om de går emot deras egna intressen eller den doxa som styr fältet de befinner sig i?

Något som inte heller bör underskattas är implikationerna av New York Posts avstängning från Twitter (Manskar, 2020). Twitter har blivit ett mycket viktigt verktyg för många journalister. Det är ett sätt att nå ut till läsare, snabbt få reda på nyheter att skriva om samt se vad politiker pratar om. I stor utsträckning har även det som sägs på Twitter blivit nyheter i sig själva. Vad händer då när ett nyhetsbolag blir avstängt för att rapportera på en story? Storyn som New York Post rapporterade borde, under normala förhållanden, vara en guldgruva för ett nyhetsbolag. Potentiell korruption relaterat till en presidentkandidat strax innan ett val (Morris & Fonrouge, 2020). Men, när ett bolag blir

avstängt måste alla andra bolag överväga huruvida det är värt att riskera att hela bolaget tappar sin tillgång till Twitter för en enda story. Inga rapporter kunde hittas om avstängningar för andra perspektiv på storyn till exempel att skriva att rapporteringen potentiellt var del av en desinformationskampanj (Bertrand, 2020). Risken som kommer med detta, är att det kan räcka för Twitter att stänga av ett nyhetsbolag för att de rapporterat ett perspektiv som Twitter inte tycker om. Sen blir det enda perspektivet som kan rapporteras angående storyn det motsägande perspektivet. Än värre blir det av att Twitter och andra sociala medieföretag ibland använder sig av nyhetsbolagen som källa för att faktakolla information på sin plattform. På så sätt kan de sedan sätta en stämpel på tweets som repeterar den informationen de från början stängde av, och på så sätt mer eller mindre tvätta information till sanning som sträcker sig långt utanför plattformen. Detta gör att Twitter och andra plattformar har en enorm potential till informationskontroll.

8. Slutsats

När det kommer till rättsliga principer påvisar Twitter tydliga brister i sitt regelverk. Otydligheter i regler, policys och *ToS*, med ett väldigt stort tolkningsutrymme tillsammans med en kraftig brist i procedurrell transparens, lämnar ett regelverk som är svårt att förhålla sig till. Twitter befinner sig dock i en nästan omöjlig position där en konstant påtryckning från flera motstridiga politiska håll, annonsörer som är rädda om vad som associeras med deras varumärken samt egna ekonomiska intressen som alla konstant sätter press på hur regelsystemet ska vara utformat och appliceras. Det största problemet i regelverket är dock dess despotiska karaktär där Twitter ger sig rätt att ta bort vad de vill av ingen och alla anledningar. Det lämnar ett regelverk där allt annat är mer eller mindre meningslöst och är mindre regler och mer i karaktären av riktlinjer eller indikationer. Som Nicolas Suzor (2018) påpekade, är dessa regelverk utvecklade för att maximera plattformarnas kontroll samtidigt som de avser att minimera deras ansvar (Suzor, 2018, s. 8). Därför framstår det som att Twitters prioritet inte har varit att ta hänsyn till rättsliga principer i deras regelverk. Ett väl utformat regelverk, i förhållande till rättsliga principer, kan inte prioritera att skydda den som agerar som suverän över allt annat. Det slutar upp med ett despotiskt regelverk som därför inte kan efterfölja rättsliga principer på ett tillfredsställande sätt.

Den makten som sociala medieplattformarna har är också oroväckande för ett fritt och demokratiskt samhälle. Att tillåta dem vara sanningens grindvakter där både journalister och folkvalda politiker kan blockeras från att nå ut till befolkningen på privata företags bevåg, kommer alltid att utgöra ett hot mot demokratin. Där fria val är beroende av att väljarkåren har tillgång till den information de behöver för att göra ett informerat val.

Att de flesta plattformarna utgår från samma geografiska område gör också att man riskerar att få en dominant doxa som förtrycker andra perspektiv utan att de som utför förtrycket ens själva är medvetna om det. Yttrandefrihet har alltid varit till för att skydda de svaga, då de starka kunnat skydda sin egen rätt. Att lämna över kontrollen över yttrandefriheten online till några av de största och mäktigaste företagen i världen riskerar att de svaga tystas ned även när plattformarna utgår från ett nobelt motiv att skydda svaga och utsatta. Det visar en viss ironi i Langvardts kommentarer att plattformar den närmaste och snabbaste censur världen någonsin skådat (Langvardt, 2018, s. 1353). När plattformarna är centrerade i ett fält är frågan huruvida de som censurerar kan se utanför den doxa som styr till den grad att de faktiskt kan bedöma vem som är svag eller utsatt om det inte faller i linje med den doxa de följer.

Risken för gruppolarisering inom de fält plattformarna befinner sig blir även stor, när alla plattformar har sin utgångspunkt i områden som är dominerade av liknande företag. När sedan den algoritmiska kontroll som reglerar vad man ser på plattformarna efter ens egna preferenser. Riskerar man även att skapa så kallade filterbubblor (Helberger, et al., 2018, s. 192), där individerna som arbetar på plattformarna ytterligare isoleras från andra synpunkter. Samtidigt isoleras andra på plattformarna på ett liknande sätt i sina egna filterbubblor, vilket riskerar att polarisera samhället och skapa konflikter som uppstår på grund av en brist på en enhetlig bild på världen.

8.1. Förslag till lagstiftare

Vad kan då göras för att förbättra detta? Det är inte sannolikt att Twitter och de andra plattformarna kan eller vill fixa den problematiken som uppstår. Det är därför tydligt att ny lagstiftning behövs. Europakommissionens två lagförslag för *rättsakten om digitala marknader* (EUROPEISKA KOMMISSIONEN, 2020) och *rättsakten om digitala tjänster* (EUROPEISKA KOMMISSIONEN, 2020) är ett bra steg i rätt riktning där *rättsakten om digitala marknader* kan öppna upp för marknadskonkurrens som skulle kunna öppna upp för företag som inte är centrerade i Silicon Valley och därför ge inflytande från fler fält. *Rättsakten om digitala tjänster* har möjligheten att göra ännu mer. Större tydligheter i *ToS* är ett stort plus även om det skulle behöva vara bredare och även gälla regler och policys. Sen att öppna upp för högre transparens och mer ansvar från plattformarna samt ge användare större möjlighet att överklaga felaktiga beslut är ett steg i rätt riktning. Det är dock det minsta möjliga efter att inte göra något alls enligt Langvardts lista över potentiella åtgärder (Langvardt, 2018). Rättsakten skulle även tillåta för en liten del av nästa punkt på Langvardts lista genom att låta användare välja bort profilbaserad annonsering (Langvardt, 2018). Vi skulle dock behöva vissa inslag av de striktare punkterna på listan. Att tillåta användare att välja hur algoritmerna ska appliceras skulle kunna tillåta att plattformarna har vissa mindre strikta regler. Det är dock inte sannolikt att de skulle ge upp den kontroll de har nu. Därför skulle det behövas regleringar i hur de tillåts att moderera plattformarna. Det skulle inte enbart stärka användares rättigheter. Det skulle även kunna hjälpa plattformarna själva att stå emot en del av påtryckningarna utifrån som leder till förhastade policys, då de har lagstiftning att lägga skulden på för varför de inte gör något åt det som är mest upprörande den här veckan. Det behövs även ett förbud mot att få sanktionera av alla och ingen anledning.

9. Framtida forskning

Som utveckling på den här uppsatsen hade det naturligtvis varit av stort intresse att undersöka andra plattformars regelverk för att se huruvida de har liknande problem med deras regelverk. Exempelvis en komparativ analys av de olika plattformarnas regelverk i syfte att undersöka hur de förhåller sig till varandra. Om flera har liknande brister, hade det även varit av intresse att se huruvida bristerna verkar vara av samma karaktär och därmed reflekterar samma eller liknande fält. Det finns även utrymme att utforska hur användare förhåller sig till dessa regler där man genom exempelvis enkätundersökning eller intervjuer kan bygga upp ett empiriskt underlag för att se huruvida dessa brister anses problematiska bland användare samt i vilken utsträckning användare är medvetna om hur regelverket är utformat och fungerar. Om *rättsakten om digitala tjänster* (EUROPEISKA KOMMISSIONEN, 2020) röstas igenom och öppnar upp för en större insikt för forskare, skulle det även vara av stort intresse att återkomma för att undersöka de mer dolda delarna av regelverket. Exempelvis hade det varit av stort intresse att kunna undersöka i vilken utsträckning algoritmerna används för att genomdriva regler. Även, om möjligt, att kunna få en insikt i hur beslut fattas och vem som fattar dem för att kunna vidare utforska Fullers desiderata som riktar sig mer mot genomdrivandet av regler snarare än hur reglerna är utformade.

Referenser

Andersen, H. & Kaspersen, L. B., 2007. *Klassiska och modern samhällsteori*. 3:9 red. Lund: Studentlitteratur AB.

Baggini, J., 2018. *A SHORT HISTORY of TRUTH: Consolation for a Post-Truth World*. Milton Keynes: Clays Ltd, Elcograf S.p.A.

Bertrand, N., 2020. Hunter Biden story is Russian disinfo, dozens of former intel officials say. *Politico*, 19 10.

BIDEN v. KNIGHT (2021) FIRST AMENDMENT INSTITUTE AT COLUMBIA UNIV..

Bourdieu, P., 1999. *Praktiskt förnuft: Bidrag till en handlingsteori*. 1 red. Göteborg: Bokförlaget Daidalos AB.

Bryman, A., 2018. *Samhällsvetenskapliga metoder*. 3:1 red. Stockholm: Liber AB.

Department of Elections, 2020. *Department of elections*. [Online]

Available at: <https://selections.sfgov.org/november-3-2020-election-results-summary> [Använd 17 05 2021].

Detrow, S., 2018. What Did Cambridge Analytica Do During The 2016 Election?. *npr*, 20 03.

Europaparlamentet, 2021. *Europaparlamentet*. [Online]

Available at:

<https://www.europarl.europa.eu/news/sv/headlines/society/20210204STO97129/social-a-medier-och-demokrati-vi-behover-lagar-inte-riktlinjer>

[Använd 17 05 2021].

EUROPEISKA KOMMISSIONEN, 2020. *Förslag till: EUROPAPARLAMENTETS OCH RÅDETS FÖRORDNING: om en inre marknad för digitala tjänster (rättsakten om digitala tjänster) och om ändring av direktiv 2000/31/EG*, Bryssel: EUROPEISKA KOMMISSIONEN.

EUROPEISKA KOMMISSIONEN, 2020. *Förslag till: EUROPAPARLAMENTETS OCH RÅDETS FÖRORDNING: om öppna och rättvisa marknader inom den digitala sektorn (rättsakten om digitala marknader)*, Bryssel: EUROPEISKA KOMMISSIONEN.

- Foer, F., 2017. Facebook's war on free will. *The Guardian*, 19 09.
- Frenken, K. & Fuenfschilling, L., 2021. The Rise of Online Platforms and the Triumph of the Corporation. *Sociologica*, 14(3), ss. 101-113.
- Fuller, L., 1969. *The Morality of Law*. Revised edition red. New Haven; London: Yale University Press.
- Gadde, V., 2020. *Twitter*. [Online]
Available at: <https://twitter.com/vijaya/status/1316923549236551680>
[Använd 15 05 2021].
- Gillespie, T., 2015. Platforms Intervene. *Social Media + Society*, Volym 1.
- Gilovich, T., Keltner, D., Chen, S. & Nisbett, R. E., 2019. *Social Psychology*. 5:e red. New York: W. W Norton & Company, Inc..
- Helberger, N., Karppinen, K. & D'Acunto, L., 2018. Exposure diversity as a design principle for recommender systems.. *Information, Communication & Society*, 21(2), ss. 191-207.
- Kingson, J. A., 2020. Exclusive: \$1 billion-plus riot damage is most expensive in insurance history. *Axios*, 16 09.
- Langvardt, K., 2018. Regulating online content moderation. *GEORGETOWN LAW JOURNAL*, 106(5), ss. 1353-1388.
- Lessig, L., 2006. *Code - Version 2.0*. 1:a red. New York: Basic Books.
- Levinson, T., 2021. Washington, DC protests in pictures: Shocking photos of the chaos. *New York Post*, 06 01.
- Lindstedt, I., 2019. *Forskningens Hantverk*. 2:2 red. Lund: Studentlitteratur AB.
- Manskar, N., 2020. Twitter, Facebook censor Post over Hunter Biden exposé. *New York Post*, 14 10.
- McMillan, R., 2020. Twitter Unlocks New York Post Account After Two-Week Standoff. *The Wall Street Journal*, 30 10.

Morris, E.-J. & Fonrouge, G., 2020. Smoking-gun email reveals how Hunter Biden introduced Ukrainian businessman to VP dad. *New York Post*, 14 10.

Robinson, E. P. & Zhu, Y., 2020. Beyond "I Agree": Users' Understanding of Web Site Terms of Service. *Social Media + Society*, Volym 6.

Scola, N., 2020. How Covid-19 pushed Twitter to fact-check Trump's tweets. *Politico*, 27 05.

Sharma, S., 2021. Sheryl Sandberg faces criticism over denial of Facebook's role in violence at Capitol riots. *The Independent*, 15 01.

SFS 2018: 1744. Brottsbalk.

Suzor, N., 2018. Digital Constitutionalism: Using the Rule of Law to Evaluate the Legitimacy of Governance by Platforms. *Social Media + Society*, 4(3).

TT, 2021. Merkel ifrågasätter att Trump Twittertystas. *Aftonbladet*, 11 01.

Twitter Inc., 2019. *Glorification of violence policy*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies/glorification-of-violence>
[Använd 17 05 2021].

Twitter Inc., 2019. *Twitter*. [Online]

Available at:

https://blog.twitter.com/en_us/topics/company/2019/worldleaders2019.html

[Använd 15 05 2021].

Twitter Inc., 2020. *Ban evasion policy*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies/ban-evasion>

[Använd 17 05 2021].

Twitter Inc., 2020. *Child sexual exploitation policy*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies/sexual-exploitation-policy>

[Använd 17 05 2021].

Twitter Inc., 2020. *Twitter Terms of Service*. [Online]

Available at: <https://twitter.com/en/tos>

[Använd 17 05 2021].

Twitter Inc., 2021. *About specific instances when a Tweet's reach may be limited*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies/twitter-reach-limited>

[Använd 17 05 2021].

Twitter Inc., 2021. *Company: It's what's happening..* [Online]

Available at: https://blog.twitter.com/en_us/topics/company.html

[Använd 17 05 2021].

Twitter Inc., 2021. *Coordinated harmful activity*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies/coordinated-harmful-activity>

[Använd 17 05 2021].

Twitter Inc., 2021. *COVID-19 misleading information policy*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies/medical-misinformation-policy>

[Använd 17 05 2021].

Twitter Inc., 2021. *Curation style guide*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies/curationstyleguide>

[Använd 17 05 2021].

Twitter Inc., 2021. *General guidelines and policies*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies#twitter-rules>

[Använd 17 05 2021].

Twitter Inc., 2021. *Hateful conduct policy*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies/hateful-conduct-policy>

[Använd 17 05 2021].

Twitter Inc., 2021. *Our range of enforcement options*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies/enforcement-options>

[Använd 17 05 2021].

Twitter Inc., 2021. *Previous Terms of Service*. [Online]

Available at: <https://twitter.com/en/tos/previous>

[Använd 17 05 2021].

Twitter Inc., 2021. *Release Notes*. [Online]

Available at: https://twitter.com/i/release_notes

[Använd 17 05 2021].

Twitter Inc., 2021. *Synthetic and manipulated media policy*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies/manipulated-media>

[Använd 17 05 2021].

Twitter Inc., 2021. *Twitter*. [Online]

Available at: https://blog.twitter.com/en_us/topics/company/2020/suspension.html

[Använd 17 05 2021].

Twitter Inc., 2021. *Twitter Rules and policies*. [Online]

Available at: <https://help.twitter.com/en/rules-and-policies#twitter-rules>

[Använd 17 05 2021].

Twitter Inc., 2021. *Twitter.com*. [Online]

Available at: [Twitter.com](https://twitter.com)

[Använd 17 05 2021].

Twitter Inc, 2021. *Trust and Safety Council*. [Online]

Available at: <https://about.twitter.com/en/our-priorities/healthy-conversations/trust-and-safety-council>

[Använd 18 05 2021].

Twitter Safety, 2021. *An update following the riots in Washington, DC*. [Online]

Available at: https://blog.twitter.com/en_us/topics/company/2021/protecting--the-conversation-following-the-riots-in-washington--.html

[Använd 17 05 2021].

Virginia Tech Libraries, 2020. *Internet Archive Wayback Machine*. [Online]

Available at:

<https://web.archive.org/web/20201204000218/https://help.twitter.com/en/rules-and-policies/coordinated-harmful-activity>

[Använd 17 05 2021].

Watt, L., 2020. Taiwan Says It Tried to Warn the World About Coronavirus. Here's What It Really Knew and When. *Time*, 19 05.

Wendling, M., 2021. QAnon: What is it and where did it come from?. *BBC*, 06 01.

World Health Organization (WHO), 2021. *Twitter.com*. [Online]

Available at: <https://twitter.com/who/status/1217043229427761152?lang=en>

[Använd 15 05 2021].

Bilaga 1: Litteratursökningar

Peer review är valt för alla sökningar och sökningarna är gjorda i abstract för att reducera antalet träffar till de mest relevanta.

Sökning 1

LUBsearch

"community guidelines" OR "terms of use" OR "Terms of service"

AND law OR legal OR legislation

AND "social media"

25 träffar

Sökning 2

LUBsearch

"community guidelines" OR "terms of use" OR "Terms of service"

AND "freedom of speech" OR "Free speech" OR "Freedom of opinion" OR "Free opinion" OR

"freedom of expression" Or "free expression"

AND "Social media"

3 träffar

Sökning 3

LUBsearch

"community guidelines" OR "terms of use" OR "Terms of service"

AND "social media"

AND censor OR censorship OR censored OR banned

2 träffar

Sökning 4

LUBsearch

AI OR algorithm*

AND "social media"

AND censor OR censorship OR censored OR banned

22 träffar

Sökning 5

LUBsearch

AI OR algorithm*

AND law OR legal OR legislation

AND "social media"

92 träffar

Sökning 6

LUBsearch

AI OR algorithm*

AND "freedom of speech" OR "Free speech" OR "Freedom of opinion" OR "Free opinion" OR
"freedom of expression" Or "free expression"

AND "social media"

6 träffar

127 träffar kvar efter hopslagning av dubletter och borttagande av konferenspapper

120 träffar kvar efter språkrensning

67 träffar kvar efter initial genomgång av abstract

34 träffar kvar efter en andra genomgång av abstract

5 topciterade verk (över 50 citeringar)

2 kvar efter kursivläsning

7 toprelevanta (över 20 citeringar, skrivna inom senaste året eller extremt relevant till uppsatsen)

2 kvar efter kursivläsning

1 tillagda från referenser

Totalt 5 texter.

Bilaga 2: Teman

Alla använda teman med numrering:

- (1) Att misslyckas skapa regler alls.
- (2) Att misslyckas med att publicera eller åtminstone att misslyckas informera den som förväntas följa regeln.
- (3) Missbruk av retroaktiv lagstiftning
- (4) Misslyckande att göra regler som går att förstå
- (5) Användande av motsägande regler
- (6) Regler som kräver agerande som ligger utanför förmågan hos den som förväntas följa reglerna.
- (7) Regler som ändras så ofta att det inte går att förhålla sig till dem
- (8) Ett misslyckande i att överensstämman angivna regler med hur de utförs
- (9) Vederbörlig process
- (10) Lika inför lagen
- (11) Nomokrati