

intellecta corporate

#TWITTERCENSUS 2012
@intellecta

Sammanfattning

Svenskspråkiga twitter ökar snabbt. Av de 299 000 konton vi kunnat identifiera har mer än hälften registrerats det senaste året. Från majoriteten av kontona, 171 000, har det skickats minst ett tweet den senaste månaden. De mest aktiva – de som skrivit minst ett tweet om dagen – ökar också och är i dag 53 000.

Median och medelantalet följare, konton som följs och antalet statusuppdateringar har alla ökat sen förra Twittercensus (mätning december 2010).

Nya grupper har börjat twittra och den tidigare starka medie- och journalistnärvaron håller på att spädas ut. Bland de grupper som ökar de senaste åren återfinns sportkluster, kändiskluster såväl som olika geografiska kluster – åtminstone bland de mer aktiva twittrarna.

Bakgrund

2011 presenterade Intellecta Corporate den första systematiska genomgången av svenskspråkiga twitterkonton. Tidigare fanns estimat på antalet konton baserade på svar i enkätundersökningar. Något mindre försök med manuell datainsamling hade också gjorts. Resultaten från dessa olika mätningar skiljde sig mycket åt. Därför beslutade vi oss för att genomföra en undersökning med målet att kunna presentera en någorlunda tillförlitlig siffra på antalet svenskspråkiga twitterkonton. Mätningen genomfördes under december 2010 och resultaten presenterades i början av 2011.

I slutet av 2012 presenterade företaget Aitellu en mätning (Svenska Twitter¹) inte helt olik Twittercensus. Den exakta metoden skiljde sig dock något från den som används i Twittercensus vilket gör att siffrorna inte är direkt jämförbara. För att få siffror som är jämförbara bestämde vi oss därför för att göra en ny mätning, med samma metod som i den första Twittercensus.

¹ <http://aitellu.com/svenskatwitter2011.pdf>

Metod

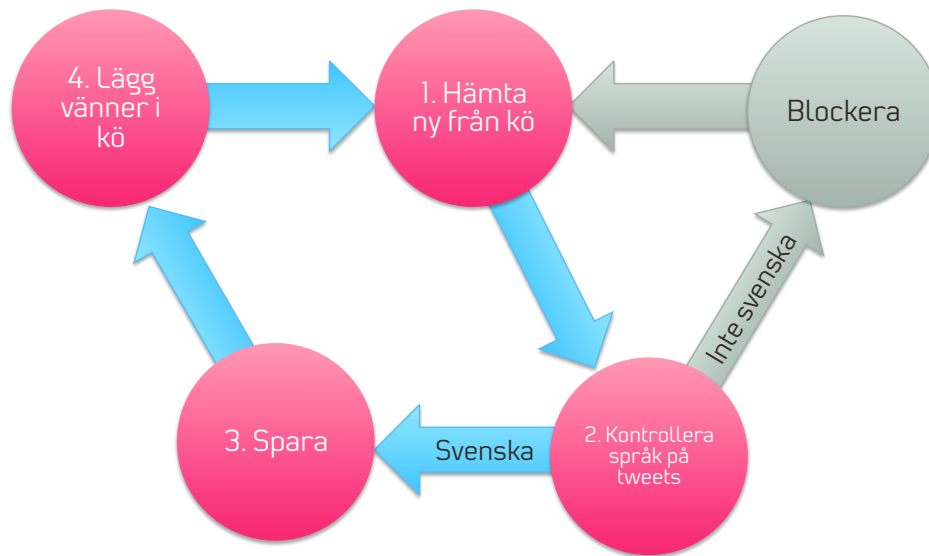
På ett twitterkonto behöver det inte anges om kontot är svenskt, registrerat i Sverige eller om innehållet i meddelanden är på svenska. Användaren kan manuellt ange en geografisk plats, och det finns en "språkinställning" för användaren som bland annat styr vilket språk twitter.com ska visas på. Standardinställningen har länge varit engelska och det är bara en minoritet av de konton vi identifierat som svenskspråkiga som faktiskt har svensk språkinställning på twitter. Vad gäller "plats" är den ett fritextfält och det går utmärkt att fylla med vad som helst: "Södermalm" går lika bra som "Stockholm" eller för den delen "hemma" eller "universum". Cirka hälften av de konton vi har identifierat som svenska har inte angivit någon plats.

Att använda angiven ort eller språkinställning är alltså i sig själv inte en tillräcklig metod för att hitta alla svenskspråkiga twittrare. Istället valde vi en metod som bygger på språkigenkänning. Metoden har visat sig fungera bra vid språkidentifiering i stora datamängder, och anses vara tolerant mot felaktigheter.²

För varje twitterkonto vi identifierade som svenskspråkigt hämtades alla konton som följer eller följs av det konto. För var och en av dessa analyserades³ de 100 senaste tweetsen med avseende på språk (om kontot hade skrivit färre än 100 tweets analyserades de tweets som fanns tillgängliga). De konton som twittrade på svenska fick sedermera sina följare analyserade på samma sätt. Om kontot i fråga inte hade tweets på svenska flyttades det till en blockeringslista för att undvika att analyserades igen.

² <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.53.9367>

³ Språkanalysen gjordes med N-gram-baserad textkategorisering. Cavnar & Trenkle (1994): "N-Gram-Based Text Categorization".



Mätningen gjordes under 25 mars och 14 april. Totalt analyserades över 2,5 miljoner twitterkonton, och för varje konto lästes de senaste 100 tweetsen in. De som valt att skydda sina statusuppdateringar kunde, av naturliga skäl, inte analyseras. Inte heller de som registrerat ett konto men inte gjort några uppdateringar analyserades. Av de skannade kontona hade 12 procent inte skrivit några meddelanden alls och 4 procent har valt att vara skyddade. Hur många av dessa som är svenskar går dock inte att säga, men vart och ett av dessa konton följdes eller följde åtminstone ett svenskspråkigt konto.

Konton som inte följs av någon, och inte heller följer någon har inte analyserats.

Metoden som har använts är samma som vid mätningen 2010/2011. De förändringar som har gjorts har varit av optimeringsskäl, och vissa delar av insamlingskoden har förändrats för att kunna hantera den större datamängden. Även vissa kodändringar för att kunna möta förändringar i Twitters API har genomförts. I årets datainsamling har även alla relationer mellan konton sparats vilket möjliggör nya former av nätverksanalyser.

Ett problem med den använda metoden är att om en grupp konton på twitter bara följer varandra och inte heller följs av någon annan, så upptäcks inte dessa konton. Det kan alltså förekomma mindre eller större grupper av konton som inte har någon kontakt med de konton som finns med i denna datamängd. Storleken eller omfattningen av sådana grupper är okänd, men efter förra Twittercensus gjordes ett stickprov där alla som skrev under hashtaggen #melfest (melodifestivalen) under några timmar kontrollerades automatiskt. Inga ytterligare konton hittades då.

Resultaten

Hur många twittrar?

Svaret ger sig inte så enkelt: först måste vi svara på vem som är en twittrare. Om man räknar alla konton på twitter som skriver på svenska och som gjort minst en uppdatering så blir siffran 299 000. Detta är alltså bruttoantalet svenskspråkiga twitterkonton⁴. Aktivitetsgraden på dessa, och när de senaste uppdaterade sitt konto, varierar.

Vid förra mätningen var motsvarande siffra 91 000 – en mer än tredubbling av antalet konton kan alltså noteras, även om siffran från 2011 kan behöva revideras upp något (se nedan).

Antalet aktiva

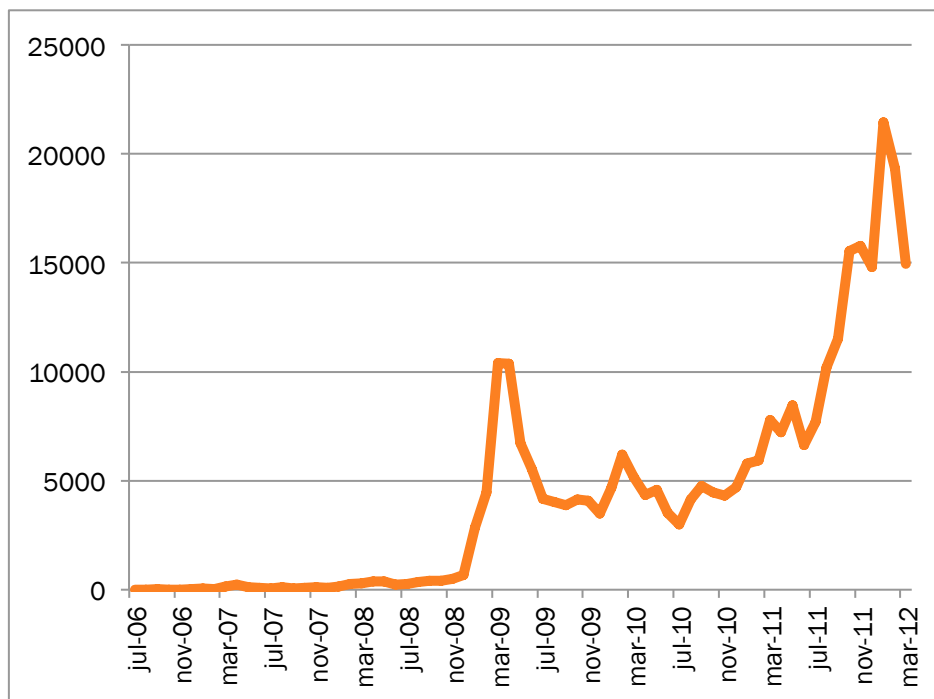
Om vi definierar en aktiv twittrare som någon som skrivit **ett** tweet under de senaste 30 dagarna så blir siffran 171 000 (2010 var motsvarande siffra 37 000). Vi vet inte i vilken utsträckning en användare *läser* tweets. Denna siffra är inte direkt jämförbar med exempelvis Facebooks användarstatistik som presenterar antal *inloggade* under den senaste månaden.

Grupp som skrivit minst ett tweet den senaste månaden refereras nedan till som "aktiva twittrare". Om man tillämpar ett ännu hårdare kriterium, att man ska ha twittrat i *snitt* minst en gång per dag blir antalet 53 000 (mot 2010 års siffra 11 000). Dessa refereras nedan till som de "mycket aktiva".

Andelarna aktiva och mycket aktiva är högre nu än vid förra mätningen. Detta kan möjligen förklaras av att twitter i dag till större del består av mer nyregistrerade.

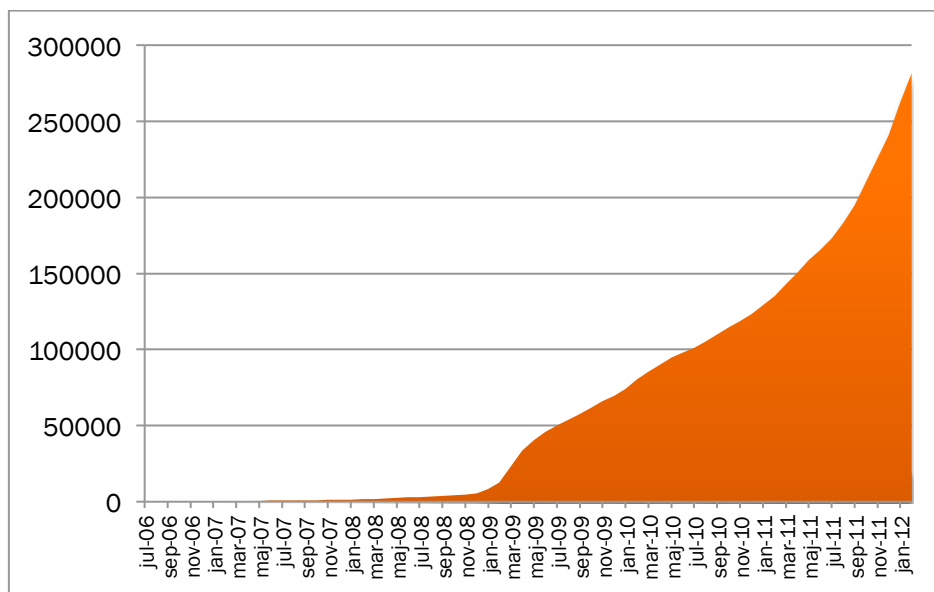
⁴ För diskussion om vilka som exkluderas från denna siffra, se metodavsnittet.

Förändring över tid



Figur 1 registreringar per månad

Som synes är registreringstakten, och alltså tillväxttakten, just nu väldigt hög. Det totala antalet konton växer väldigt snabbt och på ett år har antalet konton dubblats.

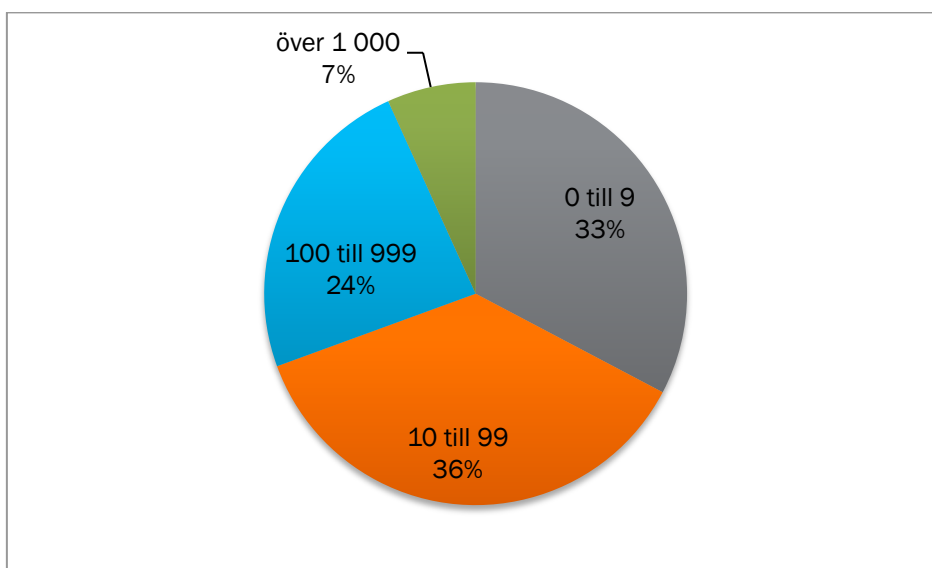


Figur 2 antal konton

Notera att antalet twitterkonton vid december 2010 är högre av vad som identifierades vid förra Twittercensus. Denna skillnad beror på att konton som registrerats innan mätningen inte fångades upp vid förra mättillfället. Det kan bero på att de inte var aktiva (inte hade skrivit något), att de var skyddade, att de skrev på annat språk eller att inte de eller någon de följde eller följdes av var länkad till något av de andra identifierade kontona.

Hur mycket twittrar vi?

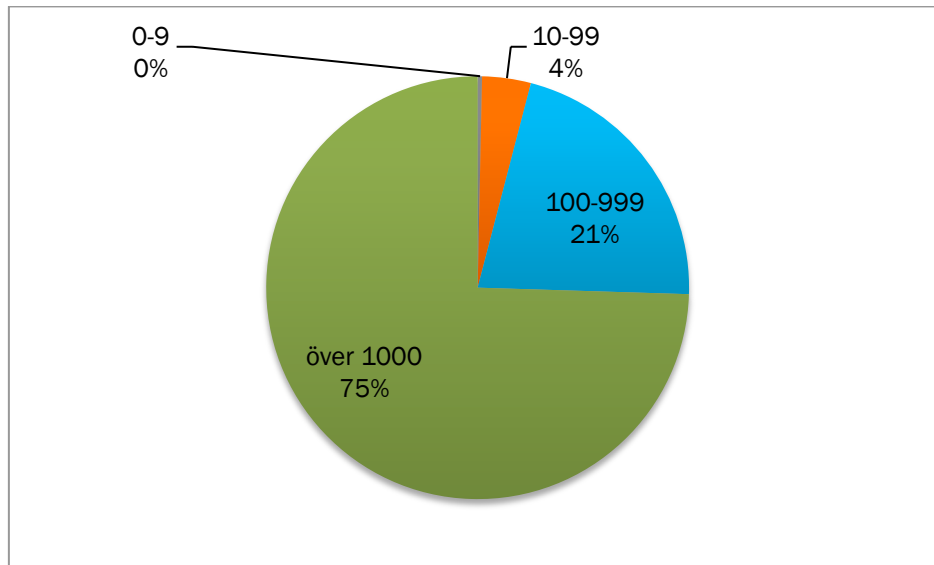
Antalet tweets olika konton skrivit fördelar sig ojämnt. Figur 3 visar procentandelen av konto som har skrivit ett visst total antal tweets.



Figur 3 Andelen som har skrivit ett visst antal tweets

Grafen ska läsas som att 33 procent av användarna har skrivit mellan ett och nio tweets. Siffrorna är i stort sett oförändrade sedan 2010.

Grafen för hur stor andel av den totala mängden tweets som respektive grupp står för (Figur 4) visar tydligt hur fördelningen är: de 7 % av kontona (ovan) som skrivit fler än 1000 uppdateringar står för 75 % av antalet statusuppdateringar.

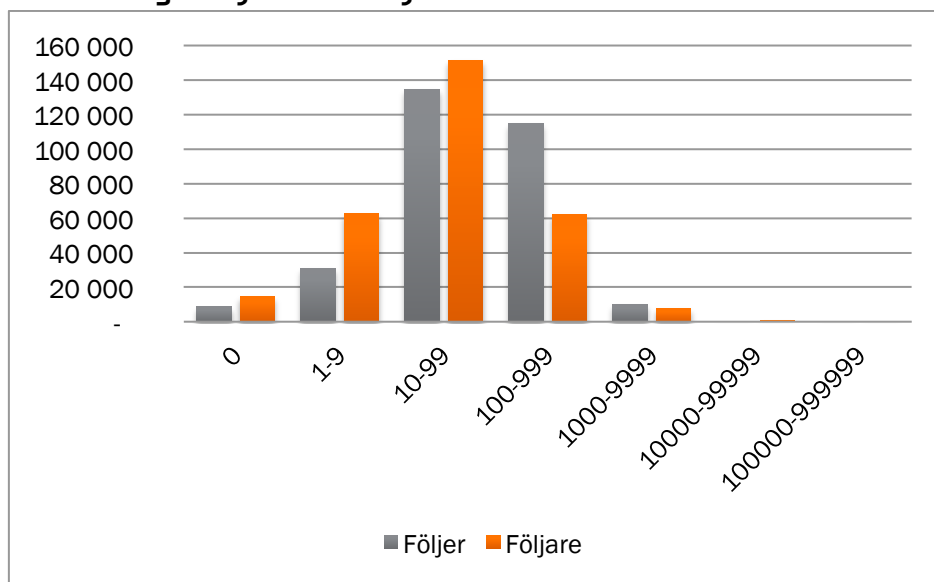


Figur 4 Andelen av tweets/antal statusuppdateringar

Dessa siffror har inte nämnvärt förändrats sedan 2010.

Snittantalet uppdateringar är 369 för alla, medan medianvärdet ligger på 28 – en ökning från 274 respektive 23 år 2010.

Hur många följer och följs?



Den stora mängden följer alltså mindre än 100 personer.

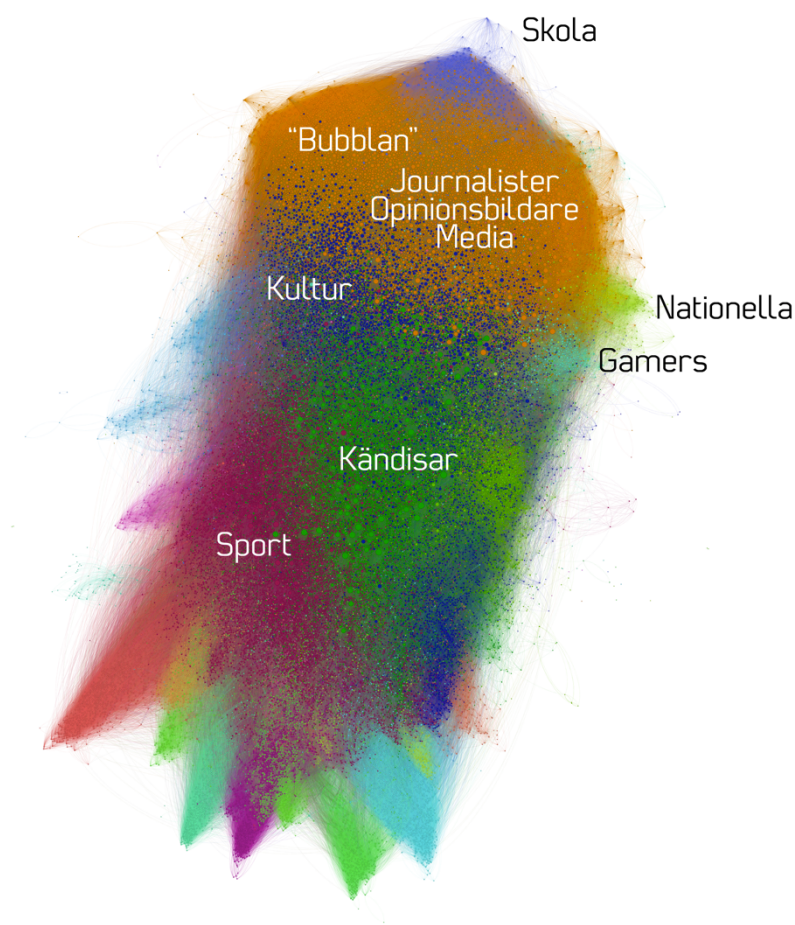
Snittantalet ett konto följer är 63, medianen är 23 (2010: 47 resp. 11).

Snittantalet följare är 62 med en median på 10 (2010: 47 resp. 8). Notera den stora förändringen av medianen på antalet ett konto följer, en ökning från 11 till 47.

Kluster

En klusteranalys har gjorts på de mest aktiva twitterkontona⁵. En sökbar och interaktiv visualisering/karta finns på twittercensus.se.

Ett kluster är en grupp av konton som delar egenskaper – de följer och följs av samma konton i högre utsträckning än andra konton. Tydliga kluster träder fram när man tittar på de mest aktiva twitterkontona, men att tolka innebörden av dem (vad de olika klustren "är") är en subjektiv övning. Att tillhöra ett visst kluster innebär inte att kontona delar andra egenskaper.



En särskild bloggpost om olika kluster kommer att publiceras på twittercensus.se.

⁵ För beskrivning av metoden, se <http://arxiv.org/abs/0803.0476>

De mycket aktiva

De mycket uppvisar en rad karakteristika. De följer fler, följs av fler och skriver mycket mer än snittanvändaren.

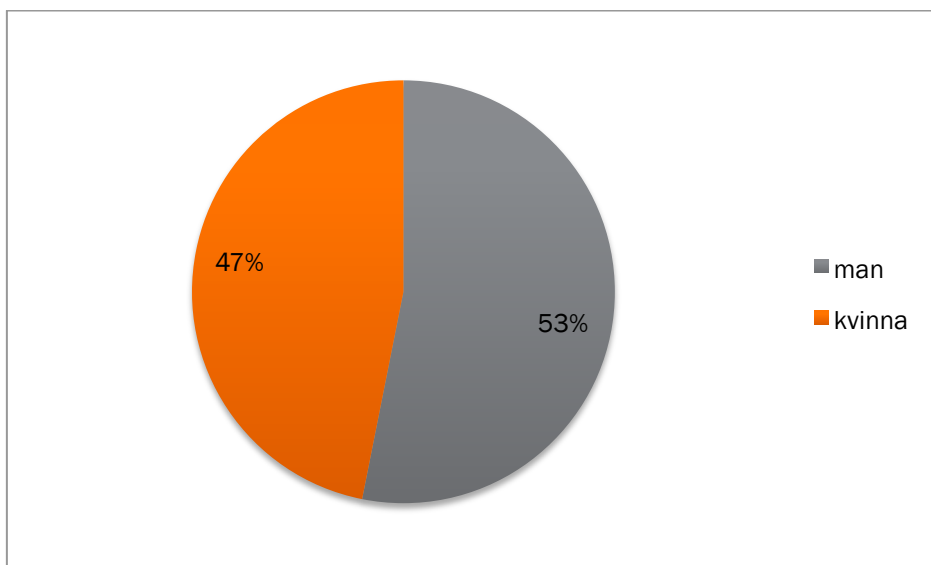
	Följer	Följs av	Medelantal uppdateringar
Mycket aktiva (n=52887)	158	216	1694
Aktiva (n=163547)	93	99	631
Alla (n=91316)	63	62	369

Tabell 1 Medelvärden på följer, följs av och uppdateringar

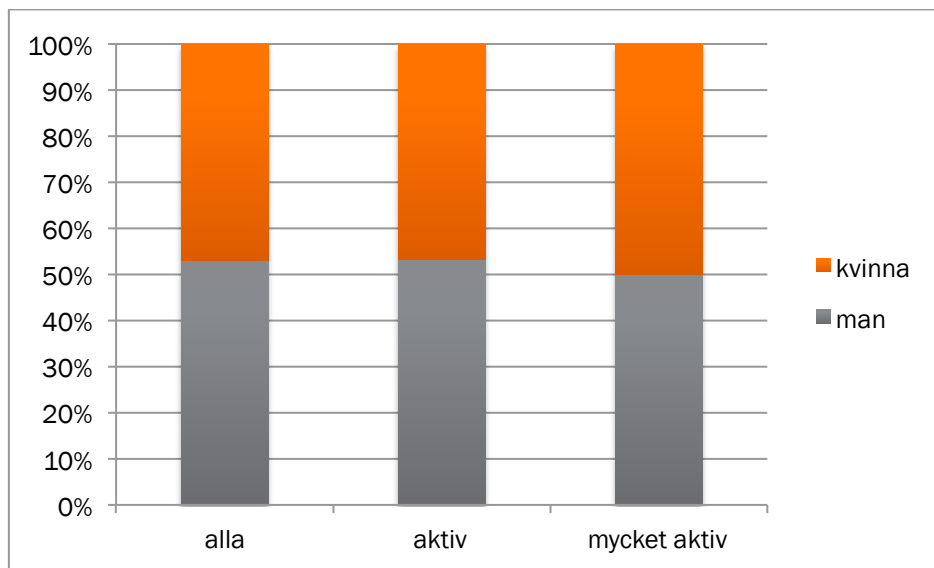
Kön

Genom att matcha de 500 vanligaste namnen på män respektive kvinnor mot det angivna namnet på Twitter kan ett estimat om kön göras.

Observera att en 27 % av twitterkontona inte kunnat analyseras. Detta kan bero, bland annat, på att konton inte anger namn, att namnet inte finns bland de vanligaste namnen eller att kontot är ett företags- eller organisationskonto.



I grupperna aktiva respektive mycket aktiva är förhållandena snarlika.



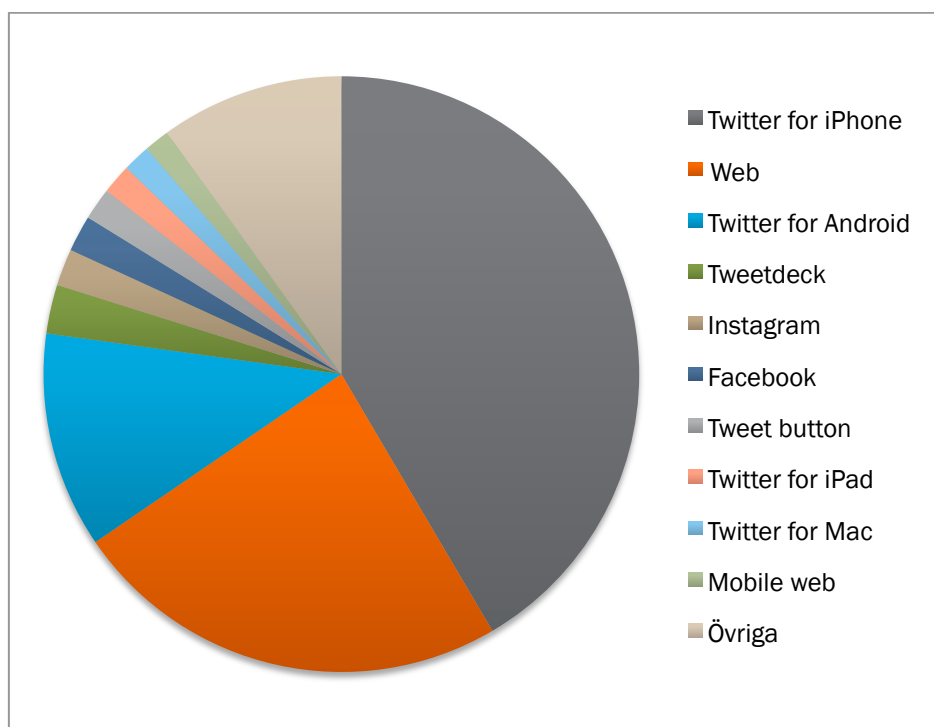
Figur 5 fördelning antal konton, per kön

Om man tittar på aktivitet visar sig en viss skillnad: män twittrar något mer (medeltal bland de aktiva 585 resp. 522).

Twitterklient

Varje twittermeddelande identifieras med vilken enhet de skrivits på.

Uppdateringar via mobila enheter utgör tillsammans över 50 % av tweetsen.



Endast tweets skrivna 2012 är analyserade, max tweets 100 per konto.

Vanligaste orter

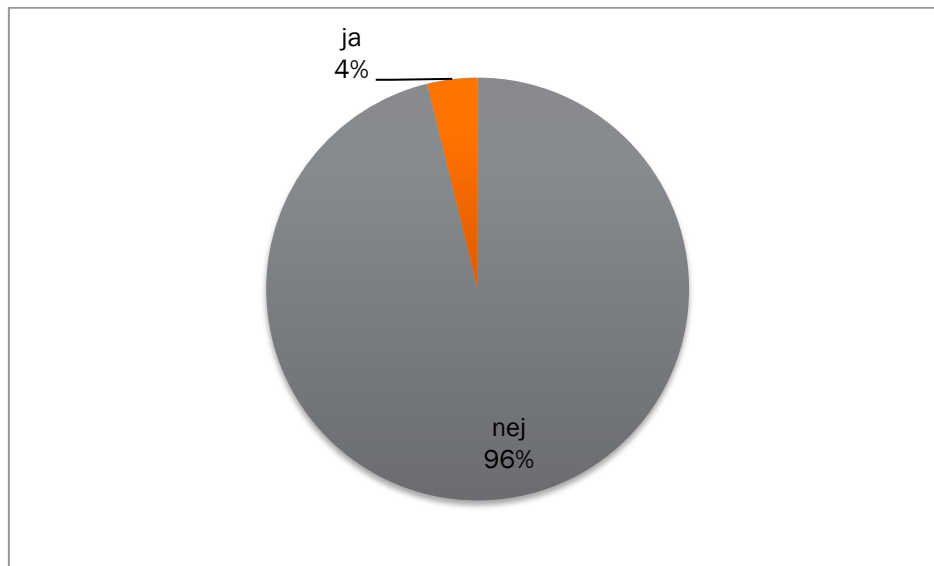
Vanligaste orterna i "location"-fältet.

Ort	Antal
(tomt)	155445
Sweden	25954
Stockholm	19512
Stockholm, Sweden	6191
Göteborg	4532
Sverige	3557
Malmö	2390
Uppsala	1666
Göteborg	1446
Jönköping	1134
Göteborg, Sweden	1065
Örebro	1065
Malmö, Sweden	1063
Umeå	1051

Linköping	991
Karlstad	907
Uppsala, Sweden	795
Göteborg, Sweden	781
Lund	773
Sundsvall	765
Helsingborg	743
västerås	730
Norrköping	621
Kalmar	612
Växjö	569
Lund, Sweden	567
Borås	529
Halmstad	513
Gävle	478
Sthlm	461

Platsangivelse

Geokodning med GPS eller liknande av tweets stöds av twitter, men används bara i liten utsträckning. 4 % av tweetsen har platsinformation. Endast tweets skrivna 2012 är analyserade, max tweets 100 per konto.



Figur 6 andel tweets med platsangivelse

Det går också att ange om platstjänster ska vara på eller avstängda på kontonivå. 18 % av kontona har denna funktion påslagen. Bland de mest aktiva är samma siffra 30 %.

Diskussion

Twitter har förändrats

Twitter har förändrats. På ett år har antalet användare dubblats. Tillväxten kan inte tillskrivas en enskild grupp, utan verkar ske ganska brett. Twitter har tidigare beskrivits som en elitkanal där media och opinionsbildare var aktiva. Det är säkert fortfarande sant, men tillväxten den senaste tiden sker även i andra kluster. De senaste åren kan vi exempelvis se stark tillväxt bland användare i "sportklustret". Andelen nyregistreringar på konton med sportrelaterade termer i sin beskrivning har också ökat rejält. Om man analyserar nätverksgraferna och tidsserierna på densamme framträder en tillväxt i vissa ganska avskilda kluster. Exakt vilka dessa är kvarstår att analysera.

Om man tittar på klusterbildningarna bland de mest aktiva twitterkontona, och sedan lägger på ett tidsfilter kan man utläsa hur twitter har spridits.

Från 2006 till 2008 sker tillväxten framförallt i den så kallade "bubblan" – bland teknik och sociala medier-intresserade. Från 2008 och 2009 expanderar dessa in i ett journalist/opinionsbildarområde (som är väldigt integrerat med det förstnämnda klustret). Vi ser överhuvudtaget en stark tillväxt – under 2009 tiodubblas antalet svenskspråkiga konton. Fram växer "kändisklustret", "hiphopklustret" med mera. 2010 växer också ett "nationellt" kluster. "Sportklustret" börjar växa redan 2009, men vi ser dess starkaste tillväxt under 2011.

2011 och 2012 sker bred tillväxt och det tillkommer helt nya kluster, som åtminstone delvis verkar vara geografiskt avgränsade.

Twitter har ett stort genomslag

Twitters genomslag i samhällsdebatt och media är stor. Referenser till vad som sägs på twitter, eller vad enskilda personer säger på twitter, görs dagligen i "traditionell" media. Antalet omnämnanden av twitter i svensk media är ökande. Omnämnande om twitter sker nästan hälften så många gånger som Facebook⁶ - fast Facebook i antalet aktiva användare kanske är tio gånger större i Sverige. Förklaringarna kan vara flera, men Twitters

⁶ Mediearkivet, sökning på "twitter" resp. "facebook"

öppenhet såväl som att många journalister är aktiva kan vara två bidragande faktorer.

Passiva användare

Många vittnar om att de har konton på twitter för att omvärldsbevaka eller för att få inblick. Ytterligare fler kanske tittar på twitter.com eller någon av alla aggregeringstjänster eller webbplatser som visar twitterflöden. Den här rapporten undersöker inte hur många som läser twittermeddelanden – utan hur många som skriver.

Få står för nästa allt

Det är en väldigt liten grupp som står för den absolut största delen av aktiviteten. Skillnaderna mellan medel och medianvärden på antal uppdateringar, antal som följer eller följs, talar alla sitt tydliga språk. Några av de mest aktiva när det kommer till att skicka ut tweets är automatpostade konton. Men även om man tar bort dessa är det en liten andel som av kontona som är mycket mer aktiva. Mediananvändaren däremot har varit på twitter mindre än ett år, följs av 10, följer 23 och har gjort 28 uppdateringar.

Fortfarande hög tröskel?

De flesta som börjar twittra skriver bara ett fåtal tweets. Hälften av kontona har skrivit mindre än 28 tweets. En väldigt aktiv twittrare skriver den mängden tweets på en dag.

Majoriteten följs också av väldigt få. Det är inte en långsökt analys att många tröttnar på att tala ut i tomma intet: nästan 23 000 följs bara av en annan person, ytterligare 15 000 följs av ingen. Totalt följs 110 000 av 5 eller mindre personer.

Förhållande till Svenska Twitter

Resultaten är snarlika de siffror som Aitellu presenterat i sin rapport Svenska Twitter. Deras metod skiljer sig något, bland annat inkluderar de konton som uppgivit ort i Sverige, oavsett om de skriver på svenska eller inte. Det förklarar de högre siffror de presenterar på flera ortsnamn. Aitellu inkluderar även analys av beskrivningen, vilket möjliggör att hitta konton som inte skrivit några tweets, vilket kan förklara deras något lägre siffra på konton som uppdaterats den senaste månaden. Vi har funnit fler konton totalt (även om Aitellu har reviderat upp sin totalsiffra till mer än 300 000), vilket förklaras av att vi även inkluderar konton som inte angivit ort eller beskrivning. Sammanfattningsvis kan sägas att skillnaderna mellan resultaten kan förklaras med avgränsningen av vilka konton som ska ingå – d.v.s. definitionen av "svensk".

	Twittercensus	Aitellu
Inga tweets		Via språkanalys av beskrivning
Ingen beskrivning	Via språkanalys av 100 tweets	Via språkanalys av beskrivning
Ingen angiven plats	Via språkanalys av 100 tweets	Via språkanalys av beskrivning
Angiven plats	Via språkanalys av 100 tweets	Inkluderas om svensk ort
Angiven beskrivning	Via språkanalys av 100 tweets	Via språkanalys av beskrivning

Kort om några begrepp

Twittermeddelande, tweets och uppdateringar används liktydigt och syftar på ett skrivet meddelande på twitter.

Twitteranvändare, twittrare, person, användare och konton refererar alla till en unika registrerade konton på twitter.com.

Svenska, svenska, svenskar på twitter refererar alla till de som i den här undersökningen identifierats som twittrande på svenska.

Frågor?

Kontakta Hampus Brynolf, hampus.brynolf@intellecta.se eller på telefon 0708-73 33 34. Eller förstås på twitter: @hampusbrynolf.