LUND UNIVERSITY

Algoritmer i samhället

Haider, Jutta; Sundin, Olof

2016

Document Version:
Förlagets slutgiltiga version

Link to publication

Citation for published version (APA):
Haider, J., & Sundin, O. (2016). *Algoritmer i samhället*. Kansliet för strategi- och samtidsfrågor, Regeringskansliet.

Total number of authors:
2

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 07. Nov. 2026

11 mars 2016

Algoritmer i samhället

Jutta Haider & Olof Sundin
Institutionen för kulturvetenskaper
Lunds universitet

En rapport skriven på uppdrag av Kansliet för strategi- och framtidsfrågor, Regeringskansliet.

Innehåll

Inledning och syfte 3
Nedslag i den internationella diskussionen: sammanhang och aktörer 5
Hur fungerar algoritmer? 6
Google: länkanalys 6
Facebook: sociala nätverk 7
Netflix: kollaborativ filtrering 7
Algoritmer och användardata 8
Kakor 9
Kategorier 9
Informationsinfrastruktur och betydelsen av synlighet 10
Att vara synlig och att vara sökbar 10
Skapande av synlighet: sökoptimering 12
Driva opinion 13
Knowledge Graph 13
Flödesoptimering 14
Hur påverkas samhället av algoritmernas styrning? 15
Aktiemarknaden och organisering av arbete 15
Nyhetsmedier 16
Bibliotek 17
Vetenskap och akademin 17
Internet of things / The quantified self 18
Insyn, kontroll och transparens 18
Algorithmia 18
Kontroll över personlig data 19
Inlåsning 19
Google: Osynlig och utan insyn 20
Informationsbubblor och det offentliga samtalet 21
En offentlig sfär? 21
Algoritmer förstärker trender 22
Delningslogiken 22
Personalisering 24
Hur kan det offentliga samtalet stöttas? 25
Journalistik & massmedier 25
Skolan 26
Bibliotek 28
Avslutning 29
Referenser 32

Inledning och syfte

I dagens samhälle fattas fler och fler beslut baserade på algoritmer. Om någon idag vill veta mer om vaccinering är det troligt att denna vänder sig till sökmotorer och sociala nätverkstjänster. Hur kunskap om vaccin, och andra kunskapsfält av samhällelig relevans, representeras och ordnas i dessa tjänster är därför av stor betydelse. När vi som författat denna rapport i januari 2016 sökte i Google hamnar en vaccinkritisk webbplats på andra plats. I YouTube varvas kritiska filmklipp med samhälleligt sanktionerade kunskapsanspråk. I Facebook tycks de vaccinkritiska grupperna dominera, även om våra respektive sökresultat skiljer sig en del åt beroende på vilka vänner vi har. I Twitter, slutligen, har den etablerade kunskapen ett tydligt företräde för oss båda. Tjänsterna ger olika slags resultat delvis p.g.a. de algoritmer som används. Exakt hur algoritmerna fungerar, varför vissa länkar rangordnas och görs synliga, är svårt att fastställa, men i denna rapport diskuterar vi principer för hur samspelet mellan algoritmer (liksom mellan algoritmer) och samhället fungerar och hur information, kultur och kunskap cirkulerar i samhället, samt vilka konsekvenser det kan få.

Med bärbara datorer, läsplattor och de smarta telefonernas framväxt har vi – i Sverige och stora delar av Europa – tillgång till snabbt internet i det närmaste överallt och när som helst. Ju mer vi lever våra liv medialiserade, dvs. hur medier och nätverksbaserad information spelar en allt större roll för våra liv, desto större betydelse får algoritmer. En förståelse av algoritmers funktion och konsekvenser i samhället blir då också väsentlig. Vad är då en algoritm?

> *Informally, an algorithm is any well-defined computational procedure that takes some value, or set of values, as input and produces some values, or set of values, as output. An algorithm is thus a sequence of computational steps that transform the input into the output. (Cormen et al., 2009, s.5)*

Algoritmer ingår i regel i programvaror och består av instruktioner om vad som ska utföras, med vad och i vilken ordning. De kan jämföras med recept som beskriver i vilken ordning ingredienser ska föras samman och hur de ska hanteras. Även om algoritmer främst förstås i relation till datorer, där de alltid har använts, kan de också användas av människor för strukturerad problemlösning. Centrala algoritmer som möjliggör dagens nätverkssamhälle är olika algoritmer för att sortera element, för att fastställa den kortaste vägen mellan två punkter, för att packa upp och packa ihop signaler, för att möjliggöra säker dataöverföring, för att möjliggöra länkanalys eller för att komprimera data (Otero, 2014; Cormen et al., 2009, s. 9-10).

Eftersom datorer finns så gott som överallt, finns numera också algoritmer nästan överallt. I rapporten avgränsas algoritmer främst till sådana som förekommer i sociala nätverkstjänster, i kollaborativa rekommendationstjänster och i sökmotorer. I dessa sammanhang utgör användargenererad data en central ingrediens. Algoritmer skapar tillgång till information samt möjliggör kommunikation mellan människor och mellan människor och maskiner. Algoritmer utgör, tillsammans med metadata, protokoll/regler, gränssnitt och tjänsternas default-inställningar, vad som kan kallas för plattformar som kodar in sociala värderingar i den digitala arkitekturen som omger oss samt på så sätt medskapar sociala och kulturella handlingsmönster (van Dijck, 2013, s. 29ff.). Rapportens avgränsning fångar mycket av det samtida intresse

som ges algoritmer samt sätter fokus på de mest populära tjänsterna på webben. Enligt webbanalytikerföretaget Alexa.com (2015-12-20) är de tio globalt mest populära webbplatserna Google.com, Facebook.com, Youtube.com, Baidu.com, Amazon.com, Wikipedia.org, Qq.com, Twitter.com, Google.co.in. Vi ser hur sociala medier och sökmotorer dominerar listan. Vi ser också hur de mest populära webbplatserna i de flesta fall är exempel på en ny affärsmodell där människor inte behöver betala för tjänsterna med pengar utan genom sina digitala spår.

De sätt som algoritmer i allt större utsträckning är en del av människors vardag gör att det finns en oro för hur algoritmer påverkar samhällets demokratiska förutsättningar. I rapporten belyses därför, i enlighet med uppdragsbeskrivningen, hur digitala medier, genom bland annat algoritmer och filterbubblor, påverkar det offentliga samtalet och i förlängningen demokratin. Rapporten har skrivits med utgångspunkt i nedanstående frågor:

- Hur styr algoritmer det informationsflöde som når människor genom sociala medier och andra digitala plattformar?
- Vilket material används algoritmerna på, och vem kontrollerar det?
- Vilka centrala parametrar används i algoritmerna? Hur ökar algoritmerna företagens intäkter – och finns det andra drivkrafter än de rent kommersiella?
- Vad skapar den informationsinfrastruktur som de kommersiella aktörerna bygger upp för förutsättningar för andra kommersiella och icke-kommersiella aktörer?
- Hur ser marknaden ut för de företag eller andra aktörer som vill påverka människors sökningar, för att vinkla algoritmerna i en önskad riktning?
- Vad innebär sökmotorsoptimering och vad kan det den få för konsekvenser? Hur mycket skiljer sig algoritmer mellan olika centrala plattformar eller sökmotorer?
- Vilken insyn finns i algoritmernas utformning? Vilka får ta del av algoritmernas utformning? Hur transparenta är algoritmer i olika former av media?
- Hur påverkas samhället av algoritmernas styrning? Riskerar slutna informationsbubblor minska möjligheten till ett brett offentligt samtal och leda till ökad åsiktspolarisering, eller är dessa farhågor överdrivna?
- Hur ser den internationella diskussionen ut på området? I vilka internationella sammanhang drivs frågan framåt och vilka aktörer är engagerade i diskussionen?
- Vilka frågor bör besvaras för att bättre förstå hur samhället formas av de nuvarande algoritmerna?
- Hur skulle man kunna gå tillväga för att påverka utvecklingen i en riktning som gynnar det offentliga samtalet, en mångfald av åsikter och offentlighetens möjlighet till källkritisk granskning?

Rapporten är delvis strukturerad utifrån frågorna, men för att göra texten läsvänlig och för att undvika upprepningar används en delvis annan rubriksättning och struktur. Det har de senaste åren publicerats flera rapporter och delbetänkanden från Digitaliseringskommissionen (SOU 2013:31; SOU 2014:13; SOU 2015:28; SOU 2015:65), Medieutredningen (SOU 2015:94) och Demokratiutredningen (t.ex. Bäck,

Bäck & Gustafsson, 2014; Strömbäck, 2014). Denna rapport relaterar till dessa där det är relevant. Vår förståelse av algoritmer är grundad i ett sociomateriellt synsätt där algoritmers samverkan med omgivningen, inklusive andra algoritmer, kommer i fokus. Vi strävar efter att ge referenser enligt gängse akademisk praxis, men texten är skriven utifrån den allmänkunskap vi utvecklat under en längre tid av omvärldsbevakning och alla påståenden är därför inte belagda med referens.

Nedslag i den internationella diskussionen: sammanhang och aktörer

Algoritmer som ämne i vetenskapen är utforskat sedan länge. En sökning i Web of Science i januari 2016 på algoritmer som ämne eller som en del av titeln ger 169836 artiklar varav den äldsta publikationen är från 1907. Circa en tredjedel av alla publikationerna är från de senaste fem åren, vilket visar på att ämnet är aktuellt och i utveckling. Forskningen domineras av naturvetenskap/teknik, men är omfattande också inom samhällsvetenskap/ekonomi/humaniora där den ökar starkt i antalet publikationer. Den forskning som denna rapport främst lutar sig emot kan beskrivas som kritisk algoritmforskning på så sätt att den diskuterar algoritmernas samhälleliga förutsättningar och konsekvenser snarare än hur algoritmer utvecklas och testas. En relevant sammanställning av litteratur på området är gjord av Social Media Collective vid Microsoft Research New England.[1]

Kunskapsutvecklingen om algoritmers samhälleliga förutsättningar och konsekvenser är dynamisk och viktiga kunskapsbidrag återfinns utanför forskningen. Det finns en uppsjö av bloggar drivna av forskare, lobbyister, aktivister och branschaktörer. Även nyhetsmedia och många journalister är aktiva i diskussionen. Eftersom företaget Google är en så dominant aktör är det vanligt att diskussionen förs i form av en positionering i relation till just Google. I lobby-organisationen Fair Search [2] organiserar sig etablerade branchaktörer såsom Microsoft (fram t.o.m. 2015), Nokia eller Oracle för att förvara möjligheterna att konkurrera på sökområdet. Organisationen har bl.a. anklagat Google för att utnyttja sin monopolsituation. Också forskare har engagerat sig utanför forskning i snäv bemärkelse. Exempelvis utgör The Society of the Query[3] en sammanslutning med europeiska forskare som bedriver kritisk algoritmforskning med samhälleligt fokus. Ett annat exempel på en sammanslutning för diskussioner kring algoritmers samhälleliga förutsättningar och konsekvenser är World Information Institute [4] som knyter samman forskare, konstnärer, aktivister och den intresserade allmänheten och de har organiserat en serie workshops och konferenser där den senaste var Algorithmic Regimes som ägde rum 2015.

[1] http://socialmediacollective.org/reading-lists/critical-algorithm-studies/
[2] http://fairsearch.org
[3] http://networkcultures.org/query/
[4] http://world-information.net

Hur fungerar algoritmer?

De senaste tio åren har det skett en kraftfull utveckling av sociala medier och sökmotorer har fått en ökad betydelse, samtidigt som marknaden har koncentrerats till några få starkt dominerande aktörer. Några av de mest kända är de amerikanska företagen Google (numera Alphabet), Facebook och Netflix. Dessa tre företag erbjuder tjänster där algoritmer spelar en avgörander roll, men på något olika sätt. Hur fungerar de då?

Google: länkanalys

Google är det mest dominerande företaget på internet och dess sökmotor toppar också den ovan nämnda listan på de mest populära webbplatserna på nätet. Google utgörs av en mängd olika tjänster, men vi avser främst sökmotorn Google Search även om denna måste förstås tillsammans med Googles övriga tjänster. Det som har gjort Google Search mer framgångsrikt än konkurrenterna, såsom en gång Altavista, är deras PageRank-algoritm. I en starkt förenklad framställning, stipulerar den att webbplatser med många in-länkar sorteras högre upp i träfflistan än webbplatser med få in-länkar. Vidare viktas de in-länkar från webbplatser som i sin tur har många in-länkar högre, dvs. länkar från vad som av många anses vara auktoritativa sidor väger mer än länkar från mindre viktiga sidor (Page et al., 1999). Grundprincipen är hämtat från bibliometrins citeringsanalys (Halavais, 2009, s. 67; Ma, Guan & Zhao, 2008) och kvantifiering av vetenskapliga tidskrifters betydelse och ”impact” som baseras på idén att en välciterad artikel eller tidskrift har högre relevans. Den kunskapen har bl.a. använts som underlag för inköp av tidskrifter till forskningsbibliotek. Principen är att en länk – liksom ett citat i vetenskapliga sammanhang – är ett utryck för relevans och ju fler länkar desto högre relevans.

Sedan den introducerades under slutet av 1990-talet är PageRank enbart en faktor i en uppsjö av andra faktorer som styr sökningens resultat i Google Search. Andra faktorer är exempelvis plats, profilens sökhistoria, aktualitet, mobilanpassning. Google Search ändrar varje år sina sökalgoritmer 500-600 gånger per år.[5] Det är viktigt att komma ihåg att Google Search är en sökmotor vars främsta uppgift är att återfinna information på den öppna webben. Utgångspunkten är att matcha sökord med ord som finns på en sida. Miljontals träffar är möjliga för de allra flesta sökningar, i alla fall om man söker på engelska. Problemet som måste lösas, relaterat till deras algoritm, är att presentera länkar till webbplatser för användaren i en ordning där de mest relevanta hamnar överst. Det kan utryckas i termer av ”recall” (hur många dokument som hittas) och ”precision” (hur exakt dokumenten som sökts fram svarar på sökfrågan), där det ena ökar när det andra minskar. Google Search har löst det med hjälp av sina algoritmer som utgår från att relevans är något personligt, relaterat till sökaren, och inte något absolut. Tanken fanns redan med i slutet på 1990-talet när Googles grundare utvecklade PageRank (Page et al., 1999) och har sedan dess lagt grunden för Googles enorma genomslag.

År 2015 publicerade forskare vid Google en rapport där det framgår att företaget utforskar möjligheterna att komplettera andra algoritmer med en Knowledge-Based Trust (KBT) Score som testar faktapåståendens sanningshalt (Dong et al., 2015).

[5] https://moz.com/search-ranking-factors/correlations

Istället för att låta rankningen av sökresultat baseras på webbplatsers popularitet mätt genom antalet in-länkar argumenterar forskarna för att låta webbplatser hamna högt om de har få felaktiga faktapåståenden. En sådan algoritm bygger på att Google skapar en databas med fakta som de extraherat från sidor de anser vara särskilt trovärdiga. Enligt forskarna är fakta ”accurate if its facts are correct, and we believe the facts are correct if they are extracted from an accurate source” (Dong et al., 2015). Sedan vill de jämföra dessa fakta med faktapåståenden från webbplatser för att bedöma om dessa webbplatser ska anses vara trovärdiga eller inte.

Facebook: sociala nätverk

Facebook är idag kanske vår mest kända sociala nätverkstjänst, i alla fall avseende popularitet. På enbart 10 år har Facebook blivit en fast punkt i många svenskars vardag och ett centralt sätt att kommunicera på och organisera det sociala livet genom. Utgångspunkten för tjänsten är att föra samman människor som känner till varandra och möjliggöra för dem att visa innehåll som sedan kan kommenteras, gillas och delas. Synlighet är det som eftersträvas, vilket också skulle kunna ses som en form av övervakning av varandra där eventuell osynlighet blir till ett problem (Bucher, 2012). I början var det amerikanska studenter som skapade och använde tjänsten, därav namnet Facebook som betyder årsbok, vilket är ett vanligt förekommande fenomen på amerikanska universitet. Det kan möjligen förklara fokuset på mätbar popularitet och yttre synlighet. Numera har Facebook, enligt Wikipedia, mer än 1,18 miljarder aktiva användare per månad (Wikipedia, 2016-01-08). De flesta företag liksom ideella organisationer har en Facebook-sida, för att inte nämna publika arrangemang och kulturella föreningar.

Många användare har hundratals vänner och ett problem som Facebook måste lösa är att bestämma vilket innehåll som ska komma överst i en användares så kallade nyhetsflöde [news feed]. Vems uppdatering ska visas, vems kommentar ska synas och vilken reklam ska vara synlig däremellan? Till en början hade man en någorlunda enkel algoritm som gick under namnet EdgeRank och som stort sett använde sig främst av tre faktorer: närhet [affinity], vikt [weight] och aktualitet [time decay] (Bucher, 2012). År 2013 bytte Facebook från EdgeRank till en mer avancerad algoritm (ibland kallat för newsfeed algoritm) som fortfarande använder sig av dessa tre faktorer, men utöver dem även tar in ett tusental andra faktorer. Tillsammans skapar det en mycket komplex algoritm, som möjliggör ett stort mått av individanpassning. Här ingår vilka sorters poster en användare brukar klicka på eller gillar, om någon har gillat en uppdatering utan att klicka på den eller hur länge man läste den, vilka sidor man gillar, vilka sidor ens vänner gillar, vilken reklam man eventuellt valt bort, om en uppdatering är en text, en länk eller ett foto. Liksom Google uppdaterar Facebook sina centrala algoritmer löpande.[6]

Netflix: kollaborativ filtrering

Den tredje tjänsten som här får illustrera algoritmers olika funktionaliteter är streamingtjänsten Netflix som ger tillgång till serier, filmer och dokumentärer. Netflix använder sig av en algoritm för rekommendationer där man försöker förutse vad en användare vill se baserat på egna och andra användares preferenser, s.k.

[6] http://newsroom.fb.com/news/category/news-feed-fyi

kollaborativ filtrering. Netflix utannonserade 2009 ett pris bland programmerare och matematiker för att förbättra streamingtjänstens algoritm. Teamet som vann fick 1 000 000 dollar för att förbättra tjänstens ursprungliga rekommendationssystem (Cinematch) med 10 %.[7] Särskilt intressant är att Netflix även har börjat producera egna TV-serier utifrån algoritmen för rekommendationer (Hallinan & Striphas, 2014). Det betyder att företaget med hjälp av sin algoritm inte enbart förser användarna med rekommendationer på filmer och serier, utan bolaget gör egna produktioner baserade på sin algoritm om vad användarna förväntas vilja se. Netflix kombinerar därmed en avancerad algoritm med av dem intuitivt skapade kategorier.

Den enklaste och mest välkända formen av kollaborativ filtrering påträffas i: ”andra som köpte NN, köpte även XX”, men det finns en mängd ytterligare faktorer som mer avancerade algoritmer för kollaborativ filtrering använder sig av. Användare uppmuntras att betygsätta, skriva recensioner eller rekommendera tjänster och produkter. Hänsyn tas bl.a. till användares sociala nätverk, profilinställningar och tidigare konsumentbeteenden, veckodag och tid på dygnet. Kollaborativ filtrering finns inte minst i andra tjänster som säljer produkter såsom Amazon, Spotify, Youtube, Amazon och Airbnb. Även Facebook använder sig av rekommendationstjänster. I Google ser man exempel på rekommendationstjänster i förslag på relaterade sökningar samt i ”autocomplete suggestions”, som baseras på vad andra har sökt på. Vi möter också rekommendationstjänster i biblioteksdatabaser (t.ex. rekommendationstjänsten bX) där användare kan få tips på vad andra användare intresserade av samma artiklar har läst.

Algoritmer och användardata

Algoritmer innehåller, som vi nämnt tidigare, regler för att bearbeta data i någon form, och i de flesta tjänster som behandlas i denna rapport är en stor del av denna data skapad av användarna genom användningen. Det kan handla om digitala spår vid lämnar vi sökningar i Google, "likes" vi gör i Facebook eller betygssättning vi gör i Netflix. Man brukar skilja mellan data och metadata, där metadata beskriver olika aspekter av data, men gränserna mellan dem är inte alltid tydliga. Metadata kan beskrivas som data som innehåller ”structured information to describe, explain, and locate information resources or otherwise make it easier to retrieve, use, or manage them” (van Dijck, 2013, s. 30).

Beroende på vad en tjänst har som primär inkomstkälla så påverkas vad en algoritm analyserar, dvs. vad den används på. Det är skillnad på en tjänst som Netflix som säljer visningar på abonnemangbasis och en tjänst som Facebook som säljer reklamplatser. I båda fallen används användargenererad data, i fallet Netflix kombineras de med metadata om serier och filmer och i fallet Facebook kombineras de med metadata om reklam (målgrupp etc). Det är företagen som kontrollerar datan, men samtidigt skyddas datan av lagstiftningen. Det kan exemplifieras med den ovanstående nämnda pristävling som Netflix arrangerade där företaget ville göra en uppföljning, men då kom i konflikt med amerikansk integritetsskydd. Netflix hade

[7] http://www.netflixprize.com

tillgängliggjort testdata-sets där verkliga användare kunde identifieras, vilket var i strid med lagstiftningen och därför fick uppföljningen utebli (Singer, 2009).

Kakor

Något som kännetecknar de slags algoritmer vi diskuterar i rapporten är att de ofta används på användargenererad data i stora kvantiteter, ofta kallad ”big data”. En del av denna data delar användare med sig medvetet, en del av datan samlar tjänsterna in utan att användare är medvetna om det och en del av den insamlade datan befinner sig där emellan. Exempel på medveten användargenererad data är information användare delar med sig av genom att fylla i sin användarprofil; betygssättning av produkter, tjänster eller personer; eller genom att användare i Facebook klickar bort reklam och förklarar varför.

Användare levererar också data när s.k. kakor [cookies] lagras på användarnas datorer och som därefter levererar meddelanden om vad en webbläsare används till. Sedan 1 juli 2011 ska användare i Sverige samtycka till att kakor lagras ner på den egna datorn. Denna lag, som bygger på ett EU-direktiv, är tänkt att öka människors medvetenhet, men i praktiken är det troligen få som väljer bort en webbplats p.g.a. de kakor som installeras. En studie av Altaweel, Good och Hoofnagle (2015) visar att en användare som besöker de 100 mest populära webbplatserna får över 6000 kakor sparade på sin dator. De allra flesta av dessa är s.k. ”third party cookies”. Exempelvis finns det Google-kakor, och därmed Googles ”tracking”-infrastruktur, på 923 av de 1000 mest populära webbplatserna. Altaweel et al. (2015) drar slutsatsen att ”Google’s ability to track users on popular websites is unparallelled, and it approaches the level of surveillance that only an Internet Service Provider can achieve”. Googles ”tracking” fungerar så att när man exempelvis söker efter resväskor på Google, kan därefter reklam för resväskor dyka upp på en dagstidningswebbplats genom att tidningen köpt användningen av Googles tjänst (s.k. doubleklick tracking). Kakors osynliga bakgrundsaktiviteter är en förutsättning för hur de dominanta algoritmerna fungerar i samtiden.

Kategorier

Algoritmer får sina data levererad på ett sätt som inte är tydligt medvetet, men heller inte helt omedvetet för de flesta användare. Facebook uppmuntrar användare att gilla reklam eller att kommentera och gilla andra medlemmars statusuppdateringar. Vidare ges användarna möjligheter att göra olika slags tester. Genom att göra testerna ger användarna då ofta omedvetet tillstånd att deras Facebook-data lagras och används i reklamsyfte (SR 2014-07-24). I dessa fall levereras data om användarna och deras beteenden, ofta utan att användarna har läst villkoren, till Facebook och andra företag. I Google Search sker denna form av dataleverans genom alla de sökningar människor gör när de är inloggade, var dessa görs, när de görs osv. Genom den data Google på detta sätt får, kan sökmotorn också förbättra sin tjänst. Men Google äger även andra tjänster som YouTube, Gmail, Google Maps, eller en annan typ av tjänster riktad till producenter som Adsense, och inte minst Google Analytics. Genom att användares alla aktiviteter i Googles olika tjänster genererar data åt Google kan bilden som Google får av användare bli nästan komplett.

Samtidigt använder sig reklamindustrin samt den reklam- och algoritmbaserade tech-industrin, dvs. Google och andra liknande företag, också av marknadsföringens traditionella indelning av konsumenter i fördefinierade segment. Bolin och Andersson Schwartz (2015) talar här om en algoritmens heuristik där användardata översätts till sociala kategorier för marknadsföring. Dessa kategorier och hur de används, menar de, skapar en distans mellan enskilda människor och hur de klassificeras. Det går att jämföra med hur man i USA har börjat med att använda algoritmer som sammanför sociala kategorier med persondata för att fatta beslut om hur en misstänkt eller dömd person ska behandlas inför rätten, t.ex. beräknas med hjälp av algoritmer vilka risker det finns för att hen ska återfalla i brottslighet eller försvinna när hen släpps genom borgen (Christin, Rosenblat & boyd, 2015). På så sätt kan de kategorier som algoritmer bidrar till ses som ett slags klassifikationssystem som får direkta konsekvenser för människors liv.

Informationsinfrastruktur och betydelsen av synlighet

Den som vill bli läst har alltid anpassat sig till informationsinfrastrukturens förutsättningar, vare sig denna bygger på tryckta medier med traditionella förlag eller på digitala medier. Vad som är förändrat är dock att i allt större utsträckning är information kanaliserad genom ett mindre antal globala företag och deras tjänster. Det finns en paradox i att vem som helst numera kan publicera sig genom tweets, Facebook-inlägg eller bloggar samtidigt som denna informations synlighet är beroende av ett fåtal aktörer. Samtidigt formar det som är synligt online vad det finns att veta om något och därmed själva problemet (Haider, 2014). I Sverige är det Google och Facebook som dominerar, men även andra sociala medier som Twitter, Instagram och SnapChat är viktiga. Det har stora konsekvenser för både konsumenter och producenter av information. Allt måste anpassas för att kunna synas i och genom dessa de-facto ”gatekeepers”. Det är möjligt att surfa direkt till en webbsida vars adress man känner till eller att använda en mindre, alternativ sökmotor, men det blir ovanligare och ovanligare. Ju mer tjänster blir inbyggda i tekniken som infrastruktur desto svårare är det att välja bort tjänsterna. Exempelvis smarta telefoner gör det i praktiken svårt att logga ut ur Google, särkilt i Android-telefonerna, vars Linux-baserade operativsystem (Android) numera ägs av Google.

Att vara synlig och att vara sökbar

För att visa på infrastrukturens betydelse och förändring kan vi ta uppslagsverk som ett exempel. Uppslagsverk utgör en slags ryggrad för den offentliga kunskapen i samhället. Sedan de tryckta uppslagsverken kom att ersättas av digitala har det varit svårt för kommersiella aktörer såsom Encyclopaedia Britannica och Nationalencyklopedin att kunna konkurrera med paret Google/Wikipedia. Det fria uppslagsverket Wikipedia hamnar nästan alltid överst i Google (Lewandowski & Spree, 2011), vilket har fått många kommentatorer att spekulera om det gynnas särskilt av Googles algoritm. Kommersiella uppslagsverk hamnar i regel långt ner i träfflistan. En anledning till att Encyclopaedia Britannica och Nationalencyklopedin i de flesta fall sorteras längre ner är troligen det faktum att de är betaltjänster och därmed i mindre utsträckning länkas till från andra webbsidor, vilket ger dem en sämre placering än en öppen tjänst som Wikipedia med sina många in-länkar. Google

nedpriorirerar dessutom webbplatser där en betalvägg hindrar användaren att komma direkt till innehållet från Googles sökresultat. Trots detta kommer den dominerande trafiken till Nationalencyklopedin från sökmotorer, med andra ord från Google. Ungefär 70 % hittar till Nationalencyklopedin via sökmotorer.[8] Här kan nämnas att motsvarande encyklopedi i Norge – Store Norske Leksikon (SNL) – numera delvis finansieras med medel ur stadsbudgeten, vilket möjliggör fri tillgång. SNL arbetar med stor framgång med att sökoptimera encyklopedin och anpassa den för en nätverksmiljö och 94 % av trafiken till SNL.no kommer från organisk sökning, dvs. icke-sponsrade länkar, medan endast 3,5 % går in direkt på webbplatsen via adressfältet.[9]

Att vara sökbar är helt avgörande för i princip alla företag, organisationer, kulturinstitutioner och även myndigheter eller privatpersoner. Det gäller sökmotorn Google, men även Google Maps, YouTube och andra Google-tjänster (som integreras alltmer i Android). Det räcker inte med att vara sökbar, utan man måste också hamna högt i träfflistan. De allra flesta klickar på länkar som finns på Googles första resultatsida (Höchstötter & Lewandowski, 2009). Att vara delningsbar i Facebook (eller andra sociala medier) är ytterligare en framgångsfaktor. Till och med organisationer som Polisen arbetar numera med sociala medier strategiskt där man strävar efter att öka polisens trovärdighet och popularitet genom att synas i dessa sammanhang. Vi får ta del av polisens arbete genom känslosamma uppdateringar och tweets, väl anpassade för den kanal som används, som många gånger delas i hög utsträckning och blir virala (Stakston, 2013). Det kan uttryckas som att Polisen anpassar sin kommunikation gentemot allmänheten utifrån hur algoritmerna fungerar. Sociala nätverkstjänster och sökmotorer kan sägas utgöra en informationsinfrastruktur som filtrerar och medierar den största delen av den information användare tar del av och det gäller kommersiell liksom allmännyttig information. Googles främsta produkt är Google Search, men deras främsta affär är reklam (Halavais, 2009, s. 83) och det som man säljer är användarnas uppmärksamhet (Halavais, 2009, s. 56 ff.). Det samma kan sägas om alla online-tjänster vars affärer bygger på reklamplacering. Användarnas uppmärksamhet är produkten man säljer i dessa företag och den produkten gäller det att förstå och att styra, vilket görs med hjälp av algoritmer som informationsproducenterna sedan måste förhålla sig till.

Att producera information betyder att anpassa utformningen för att synas och för att vara hittbar liksom att vara tekniskt och grafiskt anpassad för den plattform den ska konsumeras genom. Det görs på många olika sätt och professionella webmasters, app-utvecklare, online redaktörer eller andra som jobbar med att skapa dagens innehåll är utbildade för detta. En stor del av jobbet är att vara à jour med de olika uppdateringar som t.ex. Googles eller Facebooks algoritmer genomgår. Google tillhandahåller särskilda verktyg – Google Webmaster Tools och Google Analytics – för detta. Google Analytics har blivit ett standardverktyg för att många webbplatser. Underförstått är att en sida inte ”finns” om den inte återfinns högst upp i Googles träfflista. Även Facebook har liknande tjänster.

[8] E-post 5 oktober 2015 från Jonas Gruvö, chefredaktör vid NE.
[9] E-postbrev 6 oktober 2015 från Georg Kjøll, utvecklingsledare och redaktör vid SNL.

Skapande av synlighet: sökoptimering

Det finns företag som specialiserar sig på att göra innehåll (webbsidor och andra online produkter) sökbart, hittbart, och delningsbart – dvs. synligt – på internet. Dessa kallas för SEO (Search Engine Optimisation) företag. SEO handlar inte om att betala sökmotorföretagen för placering av reklam, utan om att skapa online-produkter (oftast webbsidor), på ett sätt som gör att de kommer högt upp i resultatlistorna för relevanta sökningar. Syftet är att därigenom få trafik till en specifik sida från det organiska sökresultatet. SEO kan ses som en del av en mer omfattande Search Engine Marketing (SEM), som innefattar även betal-reklam och andra tjänster som rör marknadsföring online. Den största och mest kända betaltjänst för online reklam är Googles AdWords, där man – förenklad uttryckt – köper reklamplacering för vissa sökord och anpassning till målgruppen som man vill nå. När man jobbar med sökmotorsoptimering inriktad på organiska sökresultat finns många olika faktorer som det tas hänsyn till för att förbättra hittbarheten, t.ex. vilka ord som används på en sida, vilken titel en sida har, hur många in-länkar en sida har och hur trovärdiga de länkande sidorna är, en sidas informationsarkitektur och användarvänlighet.[10]

Vissa av de ovanstående faktorerna har varit ganska stabila genom åren, även om deras viktning har förändrats. Idag spelar personalisering, lokalisering, mobilanpassning, s.k. social media metrics (dvs. hur framgångsrikt ett innehåll är i sociala medier) och många andra faktorer in. Vissa sätt att försöka påverka rangordningen straffas av sökmotorn. Så kallad ”link-farming”, som var vanlig när SEO började, där man försökte att öka en sidas PageRank genom många i princip innehållslösa länkar till den, är exempelvis inte längre användbar utan tvärtom leder den till lägre rangordning. Google Webmaster Guidelines förklarar de viktigaste förhållningssätten och reglerna för att skapa Google-konformt innehåll och där framgår det tydligt att försök att överlista algoritmen straffas.[11] Var gränsen mellan ”överlista” och ”anpassa” går är otydlig. Information om alla uppdateringar är mycket begränsad och varje större förändring omges av omfattande spekulationer och rykten. SEO görs alltså för att anpassa innehållet efter kriterier som algoritmerna använder för att leta upp och rangordna webbsidor. Algoritmerna förändras och uppdateras kontinuerligt, vilket betyder att vissa webbsidor måste göras om – åtminstone delvis – för att fortsatt kunna hävda sig i sökresultaten. Sökmotorsföretagen, dvs. i praktiken Google, brukar dock vara knapphändiga med information om hur förändringarna kommer att påverka sökresultaten och en del av arbetet som SEO/SEM-företagen utför handlar just om att ”testa” algoritmen och med ”trial and error” metoder förstå utformningen av algoritmerna.[12]

Det medför en särskild utmaning när företagen som står för centrala delar av samhällets informationsinfrastruktur själva börjar tillhandahålla tjänster (och produkter) som de hittills enbart har förmedlat andras information om. Google har gått från att enbart agera informationsförmedlare till att numera även vara

[10] http://static.googleusercontent.com/media/www.google.com/sv//webmasters/docs/search-engine-optimization-starter-guide.pdf

[11] https://support.google.com/webmasters/answer/35769?hl=en

[12] https://moz.com/search-ranking-factors/correlations

informationsproducent och i förlängning till att också sälja produkter och tjänster,[13] inte minst genom Google Play. Produkterna Google Shopping, Google Books och Google Flights är exempel på denna förändring. Ett mycket omtalat exempel är Google Flights, där Google även agerar resebyrå. Söker man på Google efter ”Flyg till London” får man numera överst i listan av det organiska (icke-sponsrade) resultatet en sammanställning av flyg direkt från flygbolagen där priser, restider och platser jämförs. Andra, konkurrerande, tjänster såsom Momondo, Expedia, Travellink hamnar längre ner på listan om man inte söker specifikt efter deras namn eller om de har köpt reklamplacering i den övre icke-organiska delen av resultatsidan. På EU-nivå har detta lett till en lång rad av klagomål sedan flera år tillbaka.[14] Google anklagas för att utnyttja sin marknadsdominanta position och Android-systemet, där Google finns inbyggt i infrastrukturen, till sin egen fördel. Ett antal juridiska processer som undersöker Googles roll i dagens informationsinfrastruktur och ekonomi pågår i EU.[15]

Driva opinion

En slags sökoptimering kan också göras av enskilda mer eller mindre organiserade användare för att driva opinion. I en svensk kontext ger exempelvis en sökning i YouTube 2016-01-18 på statsminister Stefan Löfven ett sökresultat där ett stort antal av de översta träffarna är upplagda och kommenterade i ett syfte att placera Stefan Löfven i ett negativt sammanhang. Orsaken torde vara, att döma av innehållet i många kommentarer, den stora aktivitet bland användare långt ut på den politiska högerkanten som ser på videor, skriver kommentarer och gillar varandras kommentarer. Youtube belönar dessa aktiviteter med högre synlighet (som i sin tur möjliggör ökad exponering för tjänstens kunder, dvs. företagen som placerar reklam i t.ex. YouTube). Videor där Stefan Löfven framställs i ett positivt eller mer neutralt sammanhang får då en lägre rangordning. En sådan tolkning stöds indirekt av resultaten som redovisas av Schwartz Anderssons et al. (2015) om ”Sverigedemokratiska politiska klustret” höga aktivitet i Twitter vilket vi menar kan peka på en strategisk och algoritmanpassad aktivitet även i andra sociala medier (jmf. O’Callaghan et al., 2015).

Knowledge Graph

Ett intressant fall inom samma område – dvs. en förskjutning från förmedling av information till produktion (curation) – som troligen kommer att vara av stor framtida betydelse är Googles Knowledge Graph. Genom Knowledge Graph vill Google inte bara förmedla länkar, utan skapa och förmedla data direkt till användaren. Vid en sökning på exempelvis Ukraina presenteras numera inte endast sökresultat utan också en faktaruta till höger med uppgifter hämtade från Världsbanken och Wikipedia. Därmed behöver de användare som endast är ute efter snabb fakta inte gå vidare till dessa källor utan kan använda Google Search som både en startpunkt och en slutdestination. Det finns analytiker som hävdar att detta kan få effekter på trafiken till Wikipedia (DeMers, 2015-09-03). Knowledge Graph är en form av semantisk sökfunktion, där man inte matchar ”character strings”

[13] http://www.slideshare.net/lutherlowe/wu-l
[14] http://www.fairsearch.org/httpmediafse-150605-chronology-pdf/
[15] http://europa.eu/rapid/press-release_IP-15-4780_en.htm

(teckensträngar), som görs vid vanlig fulltextsök, utan man letar efter och visar upp data i relation till ett begrepps semantiska betydelse. För att detta ska fungera måste innehållet (”datan”) vara uppmärkt enligt särskilda regler (kallas ofta för semantisk webb). I nuläget är det därför främst stora öppna databaser (t.ex. Världsbanken, Wikipedia, Wikidata, CIA World factbook) som används i Knowledge Graph, men i en nära framtid kommer troligen mer och mer innehåll att vara uppmärkt enligt reglerna som finns att tillgå på schema.org.

Att söka information kan därför bli mindre en fråga om att matcha sökord med webbsidor, utan snarare sökord med data som finns uppmärkta i enskilda dokument eller i databaser. Med en sådan utveckling kommer Google Knowledge Graph att extrahera data som märkts upp av andra och presentera data och dess relationer, som i det ovan nämnda exemplet Ukraina. Schema.org är sponsrat av Google, Microsoft, Yahoo och Yandex (en rysk sökmotor). Ett projekt som liknar Knowledge Graph till en viss del är Wolfram Alpha. Även här är fokus på data och fakta, fast mer i syfte att ge tydliga svar på direkta frågor. Ytterligare ett intressant projekt genomförs av ett företag som heter Diffbot som har flera av Googles konkurrenter som kunder, t.ex. Microsoft (Bing), DuckDuckGo och Amazon. Till skillnad från i Knowledge Graphs fall, där principen är att andra märker upp datan, bygger Diffbots projekt på att de har en algoritm för att först dammsuga ”vanliga” webbsidor efter innehåll som sedan märks upp för att skapa den semantiska strukturen. Ett intresse för och fokus på data istället för på dokument kommer naturligtvis även förändra förutsättningarna för sökmotorsoptimering och -marknadsföring.

Flödesoptimering

I Facebooks fall handlar det mindre om att vara sökbar, som i fallet med sökmotorer, men ändå om att vara synlig. Facebook skiljer mellan ”organic reach” och ”paid reach” där ”Organic reach is the total number of unique people who were shown your post through unpaid distribution. Paid reach is the total number of unique people who were shown your post as a result of ads.”[16] När Facebook ändrade sin algoritm för nyhetsflödet från EdgeRanks kronologiska ordning till en ordning baserad på en mångfald faktorer så klagade exempelvis icke-kommersiella intresseorganisationer (NGOs) över drastiskt minskad synlighet för de som inte har en budget för att betala för synligheten. Man kan här prata om en ”flödesoptimering” som görs för att säkerställa och öka synligheten i följarnas nyhetsflöde.

Genom att posta en viss sorts innehåll belönas det av algoritmen med ökad synlighet. Facebook ger förslag på hur man kan öka räckvidden för sina inlägg genom att posta bilder eller komma med erbjudanden och frågor som uppmuntrar interaktion.[17] Vissa sidor tar till något "drastiska" medel och gör algoritmen synlig för sina följare. Här har vi ett exempel från Rättviseförmedlingens Facebooksida, postad i januari 2016:

> *Saknar du oss? Hej alla som gillar Rättviseförmedlingen på Facebook och speciellt du som har gillat oss ett tag men inte ser våra efterlysningar så ofta i nyhetsflödet längre! Så här funkar det: Facebook håller koll på vilka personer eller sidors statusar du gillar och kommenterar ofta och antar att du vill se mycket av det i ditt flöde. Så försök tänka på*

[16] https://www.facebook.com/help/285625061456389
[17] https://www.facebook.com/help/215169031896481

> *Rättis som nån du är kär i och likea loss även när du inte har något att kommentera. På det sättet får du se mer av oss! Dessutom är det så fiffigt att genom att stödlikea våra statusar ofta så är chansen ganska stor att dina vänner (som brukar gilla dina statusar) också får se att du gillar oss och i sin tur hittar hit och trycker på gilla-knappen. The Circle of Like helt enkelt! Så passa på att gilla den här och våra typ tio senaste statusar så bidrar du till kretsloppet av fler likes fler rättviseförmedlare fler tips på personer till våra efterlysningar ännu mer rättvisa. Tack för att du är med!* [18]

Hur påverkas samhället av algoritmernas styrning?

Algoritmerna är som vi visat en djupt integrerad del av vårt samhälle. Algoritmerna är ett utryck för dominanta föreställningar och värderingar i vår kultur. Det går att uttrycka som att algoritmer kodiferierar dessa föreställningar och värderingar. Moralen som begränsar bilder på nakna bröst leder till att Facebook censurerar alltför ”explicita” bilder på amning. Detta får genomslag världen över genom de värderingar som finns inbyggda i Facebooks algoritm, som letar upp och döljer nakenhet. Ett annat exempel är att Google har behövt be om ursäkt för att en algoritm för bildbeskrivning taggade två svarta personer med ”gorillas”.[19] Dessa är bara exempel på hur sexism och rasism blir del av algoritmiska beslut, men de förtydligar hur kulturella värderingar och etiska förhållningssätt är inskrivna i algoritmer samt möjliga konsekvenser av dessa som utöver implikationer för hur ett innehåll sprids även formar själva innehållet. Algoritmer och samhälle formar därmed tillsammans varandra.

Det finns idag få verksamheter där inte algoritmer spelar stor roll och det gäller även vår vardag. Samtidigt är algoritmerna ofta osynliga i vardagens praktiker och det är inte tydligt i vilken omfattning organisationer faktiskt baserar sina beslut och sin verksamhet kring egna eller andras algoritmer. Nedan följer några exempel på områden där algoritmer spelar avgörande roll för verksamheter eller beslut.

Aktiemarknaden och organisering av arbete

Det är uppenbart att algoritmer är avgörande för många företags kommersiella verksamhet och att de är en viktig del av hur man försöker att maximera sin vinst. Om vi vidgar perspektivet till att också omfatta ett något bredare fokus än denna rapport finns det en uppsjö av olika sätt att öka företags intäkter på genom algoritmer: organisering av skiftarbete, produktionsprocesser, mätning av prestationer och övervakning av anställda och så vidare.

Ett särskilt exempel, som kan vara mycket problematiskt, är hur aktiemarknaden till stor del är styrd av algoritmer, i sin mest extrema form i form av s.k. ”high-speed trading”. Där gör algoritmer affärer med varandra, utan att människor är inblandade och med en hastighet som långt överstiger människors reaktionstid. Fördelarna är uppenbara på en marknad där snabba beslut premieras, men de negativa konsekvenserna kan vara betydande. Ett exempel på möjliga negativa följder är s.k.

[18] https://m.facebook.com/story.php?story_fbid=10154022091854750&id=361481469749
[19] http://www.theguardian.com/technology/2015/jul/01/google-sorry-racist-auto-tag-photo-app

”flash crash”. År 2013 försvann t.ex. nästan 7 miljarder dollar i Singapore på ett ögonblick och orsakade stor ekonomisk skada. Även samspelet mellan sociala medier och trading-algoritmer kan vara problematiskt. Karppi och Crawford (2015) diskuterar hur det år 2013 spreds ett falskt tweet om en terrorattack på Vita huset från ett hackat twitterkonto från Associate Press. Detta tillsammans med trading-algoritmernas tempo kan kopplas samman med en krasch på aktiemarknaden som raderade 136 miljarder dollar på ett ögonblick.

Andra nya aktörer som Uber, Lyft, Rideshare, AirBnB och inte minst Amazon använder algoritmer inte enbart för att ordna och tillgängliggöra digital information och sälja reklamplatser utan även för att organisera verksamheten utanför nätet. Dessa företag och deras tjänster delar in, styr och kontrollerar sina arbetstagare och avtalspartners själva arbete med hjälp av algoritmer (Rosenblatt & Stark, 2015). Mätbar effektivitet, högt tempo, användares recensioner – helst kvantifierbara – matar algoritmerna med data och styr arbetet. Allt detta kan ha stora konsekvenser för marknaden, staten och arbetslivet för många löntagare, som i en så-kallad ”gig economy” blir mer och mer till avtalspartner snarare än arbetstagare, men det sker på plattformarnas villkor.

Nyhetsmedier

För att komma tillbaka till rapportens mer specifika fokus – dvs. algoritmer som förekommer i sammanhang och tjänster där användargenererad data utgör en central komponent, och vars funktionalitet skapar tillgång till information och möjliggör kommunikation – så är kakorna som nämns ovan ytterst relevanta. Genom att möjliggöra för företag att få kunskap om användares beteenden och preferenser uppfyller de en central funktion för kommersiella intressen. Företag använder data som samlas in med hjälp av kakor till att skräddarsy erbjudanden om tjänster, produkter och reklam. Ett intressant exempel är hur nyhetsmedier, som med undantag av public service media är kommersiella men med ett traditionellt sett samhälleligt uppdrag, använder sig av algoritmer för att synliggöra sina nyheter och för att anpassa de efter publiken (Strömbäck, 2014). Strömbäck (2015) lyfter fram hur nyhetsmedier är hårt drabbade av förflyttningen av annonsmarknaden till nya aktörer, såsom Google och Facebook. Dessa företag är både konkurrenter med nyhetsmedier om annonsörerna samtidigt som deras tjänster utgör den infrastruktur som fler och fler människor använder för att ta del av nyhetsmediers innehåll. På olika sätt har algoritmer tillsammans med (big) data därtill börjat att användas i redaktionerna, antingen för att generera nyheter ur data (s.k. robot-journalism) eller för att anpassa nyheterna efter läsarna. Westlund och Lewis (2014; se också Lewis & Westlund, 2015) lyfter fram hur teknikernas och teknikens roll i redaktionerna växer medan redaktionellt arbete – som kräver finansiering av journalister – minskar. Följden är, vilket även Strömbäck (2015, s. 360f) diskuterar, att ”andelen journalistik av det totala medieutbudet har minskat och kommer att fortsätta minska”. Istället ges mer utrymme åt användargenererat och automatgenererat innehåll. Därtill anpassar sig nyhetsmedier till strukturerna som den digitala reklammarknaden erbjuder och där popularitet som är likställd med synlighet eller delningsbarhet belönas.

Bibliotek

Det är inte enbart kommersiella företag som använder sig av algoritmer och det finns andra drivkrafter än rent kommersiella för att vilja öka synlighet. Exempelvis så utgör algoritmer en viktig ingrediens i synliggörandet av många bibliotek. Fast eftersom dagens mest dominanta tjänster omsätter en kommersiell logik i sina algoritmer så kan man också se hur det skapar konflikter för biblioteken. Vi ser i folkbiblioteken hur den kommersiella logiken utmanar det traditionella uppdraget. Ett exempel är hur folkbiblioteken i Sverige köper tillgängliggörande av e-böcker av företaget Elib. Folkbiblioteken är här beroende av ett externt företag som inte bara tillgängliggör litteraturen utan också presenterar den för folkbibliotekets användare i enlighet med den kommersiella logiken där popularitet lyfts fram. För att som svensk låntagare kunna få tillgång till en e-bok via ett svenskt folkbibliotek måste du utöver hårdvaran ha tillgång till ett Apple eller Google Play konto samt ett så-kallat läsarprogram. För exempelvis Ipad/Iphone krävs Blue Fire Reader, vilket då kräver ett Adobe ID. Genom Blue Fire Reader, eller motsvarande läsarprogram, förser svenska e-bokslåntagare det internationella företaget Blue Fire Reader med både personlig data och data som inte kan kopplas direkt till en person. För den förstnämnda garanterar företaget att den inte säljs vidare till ”third parties”, men de kan inte utesluta att datan lämnas ut till staten om lagstiftningen kräver det.[20] Vidare använder sig många folkbibliotek av topplistor i form av flest lån, användares betygssättning och andra kännetecken på inflytande från algoritmerna i programvaran. Algoritmer har på ett sätt utan tvekan förbättrat tillgången för folkbibliotekens användare, men på ett annat sätt har de kommit att begränsa användarnas uppmärksamhet i någon mån till det som är populärt.

Vetenskap och akademin

Även inom vetenskaplig publicering utgör algoritmer en central del av hur genomslag eller relevans bedöms. Exempelvis Google Scholar är en sökmotor för vetenskapliga publikationer som bygger på liknande principer som Google Search. Publikationer hamnar högt upp i träfflistan baserat bl.a. på antalet citeringar som artiklar får från andra publikationer. Detta har kritiserats då Google Scholars algoritm inte är transparant på samma sätt som bibliotekens traditionella system (van Dijck, 2010). Metriska mått avgör vad som räknas som god forskning och de kvantitativa måtten på forskningsframgång blir viktigare än det kvalitativa (Sundin, 2014). Anledningen till varför man började använda dessa mått är att det ofta finns ett samband mellan kvalitet/relevans och hur ofta en text läses eller citeras. Men det är statistiskt samband och kan därmed inte användas för att utvärdera och bedöma individuella artiklar och än mindre forskare. I Google Scholar utgör forskarens vilja eller snarare behov av synlighet en av de främsta underliggande drivkrafterna för sökmotorns funktionalitet. Detta kan enbart indirekt länkas till kommersiella drivkrafter, utan ligger snarare till grund för forskares akademiska prestige, akademiska status och därmed möjlighet till karriärsutveckling. Vi ser motsvarande funktion i akademiska tjänster som ResearchGate, Academia eller Mendeley. Det är viktigt att påpeka att det inte är algoritmerna själva som enbart skapar ett ökat fokus på kvantitativa mått för kvalitet (t.ex. popularitet), utan algoritmerna bör ses i ljuset av en bredare förändring av offentlig verksamhet som ofta brukar benämnas för en mätbarhetens kultur som

[20] http://www.bluefirereader.com/internal/privacy-and-license/privacy.html

tar sitt utryck t.ex. i management metoder med fokus på ansvarsskyldighet och utvärdering.

Internet of things / The quantified self

Det är främst genom våra telefoner, läsplattor och persondatorer som det blir tydligt hur algoritmer blir en del av vår vardag, men de flesta av våra apparater har idag också små datorer i sig – klockor (Google Watch), tvättmaskiner, gräsklippare, dammsugare och bilar. Det som kallas för *Internet of things*, en värld där fler och fler föremål är uppkopplade, håller på att bli verklighet. Även produkter och tjänster för att övervaka kroppens funktioner (sömn, träning, sjukdom etc.) vars användning ofta samlas under begreppet *The Quantified Self Movement*, växer fram. Stora mängder data samlas in av dessa föremål och tjänster och allteftersom dessa blir uppkopplade, inte minst till varandra, kommer algoritmer i högre grad påverka samhället och dess styrning. Inte minst har försäkringsbolagen visat stort intresse för att utnyttja möjligheterna som öppnas med *Internet of Things* och *The Quantified Self* för att anpassa sina produkter med hjälp av algoritmer för riskbedömning.

Insyn, kontroll och transparens

Eftersom algoritmer styr stora delar av hur vårt samhälle fungerar kan frågan om insyn i deras utformning tyckas berättigad. Samtidigt är de mest dominanta algoritmerna grunden för företagens kommersiella existens och mycket av deras utformning är hemlig, inte minst på grund av konkurrenssituationen. Det är närmast omöjligt att studera algoritmers påverkan på ett exakt sätt då kod och människor är i ständig rörelse (Introna, 2015). Det finns helt enkelt ingen neutral punkt när algoritmer kan studeras utan olika tjänster och system interagerar hela tiden med varandra, bl.a. beroende på den data som de bygger sina tjänster på och de plattformar tjänsten används genom (Seaver, 2013). Algoritmer är dynamiska och förändras kontinuerligt samtidigt som de förfinas och därmed blir mer omfattande. De flesta stora tjänster kan idag sägas operera med algoritmiska system snarare än med enstaka algoritmer, vilket gör fullständig insyn närmast omöjlig (Introna, 2015). Algoritmer kan inte studeras isolerade utan måste ses som en del av en teknisk och social helhet där algoritmer ingår i nätverk av möjligheter och begränsningar som formar och formas av varandra.

Algorithmia

Med detta sagt, en viss insyn kan erhållas genom att studera begränsade effekter av algoritmers verkan inom ett avgränsat område på så sätt som det görs inom SEO branschen eller inom samhällsvetenskaplig och humanistisk forskning om algoritmer (critical algorithm studies). Inom forskningen finns en medvetenhet om betydelsen av grundläggande kulturella premisser för algoritmers utformning, ofta grundade i konsumtionssamhällets ideal och med syfte att öka synlighet utifrån popularitet, spridning och effektivitet samtidigt som olika etiska värderingar om vad som utgör relevans i olika sammanhang finns inskrivna i dem. Å andra sidan existerar även algoritmer med öppen källkod som kan återanvändas. En intressant företeelse som

kan nämnas i detta sammanhang är ”Algorithmia: Market place for algorithms”.[21] Webbplatsen är ett forum för att köpa och sälja algoritmer. Algoritmerna är ”open source” och forumet utgör en slags gemenskap för programmerare och de användare som söker en specifik algoritm eller en utvecklare för lösningen av ett specifikt problem som företaget, organisationen eller privatpersonen har. Det rör sig inte sällan om betydligt mindre kostnader än det ovan nämnda exemplet om Netflix pristävling. För 150 dollar söker i december 2015 en användare en algoritm för att lösa problemet med kommatering i texter. Här finns alltså insyn i utformningen av individuella algoritmer. Hur dessa sedan blir en del av ett algoritmiskt system (jmf Introna, 2015) och hur de uppdateras allteftersom de används är däremot inte möjligt att ha insyn i.

Kontroll över personlig data

Många algoritmer som nämns i denna rapport samlar in data i olika former och processar dem för att ge önskat resultat, men även för att utvecklas och förbättras. Dataskydd och integritetsfrågor är inte i rapportens egentliga fokus, men datainsamling och hur data måste hanteras är reglerad genom lagen på olika sätt. Trots detta är en användares kontroll över data som avser ens person – även om det finns lagstiftning gällande dataskydd och integritet - i praktiken mycket begränsad. Endast i liten utsträckning kan man säga att de data som algoritmer arbetar med kontrolleras av användarna, eller för den delen av demokratiska instanser. Istället ägs de av kommersiella företag som till stor del finansieras genom placering av reklam. Ändå finns det på olika sätt metoder som den medvetna användaren kan använda sig av för att i större utsträckning kontrollera vilka data som samlas in om en själv, eller hur de används, främst genom att anpassa privacy-inställningarna i sina användarprofiler. Möjligheterna är dock inte särskilt flexibla och de är ofta begränsade till att kryssa för ett fåtal boxar utan vidare förklaring av konsekvenserna.

Googles olika tjänster, i allt från Google Play, till Youtube, Gmail eller Calender och till och med olika Android-produkter är ofta kopplade till ett och samma Google-konto. Genom Google Dashboard möjliggörs för användarna att se vilken information som samlas in om dem eller närmare bestämt om deras profil. Ett annat exempel är att man genom särskilda ”plug-ins” till sin browser kan synliggöra vilka kakor som sparas ner. Ytterligare ett sätt att kontrollera vilka data användaren lämnar ut är att man i Google, eller för den delen Apple, låter bli att logga in. Samtidigt är det svårt att låta bli att vara inloggad när smarta telefoner blir allt vanligare. Vi lämnar data till olika tjänster. Vid varje enskilt tillfälle kanske det inte upplevs som problematiskt, men när datan sammanställs kan de enskilda delarna bilda ett mönster som vi inte är beredda på och utgöra en ”algoritmic harm” (Tufekci, 2015b).

Inlåsning

Både Google och Facebook vill vara startpunkten för människors tillgång till internet. José van Dijck (2013, s. 154ff.) visar hur plattformar strävar efter att behålla sina användare genom att låsa in sina användare och stänga ute konkurrenter där det är möjligt. Det görs genom att göra konkurrenters applikationer icke- eller åtminstone svår-kompatibla eller genom att försöka skapa ett komplett ekosystem för internet där en kedja av plattformar gör en sömlös användning möjlig. Google har t.ex. partnerskap med Twitter, en egen webbläsare (Chrome) och äger numera

[21] https://algorithmia.com/

Android, YouTube samt en hel portfölj av olika kommunikationstjänster[22]. Facebook har samarbete med Microsoft och Spotify och har t.ex köpt up WhatsApp och Instagram. Microsoft och Apple är andra som erbjuder eller försöka erbjuda närmast kompletta ”paket” som uppfyller de flesta behov. De olika relationerna förändras hela tiden, inte minst genom uppköp. Men poängen är att tillsammans skapar de konkurrerande kedjor av plattformar och en social och kulturell verklighet som det är svårt att stå utanför. När alltfler kulturella arrangemang synliggörs på Facebook, när närmast alla organisationer och företag använder Facebook som främsta kanal för att nå ut, när även offentligt finansierade institutioner som universitet, bibliotek, eller public service radio informerar genom Facebook är det då möjligt att välja bort Facebook utan att gå miste om mycket som krävs för att vara delaktig i samhället?

Google: Osynlig och utan insyn

Sökning har i det närmaste blivit synonymt med att googla. Ju mer integrerad sökning blir i vardagen, desto mindre ifrågasätts Googles rangordning av sökresultat. Vår egen pågående forskning visar att människor har svårt att begreppsliggöra och, kanske framförallt, konkretisera alternativ till Google och dess sökresultat. Google blir en del av vad som tycks vara en neutral infrastruktur och därmed osynlig för användaren (jmf Hillis, Petit & Jarrett, 2013). Googles sökresultat tenderar att ses som självklara av många användare.

Vid sökning är kontroll över indexet minst lika viktig eller kanske till och med viktigare än kontroll över algoritmen för att söka på det. Sökmotorers algoritmer arbetar inte enbart på användargenererad data, utan framförallt på det webbindex som utgör grunden för t.ex. Google Search. Sökrobotar samlar in data från världens webbplatser och bygger av detta det index som sedan användarna söker i. Också insamlingen av datan till Googles index är styrd med hjälp av algoritmer som talar om när data ska samlas in, hur mycket och i vilken följd. Det finns enbart ett annat index som matchar Googles och det ägs av Microsoft och används av såväl Bing som Yahoo. Även några mindre alternativa sökmotorer (t.ex. DuckDuckGo) har tillgång till Microsofts index. Yandex, en rysk sökmotor, liksom Baidu, Kinas dominanta sökmotor, har också egna index. Yandex och Baidu har numera större marknadsandelar än Microsofts Bing. Det finns dock inget offentligt tillgängligt index av webben. Det betyder att möjligheten till utveckling av nya alternativa, kommersiella eller icke-kommersiella, sökmotorer, vars algoritmer omsätter andra värderingar än de som dominerar dagens digitala samhälle, är starkt begränsad, vilket kan ses som ett problem för det demokratiska samtalets mångfald. Forskare, såsom sökmotorsforskaren Dirk Lewandowski, argumenterar därför för upprättandet av ett oberoende index så att flera sökmotorer kan använda samma underlag.[23]

[22] https://www.google.com/intl/en/about/products/

[23] http://networkcultures.org/query/2013/11/11/dirk-lewandowski-why-we-need-an-independent-index-of-the-web/

Informationsbubblor och det offentliga samtalet

Eli Pariser (2011) myntade uttrycket filterbubbla, med vilket han avser hur vår tillgång till information tenderar att bli allt mer personaliserad så att två människor riskerar att ta del av olika information trots att de letar efter samma ämne. Personalisering bygger bl.a. på vad en användare har klickat på, platsen varifrån användare når nätet, vilka vänner en användare har (t.ex. i Twitter) och användarens sökhistorik. Genom denna information försöker algoritmen förutsäga vilken information användaren vill ha, såsom länkar, nyheter eller förslag på böcker. Fördelen med personalisering är uppenbar när den fungerar, men den nackdel som brukar lyftas fram är att den kan leda till en bekräftelse av användarnas världsbilder genom att de nyheter eller webbplatser med innehåll som användarna håller med om prioriteras. Personalisering finns i olika grader i en rad olika webbaserade tjänster, men i rapporten beskrivs fenomenet genom Google Search, Facebook, Twitter och Amazon. Det är utan tvekan så att personalisering i sociala medier och sökmotorer vanligtvis underlättar avsevärt för användare. Om vi söker på pizzeria i Google är det en fördel om vi får tips på restauranger i vår geografiska närhet. Det är ett mervärde för en tjänst som Amazon, Netflix eller Spotify att kunna föreslå böcker, filmer eller musik som vi annars kanske inte hade vetat fanns. Med detta sagt är det ett problem om vi inte vet att ett visst innehåll prioriteras högre än ett annat, eller i vilken grad som personaliseringen äger rum.

En offentlig sfär?

Dagens dominerande sociala medier ägs av globala kommersiella företag där konsumtionssamhällets logik många gånger finns inbyggda i algoritmerna. Detsamma kan också sägas gälla sökmotorer (Mager, 2012). Därigenom har sociala medier och sökmotorer i regel en annan grundläggande utgångspunkt än den åsiktspluralism som ofta sägs vara en utgångspunkt för det demokratiska samhället. Istället för att skapa förutsättningar för en mångfald av synsätt och åsikter stärks ”det populära” och det individuellt anpassade. Medias roll i samhället diskuteras många gånger i relation till det teoretiska begreppet offentlig sfär. Sociologen Jürgen Habermas, som myntade begreppet offentlig sfär, gör en distinktion mellan en offentlig plats och en offentlig sfär (se Papacharissi, 2010, s. 115). En offentlig sfär tjänar till att stödja den representativa demokratin genom att medborgare deltar i rationellt offentligt samtal där en mångfald av röster blir hörda, medan en offentlig plats visserligen kan skapa förutsättningar för diskussion, men inte på ett för demokratin gynnande vis. Begreppet offentlig sfär har problematiserats i relation till huruvida det egentligen är möjligt att uppnå en sådan sfär eller om det inte snarare är ett utopiskt ideal. Oavsett vilket kan idéen om en offentlig sfär fungera för att diskutera algoritmers roll i samhället.

Internets potential för att kunna utveckla samhällets offentliga sfär är tydlig. Medborgarna ges tillgänglighet till information och möjligheter att föra fram argument och på så sätt delta i ett offentligt samtal som syftar till att skapa goda förutsättningar för demokrati. Samtidigt visar forskningen (Papacharissi, 2010, 120ff) att internet visserligen skapar förutsättningar för diskussioner (plats), men inte nödvändigtvis på ett sådant sätt att de stöder demokrati (sfär). Tillgängligheten till information leder inte nödvändigtvis till att människor faktiskt tar till sig information, möjligheter till samtal betyder inte alltid att samtalet är ömsesidigt och den ökande

kommersialiseringen av internet minskar förutsättningar för det öppna samtalet (ibid). I ljuset av det vi diskuterar i vår rapport vill vi komplettera bilden med hur sökmotorer, sociala medier och de algoritmer som dessa bygger på, påverkar den information vi tar del av och de samtal vi faktiskt tar del av. Om en webbplats existerar betyder inte det att den är tillgänglig i praktiken om den rankas långt ner av Google. Det finns gott om möjligheter att göra sig hörd, men om filterbubblor gör att människor i liten utsträckning blir åsiktsmässigt utmanade blir det svårt att tala om en offentlig sfär.

Algoritmer förstärker trender

Algoritmer ses i rapporten som sociomateriella och det är då viktigt att betona att samhället också skapar algoritmernas utformning. Algoritmerna är en del av vår kultur som i allt större utsträckning är en individualiserad konsumtionskultur byggd på en kapitalistisk logik av ackumulation och konkurrens, efterfrågan och utbud. På olika sätt är detta inbyggt i algoritmerna, ibland tydligare som med Facebook, Amazon, Netflix eller Uber, ibland undangömd som i fallet Google. Kunskap skapas och cirkuleras genom dessa tjänsters algoritmer, personer blir profiler och synlighet belönas. Företag som Uber som effektiviserar hela sin verksamhet baserad på olika algoritmer påverkar samhället och hur arbetet organiseras. Algoritmer som kan omsätta en samhällstrend belönas. Det är inte algoritmer som enskilt skapar detta, utan de accelererar och effektiviserar pågående samhällstrender. De bidrar även till att sprida dessa trender till verksamheter som tidigare befann sig utanför en individualiserad konsumtionskultur. Genom algoritmer förstärks t.ex. denna kultur även i bibliotekens, i encyklopediers, i public service och i forskningens verksamhet. När biblioteken tar hjälp av sociala medier och popularitetslistor för att ”marknadsföra” sin verksamhet eller styra läsares val, så förändrar det bibliotekens sätt att fungera och biblioteken riskerar att bli till en plats för möjlig kulturell konsumtion att välja bland andra. Uppslagsverk bygger sitt arbete i större utsträckning på Google Analytics och vad människor söker efter (Sundin & Haider, 2013). Istället för ta utgångspunkt i hur forskningen förändrar kunskapen blir utgångspunkten för uppslagsverkets innehållsliga arbete mer än tidigare vilka ämnen människor söker på i uppslagsverket. Algoritmer har även stora konsekvenser för hur nyhetsmedier skapar innehåll och vad som flyter upp till ytan i offentligheten eller vad som sjunker ner (SOU 2015:94). Mätbar popularitet – mest "likes", mest delat, mest läst – belönas och därmed skapas en marknad för det statistiskt populära. Även forskning mäts allt oftare i form av impact och nya former av att mäta detta genomslag baserad på sociala medier håller på att utvecklas (Sundin, 2014).

Delningslogiken

Twitter medför en personalisering av flödet, vilket torde vara självklart då grundprincipen bakom Twitter är att du främst tar del av de tweets som publiceras av personer/profiler som du har valt att följa. Samtidigt är det uppenbart hur Twitter kan förstärka ens världsbild avseende exempelvis nyheter. Många tweets är länkar till nyhetsartiklar eller analyser från andra medier och om du delar politisk hemvist med de som du följer ser du i regel enbart tweets som överensstämmer med dina värderingar. Det går att twittra också till de som inte är ens följare, men flera studier har synliggjort styrkan av Twitters personalisering (Abel et al., 2011). Det finns ett fåtal funktioner (top tweets, trending tweets, rekommenderade användare) i Twitter

som utnyttjar popularitetsmått för att synliggöra visst innehåll och dessa bygger på ens vidare nätverk av följare, men ett twitterflöde är främst organiserat i kronologisk ordning och påverkas inte av en användares interaktioner som exempelvis nyhetsflödet i Facebook nyhetsflödet gör.

I Facebook får du ta del av dina vänners nyheter och delningar, men Facebook är en betydligt större tjänst än Twitter när det gäller antalet användare och dess algoritm fungerar annorlunda och tar hänsyn till betydligt fler faktorer. Såsom vi beskriver ovan skapas personaliseringen i Facebook av en lång rad faktorer: vänner, profiluppgifter, "likes", klick på länkar, spel och tester användaren deltar i och användarens egna uppdateringar. Nyligen publicerade forskare från Facebook en rapport (Bakshy, Messing & Adamic, 2015) där författarna hävdar att bubbel-fenomenet i Facebook främst beror på vad användarna medvetet väljer att klicka på. Studien har kritiserats av många forskare. T.ex. Tufekci (2015a) hävdar istället att Facebook faktiskt gömmer nyheter i flödet som Facebooks algoritm tror du inte skulle hålla med om. Istället argumenterar Tufekci i samma text för att den enskilde användaren borde kunna ställa in villkoren för vad som ska visas i nyhetsflödet så att t.ex. mer av sådant som hon inte håller med om visas, dvs. en ökad kontroll för användaren att vara med och påverka algoritmen.

Som Bäck, Bäck och Gustafsson (2014; jmf SOU 2015:94, s. 225 ff.) skriver kommer ”en stor del av trafiken till svenska tidningars webbsidor via Facebook” och troligtvis är det flera andra sociala medier som står för ytterligare en stor del. Personliga rekommendationer tillsammans med algoritmens tolkning av användarnas preferenser och intressen styr alltså idag till en hög grad vad som läses i medierna: ”Denna typ av mediekonsumtion leder till att redaktionellt arbete och traditionell nyhetsvärdering får relativt mindre inverkan på vilket innehåll som konsumeras, medan de mekanismer som får människor att dela vidare innehåll får större inverkan: individens sociala nätverk kommer at fungera som ett mediefilter där vissa nyheter förstärks och andra sållas bort” (Bäck et al., 2014, s. 11).

Det är viktigt att komma ihåg att det personliga nätverket enbart är en faktor i hur algoritmerna styr synligheten av innehåll. Som vi diskuterar ovan säljer dessa företag användarnas uppmärksamhet och algoritmerna är utformade i syfte att mäta och styra uppmärksamheten. Gilla, dela, kommentera är vad man kan göra på Facebook. Alla dessa aktiviteter är uttryck för ett intresse som ökar spridningen och synligheten av ett innehåll. Popularitet, dvs. omfattande spridning och delning genom nätverket, premieras och belönas eftersom den är likställd med stor räckvidd och hög uppmärksamhet. Att bli viral – dvs. att sprida sig snabbt och till många – är idealet som eftersträvas i många fall. I en omfattande studie har det uppmärksammats att icke fakta-baserat innehåll ofta har tendens att spridas snabbare än fakta-baserat innehåll (Silverman, 2015). Rykten blir lättare virala än sanningar. Dessutom sprids de under en längre tid medan rättelserna sällan lever lika länge och inte blir delade lika ofta. Hur snabbt det kan gå visar det mobilanpassade mediebolaget KIT när det steg-för-steg redogör för hur ”nyheten” om åldersgräns för sociala medier blev viral och feltolkad 29 januari 2016.[24]

[24] https://kit.se/2016/01/29/33055/studie-i-4-steg-sa-har-far-en-nyhet-eget-orimligt-liv/

Det är redan nu tydligt att medierna anpassas efter den logiken i och med att kategorier som ”mest läst just nu” eller ”mest delat just nu” har blivit vanliga i de flesta nyhetsmediers online-tjänster. Man anpassar alltså inte enbart det redaktionella arbetet efter de sociala mediernas algoritmer utan synliggör detta även för läsarna, vilket riskerar att skapa ett slutet kretslopp av bekräftande. Bolin och Schwarz Andersson (2015, s. 9) skriver om ett algoritmisk tänkande som har tagit sig in i journalistiken och vilket, genom att försöka förutspå viralitet och genom att reagera på delningsmönster i sociala medier (t.ex. Twitter eller YouTube), förstärker algoritmernas kommersiella logik och de stora infrastrukturaktörernas dominanta ställning. Den mätbara populariteten tillsammans med betydelsen av personliga nätverk, en kategorisering av användare enligt marknadsforskningens segment och en tendens till fragmentering av innehåll (t.ex. i individuella artiklar eller filmer), förändrar inte enbart hur medierna konsumeras utan även hur de tillgängliggörs och därmed produceras.

Samtidigt motverkat det innehåll som blir viralt genom delningar i sociala nätverkstjänster i någon mån informationsbubblor. När ett innehåll har blivit delat i sådan utsträckning bryter det ut ur informationsbubblor. Det gäller särskilt när nyhetsmedier fångar upp virala fenomen och ytterligare cementerar deras status som kulturella fenomen som binder samman människor i en samhällelig gemenskap (Gillespie, in press 2016). Det kan exemplifieras med när barnen på förskolor dansar Gangnam Style samtidigt som SVT diskuterar fenomenet i kulturprogram och rapporterar om Gangnam Style flashmob i nyhetssändningar. YouTube fenomenet blir då i det närmaste en gemensam global kulturell referenspunkt som är medskapad av den algoritmiska delningslogiken. Gangnam Style är nu flera år gammalt, men har tagit steget till ett kollektivt kulturellt minne och referenspunkt.

Personalisering

Googles algoritmer arbetar, som vi diskuterar ovan, med olika faktorer för personalisering. I algoritmen finns det ett inbyggt användarperspektiv, vilket innebär att det är mycket svårt att undgå någon form av personalisering, inklusive geolokalisering. Syftet är att öka sökresultatens relevans med utgångspunkt att relevans är personlig och inte objektiv och universell. Feuz et al. (2011) gjorde en empirisk undersökning av personaliseringens effekt vid webbsökning och kom fram till att personalisering handlar om mer än att enbart öka relevansen för användarna. När de gjorde sin studie konstaterade de att omfånget av personaliseringen är omfattande men att resultaten ofta blir ganska triviala. Det mest spännande i studien är hur det även i Googles fall handlar om att klassificera användare i enlighet med statistiskt skapade grupper, vilket resulterar i att människor inte enbart får ”the results they want (based in what Google knows about them for a fact), but also generates results that Google thinks might be good to users (or advertisers) thus more or less subtly pushing users to see the world according to criteria pre–defined by Google” (Feuz et al., 2011).

I en studie från 2013 visar forskare (Hannak et al., 2013) att det i genomsnitt är knappt 12 % av resultaten som är olika p.g.a. personalisering, men att det skiftar mycket beroende på sökterm och resultatens rangordning. I en senare studie visar forskare (Kliman-Silver et al., 2015) att geolokalisering har stor betydelse för

personalisering av webbsök, men att det även här finns stora skillnader beroende på sökterm. Det påminner oss om att geolokalisering, som även görs på en lokal nivå, för med sig en rad andra antaganden om t.ex. etnicitet, politisk inställning eller socio-ekonomiska förutsättningar på gruppnivå som tillkommer och kompletterar individbaserade faktorer (se även Tufekci, 2015b). Användare bidrar till att skapa dessa kategorier genom att förse Google med användardata samtidigt som kategorierna bidrar till att skapa användarprofiler. Relationen är alltså ömsesidig. Informationsbubblor kan skapas även i YouTube. O'Callaghan et al. (2015) visar t.ex. hur YouTubes rekommendationssystem tenderar att skapa ideologiska bubblor. Forskarna visar att algoritmer och tjänster, i detta fall YouTube, kan ses som de facto medskapare av ett ideologiskt innehåll eftersom rekommendationstjänsten har en så viktig funktion för användarna. Vi ser här hur en förståelse av internet som en offentlig sfär i betydelsen att alla har tillgång till information och att alla kan delta i ett åsiktspluralistiskt samtal stämmer dåligt överens med den fragmenterade verkligheten på internet.

Personalisering görs förstås också av Amazon och andra e-handlare. Dessa tjänster använder sig av algoritmer som gör att de produkter användarna får se matchar det som tjänsten tror att de är intresserad av, baserat på vad de har köpt tidigare, profiluppgifter och vad andra konsumenter har köpt. Ofta är det till konsumentens fördel, t.ex. när vi blir föreslagna vissa böcker (eller andra produkter) i Amazon baserat på vad vi köpt tidigare genom kollaborativ filtrering. Även musik-streaming-tjänsten Spotify använder personalisering tillsammans med kollaborativ filtrering för att öka relevans. Tjänsten använder sig av användarnas spellistor och "smakprofiler" för att förfina algoritmen som främst baseras på kollaborativ filtrering tillsammans med t.ex. genreinformation för att rekommendera musik till användarna. På det sättet tillför Spotify mänsklig "handpåläggning" till urvalet som har visat sig vara nödvändigt för rekommendation av musik. Andra tjänster (t.ex. Apple) använder faktiskt mänskliga kuratorer för att balansera rent algoritm-baserade rekommendationer som kan upplevas som väl generiska.[25] Hannak et al. (2014) visar också hur personaliseringen kan leda till att olika användare får olika priserbjudande för samma produkt, t.ex. hotell, vilket kan få negativa effekter för konsumentförtroendet.

Hur kan det offentliga samtalet stöttas?

Det offentliga samtalet kan stöttas på olika sätt och på olika nivåer. Vi vill särskilt lyfta fram tre viktiga aktörer: journalistik & medier, skola och bibliotek.

Journalistik & massmedier

Massmedier har en central funktion när det gäller att skapa ett öppet samtal, en offentlig sfär, i samhället. Ett sådant samtal är en nödvändig förutsättning för demokratin. Samtidigt visar vi i rapporten de svårigheter som finns för att upprätthålla det öppna samtalet. I *Medborgarna och medierna* (SOU 2015:94) ges flera

[25] http://www.theverge.com/2015/9/30/9416579/spotify-discover-weekly-online-music-curation-interview

exempel på hur algoritmer har kommit att påverka nyhetsinnehållet. Nyhetsartiklar kan skrivas med hjälp av algoritmer, tidningar kan individanpassa hur nyheterna sorteras, massmedias sökoptimering skapar synlighet i sökmotorer och sociala medier. En konsekvens är de informationsbubblor som leder till att den egna världsbilden tenderar att förstärkas. Strömbäck (2015, s. 361) skriver att en ”ökad digitalisering riskerar att leda till att stärka de centrifugala krafterna”, med vilket avses de krafter som polariserar ett samhälle.

Det finns en intressekonflikt mellan de kommersiella intressen som finns inbyggda i algoritmer och det demokratiska ansvar som många massmedier har. Människor tar till sig nyhetsmedier på ett alltmer fragmenterat sätt och de nyheter som skapar mycket uppmärksamhet i form av klick i sociala medier prioriteras av nyhetsjournalistiken framför nyheter av mer komplex karaktär (Strömbäck, 2014). Vi stödjer den analys som Strömbäck (2014) gör när han förespråkar stöd till kvalificerad och bred journalistik. Framför allt har public service en mycket viktig funktion att fylla. Det är av stor vikt att det ges resurser nog till denna att fatta redaktionella beslut som inte är grundade i en enkel tolkning av de dominanta algoritmernas logik. Public service, liksom annan kvalitetsjournalistik, bör arbeta med att hitta nya vägar för att vända på den algoritmiska synlighetsfällan och därmed göra det idag osynliga synligt. Att bli viral bör inte vara konspirationsteoriernas privilegium. Det är alltså viktigt att denna typ av journalistik når ut brett genom att synas i just sociala nätverkstjänster och sökmotorer.

Massmedier har också en viktig roll att fylla med att aktivt bedriva upplysningsarbete för att motverka exempelvis viral ryktesspridning. Viralgranskarna arbete med att synliggöra ryktesspridning är ett utmärkt exempel på detta liksom annan journalistik som syftar till att förmedla kunskaper om hur algoritmer påverkar oss. Public service har ett särskilt ansvar att producera utbildningsprogram som stödjer människor medie- och informationskunnighet (se nästa stycke). Här har inte minst Utbildningsradion en viktig roll. Lewis och Westlund (2014) diskuterar data journalism och lyfter särskilt fram behovet av att omvärdera journalistiska normer i ljuset av samspelet mellan big data och algoritmer. De efterlyser även ökad kunskap hos journalisterna om algoritmernas inbyggda värderingar och även ökade färdigheter att hantera data, kod och algoritmer för att överbrygga ”the skills gap” mellan journalister och teknikerna.

Skolan

En viktig aktör för en positiv utveckling för det offentliga samtalet är skolan som har potential att verka utjämnande avseende förmåga och kompetens att på ett kritiskt sätt söka och granska digital information. Grundskolans läroplan innehåller också många kunskapsmål som har med sökning och källkritik att göra, men som Sundin (2015) visar behandlar läroplanen sökning som en neutral teknik och de speciella förutsättningar för källkritik som sociala medier eller sökmotorer för med sig berörs i mindre utsträckning. Forskning visar på att skolan inte lyckas med sitt uppdrag att fungera som utjämnande på IT-området (Samuelsson, 2014). De förmågor och färdigheter som brukar lyftas fram på detta område är bl.a. digital kompetens, informationskompetens och mediekompetens. De senaste åren har begreppet medie- och informationskunnighet (MIK) vuxit fram som ett samlande begrepp, inte minst genom Unesco, och på svenska har de givits ut flera skrifter som diskuterar MIK

(Carlsson, 2013; Rivano Eckerdal & Sundin, 2014). MIK är det begrepp som vi föredrar och som också Medieutredningen har valt att använda (SOU: 2015:94, s. 60).

Den svenska skolan utgör idag ett stort laboratorium där det går att hitta såväl goda som mindre goda exempel på undervisning för att stödja elevers MIK (inklusive digital kompetens). Många skolor har en progressiv syn på teknikens möjligheter i undervisningen. Det är dock viktigt att försäkra sig om att mediaföretags sponsring av skolor inte påverkar bredden i medieutbudet och att MIK inte enbart blir en fråga om att lära sig använda specifika program. Det finns annars en risk att skolorna, trots goda föresatser, låser in användningen av internet i särskilda företag och deras algoritmer. Vi menar att undervisning i MIK även måste inkludera en kritisk medvetenhet om vad algoritmer för med sig, tillsammans med vissa kunskaper i programmering. Att kunna söka, kommunicera och producera information är viktigt (liksom en förståelse för andra sammanhang där algoritmer verkar i samhället), men för att det ska kunna göras på ett sätt som tillförsäkrar ett öppet samhälleligt samtal måste en teknisk förståelse av algoritmer vara grundad i en kulturell förståelse av teknik. För detta arbete krävs samarbete inom och utanför skolan och på olika nivåer. Att ställa kritiska frågor om vilka värderingar som finns inbyggda i sökmotorer, sociala nätverkstjänster eller andra algoritmbaserade system borde vara en självklar, men svår, utgångspunkt. För den sakens skull bör inte användning av exempelvis Google i skolan ses som fel. Det är genom att låta dessa tjänster vara en självklar del av skolan – utan att de ges en monopolposition – som de kan diskuteras och problematiseras. Inom skolan har lärare och skolbibliotekarier, men även annan undervisande personal så som fritidspedagoger, en viktig funktion. Därtill är externa aktörer som Webbstjärnan och Statens Medieråd viktiga.

Skolans läroplaner bör utvecklas så att de tydligare innefattar kunskaper om dagens informationsinfrastruktur. Behovet av att utveckla läroplaner avseende digitala kompetenser framgår också i *En digital agenda i människans tjänst: En ljusnande framtid är vår* (SOU 2014: 13, s. 189 ff.), men vi vill särskilt lyfta fram att en utveckling av läroplaner även måste innefatta en ökad kritisk förståelse av den underliggande informationsinfrastrukturen och inte minst av algoritmerna som styr synlighet. I Digitaliseringskommissionens delbetänkanden diskuteras inte sällan skolan med ett fokus på hur man ska åstadkomma ökad användning genom digital kompetens och ökad tillgänglighet till IT och internet. Det görs fr.a. i *Gör Sverige i framtiden: Digital kompetens* (SOU 2015:28) där just digital kompetens är i blickfånget. Vi vill understryka behovet av att, utöver god tillgänglighet samt digitala kompetenser att använda och utveckla digitala applikationer, även betona de kritiska perspektiven som vi menar är en väsentlig del av MIK. Utöver grund- och gymnasieskolans läroplaner bör utbildningar på universitetsnivå ses över så att dessa motsvarar de krav som det förändrade informationslandskapet för med sig. Det gäller inte minst utbildningsplaner (och undervisningsinnehåll) för universitetsutbildning till lärare, journalister och bibliotekarier, men det är också viktigt för andra professionella yrken såsom jurister, fritidspedagoger, kommunikatörer, kuratorer m.fl.

Vi vill även betona att MIK inte bör ses isolerad, utan som en integrerad del av den kunskap vi medborgare bör ha för att kunna ta aktiv del i samhället. Att isolera MIK från andra ämnen, exempelvis kunskaper i svenska eller samhällskunskap, kan vara

direkt farligt. Att ha kunskap om algoritmers betydelse för informationens tillgänglighet är bara användbart om den existerar tillsammans med annan kunskap och en allmänbildning. Det spelar mindre roll om människor har en kritisk förståelse av t.ex. Google om de inte har förmågan att tolka sökresultat i ljuset av en allmän kunskap om olika fenomen och deras historier i samhället. En medborgare som söker kunskap på nätet om exempelvis invandring måste kunna relatera de länkar som sökmotorn förser henne med till en grundläggande kunskap om just invandring och varför den förklaras på olika sätt. En person som vill leta upp information om invandringens samhällskonsekvenser eller historia på nätet måste kunna positionera de olika resultatens avsändare på en politisk skala och i samhällets politiska debatt för att kunna dra slutsatser i enlighet med demokratiska värderingar. En övertro på informationsinfrastrukturens möjlighet att enkom och utan ifrågasättande kunna fungera som människors neutrala, externa minne är problematisk.

Bibliotek

Om skolan ansvarar för den formella undervisningen avseende MIK har biblioteken en central funktion när det gäller tillgång till information och stöttande av medie- och informationskunnighet vid sidan om det formella utbildningssystemet. Medieutredningens första delbetänkande (SOU 2015:94) belyser betydelsen av samhälleliga institutioner som kan tillvarata det allmännas intressen. Delbetänkandets i huvudsak positiva grundton i fråga om ökad tillgänglighet balanseras genom att de nämner att ett fåtal globala aktörer har skaffat sig en kraftfull dominans på marknaden. Delbetänkandet lyfter härvidlag fram bibliotekens roll som tillgängliggörare av ”avgiftsfri kunskap” och som skapare av ”neutrala arenor med digitalt kunniga guider” (s. 16, 295f). Det handlar alltså om två nära relaterade roller: 1) tillhandahållandet av en informationsinfrastruktur som, i alla fall delvis, bygger på annan logik än den som diskuteras i denna rapport samt 2) stödja människors medie- och informationskunnighet.

I den förstnämnda rollen arbetar biblioteken sedan länge med att utveckla och tillgängliggöra informationsinfrastrukturen samt förse den med ett kvalitativt innehåll. Här vill vi särskilt lyfta fram behovet av att de offentligt finansierade biblioteken utvecklar alternativ till den inlåsning av information som de kommersiella aktörerna på nätet skapar. Det handlar både om en juridisk inlåsning (digital right management) och den inlåsningseffekt eller filterbubblor som kommersiella aktörers algoritmer skapar. Dessutom handlar det om att biblioteken ges förutsättningar och resurser att motstå den lockelse som finns i att presentera information utifrån den kommersiella logik som finns inbyggd i många algoritmer där popularitet premieras. Dessa är utvecklade för att öka försäljning av reklamplats eller produkter, vilket biblioteken inte gör.

I den ovan sistnämnda rollen lyfts bibliotekens pedagogiska funktion som digitala guider fram. I detta sammanhang vill vi lyfta fram att utveckla och tillgängliggöra informationsinfrastrukturen inte räcker. Det handlar också om att synliggöra infrastrukturen som i många sammanhang blir osynligare och osynligare. Bibliotekarier har sedan länge stöttat sina användare genom pedagogiska insatser och den rollen blir än viktigare framöver, gärna i samarbete med andra aktörer. Ett positivt exempel är den MIK-vecka som Statens medieråd arrangerade 2014 och 2015, bl.a. vid folkbibliotek, i syfte att stärka människors medie- och

informationskunnighet. Målet måste vara att öka kunskapen om algoritmernas logik i hela samhället, där kritiska perspektiv behandlas tillsammans med synliggörande av infrastrukturen. Folkbiblioteken är en icke-kommersiell och öppen mötesplats och det finns redan idag positiva exempel på utbildningsinsatser i den riktning som här föreslås. I Medieutredningens delbetänkande (SOU 2015:94, s. 297) lyfts även fram att den mångfald användare som folkbiblioteken har kan påverka sökresultaten vid användning av sökmotorer på ett positivt sätt genom att motverka personalisering av sökresultaten. Vi vill också särskilt lyfta fram skolbibliotekariernas viktiga roll som kunskapsmässigt och pedagogiskt stöd till lärare i arbetet med MIK.

I Digitaliseringskommissionens delbetänkande *Gör Sverige i framtiden: Digital kompetens* (SOU: 2015:28) föreslås digitala servicecenter i kommunal regi som stöd för invånarna att utnyttja grundläggande samhällstjänster som flyttat ut till internet. Vi föreslår att sådana servicecenter knyts till folkbibliotekens redan existerande infrastruktur och inkluderar även lärandeaspekter. Folkbiblioteken har ett högt förtroende bland medborgarna och utgör redan idag ett stöd för dessa frågor för många människor.

Avslutning

Algoritmer är en del av vår vardag och de underlättar våra liv samtidigt som de också skapar begränsningar. När t.ex. nyheter, kunskap och kultur kommuniceras digitalt kommer frågor som sökbarhet och delningsbarhet i förgrunden. Vem som helst kan idag publicera nästan vad som helst, men det innebär inte nödvändigtvis att det som publiceras återfinns av andra. Den andra sidan av myntet handlar om människors förmågor att bemästra algoritmbaserade system på ett sådant sätt att potentialen utnyttjas samtidigt som vi förstår vilket inflyttande algoritmer har och vilka beslut som baseras på algoritmer. I rapporten redogör vi framförallt för hur algoritmer påverkar samhällets demokratiska förutsättningar samt, i någon mån, hur vi kan påverka utvecklingen i en riktning som gynnar det öppna samtalet.

Det krävs en balansgång mellan att bejaka en utveckling som på många sätt är positiv samtidigt som vi inte döljer de negativa konsekvenser som finns. Till viss del kan, även om det inte är i fokus för rapporten, det sistnämnda göras genom regleringar kring återbruk av användardata och motverkan mot monopol. Därtill behövs det aktiva offentliga aktörer, såsom bibliotek, public service och skola. Det finns ingen algoritmernas logik, men däremot finns det en algoritmernas kommersiella logik som inte nödvändigtvis behöver vara den enda logiken, utan en algoritm kan också tänkas vara uttryck för andra värderingar. Slutligen behövs en kraftfull kompetensutveckling för både medborgare i allmänhet och för ett flertal professionella yrkesgrupper – exempelvis lärare, bibliotekarier, journalister och fritidspedagoger – i synnerhet. Det handlar om en ökad medie- och informationskunnighet som bl.a. tar sikte på hur makten förflyttats i samhället genom en förändrad informationsinfrastruktur, vilket är långt ifrån enbart en teknisk fråga.

Det är en stor utmaning att låta en kritisk diskussion inte bli en generationsfråga där en äldre generation ”varnar” den efterkommande. Samtidigt är det centralt att ha denna kritiska diskussion. I dagens samhälle fattas många beslut med hjälp av

algoritmer som berör de flesta områden, från medier, kultur och utbildning till arbetslivet, trafik, rättsväsendet, vård, försäkring och marknad. Områden och frågor där algoritmer utgör grunden för beslut kommer bara att bli fler. Det handlar därför om att synliggöra algoritmer och deras inbyggda värderingar. Algoritmer är för centrala för att förbli osynliga. Först då kan vi ha en allvarlig diskussion om algoritmernas etik och en sådan diskussion är oundviklig om vi vill förnya kalibreringssystemet för det demokratiska samhällets kunskap och värderingar.

Det finns en rad viktiga frågor inför framtiden. Dessa är av teknisk, social och kulturell art. Det krävs en specifik kunskap om hur algoritmer samspelar med en lång rad sektorer i samhället: kultur, vård, forskning, utbildning, arbetsliv, affärsvärlden, rättsväsendet, journalistik osv. Denna rapport har enbart kunna beröra dessa på ett mer generellt plan. Utifrån vår erfarenhet, som måste kompletteras med andra erfarenheter, vill vi särskilt lyfta fram nedanstående frågekomplex som bör utforskas för att bättre förstå hur samhället formas av de nuvarande algoritmerna, liksom det omvända, hur algoritmer formas av samhället:

- *Centrala aktörer:* Det finns ett behov av starka offentliga, icke-kommersiella, aktörer som bygger sin verksamhet på andra principer än de kommersiella, både för utbildning och för att skapa alternativa vägar att kommunicera och producera information, kunskap och kultur. I rapporten har särskilt public service, skola och bibliotek lyfts fram. Samtidigt verkar dessa aktörer i en kommersiell omvärld. Hur påverkas de av att delvis anpassa sin verksamhet efter hur algoritmerna i dagens dominerande online-tjänster fungerar för att få synlighet? Hur ska offentliga kultur- och kunskapsaktörer såsom skola, bibliotek och public service både kunna vara en alternativ och kritisk röst samtidigt som dessa offentliga aktörer även behöver vara synliga och aktiva i kommersiella tjänster? Vilka möjligheter finns för offentligt finansierade aktörer och kulturproducenter att skapa och tillgängliggöra innehåll som fungerar inom ett algoritm-styrt nät utan att reproducera algoritmernas kommersiella logik?

- *Källkritik, sökkritik och algoritmernas synlighet:* Digitaliseringen har lett till ett ökat intresse för källkritik, en metod som utvecklats i historievetenskapen. När vem som helst kan publicera sig har ansvaret för att granska trovärdigheten i många fall flyttats över från utgivaren till läsaren. Frågan hur denna traditionella metod kan anpassas till dagens nätkultur är av stor vikt i en delningskultur där ideologiskt färgade påståenden sprids snabbt och där rent av konspirationsteorier får stort utrymme. Samtidigt bör källkritik ses tillsammans med vad vi benämner sökkritik, dvs. ett kritiskt förhållningssätt till inte bara källor på nätet utan också till vad som gör att vi faktiskt tar del av just dessa källor? Hur blir algoritmerna synliga som kulturella artefakter? Det faller inom området medie- och informationskunnighet. Många människor har en viss generell kunskap om algoritmers betydelse, men samtidigt har denna kunskap liten betydelse för hur människor faktiskt agerar. Det kan jämföras med hur kunskaper om miljöförstöring och hälsorisker har svårt att översättas från abstrakt kunskap till konkreta handlingsmönster. Det är därför även en stor utmaning att gå från

folkbildning till handling och en viktig fråga för framtiden är hur det ska göras.

- *Föreställningar om big data:* I idén om algoritmer ingår idéer om att allting kan brytas ner och beskrivas i form av diskreta, neutrala data-objekt som sedan kan föras samman igen för att tolkas efter tydliga faktorer och kriterier. Det finns en tendens i samhället att vilja gå bortom representation (Andrejevic, 2013) och fakta, aggregerade av (ofta användargenererad) big data, anses vara lösningen och traditionella kunskapsformer ses inte sällan som förlegade. Tilltron till att fakta kan finnas som rent neutral, icke-ideologiskt, material i form av data är stor. Vad kan man egentligen veta och fatta beslut om baserad på dessa data-objekt? Vad är det för kunskap som prioriteras när big data lyfts fram som uttryck för "hur det är"? Hur kan vi förstå den förskjutning i synen på kunskap som de nya tjänsterna för att producera, kommunicera och använda kunskap för med sig? Vad får den för konsekvenser för den traditionella kunskapens auktoritet? Vilka etiska implikationer härrör från att vi lämnar avgörande beslut till algoritmer?

Referenser

Abel, F., Gao, Q., Houben, G.-J. & Tao, K. (2011). Analyzing user modeling on Twitter for personalized news recommendations. In S. Carberry, S. Weibelzahl, A. Micarelli & G. Semeraro (eds.) *User Modeling, Adaption and Personalization*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 1-12.

Altaweel I., Good N. & Hoofnagle C. (2015). Web privacy census. *Technology Science*. December 15, 2015. http://techscience.org/a/2015121502 [2016-01-12].

Andrejevic, M. (2013). *Infoglut: How too Much Information is Changing the Way We Think and Know*. New York: Routledge.

Bakshy, E., Messing, S. & Adamic, L. A, (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130-1132. DOI: 10.1126/science.aaa1160

Bolin, G. & Andersson Schwarz, J. (2015). Heuristics of the algorithm: Big Data, user interpretation and institutional translation, 2(2). DOI: 10.1177/2053951715608406

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

Bäck, E., Bäck, H. & Gustafsson, N. (2014). *Ungas politiska deltagande: Nya former och aktivitet genom sociala medier?* Rapport för Demokratiutredningen.

Carlsson, U. (red.) (2013). *Medie- och informationskunnighet i nätverkssamhället: Skolan och demokratin*. Göteborg: Göteborgs universitet, NORDICOM.

Christin, A.; Rosenblat, A. & boyd, d. (2015). Courts and predictive algorithms. *Data & Civil Rights: A New Era of Policing and Justice*. (workshop primer) Oct 27, 2015. http://www.datacivilrights.org/pubs/2015-1027/Courts_and_Predictive_Algorithms.pdf

Cormen, T.H., Leiserson, C.E., Rivest, R. L. & Stein, C. (2009). *Introduction to algorithms*. 3rd ed. MIT Press and McGrawhill.

DeMers, J. (2015-09-03). Is the Google knowledge Graph killing Wikipedia? *Forbes*. http://www.forbes.com/sites/jaysondemers/2015/09/03/is-the-google-knowledge-graph-killing-wikipedia/ [2016-01-12].

van Dijck, J. (2010). Search engines and the production of academic knowledge. *International Journal of Cultural Studies*, 13(6), 574-592. DOI: 10.1177/1367877910376582

van Dijck, J. (2013). *The Culture of Connectivity: A Critical History of Social Media*. Oxford: Oxford University Press.

Dong, X. L., Gabrilovich, E., Murphy, K., Dang, V., Horn, W., Lugaresi, C., Sun, S. & Zhang, W. (2015). *Knowledge-Based Trust: Estimating the Trustworthiness of Web Sources*. arXiv:1502.03519 [cs.DB] [2016-01-12]

Feuz, M., Fuller, M. & Stalder, F. (2011). Personal web searching in the age of semantic capitalism: Diagnosing the mechanisms of personalisation. *First Monday*, 16(2). http://firstmonday.org/ojs/index.php/fm/article/view/3344/2766 [2016-01-12].

Gillespie, T. (in press 2016). #trendingistrending: When algorithms become culture. I R. Seyfert & J. Roberge, *Algorithmic Cultures: Essays on Meaning, Performance and New Technologies*. Routledge.

Haider, J. (2014). Taking the environment online: issue and link networks surrounding personal green living blogs. *Online Information Review*. 38(2), 248-264.

Halavais, A. (2009). *Search Engine Society*. Cambridge: Polity Press.

Hallinan, B. & Striphas, T. (2014). Recommended for you: The Netflix Prize and the production of algorithmic culture. *New Media & Society*, 18 (1), 117-137.

Hannak, A., Sapiezynski, P., Kakhki, A. M., Krishnamurthy, B., Lazer, D., Mislove, A. & Wilson, C. (2013). Measuring personalization of web search. *WWW '13 Proceedings of the 22nd international conference on World Wide Web*, s. 527-538. http://personalization.ccs.neu.edu/papers/web_search.pdf [2016-01-22].

Hannak, A., Soeller, G., Lazer, D., Mislove, A. & Wilson, C. (2014). Measuring price discrimination and steering on e-commerce web sites. *IMC 14, November 5-7, Vancouver, BC, Canada.* http://dl.acm.org/citation.cfm?id=2663744 [2016-01-12].

Hillis, K., Petit, M. & Jarrett, K. (2013). *Google and the Culture of Search*. New York: Routledge.

Höchstötter, N. & Lewandowski, D. (2009). What users' see: Structures in search engine result pages. *Information Sciences*, 179(12), 1796-1812.

Introna, L. D. (2015). Algorithms, governance, and governmentality: On governing academic writing. *Science, Technology and Human Values*, 41(1), 17-49. DOI: 10.1177/0162243915587360

Karppi, T., & Crawford, K. (2015). Social Media, financial algorithms and the hack crash. *Theory, Culture & Society*. DOI:10.1177/0263276415583139

Kliman-Silver, C., Hannak, A., Lazer, D., Wilson, C. & Mislove, A. (2015). Location, location, location: The impact of golocation on web search personalization. *IMC '15 Proceedings of the 2015 ACM Conference on Internet Measurement Conference*, s. 121-127. http://www.ccs.neu.edu/home/cbw/pdf/location-imc15.pdf [2016-01-12].

Lewandowski, D. & Spree, U. (2011). Ranking of Wikipedia articles in search engines revisited: Fair ranking for reasonable quality? *Journal of the American Society for Information Science and Technology*, 62(1), 117-132.

Lewis, S. C., & Westlund, O. (2015). Big data and journalism. *Digital Journalism*, 3(3), 447–466. DOI:10.1080/21670811.2014.976418

Ma, N., Guan, J., & Zhao, Y. (2008). Bringing PageRank to the citation analysis. *Information Processing and Management*, 44, 800-810. DOI:10.1016/j.ipm.2007.06.006

Mager, A. (2012). Algorithmic ideology: How capitalist society shapes search engines. *Information, Communication & Society*, 15(5), 769-787. DOI:10.1080/1369118X.2012.676056

Manson, M. (2014). *Facebook zero: Consideing life after the demise of organic research.* https://social.ogilvy.com/facebook-zero-considering-life-after-the-demise-of-organic-reach/ [2016-01-08].

O'Callaghan, D., Greene, D., Conway, M., Carth, J. & Cunningham, P. (2015). Down the (white) rabbit Hole: The extreme right and online recommender system. *Social Science Computer Review*, 33(4), 459-478. DOI: 10.1177/0894439314555329

Otero, M. (2014). *The Real 10 Algorithms that Dominate our World: Maybe a Little More than 10*. https://medium.com/@_marcos_otero/the-real-10-algorithms-that-dominate-our-world-e95fa9f16c04#.ckpyf4fdl [2016-01-12].

Page, L., Brin, S., Motwani, R. & Winograd, T. (1999). *The PageRank Citation Ranking: Bringing Order to the Web. Technical Report*. Stanford InfoLab.

Papacharissi, Z. A. (2010). *A Private Sphere: Democracy in a Digital Age*. Cambridge: Polity Press.

Pariser, E. (2011). *The Filter Bubble: What the Internet Is Hiding from You*. New York: Penguin Press.

Rivano Eckerdal, J. & Sundin, O. (red.) (2014). *Medie- och informationskunnighet: En forskningsantologi*. Stockholm: Svensk Biblioteksförening.

Rosenblat, A. & Stark, L. (2015), *Uber's Drivers: Information, Asymmetries and Control in Dynamic Work*. Data and Society Research Institute.

Samuelsson, U. (2014). *Digital (o)jämlikhet IKT-användning i skolan och elevers tekniska kapital*. Högskolan i Jönköping. Diss.

Schwartz Andersson, J., Hammarlund, J., Di Grado, S. & Kjellberg, M. (2015). *Opinioner och ffentligheter online*. Stockholm: Internetfonden.

Seaver, N. (2013). Knowing algorithms. *Media in Transition 8*: 1–12. Unpublished paper.

Silverman, C. (2015). *Lies, damn lies, and viral content: How news ebsites spread (and debunk) online rumors, unverified claims, and misinformation*. Tow Center for Digital Journalism. http://towcenter.org/wp-content/uploads/2015/02/LiesDamnLies_Silverman_TowCenter.pdf [2016-01-12]

Singer, R. (2009). Netflix Spilled Your Brokeback Mountain Secret, Lawsuit Claims. *Wired*. http://www.wired.com/2009/12/netflix-privacy-lawsuit/ [2016-01-12]

SOU (2013:31). *En digital agenda i människans tjänst: Sveriges digitala ekosystem, dess aktörer och drivkrafter*. Delbetänkande av Digitaliseringskommissionen. Fritzes: Stockholm.

SOU (2014:13). *En digital agenda i människans tjänst: En ljusnande framtid kan bli vår*. Delbetänkande från Digitaliseringskommissionen. Fritzes: Stockholm.

SOU (2015:28). *Gör Sverige i framtiden: Digital kompetens*. Delbetänkande från Digitaliseringskommissionen. Wolters Kluwer: Stockholm.

SOU (2015:65). *Om Sverige i framtiden. En antologi om digitaliseringens möjligheter*. Wolters Kluwer: Stockholm.

SOU (2015:94). *Medieborgarna & medierna: En digital värld av rättigheter, skyldigheter – möjligheter och ansvar*. Delbetänkande av Medieutredningen. Wolters Kluwer: Stockholm.

Stakston, B. (2013). *Polisens närvaro i sociala medier*. JMW kommunikation. https://polisen.se/Global/www%20och%20Intrapolis/Rapporter-utredningar/01%20Polisen%20nationellt/Ovriga%20rapporter-utredningar/Utvardering_Polisens_narvaro_i_sociala_medier.pdf [2016-01-31]

Strömbäck, J. (2014). *Demokratin och det förändrade medielandskapet: Mot ökade kunskapsklyftor och deltagandeklyftor?* Rapport för Demokratiutredningen.

Strömbäck, J. (2015). SOU (2015:65). Framtidens medielandskap, demokratin och den sociala sammanhållningen. I *Om Sverige i framtiden: En antologi om digitaliseringens möjligheter* (SOU 2015:65). Delbetänkande av Medieutredningen. Fritzes: Stockholm, s. 341-367.

Sundin, O. (2014). Akademisk vardag på en digital referensmarknad. *Kulturella perspektiv*, 23, 49-58.

Sundin, O. (2015). Invisible Search: Information Literacy in the Swedish curriculum for compulsory schools. *Nordic Journal of Digital Literacy*, 10, 193-209.

Sundin, O. & Haider, J. (2013) The networked life of professional encyclopaedias: Quantification, tradition, and trustworthiness. *First Monday*, 18(6). http://firstmonday.org/ojs/index.php/fm/article/view/4383/3686 [2016-01-12]. DOI:10.5210/fm.v18i6.4383

SR (2014-07-24). *Testerna på Facebook lurar dig.* http://sverigesradio.se/sida/artikel.aspx?programid=125&artikel=5921760 [2016-01-22]

Tufekci, Z. (2015a). *How Facebook's Algorithm Suppresses Content Diversity (Modestly) and How the Newsfeed Rules Your Clicks.* https://medium.com/message/how-facebook-s-algorithm-suppresses-content-diversity-modestly-how-the-newsfeed-rules-the-clicks-b5f8a4bb7bab#.582z5j59o [2016-01-12].

Tufekci, Z. (2015b). Algorithmic harms beyond Facebook and Google: Emergent challenges of computational agency. *Colorado Technology Law Journal.* http://ctlj.colorado.edu/wp-content/uploads/2015/08/Tufekci-final.pdf [2016-02-04]

Westlund, O., & Lewis, S. C. (2014). Agents of media innovations: Actors, actants, and audiences. *The Journal of Media Innovations*, 2, 10–35. DOI:10.5617/jmi.v1i2.856

Wikipedia (2016-01-08). *Facebook.* https://sv.wikipedia.org/w/index.php?title=Facebook&oldid=32801600 [2016-01-31]