

Jonas Andersson Schwarz & Johan Hammarlund

Kontextförlust och kontextkollaps

Metodproblem vid innehållsanalys av sociala medier

Abstract

The article proposes that social media platforms enable large volumes of user-driven circulation of media content, and argues for a combination of qualitative and quantitative considerations when analysing data from such platforms. Issues of context are vital; context must be understood both qualitatively (cultural setting) and quantitatively (statistical reference points for comparison). The authors emphasise that the possibilities of ‘big data’ should not tilt analyses so that sensitivities to subtler meanings are lost. By examining a recent research project of our own, examples are given of how topological network analysis can be successfully combined with close readings of strategically selected parts of the data and how, by doing so, context shifts can be identified that increase the reliability of the analysis. Consequently, it is recommended that mere number crunching is not enough, and that questions of ‘how,’ ‘why,’ and ‘whether’ are required in order to understand the phenomena in their societal settings.

Keywords: Twitter, data analysis, epistemology, knowledge context, competitive intelligence

Baserat på erfarenheter från ett samarbete mellan medieforskaren Jonas Andersson Schwarz och omvärldsbevaknings- och medieanalysföretagen M-Brain och Retriever vill vi här lyfta fram ett antal metodfrågor som vi ser som centrala inom databaserade metoder, både vad gäller strikt vetenskapliga tillämpningar och branschpraxis. Vi tar vår utgångspunkt i egen tillämpad forskning där vi kombinerat analys av relationer och flöden i sociala nätverk med analys av ostrukturerad text i ett projekt finansierat av Internetfonden år 2014-2015: ”Vad gör en politisk utsaga framgångsrik: den användardrivna kommunikationens villkor” (Andersson Schwarz m.fl. 2015). Projektet undersökte olika aspekter av hur politiskt kodad kommunikation cirkulerar på svenskspråkiga Twitter samt hur press och sociala medier samspelar. Vi vill därmed bidra både till en forskningsmässig metodmedve-

tenhet och till att belysa omvärldsanalyssektorns mer praktiska tillvägagångssätt.

En tidig insikt under arbetet var nödvändigheten av kontext – och att kontext, för bästa resultat, bör förstås både kvalitativt och kvantitativt. Kvalitativt kan kontext förstås som *kulturellt sammanhang* men kontext måste också förstås som *referenspunkter för kvantitativa jämförelser*. Denna senare aspekt kan ibland glömmas bort, dels i renodlat kvalitativa studier där frågor om representativitet helt väljs bort, dels när forskare inom mer kvantitativt orienterad forskning upprättar mätvärden som i själva verket är mycket svåra att överhuvudtaget dra slutsatser ifrån, då dessa värden inte är självförklarande utan måste sättas i relation till andra värden. Hur mycket är ’mycket’?

Framför allt ser vi en risk att analysen går miste om både kvalitativ och kvantitativ

kontextualisering när tiden är knapp och resurserna begränsade. Ett relaterat problem är att analysförsök ofta görs trots att datan egentligen är allt för knapphändig avseende kontext, eller rentav inte tillräckligt vederhäftig till att börja med: Dagens kommersiella internetbaserade plattformar för socialt utbyte och mediekonsumtion (vi åsyftar primärt de Alphabet- och Facebookägda plattformarna) brister i transparens beträffande datans skick och mätapparaters tillförlitlighet. Forskare tillåts inte ta reda på huruvida resultaten är präglade av olika former av dolda bias; detta är ett akut kunskapsproblem och berör därmed även regleringen av dessa plattformstjänster (Andersson Schwarz 2016).

Vi diskuterar följaktligen kontext i både kvalitativ och kvantitativ bemärkelse, och vi kommer att koppla diskussionen till två radikalt skilda förhållningssätt som vi har observerat, beträffande empirisk datainsamling och -analys: Vi kallar den första för *närläsning* – en anekdotisk, abduktiv *ad hoc*-analys där analytikern ofta befinner sig 'inuti' skraddarsydda realtidsflöden, precis som vardagliga medieanvändare gör. Här är en kritisk faktor den algoritmiska filtrering som riskerar att uppstå, då varje användare är kringskuren av de sätt på vilka 'frontend'-gränssnitten är utformade, ett fenomen som populärt kallats för 'filterbubblor' (Pariser 2011). Dagens kraftfulla specialistsökverktyg gör det visserligen möjligt att spräcka dessa bubblor men de kan också i värsta fall göra det allt för enkelt att hitta utsagor som stöder ens hypotes eller föreställningar. Följaktligen krävs en hög grad av metodologisk medvetenhet och självreflektion för att man inte själv ska skapa en 'manuell' motsvarighet till de algoritmiska filterbubblorna. Det andra förhållningssättet kallar vi *fjärrläsning* – en sammanfattande, induktiv *post hoc*-analys, där analytikerns uppgift blir att uppskatta datamängder genom att retroaktivt söka eller på annat sätt erhålla urval ur infrastrukturens 'backend', det vill säga de underliggande databaserna, och där data-mängderna är för stora för manuell innehållsanalys så att analytikern således måste förlita sig på datadrivna, automatiserade bryt-

ningar av materialet (exempelvis algoritmisk sortering, regressionsanalys, 'deep learning'). Fjärrläsning ska inte ses som synonymt till det inom digitala humaniora så vanliga begreppet distansläsning ('distant reading'), intimt förknippat med Franco Moretti (2013), då vi här även inkluderar mindre ambitiösa kvantitativa ansatser. Vi menar att styrkan i dessa nya metoder, i en akademisk kontext, framförallt uppstår när de kombineras med andra, mer konventionella metoder då triangulering kan generera både kontext och representativitet.

Precis som Lewis m.fl. (2013) och Gaffney & Puschmann (2014) förordar vi en pragmatisk ansats, där datalogiska och manuellt tolkande metoder kombineras; en 'kvali-kvantitativ' ansats (Latour m.fl. 2012). Genom att omsorgsfullt kombinera tekniker går det att bibehålla fördelarna med konventionell innehållsanalys (systematisk stringens och kontextuell känslighet) samtidigt som storskalig kapacitet kan erhållas genom stora datavolymer och beräkningsmetoder baserade på algoritmer (vilket kan synliggöra relationer och samvariationer, samt påvisa olika aggregationers relativa storlekar).

Vi stöder oss på ett tentativt schema över omvärldsanalysens förändrade historiska karaktär, föreslagit av Boullier (2015), samt forskningsartiklar inom fältet kritiska sociala mediastudier. Vårt teoretiska ramverk är socialkonstruktivistiskt men med stor tonvikt även på den icke-mänskliga agensen; det vill säga, en cybernetiskt influerad konstruktivism där begrepp som infrastruktur och system är lika viktiga för att förstå begreppet kontext som sociokulturella kategorier. Vi har inte utrymme att ingående diskutera forskningsetik och personlig integritet men det bör framgå att en kontextkänslig analysmodell per definition aktualiserar sådana aspekter.

Varför Twitter?

Den enda av de stora, populära plattformstjänsterna som ger någorlunda goda möjligheter till översikt är Twitter, då den absoluta majoriteten av dess data är (åtminstone teoretiskt sett) offentlig, samt att den begränsade läng-

den på varje statusuppdatering (140 tecken) borgar för effektiv informationsinhämtning och jämförande analys (Tao m.fl. 2014).

Twitter är en arena där opinionsbildare, journalister och medborgare möts och interagerar med varandra, liksom även institutionella användare (organisationer, partier, företag). Det rör sig både om offentliga samtal och diskussioner inom små grupper. Som kommunikationsplattform förefaller Twitter därför ha ett påfallande samhällsinflytande; plattformen fungerar på många sätt i symbios med journalistiska massmedier (Broersma & Graham 2012, Lasorsa m.fl. 2012, Paulussen & Harder 2014). I en stor studie av användarbetenden observerades att Twitter är mer att betrakta som en medieplattform än ett socialt nätverk; 85 procent av inläggen var nyhetsrelaterade (Kwak m.fl. 2010). Samtidigt är Twitter helt annorlunda än traditionella medier eftersom användarna agerar 'sociala sensorer' (Sakaki m.fl. 2010) som kan ge förstahandsinformation om olika aspekter av nyhetskändelser. Mycket av den kommunikation som sker på Twitter och som etiketteras som politisk kommer dock från aktörer som redan är involverade professionellt i politiken (Larsson 2014, 2015, 2016).

Twitters roll som agendasättande arena, i kombination med dess öppenhet – och den relativt lättillgängliga sociala mediedata detta medför – gör att plattformen är en tämligen välanvänd datakälla inte bara bland journalister, men även i den internationella medieforskningen (Lewis m.fl. 2013, Weller m.fl. 2014), trots att det är långt ifrån ett populärt medium rent demografiskt. Uppskattningsvis sex procent av de svenska internetanvändarna använder tjänsten dagligen (Findahl & Davidsson 2015) – samtidigt visar vår forskning att Twitter tveklöst är den sociala medieplattform som det hänvisas oftast till i dagstidningarna (Andersson Schwarz m.fl. 2015: 20).

Twittercensus, en årlig kartläggning av det svenskspråkiga Twitter (Brynolf 2015), visualiserar dess olika nätverk i form av följarelationer. Dessa nätverk är samtidigt överlappande med mer eller mindre permanenta intressenätverk som formas kring olika fe-

nomen och händelser. De nätverk som lyftes fram i vår studie var objekt som i viss mening skapades ur dedikerade sökningar på vissa termer eller hashtaggar. Skärningar som görs av 'nätverk inom nätverken' av det här slaget går att använda som populationer i sin egen rätt, vilka de med tillgång till data och verktyg kan hänvisa till, rentav använda retoriskt som led i argumentation ('frågan X debatteras just nu flitigt på Twitter'). Huruvida de observerade deltagarna själva anser sig ingå i nätverket är ofta oklart, men en högst angelägen fråga. Marwick & Boyd (2011) dokumenterar olika avsändarstrategier som användarna har angående detta. Samtidigt bör också läsaren ha en strategi för att hantera att den information denne möter i sitt flöde (eller *dashboard*, se nedan) kan vara skapat utifrån en helt annan kontext än den användaren själv verkar inom och därför kan misstolkas.

Relevans för praktisk omvärldsanalys och teoretisk forskning

Nätverksplattformar såsom Twitter och Facebook gör olika former av närläsning och fjärrläsning av stora mängder information förrådiskt enkla. Enkla, eftersom potentiellt stora mängder information kan hämtas in med ett fåtal knapptryckningar. Förrådiskt, eftersom processerna och algoritmerna som styr vilken information som ges tillgång till är svårgenomträngliga; de saknar vad Diakopoulos (2014) kallar 'algorithmic accountability'. Medel i form av datatillgång, digitala verktyg såsom dashboards (kontrollpaneler för sökningar i datafloden), samt kompetensen att använda dem möjliggör strukturerade sökningar som genererar allt ifrån enstaka citat till stora mängder data om vissa grupper betenden och kommunikation. Artefakter av detta slag kan sedan omvandlas till retoriskt övertygande visualiseringar, ofta utan tillräcklig käll- och metodredovisning.

Två godtyckliga element spelar in när sociala grupperingar av det här slaget observeras: För det första är grupperingar på nätet i regel 'självrekruterade' (självalda gemenskaper) eller resultat av att aktörer med specifika intres-

sen spontant går samman. Därmed är de sällan *a priori* representativa för större populationer. För det andra innebär själva akten att välja ut en gruppering att analytikern gör en avgränsning och därmed 'skär ut' en subgrupp ur en större population. Därtill kommer frågan om huruvida datan är vederhäftig, något som blir svårt att veta på grund av avsaknaden av transparens från plattformsföretagen. Många har som sagt vittnat om en så kallad 'filterbubbla', att osynliga algoritmer styr viken information vi exponeras för (Pariser 2011). De dashboards som många omvärldsbevakningsföretag erbjuder kan verka som motmedel till denna. Men utan en hög självmedvetenhet om ens egna föreställningar riskerar analytikern att själv upprätta s.k. *confirmation bias*, i förlängningen en mental bubbla inom vilken de egna förutfattade teorierna bekräftas. Skruvar du tillräckligt länge på inställningarna i verktygen för sök, analys och visualisering så ökar chansen att du till slut får 'rätt' svar: 'hårda data' som stödjer din ståndpunkt.

Både närläsning och fjärrläsning av de här slagen möjliggör på många sätt 'nya' typer av information om mänskligt beteende – på ett enkelt och ofta kostnadseffektivt sätt. De utgör därför på många sätt en sorts tredje generationens kvantitativ metod inom samhällsforskningen (Boullier 2015, se nedan). En risk är dock att den kontext i vilken information skapats går förlorad. En sådan form av förlust är kontextkollaps (Marwick & boyd 2011), vilket åsyftar de mekanismer då olika användares föreställningsvärldar och tolkningar krockar, eftersom information idag kan röra sig mellan nätverk och grupper i svårförutsedda mönster.

Kontext blir i vår diskussion i hög utsträckning synonymt med nätverk – och *nätverksanalys* blir följaktligen ett sätt att (re)kontextualisera information som ryckts loss ur sitt ursprungliga sammanhang. Det är värt att notera att inte bara individer agerar i nätverken. Mängder av institutionella aktörer är också aktiva, automatiserade konton (s.k. *bots*) därtill, och många individer tycks växla mellan att agera åt/som sig själva och åt/som institutioner. Diskurs och (digital) praktik inom de medierade sociala nätverken samverkar

vidare med de fysiska sociala nätverken i en ömsesidig växelverkan. Till exempel kunde vi, i en studie genomförd i samarbete med Institutet för mediestudier (Andersson Schwarz & Hammarlund 2016), se hur den svenska diskursen kring migration i såväl sociala som redaktionella medier (dessa är centrala aktörer även i de sociala nätverken) påverkades när relationerna i nätverken förändrades, exempelvis när drunknande flyktingar uppmärksammades i massmedia. Nätverk är, i denna bemärkelse, samtidigt både medierade och köttsliga, precis som vår tillvaro idag. De är överlappande, ständigt föränderliga och samverkar med andra nätverk i en stor svåröverskådlig väv. För att parafrasera Herakleitos så är det ej möjligt att två gånger nedstiga i samma nätverk – följaktligen kommer vår kunskap alltid att vara bristfällig och i någon mån föråldrad. De digitala nätverken lämnar visserligen hela tiden spår (mer eller mindre öppna) och dessa spår kan analyseras. Betyder det att vi genom att studera en aspekt av den nätverkande människan, exempelvis spåren av medierad kommunikation på en viss arena, kan förstå helheten? Nej, men kanske att vi har fått några nya pusselbitar.

Datatillgång

Det finns dilemman beträffande datatillgång och kontroll som möter alla som på något vis interagerar med digitala medieplattformar (Puschmann & Burgess 2014). Dessa frågor är själva grundbulten för alla andra resonemang (empiriska, epistemologiska, regleringstekniska) vi kan föra om dessa plattformar. Som boyd & Crawford (2012) och även Zelenkauskait & Bucy (2016) har konstaterat är det hela ytterst en fråga om makt.

Sociala medieplattformar skiljer sig åt i graden av dataåtkomst de är villiga att ge, och det är en myt att forskare fritt kan samla användargenererad data från dessa plattformar (Gaffney & Puschmann 2014, Tao m.fl. 2014, Vis 2013). Vad gäller Facebook kan majoriteten av användarkonton uppskattas vara slutna/privata i olika utsträckning (exakt hur många konton och vilken grad av öppenhet behand-

lar företaget Facebook dock som en affärs-hemlighet). Twitter, däremot, kännetecknas av en strukturell öppenhet som härrör ur en tydlig norm bland användarna att låta sina konton vara offentliga. Även Twitter har dock gradvis begränsat den fullskaliga tillgången till dess databibliotek (s.k. 'firehose access') till dataleverantören Gnip som köptes av Twitter år 2014. Även Facebook har exklusivt knutit sin kommersiella datatillgång till det London-baserade företaget Datasift. Genom s.k. applikationsprogrammeringsgränssnitt (API:er) kan begränsad tillgång till Twitters data visserligen erhållas av kunniga användare, men full tillgång till företagets data sker endast på kommersiell basis, och då genom Gnips *de facto* monopol.

Den typ av slutenhet som detta leder till drabbar framförallt tredjepartsaktörer som vill få en överblick i den sociala medieekologin: omvärldsbevakningsföretag samt akademiska forskare. För att få mer heltäckande tillgång tvingas dessa tredjepartsaktörer således att tillgå ett grossistled av dataleverantörer som för närvarande dominerar av plattformsföretagen själva, en trend som kan komma att ytterligare fördjupas.

Vår undersökning

Den studie som utgör basen för vårt empiriska material (Andersson Schwarz m.fl. 2015) inleddes 2014 med huvudsyftet att få en aktuell lägesbild av vad som cirkulerar inom Twittersfären som sådan, och mellan sociala och redaktionella medier. I projektet har vi huvudsakligen använt olika kvantitativa metoder (induktiv nätverksanalys samt konventionell innehållsanalys) för att hitta mönster och bättre kunna kategorisera det som återcirkuleras inom politisk debatt och opinionsbildning, samt illustrera några viktiga aspekter av samspelet mellan Twitter och redaktionella medier.

Projektet bestod av tre separata kvantitativa studier samt den fallstudie av nyheten om Pippi Långstrump som beskrivs nedan. Projektets egenart är att det kombinerar en induktiv nätverksanalys (d.v.s. en graf som

visar det topologiska mönster som användares interaktioner genererar) med konventionell innehållsanalys (d.v.s. manuell kodning av ett stickprov). Den nödvändiga datatillgången möjliggjordes genom samarbete med omvärldsbevakningsföretagen Retriever och M-Brain. Genom detta samarbete kunde vi samla in ett råmaterial av 109 000 Twitter-inlägg (inklusive retweets) och ett tusental tidningsartiklar, vilket i analys- och visualiseringsarbetet reducerades till knappt 60 000 relevanta tweets och retweets samt drygt 400 tidningsartiklar. Studien går att upprepa, med förbehållet att vi avsiktligt studerade Twitter under 'supervalåret' 2014, en politiskt turbulent period då det politiska landskapet i Sverige polariserades (vederlagt i bland annat Bergström m.fl. 2015). Detta tycks ha cementerats även på det politiska Twitter, där beteendet att retweeta (återcirkulera) utgör följer uppenbart gruppdynamiska mönster.

En del av vår studie som illustrerar många av de fenomen vi vill lyfta fram var fallstudien där vi analyserade reaktionerna på nyheten att tv-serien *Här kommer Pippi Långstrump* skulle moderniseras (Andersson Schwarz 2015). Kortfattat konstaterade vi att olika kluster (användare av hashtaggen #svpol respektive övriga användare på svenskspråkiga Twitter) reagerade olika på nyheten, samt att deras kommunikativa beteenden i klustret/nätverket också skiljde sig åt. En högljudd minoritet motsatte sig redigeringen och en liten grupp inom denna minoritet var väldigt aktiva på sociala medier med delning av inlägg. Vi analyserade de 50 mest spridda inlägg som nämnde Pippi den dagen och var avsiktligt markerade med #svpol. Av dessa uttryckte två tredjedelar sentiment; samtliga inlägg som gjorde detta uttryckte motstånd mot moderniseringen. Omvänt fann vi att bland de 50 mest spridda inläggen i den mycket större samlingen inlägg ej markerade med #svpol var sentimentet överväldigande positivt till moderniseringen. Dock var användarna själva sannolikt i väldigt varierande grad av uppfattningen att de över huvud taget tillhörde ett särskilt nätverk eller kluster (utöver Twitter i sig). #svpol är visserligen en etablerad hashtag (som har en

påvisad överrepresentation av högerextrema/främlingsfientliga, se Larsson 2014) men många av de övriga som skrev till sina följare/vänner om Pippidebatten skulle sannolikt inte anse att de utgjorde en part av någon tänkt gruppering av 'motståndare' till moderniseringen av Pippi i den öppna debatten utan den typen av kategorisering skapas endast *a posteriori* – i vår analys, baserad på deras kommunikation på Twitter i just det observerade ögonblicket. Hade vi enbart valt att använda hela Twitter som population så hade vi här missat en central slutsats, splittringen mellan de olika kluster/gemenskaper som förekommer inom populationen. Omvänt hade vi med en ofullständig analys av #svpol-användarna kunnat konstatera att denna grupp motsatte sig redigeringen med en relativt hög volym av inlägg i förhållande till antalet individer/konton. Värdet i analysen kommer först när vi särskiljer dessa nätverk samt tittar på deras karakteristika – mer exakt, hur #svpol avviker. Då synliggörs plötsligt ett nätverk inom nätverket med en hög grad av anonyma konton, en markant avvikande åsikt i sakfrågan (enligt de utsagor som görs), samt ett kommunikativt beteende där en liten grupp är väldigt aktiva med att retweeta inlägg och därigenom blåser upp gruppens delningssiffror.

Likaledes, i vår större rapport, parallellt med Pippistudien, använde vi, för att göra urvalet som låg till grund för analysen, dels sökordsbaserade funktioner i insamlingsverktyget (n=109 000), dels klustringsfunktioner i visualiseringsverktyget som visar hur retweetandet hänger samman (här blev drygt en tredjedel av urvalet relevant, n=37 959). I analys av det här slaget skapas alltså kategorier som syftar till att situera utsagor som annars hade förblivit lösryckta. Att därefter bädda in funna excerpter i en ny miljö, vår egen forskningsrapport, innebär i sin tur en kontextväxling som de observerade deltagarna svårligen kan ha förväntat sig.

En liknande typ av kontextväxling illustreras också av den kedja i vilken informationen om redigeringen av Pippi rörde sig (utförligare beskrivet i Andersson Schwarz 2015). Den ursprungliga pressreleasen från SVT fångades

upp av Aftonbladet som skrev en nyhet om detta. Denna nyhet fångades i sin tur upp av en användare på Facebook (som opponerade sig mot redigeringen), som i sin tur hängdes ut av en användare på Twitter (genom att denne ironiserade över en skärmbild av Facebookinlägget). När denna tweet sedan retweetades av en av landets mest inflytelserika opinionsbildare (som själv är kopplad till public service) fullbordades en referenskedja av olika budskap, där räckvidden eskalerades närmast exponentiellt. Kedjor av det här slaget överbryggar plattformar, kontexter, normer och beteenden, samt även lagrum.

Sådana kedjor där information hoppar mellan kontexter och nätverk är illustrativa för dagens komplexa kommunikationslandskap där individers 'privata' beteenden i vissa lägen allt mer liknar organisationers beteenden. Individuella användare på Twitter och andra plattformar kan idag använda en rad olika tekniker och redskap för att strukturerat samla in information ämnen och personer eller organisationer de finner intressanta. Genom att använda Twitters egna 'backend'-funktion *Twitter Analytics* kan användare analysera sin egna kommunikativa effektivitet. När tekniker av det här slaget är allmänt tillgängliga har vi en situation där individer med tekniska medel automatiskt och genom sökfunktionalitet kan bevaka sina intresseområden, där information lätt kan flyttas från kontext till kontext, och där relativt nya hybridtillämpningar som 'citizen journalism' och 'DIY-intelligence' (och i viss mån datajournalistik) möjliggörs. Den spontana individuella omvärldsbevakningspraktik som Hamrefors (1999) beskrivit i termer av 'spontaneous environmental scanning' har idag alltså kompletterats med en mer strukturerad och delvis automatiserad praktik. Teknikutvecklingen gör denna praktik tillgänglig för en stor majoritet av internetanvändarna, även om andelen som *de facto* använder sig av dessa metoder troligen är relativt liten. En liknande, men omvänd, uppluckring ses i den företagsekonomiska kunskapsbildningen om organisering i värdenätverk (Christensen 1997) som beskriver hur företag får tillgång till kunskap via relationsbyggande i organisa-

tionsöverskridande nätverk. Organisationer kan här ses växla mellan olika roller i samspel med andra organisationer (Hamrefors 2014), och det vore värdefullt att vidare utforska hur dessa tendenser förhåller sig till varandra.

Diskussion

Sociala medier som empiriska objekt

Kvantitativ innehållsanalys kan definieras som en forskningsmetod som syftar till objektiva, systematiska och kvantitativa beskrivningar av manifest [direkt observerbart] kommunikativt innehåll (Berelson 1952: 18, vår översättning). Med *manifest* menas olika texters 'uppenbara' innebörd, medan antydda eller underförstådda innebörder hänförs till den *latenta* nivån. Datorassisterade metoder kan med fördel tillämpas på manifest innehåll, då kodningen där ofta handlar om att tillämpa uppenbara logiska instruktioner (det vill säga, algoritmer), medan mänsklig arbetskraft alltjämt är oundgänglig när det gäller latenta dimensioner, då detta handlar om associativ tolkning (Sjøvaag & Stavelin 2012: 219).

I vårt projekt var t.ex. bevekelsegrunderna för twittrarna i fråga ej manifest. Vi kunde tydligt se omfattningen av den spridning som hade skett, men vi kunde inte med säkerhet veta varför upphovspersonerna bakom inläggen hade valt att skriva just dessa inlägg, och vi kunde inte heller med säkerhet veta varför de som hade retweetat dessa inlägg gjorde det.

Trots tjuvningen med datorbaserade beräkningsmetoder är mänsklig tolkning alltså fortfarande oundgänglig för innehållsanalys; både när det kommer till förförståelse och förmågan att inse sammanhang ('contextual sensitivity'; se Lewis m.fl. 2013: 48) och deras placering i den större, omgivande medieekologin. Medvetenheten och känslan för kontexten blir extra viktig i studier av *big data*, då de hisnande datavolymer riskerar att sporra forskare till att leta efter mönster utan att se de subtila dimensioner som finns inbäddade i medietexternas tillblivelse och cirkulation, samt att lockelsen ofta är stor att få de olika brytningar som görs av datan att passa färdiga

teoretiska modeller (boyd & Crawford 2012: 671).

Tydligt är att nätverkssamhället inte ger upphov till en jämn, ensartad population utan att mängder av olika, delvis överlappande klusterformationer och grupperingar uppstår, ofta med radikalt ojämlika styrkeförhållanden vad gäller enskilda deltagares inflytande. Huruvida de sociala gemenskaper som möjliggörs på nätet bidrar till starkare skillnader mellan grupper är fortfarande en i hög grad obesvarad fråga (Wadbring 2016). Forskning indikerar att olika åsiktskluster uppstår (Garimella m.fl. 2015). Man kan prata om fragmentisering (Bergström m.fl. 2015, Strömbäck 2015) eller rentav enklavisering (Sunstein 2007), men mycket beror på vad man lägger i begreppen. Hur stor är överlappningen mellan olika gruppformer? I vilken grad förekommer gruppöverskridande åsiktsutbyten? Även här blir kontexten avgörande för bilden av den större samhällsordning i vilken handlandet äger rum.

Det är en kvalitativ ny situation att självvalda åsikts- och intressegemenskaper kompletteras av algoritmisk filtrering där flöden moduleras (Andersson Schwarz 2015) genom infrastrukturella popularitets- och personaliseringfilter som varken är transparenta eller självvalda (Pariser 2011, Webster 2014, Zuiderveen Borgesius m.fl. 2016). I en pionjärstudie av personaliseringen av nyhetskonsumtionen fann Flaxman m.fl. (2016) motsägelsefulla resultat. Samtidigt som polariserande effekter konstaterades tycktes sociala medier bidra till gruppöverskridande exponering. Nischat medieinnehåll konsumeras normalt i tillägg till en mainstreamdiet (Bimber & Davis 2003, Trilling & Schoenbach, 2015). Kruket är att politik helt kan väljas bort (Flaxman m.fl. 2016, Zuiderveen Borgesius m.fl. 2016) med följd att den mer uttalat partiska mediadebatt som kvarstår kännetecknas av ömsesidigt repellerande grupper i högre grad än vad som annars vore fallet. De slutsatser som går att dra om huruvida medielandskapet verkligen förstärker fragmentisering är dock beroende av metod, val av fokus och val av kriterier/gränsvärden (Webster & Ksiazek 2012).

Trots heterogena populationer och svårjämförda grupper finns ett sätt att skapa åtminstone teoretiska jämförelsepunkter: att följa en och samma parameter konsekvent över tid. Oavsett mätvärdets eventuella övriga brister uppstår en jämförande tidsserie för mätvärdet i fråga. En annan jämförande dimension som ofta uppstår är möjligheten att jämföra likvärdiga artefakter. Inom kvantitativ medieanalys kan det till exempel handla om graden av förekomster av ett och samma sökord i olika miljöer. Därutöver, genom att göra kompletterande urval, till exempel av helt andra men likväl jämförbara sökord, går det att sätta förekomsterna i ett kvantitativt sammanhang där de uppmätta mängder man valt att fokusera på jämförs med andra uppmätta mängder under samma tidsperiod.

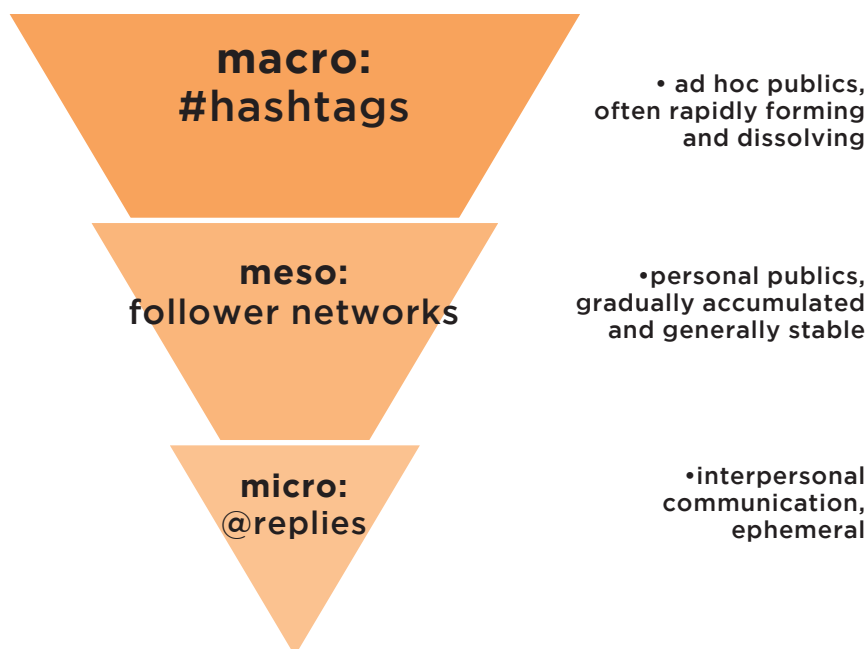
Detta förklarar våra egna metodval, då syftet var att upprätta en form av triangulering för att kunna belysa samma fenomen ur olika synvinklar, något vi generellt vill propagera för. Ju fler jämförande dimensioner som upprättas, desto rikare kontext och bättre validitet.

I studiet av Twitter kan tonvikt läggas på olika nivåer av närhet, där skilda sociala prak-

tiker hamnar i fokus beroende på vilken analysnivå som väljs. Helt olika optik erfordras; i vår studie opererade den automatiserade nätverksklusteranalysen på makronivån medan vår manuella innehållsanalys av enskilda tweets opererade på mikronivån. Våra kluster skapades dock utifrån retweetande, vilket Bruns och Moe (figur 1) beskriver som mer flyktig kommunikation på mikronivå (möjlig mesonivå). Genom att samla in stora volymer över tid skapar vi dock en aggregerad makrostruktur där vi kan se stabila mönster vad gäller såväl beteende som relationer.

Twitter och kontextkollaps

Kontextkollaps åsyftar olika konfliktartade skeenden beträffande de sociala normer som det sociala handlandet förväntas bygga på (Wesch 2009), inte minst beträffande människors förståelse av sig själva i relation till vad Mead (1934) en gång kallade 'den generaliserade andre' (förmågan att föreställa sig sitt beteende sett utifrån). När mänsklig samvaro determineras av internetplattformar som Facebook tenderar en rad sociala sammanhang, vilka tidigare varit rumsligt åtskilda, att kol-



Figur 1. Olika nivåer av kommunikativa utrymmen på Twitter

Källa: Bruns & Moe 2014: 20 (Faksimil)

lapsas till ett enda heterogent utrymme (Hogan 2012).

I onlinemiljöer försvaras möjligheterna att skilja på de sammanhang i vilka handlandet sker; digitala delpersonligheter tillåts smälta samman och/eller överlappa på andra sätt än offline. Aktörer i semi-offentliga sociala nätverk tvingas därför hantera möjligheten att deras publik rymmer ohanterbart många olika 'generaliserade andra'. Agerande som en själv anser självklart kanske ej är det i vissa andras ögon. Twitter tillåter dock en friare presentation av identitet än Facebook och Google; exempelvis kan användare ha multipla konton och agera i enlighet med de många olika delpersonligheter som alla människor har. Twitterflöden är i hög grad intressebaserade snarare än identitetsbaserade (Marwick & boyd 2011); många användare tycks ha en funktionell inställning där Twitter används för omvärldsbevakning snarare än expressivt som en skildring av ens egna vardag, även om det senare också förekommer på Twitter.

Kontext kan förstås spatialt; en plats som avgränsas i tid och rum med igenkännbara roller och sociala normer (Hogan 2012: 300). Miljöpsykologer talar om beteenderegioner (Barker 1968), medan Goffman (1959) använder scenografiska metaforer ('frontstage', 'backstage') och Feld (1981) etablerar begreppet 'fokus' inom social nätverksanalys; sociala, psykologiska eller fysiska företeelser kring vilka nätverksaktörer kan organisera gemensamma aktiviteter (ibid.: 1025). Sådana fokus begränsar i regel handlandet, men ju fler aktörer som delar samma fokus desto friare är det normativa utrymmet (Melbeck 1996: 261). Att kontexter delvis överlappar har länge varit ett sociologiskt faktum (se exv. Simmel 1922). Den rent fysiska platsen för cirkulation och reproduktion av mediediskurser påverkar förväntningarna på genre, modalitet och sanningsanspråk (Fairclough 2012). Materiellt villkorade förväntningar är förstås också sociokulturellt villkorade; de kollektivt delade koderna för läsningen och förståelsen av medietexter är starkt normativa och ideologiskt präglade (Hall 1980). Notera att detta även gäller analytikern, inte minst beträffande

de rådande normer om information som varuform (Bolin & Andersson Schwarz 2015) som påverkar läsningen av social mediedata, dashboards, infographics och visualiseringar.

Bounegru (2009) menar att digitala sociala nätverkstjänster som Twitter tycks ge upphov till en form av 'sekundär oralitet' (Ong 1982) typisk för sammansmältningen av skriven, muntlig och elektronisk kultur, där deltagarna ständigt påminns om att de verkar i ett i någon mening offentligt sammanhang, oerhört mycket större än det som är aktuellt för primär oral samvaro (ibid.: 136). Här premieras deltagande, främjande av en gemensam känsla, ett fokus på nuet, samt även förhållandevis tillrättalagda, förutsägbara, rentav stereotypa uttrycksformer. Men i och med att mediets diskurser ständigt lagras i beständiga medier – lämnar spår – blir onlinekulturen jämförelsevis mer försiktig och självmedveten i sitt uttryck än det talade ordet. Kommunikation som lämnar spår av det här slaget har därtill egenskapen att den synliggör tydliga demarkationer av gruppstillhörighet. Jämfört med muntlig kommunikation kan skriven kommunikation därigenom framstå som entydig och kategorisk, inte minst i moral- och politikfrågor.

Länge byggde Twitter på en relativ konsensus beträffande de gemensamt delade koderna och normerna för deltagandet, menar Meyer (2015, som i sin tur hänvisar till Stewart 2014). Denna gemensamma praxis byggde i sin tur på att både läsandet och skrivandet (konsumtionen och produktionen) var inflikade i varandra, något som bidrog till att stärka den ömsesidiga förståelsen av kontexterna för skrivandet. Twitter kan skämtsamt ses som en gigantisk självrekryterad webbpanel med endast en fråga, vars tonalitet låter oss observera tre generationer av Twitter (figur 2).

Utöver de grundläggande incitamenten för att över huvud taget skriva något på Twitter – vilket determineras både av form/design och av normer i den specifika gemenskap inom vilken användaren verkar – kan vi rada upp en mängd förutsättningar och försöka kategorisera dem, även dessa med förbehåll för att de diskursiva villkoren är tätt sammanlänkade

Twitter I (2006-2009) kännetecknas av frågan "What are you doing?" vilket appellerar till urban livsstil, och vardagliga statusuppdateringar.

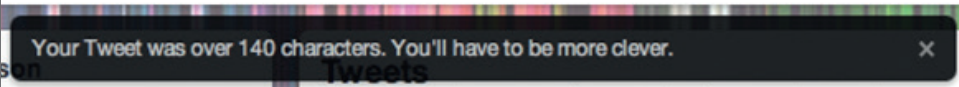
Twitter II (2009-2012) kännetecknas av frågan "What's happening?" och förskjuter tonvikten till världshändelser som val, katastrofer och revolutioner. Här blev mätverktyget "trending topics" av mycket stort intresse.

Twitter III (2012-) kännetecknas av uppmaningen "Compose new tweet" vilket avspeglar hur plattformen blivit ett generiskt verktyg för omvärldsbevakning (valrörelser, börskurser, celebriteter). Talande för denna utveckling är funktionen Twitter Analytics – ett dashboard integrerat direkt med Twitter som användarna själva kan tillgå.

Figur 2. Tre generationer av Twitter

Källa: Rogers 2013 (vår översättning och sammanfattning)

- **mediets form**
 - begränsningen i längd (140 tecken)
 - hur kedjor av tweets fungerar (det är ej alltid tekniskt uppenbart att ett inlägg utgör svar på ett annat inlägg eller är del av en längre "tråd" av inlägg)
- **villkoren för skrivandet som handling**
 - tidsliga villkor (dels att formulera sig kort, dels att skriva sitt inlägg medan debatten är aktuell)
 - skribentens upplevda behov av att vara slagfärdig för att bli retweetad och nå ut
- **kollektiva normer i gemenskapen av användare**
 - förväntningen att Twitter är ett jämförelsevis "smart" medium, för insatta användare (både tekniskt kunniga och intellektuellt kapabla att förstå ironier och andemeningar), något som under Twitters historia faktiskt rent konkret har uppmuntrats av plattformens diskursiva utformning:



En rutinmässig uppmaning till användarna vilken säger något om den förväntade interaktionen; en bokstavlig förväntning om "cleverness" hos användarna (skärmdump från 2012).

Figur 3. En typologi över Twitterspecifika kontextuella element

Not: Författarnas egna sammanställningar.

med varandra både formmässigt (det vill säga, tekniskt och språkligt) och normmässigt (socialt inom gruppen och beträffande bredare samhälleliga konventioner) (figur 3).

Om en användare har en förförståelse beträffande dessa villkor för skrivandet så påverkas rimligen dennes läsart av detta. Meyers och Stewarts resonemang är att denna delade praxis på senare år blivit allt mer ansträngd, främst på grund av Twitters expansion i antal användare. Denna expansion har inte bara gjort att enstaka kändisar får en närmast massmedial närvaro på Twitter, med miljontals följare, utan innebär också att Twitter är fragmentiserat (vilket Brynolf 2015 visar empiriskt) samt att det kommit att ske ett idogt hänvisande till Twitter inom massmedia, vilket vi visar empiriskt (Anders-

son Schwarz m.fl. 2015). Citat lyfts ofta ur sin sociala kontext på Twitter och sätts i en massmedial kontext. Enskilda uttalanden som egentligen är del av större konversationer tas ur sitt sammanhang och behandlas därefter av människor utanför samtalet som vore de entydiga politiska uttalanden. Det är inte bara twittrarens egna ideologiska motståndare som kan plocka billiga poänger på att ta svarigt prat ur sitt sammanhang; även dennes egna supporters kan ta enskilda uttalanden ur sin kontext, till exempel för att visa på förmodad oärlighet eller illojalitet.

Utöver krockarna mellan konversionsbaserad förväntan och textbaserad tolkning sker alltså även olika former av kontextkollaps beträffande de avsedda mottagargrupperna för olika budskap, och de avsedda graderna av

formalitet som uttalandet är tänkt att läsas genom.

Slutsatser

Nya digitala verktyg kan låta analytikern dyka ner i dataflödet och med hjälp av närläsning hitta anekdoter som egentligen inte är det minsta representativa för den större popu-

lationen. Likafullt kan analytikern, efter att ha hämtat ut stora mängder 'backend'-data genom olika typer av digitala gränssnitt, göra fjärrläsningar som kan visa på samband och mönster som kan tyckas mycket egenartade i jämförelse med konventionell deskriptiv statistik.¹ Det blir uppenbart att läsningar i det Boullier (2015) identifierat som tredje generationens kvantitativa förklaringsmodeller

	Första generationen	Andra generationen	Tredje generationen
<i>Föreställning om det sociala</i>	Samhälle(n)	Opinion(er)	Vibration(er)
<i>Insamlingsapparatur</i>	Folkräkning (census)	Enkäter (survey)	Spår (big data)
<i>Valideringsprincip</i>	Uttömlighet	Representativitet	Spårbarhet
<i>Kunskapsgenererande verktyg</i>	Register, förteckningar, bokföring	Publikmätningar, opinionsundersökningar	Sökmotorer, dashboards
<i>Social aktivitet i fokus för mätningen</i>	Förekomst	Uttryck	Engagemang
<i>Intressenter / finansärer</i>	Nationsstater	Massmedia	Varumärken
<i>Operativa aktörer</i>	Institut	Opinionsundersökningsföretag	Plattformsföretag
<i>Centrala intellektuella gestalter</i>	Émile Durkheim	Paul Lazarsfeld, George Gallup	Gabriel Tarde, Bruno Latour, Lev Manovich, Richard Rogers
<i>Huvudproblem som analyseras</i>	Arbetsfördelning och välfärdsstat (demografi)	Massmedialt inflytande och propaganda (publikmätning)	Spridning av kunskap (scientometri), samband av preferenser (sociometri, psykometri)
<i>Tekniska villkor</i>	Hålkortsmaskin / tabulator (Hollerith)	Informationsteknik (Turing, Shannon, von Neumann)	Internet, webb, mobila sensorer
<i>Semiotiska format</i>	Korstabeller, topografiska kartor	Normalkurvor, histogram, stapeldiagram, cirkeldiagram	Nätverksgrafer, infographics, tidslinjer
<i>Mätmetod</i>	Register över hela populationer (deskriptiv statistik)	Representativa stickprov/urval (induktiv statistik)	Topologi, frekvens, paretofördelade ej normalfördelade fenomen (Bayesiansk statistik)
<i>Tekniska kriterier för datakvalitet</i>	Relevans, exakthet, aktualitet, tillgänglighet, jämförbarhet, sammanhang	Konfidensintervall, probabilitet	Volym, variation, hastighet (big data)

Figur 4. Tre generationer av kvantitativa förklaringsmodeller

Källa: Boullier 2016: 13 (vår översättning och bearbetning)

inom samhällsvetenskapen (figur 4) ständigt behöver kompletteras med insikter genererade ur första och andra generationens.

Denna tänkta generationsväxling beträffande statistiska metoder och världsbilder återfinns även i omvärldsbevakningsbranschens diskurser: Demonstrationer av *dashboards* och *big data* åtföljs av lockande appeller till affärsdata vars tillblivelse är 'cool, easy, fast, fun', och där en sorts postmodern vändning tycks ha inträffat där närhet föredras framför distans, subjektivitet framför objektivitet och relevans framför representativitet (figur 5).

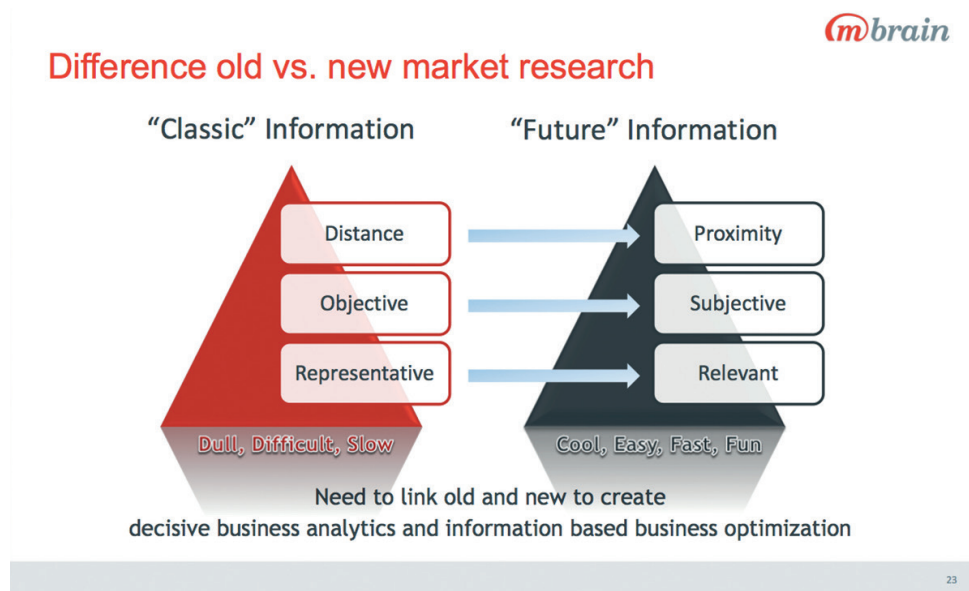
Även inom finansmarknader och politiska opinionsmätningar har sentimentanalys – m.a.o. realtidsindikationer av stämningar och känslolägen snarare än registrering av faktum eller riskmodellering baserad på statistiska normalfunktioner – kommit att dominera, något som yttrade sig i de felaktiga prediktionerna av 2016 års Brexit-omröstning i Storbritannien (Davies 2016).

Vi menar att det inte är så enkelt. Ordentlig kontextualisering innebär tyvärr arbetsmoment som faktiskt är 'dull, difficult, slow' och i regel kräver insiktsfulla analytiker. Ur

branschhänseende rör det sig snarare om förädling snarare än automatiserad vidarefordran, en förflyttning uppåt i värdekedjan snarare än sidledes eller nedåt.

Analytikern bör ge akt på datans tillblivelse (de specifika villkor för användningen som gör de spår som lämnas särpräglade), dess tillgänglighet (huruvida man arbetar med kommersiellt införskaffad data och vad denna har för tänkbara begränsningar), samt huruvida ens slutsatser dras ur det manifest synbara eller ur latent (ej uppenbara) dimensioner i datan och vilka statistiska artefakter som skapas genom analytikerns egna avgränsningar. Den förförståelse som möjliggör tolkningar av latent innebär ingår i datans kulturella kontext – som är specifik för varje social medieteknologi.

Ser vi till den distinkta sociala topologi och kontextkollaps vi observerade på Twitter, är frekvens i många avseenden minst lika viktigt som räckvidd i sociala mediemiljöer, då hög frekvens kan ge intrycket att frågor är allestädes närvarande, och att enskilda frågor blir 'akuta' när 'alla' i ens eget flöde samtidigt lyfter dem. Därtill kan räckvidd vara olika bred. En aktör som når hundra olika användare inom olika grupper kan ha större influens än en



Figur 5. Förmodade skillnader mellan 'gammal' och 'ny' omvärldsbevakning

Källa: Utdrag ur presentationen 'Social Media for Competitive and Market Intelligence', Susanna Tirkkonen, Stockholm 18 november 2015 (Faksimil)

aktör som har tusen följare inom en väldigt snäv grupp. Nätverk med hög inbördes länkning är således mindre inflytelserika än vad de manifesta räckviddssiffrorna kan ge sken av. Därför rekommenderar vi att analytiker

inte ska stirra sig blinda på siffror för 'antal följare' utan snarare bör ge akt på grundfrågor såsom *huruvida*, *hur*, och *hur ofta* ett budskap cirkulerar i olika individers flöden och vilka som därmed nås.

Not

1. Läsningar av det här slaget kan exempelvis göras med hjälp av statistik- och visualiseringsverktyg såsom Python, R, Elasticsearch, NodeXL & Gephi.

Referenser

- Andersson Schwarz, Jonas (2015). Hur flöden module-
ras: Vad som hamnar på agendan i sociala medier, i
Carlsson, Eric m.fl. (red.) *Digital politik. Sociala me-
dier, deltagande och engagemang*. Göteborg: Daidalos.
- Andersson Schwarz, Jonas (2016). Delningslogik och
plattformisering, i Westlund, Oscar (red.) *Männis-
korna, medierna & marknaden: Medieutredningens
forskningsantologi om en demokrati i förändring*.
SOU 2016:30. 133–164.
- Andersson Schwarz, Jonas; Hammarlund, Johan; di
Grado, Stefan; Kjellberg, Magnus (2015). *Opinioner
och offentligheter online: Slutrapport för forsknings-
projektet 'Vad gör en politisk utsaga framgångsrik?
Den användardrivna kommunikationens villkor'*.
Stockholm: IIS.
- Andersson Schwarz, Jonas & Hammarlund, Johan
(2016). Kända men inte erkända: Skildringen av
Avpixlat och andra invandringsfientliga nätmedier
i fem svenska dagstidningar 2013-2015, s. 172-188
i Truedson, Lars (red.) *Migrationen i medierna –
men det får en väl inte prata om?* Stockholm: Insti-
tutet för mediestudier.
- Barker, Roger G. (1968). *Ecological Psychology: Con-
cepts and Methods for Studying the Environment
of Human Behavior*. Stanford, CA: Stanford Uni-
versity Press.
- Berelson, Bernard (1952). *Content analysis in com-
munication research*. New York, NY: Free Press.
- Bergström, Annika; Johansson, Bengt; Oscarsson,
Henrik & Oskarson, Maria (2015). Fragment, s.
11-36 i Annika Bergström m.fl. (red.) *Fragment*.
Göteborgs universitet: SOM-institutet.
- Bimber, Bruce & Davis, Richard (2003). *Campaigning
online: The Internet in U.S. Elections*. New York:
Oxford University Press.
- Bolin, Göran & Andersson Schwarz, Jonas (2015).
Heuristics of the algorithm: Big Data, user inter-
pretation and institutional translation. *Big Data &
Society* 2(2). doi:10.1177/2053951715608406
- Boullier, Dominique (2016). *Big data challenges for
the social sciences: from society and opinion to re-
plications*. Social Media Lab, EPFL Lausanne, maj.
arXiv:1607.05034
- Bounegru, Liliana (2009). Secondary Orality in Mi-
croblogging: Orality versus literacy in the history
of human consciousness. Blogginlägg, 20 novem-
ber 2009. [http://lilianabounegru.org/2009/11/20/
secondary-orality-in-microblogging/](http://lilianabounegru.org/2009/11/20/secondary-orality-in-microblogging/)
- boyd, danah & Crawford, Kate (2012). Critical ques-
tions for Big Data: Provocations for a cultural, tech-
nological, and scholarly phenomenon. *Information,
Communication & Society*, 15(5): 662-679. doi:10.1
080/1369118X.2012.678878
- Broersma, Marcel, & Graham, Todd (2012). Social
media as beat: Tweets as a news source during the
2010 British and Dutch elections. *Journalism Prac-
tice* 6(3): 403-419.
- Bruns, Axel & Moe, Hallvard (2014). Structural Layers
of Communication on Twitter, s. 15-28 I Weller,
Katrin m.fl. (red.) *Twitter and Society*. New York:
Peter Lang.
- Brynolf, Hampus (2015). *Twitertensus*. <http://twitertensus.se/>
- Christensen, Clayton (1997). *The Innovator's Dilemma:
When New Technologies Cause Great Firms to Fail*.
Boston, MA: Harvard Business School Press.
- Davies, Will (2016). Thoughts on the sociology of
Brexit. 24 juni. Political Economy Research Centre
(PERC) blog, Goldsmiths, University of London.
[http://www.perc.org.uk/project_posts/thoughts-
on-the-sociology-of-brexit/](http://www.perc.org.uk/project_posts/thoughts-on-the-sociology-of-brexit/)
- Diakopoulos, Nicholas (2014). Algorithmic accounta-
bility. Journalistic investigation of computational
power structures. *Digital Journalism* 3(3): 398-415.
doi:10.1080/21670811.2014.976411
- Fairclough, Norman (2012). Critical discourse ana-
lysis. Working paper. [https://www.academia.
edu/3791325/Critical_discourse_analysis_2012_](https://www.academia.edu/3791325/Critical_discourse_analysis_2012_)
- Feld, Scott L. (1981). The Focused Organization of
Social Ties. *American Journal of Sociology*, 86(5):
1015-1035.
- Findahl, Olle & Davidsson, Pamela (2015). *Svenskarna
och internet 2015*. Stockholm: IIS.
- Flaxman, Seth; Goel, Sharad; Rao, Justin M. (2016).
Filter bubbles, echo chambers, and online news
consumption. *Public Opinion Quarterly* 80: 298-
320.
- Gaffney, Devin & Puschmann, Cornelius (2014) Data
Collection on Twitter, s. 55-67 i Weller Katrin m.fl.
(red.) *Twitter and Society*. New York: Peter Lang.
- Garimella, Kiran; De Francisci Morales, Gianmar-
co; Gionis, Aristides; Mathioudakis, Michael
(2015). *Quantifying Controversy in Social Media*.
arXiv:1507.05224v2 [cs.SI]

- Goffman, Erving (1959). *The Presentation of Self in Everyday Life*. New York: Doubleday.
- Hall, Stuart (1980) Encoding/Decoding, s. 128-138 i Centre for Contemporary Cultural Studies (red.) *Culture, Media, Language: Working Papers in Cultural Studies, 1972-79*. London: Hutchinson.
- Hamrefors, Sven (1999). *Spontaneous Environmental Scanning. Putting 'putting into perspective' into perspective*. Doktorsavhandling. Stockholm: Ekonomiska forskningsinstitutet vid Handelshögskolan.
- Hamrefors, Sven (2014). *Kommunikativt ledarskap: Den nya tidens ledarskap i värdeskapande nätverk*. Lund: Studentlitteratur.
- Hogan, Bernie (2012). Pseudonyms and the Rise of the Real-Name Web, s. 290-308 i Hartley, John m.fl. (red.) *A Companion to New Media Dynamics*. Chichester: Blackwell Publishing.
- Kwak, Haewoon; Lee, Changhyun; Park, Hosung; Moon, Sue (2010). What is Twitter, a social network or a news media?. s. 591-600 i Proceedings of the 19th International Conference on World Wide Web, Raleigh, NC.
- Larsson, Anders O. (2014). Everyday elites, citizens or extremists? Assessing the use and users of nonelection political hashtags. *MedieKultur*, 30(56): 61-78.
- Larsson, Anders O. (2015). ”Ju mer vi är tillsammans” – partiledare, samhälleliga eliter och medborgare på Twitter under riksdagsvalet 2014, s. 91-112 i Larsson, Eric m.fl. (red.) *Digital politik. Sociala medier, deltagande och engagemang*. Göteborg: Daidalos.
- Larsson, Anders O. (2016). Online, all the time? A quantitative assessment of the permanent campaign on Facebook. *New Media & Society*, 18(2): 274-292.
- Lasorsa, Dominic; Lewis, Seth; Holton, Avery (2012). Normalizing Twitter: Journalism practice in an emerging communication space. *Journalism studies*, 13(1): 19-36.
- Latour, Bruno; Jensen, Pablo; Venturini, Tommaso; Grauwil, Sébastien; Boullier, Dominique (2012). ”The whole is always smaller than its parts” – A digital test of Gabriel Tarde’s monads. *British Journal of Sociology*, 63(4): 590-615.
- Lewis, Seth; Zamith, Rodrigo; Hermida, Alfred (2013). Content Analysis in an Era of Big Data: A Hybrid Approach to Computational and Manual Methods, *Journal of Broadcasting & Electronic Media*, 57:1: 34-52, doi:10.1080/08838151.2012.761702
- Marwick, Alice & boyd, danah (2011). I Tweet Honestly, I Tweet Passionately: Twitter Users, Context Collapse, and the Imagined Audience. *New Media & Society*, 13(1): 114-133. doi:10.1177/1461444810365313
- Mead, George H. (1934). *Mind, self and society from the standpoint of a social behaviorist*. Chicago: University of Chicago Press.
- Melbeck, Christian (1996). Relationships between neighbors and between acquaintances, s. 248-268 i Auhagen, Ann E. & von Salisch, Maria (red.) *The Diversity of Human Relationships*. Cambridge: Cambridge University Press.
- Meyer, Robinson (2015). The Decay of Twitter. The Atlantic, 2 november. <http://www.theatlantic.com/technology/archive/2015/11/conversation-smoosh-twitter-decay/412867/>
- Moretti, Franco (2013). *Distant Reading*. London: Verso.
- Ong, Walter (1982). *Orality and Literacy: The Technologizing of the World*. London & New York: Methuen.
- Pariser, Eli (2011). *The Filter Bubble: What the internet is hiding from you*. New York: Penguin Books.
- Paulussen, Steve & Harder, Raymond (2014). Social media references in newspapers: Facebook, Twitter and YouTube as sources in newspaper journalism. *Journalism Practice*, 8(5): 542-551.
- Puschmann, Cornelius & Burgess, Jean (2014). The Politics of Twitter Data, s. 43-54 i Weller, Katrin m.fl. (red.) *Twitter and Society*. New York: Peter Lang.
- Rogers, Richard. (2013). Debanalizing Twitter: The Transformation of an Object of Study. WebSci '13, Proceedings of the 5th Annual ACM Web Science Conference, 356-365. doi:10.1145/2464464.2464511
- Sakaki, Takeshi; Okazaki, Makoto; Matsuo, Yutaka (2010). Earthquake shakes Twitter users: Real-time event detection by social sensors, s. 851-860 i Proceedings of the 19th International Conference on World Wide Web. Raleigh, NC.
- Simmel, Georg (1922). The Web of Group Affiliations, s. 125-195 i *Conflict and the Web of Group Affiliations*. New York: Free Press.
- Sjøvaag, Helle & Stavelin, Eirik (2012). Web media and the quantitative content analysis: Methodological challenges in measuring online news content. *Convergence: The International Journal of Research into New Media Technologies*, 18(2): 215-229. doi:10.1177/1354856511429641
- Stewart, Bonnie (2014). Something is rotten in the state of... Twitter. Blogginlägg, 2 september. <http://theory.cribchronicles.com/2014/09/02/something-is-rotten-in-the-state-of-twitter/>
- Strömbäck, Jesper (2015). Social sammanhållning och medieanvändning, s. 63-80 i Bergström, Annika m.fl. (red.) *Fragment*. Göteborgs universitet: SOM-institutet.
- Sunstein, Cass (2007). *Republic.com 2.0*. Princeton: Princeton University Press.
- Tao, Ke; Hauff, Claudia; Abel, Fabian; Houben, Geert-Jan (2014). Information Retrieval for Twitter Data, s. 195-206 i Weller, Katrin m.fl. (red.) *Twitter and Society*. New York: Peter Lang.
- Trilling, Damian & Schoenbach, Klaus (2015). Investigating people’s news diets: How online news users use offline news. *Communications: The European Journal of Communication Research*, 40(1): 67-91.
- Vis, Farida (2013). A critical reflection on Big Data: Considering APIs, researchers and tools as data makers. *First Monday*, 18(10). doi: 10.5210/fm.v18i10.4878

- Wadbring, Ingela (2016) Om dem som tar del av nyheter i lägre utsträckning än andra. I: Oscar Westlund (red.) *Människorna, medierna & marknaden: Medieutredningens forskningsantologi om en demokrati i förändring*, SOU 2016: 30, s. 463-486.
- Warner, Benjamin (2011). *Fragmentation and Social Polarization: The Character of Political Attitude Extremism in America*. Doktorsavhandling, University of Kansas.
- Webster, James (2014). *The Marketplace of Attention: How Audiences Take Shape in a Digital Age*. Cambridge, MA: MIT Press.
- Webster, James & Ksiazek, Thomas (2012). The Dynamics of Audience Fragmentation: Public Attention in an Age of Digital Media. *Journal of Communication*, 62(1): 39-56.
- Weller, Katrin; Bruns, Axel; Burgess, Jean; Mahrt, Merja; Puschmann, Cornelius (red.) (2014). *Twitter and Society*. New York: Peter Lang.
- Wesch, Michael (2009). YouTube and You: Experiences of Self-awareness in the Context Collapse of the Recording Webcam. *Explorations in Media Ecology*, 8(2): 19-34.
- Zelenkauskaite, Asta & Bucy, Erik (2016). A scholarly divide: Social media, Big Data, and unattainable scholarship. *First Monday* 21(5). doi:10.5210/firstmonday.v21i5.6358
- Zuiderveen Borgesius, Frederik; Trilling, Damian; Möller, Judith; Bodó, Balázs; de Vreese, Claes; Natali Helberger, Natali (2016). Should we worry about filter bubbles? *Internet Policy Review* 5(1). doi:10.14763/2016.1.401

Jonas Andersson Schwarz är lektor i medie- och kommunikationsvetenskap (Södertörns högskola) och intresserar sig för hur vardagsliv och samhällsstrukturer påverkas av den tilltagande digitaliseringen. Efter flera års forskning om delning och delningssamfund har han på senare tid skiftat fokus mot digitala plattformar som strukturellt fenomen, inte minst den datadrivna medieekologi som dessa möjliggör.

Johan Hammarlund är fristående medieforskare och gästlärare i medie- och kommunikationsvetenskap på Södertörns högskola. Han är även rådgivare inom omvärldsbevakning och medieanalys. I fokus för båda dessa verksamheter är det framväxande digitala medielandskapet där sociala medier och redaktionella medier samspelar på nya sätt.