



LUNDS UNIVERSITET

Ekonomihögskolan

Institutionen för informatik

Är du informerad?

En studie om hur informerade användare är om datainsamling på sociala medier.

Kandidatuppsats 15 hp, kurs SYSK16 i Informatik

Författare: Victor Jennekvist
Benjamin Idzikowski

Handledare: **Björn Svensson**

Rättande lärare: Björn Johansson
Benjamin Weaver

Är du informerad? : En studie om hur informerade användare är om datainsamling på sociala medier

ENGELSK TITEL: Are you informed? : A study on how informed users are regarding data collection on social media.

FÖRFATTARE: Victor Jennekvist, Benjamin Idzikowski

UTGIVARE: Institutionen för informatik, Ekonomihögskolan, Lunds universitet

EXAMINATOR: Odd Steen, Docent, Fil Dr

FRAMLAGD: maj, 2019

DOKUMENTTYP: Kandidatuppsats

ANTAL SIDOR: 54

NYCKELORD: data, persondata, datainsamling, sociala medier, sociala nätverk, informerad

SAMMANFATTNING (MAX. 200 ORD):

Användningen av persondata har blivit ett allt vanligare fenomen och utnyttjandet av data spelar en mer central roll i ekonomin. Data är en värdefull resurs men det har blivit svårare för konsumenterna att kontrollera hur data om dem samlas in och delas eller säljs vidare. Många aktörer är ofta involverade i datautvinningsprocessen utan att det framgår vilken roll de spelar och till vilket ändamål de lagrar och använder sig av konsumenternas data. Utvecklingen av sociala medier har gett större möjligheter för datainsamling men många av datainsamlingsmetoderna är dolda för användarna på sociala medier.

Uppsatsen avser att ta reda på hur informerade användare på sociala medier är om hur mycket av deras data som samlas in genom att utföra en kvantitativ studie på området med hjälp av en enkätundersökning. Studien resulterade i att det finns en stor variation bland användare och hur informerade de är kring de datainsamlingsmetoder som utreds i uppsatsen, men att en övervägande del av dessa metoder är okända för en majoritet av de användare som deltog i enkätundersökningen.

Innehåll

1	Introduktion.....	6
1.1	Bakgrund	6
1.2	Problemområde.....	7
1.3	Forskningsfråga	8
1.4	Syfte.....	8
1.5	Avgränsningar	8
2	Litteraturgenomgång	9
2.1	Sociala medier	9
2.2	Beskrivning av data	10
2.3	Datainsamling på sociala medier	12
2.3.1	Cookies.....	14
2.3.2	Självzensurering	15
2.3.3	Spårningsmetoder	15
2.3.4	Sociala plug-ins	17
2.3.5	Hur informerade är användare om datainsamling	18
2.4	Teoretiskt resultat	18
2.5	Litteratortabell	19
3	Metod	21
3.1	Kvantitativ metod	21
3.2	Enkätutformning	22
3.3	Urval och spridning av enkät	24
3.4	Analysmetod	25
3.5	Validitet och Reliabilitet.....	25
3.6	Etik.....	26
3.7	Felkällor i metod.....	26
4	Resultat	27
5	Diskussion.....	33
5.1	Användningsfrekvens	33
5.2	Datainsamling	33
5.2.1	Datainsamling på sociala medier.....	33
5.2.2	Sociala plug-ins, cookies och självzensurering.....	34
5.2.3	Spårningsmetoder	36
5.2.4	Diskussion om teorins begränsningar	37
6	Slutsats	38
	Appendix 1 - Enkät	39

Appendix 2 – Enkät svar	43
Referenser.....	49

Figurer

Figur 2.1: Organisationer med flest aktiva tredjepartsaktörer vid en analys av en miljon populära webbsidor (Englehardt & Narayanan, 2016, s.8).....	16
Figur 4.1: Svar på enkätfråga 1	28
Figur 4.2: Svar på enkätfråga 2	28
Figur 4.3: Svar på enkätfråga 3	29
Figur 4.4: Svar på enkätfråga 4	29
Figur 4.5: Svar på enkätfråga 5	30
Figur 4.6: Svar på enkätfråga 6	30
Figur 4.7: Svar på enkätfråga 7	31
Figur 4.8: Svar på enkätfråga 8	31
Figur 4.9: Svar på enkätfråga 9	32
Figur 4.10: Svar på enkätfråga 10	32

Tabeller

Tabell 2.1: Litteraturtabell.....	19
Tabell 3.1: Enkätutformning	22
Tabell 3.2: Tabell över värden i enkäten.....	23
Tabell 3.3: Tabell över värden i enkäten.....	23

1 Introduktion

1.1 Bakgrund

I takt med den ökade digitaliseringen av samhället har användningen av data som samlas in blivit en central del av affärsstrategin för företag som verkar på den digitala marknaden samtidigt som konsumenterna förlorar alltmer kontroll över vilken data som samlas in av vilka aktörer och hur dessa aktörer använder sig av datan (Larsson, 2018). I samband med att EU-kommissionären Viviane Reding sökte en ny överenskommelse om dataskydd för Europa 2014 uttryckte hon sin syn på data som *“Data är den moderna tidsålderns valuta. Data används av alla typer av företag, från försäkringsbolag och banker till sociala medier och sökmotorer.”* (EU-kommissionen, 2014).

Norska myndigheten Datatilsynet gjorde 2015 en studie som mätte aktiviteten på sex olika norska dagstidningars hemsidor och till vilken grad informationen kring besöken skickades vidare. I studien framgår det att omkring 100 till 200 cookies placerades i besökarens browser, deras IP-adress skickades till 356 servrar och att 46 tredjepartsaktörer var involverade i varje besök (Datatilsynet, 2015). Som användare kan det vara en utmaning att ta del av denna dataförmedling och vilka aktörer som faktiskt har tillgång till informationen, vilket ger en indikation på att den bristande insynen i hur datahanteringen går till gör att konsumenten förlorar kontroll över sin data (Pew, 2014).

I en studie gjord av Appelgren och Leckner (2016) för SOM-Institutet så säger sig 37% av 1570 tillfrågade svenskar vara negativa till insamling av personlig information vid användning av sociala medier, medans bara 1% av de tillfrågade var positiva. I samma studie så framgår det även att endast 15% av 1591 tillfrågade noga läser igenom användaravtal vid registrering på olika webbsidor och appar (Appelgren & Leckner, 2016). Bechmann (2014) menar att formaliteten vid avtal genom ett klick inte går att jämföra med ett avtal som sker ansikte mot ansikte, vilket innebär att det är tveksamt att medgivandet som ges via ett enkelt klick överhuvudtaget kan ses som informerat.

Dagens datainsamling har utvecklats och blivit mer omfattande än vad den var förr (Kshetri, 2014). Aktiviteter som tidigare var mer kvalitativa och inte var övervakade eller sparade har nu blivit loggade och mer kvantitativa, allt ifrån en användares webbhistorik, cookies, tittade videoklipp, webbsökningar, vilken geografisk position som ens enhet är på till vilka annonser som man interagerar med (Larsson, 2018). Larsson (2018) påtalar att denna information, tillsammans med demografisk information, om vilka produkter eller tjänster en användare köper och vid vilken tidpunkt, ger en väldigt stor möjlighet att profilera användare i ett marknadsföringssyfte.

Profilering går även att använda till andra saker än att marknadsföra produkter. Våren 2018 seglade Cambridge Analytica-skandalen upp, där det visade sig att 87 miljoner människor hade fått sin information insamlad av företaget, via Facebook (Badshah, 2018). Cambridge Analytica använde sig av den insamlade informationen och med hjälp utav ett kraftfullt

program så lyckades de skapa träffsäkra profiler av människor, som de sedan använde för att förutspå samt påverka människor för att på så sätt skapa opinion och därmed styra politiska val åt ett visst håll (Cadwalladr & Graham-Harrison, 2018). I Cadwalladr och Graham-Harrisons (2018) artikel så pekas två stora val ut som influerade av Cambridge Analytica: presidentvalet i USA år 2016 samt valet om Storbritanniens utträde ur EU år 2016, även kallat Brexit, och detta påvisar vikten för användare på sociala medier att verkligen hålla koll på sin data.

1.2 Problemområde

Persondata spelar en alltmer central roll i den datadrivna ekonomin och har blivit en form av valuta, där mycket av tillväxten i ekonomin och tjänsteutvecklingen baseras på datadrivna applikationer (Larsson & Ledendal, 2017). Trots detta upplever konsumenter en ökad brist på kontroll över sin persondata och har en begränsad insyn i hur data används, vilket medför att många oroar sig över att tredjepartsaktörer har tillgång till deras data och om datan används på annat sätt än vad som ursprungligen var avsett (Kshetri, 2014).

Christl (2017) beskriver digitaliseringen som starkt integrerad i vår vardag och att data samlas in i en form av "ekosystem" av plattformar och aktörer. Detta ekosystem beskrivs som mångskiftande och fragmenterat med oerhört snabb utveckling och många, ofta motstridiga, intressen (Christl, 2017). Vissa större aktörer framgår tydligt på marknaden medan andra aktörer arbetar i bakgrunden med insamling och bearbetning av data (Cooper et al., 2015). I motsats till vad många tror säljer, för det mesta, inte de större aktörerna vidare sin data till tredjepartsaktörer men däremot låter de andra företag utnyttja deras data och använda sig av deras infrastruktur för att samla in mer data (Cooper et al., 2015)

Den tilltagande övervakningen av användare och den ökande datainsamlingen är svår att överblicka och beskrivs av Frank Pasquale (2015) som ett "black-box"-samhälle, där bristen på transparens tas upp som ett av de främsta problemen. Vidare nämner Pasquale (2015) att en övergång hos företag från att influera användares beslut till att istället manipulera besluten efter egen vinning blir allt vanligare. Att man som användare har möjlighet att avregistrera eller på annat sätt ta sig ur tjänster som ämnar att övervaka är ett sätt att begränsa den data man delar men det är knappast ett rimligt alternativ i de flesta fall och kan leda till negativa konsekvenser (Christl & Spiekermann, 2016). Maciej Ceglowski (2016) summerar problematiken genom en liknelse mellan att du som individ har möjlighet att göra ett val om du vill använda dig av elektricitet eller inte, ett val som är teoretiskt möjligt men i praktiken skulle det innebära "opting out of much of modern life".

I en studie från 2009 som undersökte hur medvetna användare på Facebook är gällande hur mycket av deras information som delas vidare eller finns tillgänglig för andra, framgår det att majoriteten av användarna inte oroar sig över att den information de delar med sig av frivilligt kommer att påverka deras privatliv på ett negativt sätt (Tuunainen et al., 2009). 94% av de 210 respondenterna känner till att de kan ändra inställningarna för hur mycket information de delar med sig av, däremot kände en majoritet (73%) av respondenterna inte till att Facebook kan dela med sig av deras information till tredjepartsaktörer trots att 61% hävdar att de läst Facebooks privacy policy där det framgår att informationen delas vidare till tredjepartsaktörer (Tuunainen et al., 2009).

Facebooks ägare och skapare Mark Zuckerberg satt i april 2018 i en omfattande utfrågning av den amerikanska kongressen (Wong, 2018). I förhöret med kongressen så framkom det att Facebook sparar användardata på flera olika sätt som användare tidigare kanske inte varit medvetna om (Kanter, 2018). Kanter (2018) listar spårningsmetoder som framkom i utfrågningen, och dessa inkluderar bland annat att Facebook loggar användares musrörelser på skärmen, de ser över om webbläsaren ligger i bakgrunden eller förgrunden, och i vissa fall så övervakar de även enheter som är i närheten av användaren eller som är på samma nätverk som användaren är uppkopplad på. Tidigare forskning av Tuunainen et al (2009) har visat att en majoritet av användare är medvetna om att Facebook har datan som de själva tillhandahåller, men däremot är forskningen begränsad när det kommer till hur informerade användare är om alternativa spårningsmetoder, bland annat de som framkom i utfrågningen, det vill säga att musrörelser loggas, enheter uppkopplade på samma nätverk övervakas och om webbläsaren ligger i bakgrunden eller i förgrunden.

1.3 Forskningsfråga

Hur informerade är användare på sociala medier om i vilken utsträckning deras data samlas in?

1.4 Syfte

Uppsatsen syftar till att ta reda på till vilken grad användare på sociala medier är informerade om hur deras data samlas in på sociala medier.

1.5 Avgränsningar

Uppsatsen avgränsar sig till att inte ta reda på hur informerade användare är om sin frivilligt delgivna data som delas på sociala medier. Anledningen till detta är att studien huvudsakligen ämnar att behandla data som genereras via aktiviteter och beteende på sociala medier.

2 Litteraturgenomgång

2.1 Sociala medier

Sociala medier är ett slags elektroniskt kommunikationsmedium där människor kan skapa en gemenskap och dela information, meddelanden och annat innehåll (Merriam Webster, 2019). År 2017 hade sociala medier sammanlagt 2.46 miljarder användare över hela världen, och ökningen av användare på sociala medier har vuxit markant från dess början, år 2010 så fanns det 970 miljoner användare på sociala medier, och år 2021 så förutspås sociala medier ha så mycket som 3.02 miljarder användare (Statista, 2019a). Det största sociala nätverket är Facebook med cirka 2.3 miljarder användare (Statista, 2019b) och det används främst via smartphones och tablets, följt av bärbara samt stationära datorer (Statista, 2019c).

Oftast så är sociala medier skapade med tanken att vara tillgänglig för alla, men det är inte ovanligt att olika sociala medier blir segmenterade och används i speciella grupper baserat på exempelvis ålder eller nationalitet (Boyd & Ellison, 2007). Kulturen som skapas kring olika sociala medier är alltså varierande, men Boyd och Ellison (2007) fastslår att de tekniska nyckelelementen är relativt konsekventa. Enligt Richter & Koch (2008) så finns det sex stycken basfunktioner som karaktäriserar sociala medier och dessa är:

- Identitetshantering:

Identitetshantering innebär att en användare kan fylla i personlig information och ställa in åtkomsträttigheter till dennes profil.

- Sökfunktion:

Sociala medier bör även ha en sökfunktion där användaren kan hitta andra personer eller företag utefter vissa kriterier, genom exempelvis en sökruta, och detta innefattar även att användare skall få pro-aktiva rekommendationer på saker som användaren tycks vara intresserad av.

- Kontextmedvetenhet

Kontextmedvetenhet handlar om att användare skall bli medvetna om gemensamma vänner, gemensamma intressen eller om användare har arbetat på samma företag, och detta är i huvudsak till för att skapa tillit och ett samband mellan användare.

- Kontakthantering:

Med kontakthantering så kombineras alla funktioner som gör det möjligt för användare att underhålla sitt personliga nätverk, exempelvis att tagga andra människor eller att begränsa åtkomsten till ens profil.

- Nätverksmedvetenhet:

Genom nätverkshantering så har användare koll på sitt nätverk, med funktioner som skapar indirekt kommunikation och medvetande, exempel på sådana funktioner är ett nyhetsflöde eller notifikationer om att någon fyller år.

- Informationsutbyte:

Informationsutbyte innefattar alla olika sorters direkta och indirekta informationsutbyten, så som direktmeddelanden i en chatt eller via foton i fotoalbum.

Sociala medier har en stor varians i hur mycket det används och det har ett brett användningsområde (Bylund, 2013). Detta kan ses i en studie gjord av Whiting och Williams (2013) där det framkom att det största användningsområdet för sociala medier är social interaktion, att hålla kontakt med vänner och familj, där respondenter tillade att de har mer kontakt med folk över sociala medier än i verkligheten och att sociala medier ger dem ett socialt liv. Efter social interaktion så listades informationssökning som det näst största användningsområdet för sociala medier, för att hitta bland annat rabatter på produkter, information om företag eller hitta olika evenemang (Whiting & Williams, 2013). Andra användningsområden som nämns är tidsfördriv, underhållning, avkoppling och att övervaka eller ta del av vad andra gör (Whiting & Williams, 2013).

Men det är inte bara konsumenter som använder sig utav sociala medier, Asur och Huberman (2010) påvisade att man genom analys av sociala medier kan förutspå händelser i verkliga livet, som de visar när de förutspått intäkter av biofilmer mer precist än vad börserna i Hollywood gjort. Även företag befinner sig på dessa plattformar, och enligt en studie utförd av Baird och Parasnis (2011) så hade 79% av 351 tillfrågade företag en profil på sociala medier. Baird och Parasnis (2011) berättar även att vissa företag skapar program för hantera sociala medier, och att potentialen med att komma närmare kunderna via sociala medier är enorm och kan resultera i ökade intäkter, minskade kostnader samt förbättrad effektivitet. Företag använder även sociala medier för att göra reklam för sina tjänster och produkter, och endast på Facebook så är det tre miljoner företag som aktivt visar reklam (Facebook, 2016).

Många handlingar som användare gör på sociala medier, såsom "likes", kommentarer och vilka vänner som användaren har, samlas in av leverantörerna av dessa sociala medier och förvandlas enligt Bylund (2013) till en guldgruva för de som vill ta reda på vad individer har för intressen. Denna information som leverantörerna får kan sedan användas för att sälja riktad reklam samt beteendedata till företag (Bylund, 2013). Allt detta är möjligt genom att när en användare först skapar ett konto på sociala medier accepterar de ett avtal, något som kritiserats av Bechmann (2014) som beskriver det som en form av blint samtycke som inte kan likställas med ett formellt samtycke utanför internet, och anser att det går att diskutera huruvida det går att kalla det ett informerat samtycke överhuvudtaget. Detta styrks av Appelgren och Leckner (2016) som i sin studie kom fram till att det är väldigt få som läser igenom avtal när de registrerar sig på webbsidor och appar.

2.2 Beskrivning av data

Enligt Zins (2007) så finns det flera olika accepterade definitioner om data. Zins (2007) konstaterar dock att många forskare inom informatik anser att data, information och kunskap är delar i en sammansatt ordning, där data är det råaste materialet, likt byggstenar, som blir till en grund för information, som sedan blir underlag för kunskap. När människor använder

internet så genereras data, genom vilka webbsidor som besöks eller aktiviteter som utförs i en digital kontext (Larsson & Ledendal, 2017). Denna data kan räknas som personuppgifter, som enligt Datainspektionen (u.å.) beskrivs som *“Personuppgifter är all slags information som kan knytas till en levande person.”*

Christl och Spiekermann (2016) berättar att personlig data kan grupperas in i tre olika grupper, beroende på hur den samlas in:

- **Frivillig data:** Den frivilliga datan skapas samt delas av en individ på exempelvis profiler på sociala nätverk. Detta delas av användarna när de skriver in sig på en tjänst, och kan exempelvis vara namn, ålder eller e-mailadress.
- **Observerad data:** Den observerade datan är hämtad på så sätt att individers handlingar loggas, t.ex platsdata från mobiltelefoner. Denna data skapas och delas passivt av användarna.
- **Antagen data:** Den antagna datan är baserad på en analys gjord på frivillig och/eller observerad data, och är antaganden som exempelvis tredjepartsaktörer kan göra baserat på en sammansättning av frivillig data och observerad data.

Vidare så berättar Christl och Spiekermann (2016) att data går att dela in i ytterligare två grupper, baserat på vilken sorts relation en användare eller konsument har gentemot företaget som samlar in datan:

- **Första parts data:** Data som samlas in av första part är data som samlas in av företag som har en direkt relation med sina kunder. Datan samlas exklusivt in från kunden via interaktioner med företaget, exempelvis när en kund köper en produkt från ett företags hemsida.
- **Tredje parts data:** Data som samlas in av tredje part är data som en tredje part har köpt, bytt till sig eller förvärvat genom licensiering från en första part. Det kan också vara data som är hämtad gratis via analyser på olika sociala medier eller via allmänna handlingar. Tredje parts data kan även samlas in via deras egna cookies som är sparade på användares enheter som sedan används när användaren nyttjar en första parts hemsida.

Det går även att dela in data i fem kategorier utefter vilken karaktär den har, enligt Busby et al. (2012).

- **Demografisk data:** Innefattar information så som ålder, kön och inkomst. Många gånger är detta essentiellt för kundsegmentering hos företag.
- **Beteende eller kontextuell data:** Användares olika intressen och attityder ingår i den här klassen. Denna data beror på vilket innehåll som en användare har nyttjat, eller kan komma från exempelvis cookies. För att denna data skall ses som användbar inom exempelvis annonsering, så behövs ofta data från flera olika källor sammansättas.
- **Köpintensionsdata:** Köpintensionsdata mäter hur stora intentioner någon har över att köpa något. Denna data fås om någon har fyllt i ett formulär och visat intresse eller så är den baserad på sökningar. Den kan också vara baserad på tidigare köp. Om en

användare har kikat på en produkt men inte köpt den, så kan datan användas för att göra reklam för produkten i ett senare skede.

- **Socialdata:** Denna kategori av data är baserad på vilka relationer som människor har med varandra, och innebär en tankegång om att de som är sammankopplade på internet har liknande attribut eller köpintention. Denna data kan hämtas via sociala medier eller genom interaktioner mellan människor, exempelvis genom att någon delar en artikel med en annan.
- **Platsdata:** Historiskt sett har denna data hämtats beroende på vilken webbsida som en användare har besökt eller vart användarens IP-adress är lokaliserad. Detta kan sedan användas för att annonsera om exempelvis varor och tjänster som ligger nära användaren geografisk. Detta har blivit lättare och mer noggrant med hjälp av mobilt internet, som förutom att det ger en mer exakt bild var användaren är, så kan även den som tar del av platsdatan veta när användaren är och handlar.

Användningen av persondata har blivit en allt vanligare företeelse hos många företag (Cooper et al., 2015). I en undersökning utförd av företaget *Accenture* som frågade ungefär 600 andra företag om deras datainsamling framgick det att 79% av deltagarna i undersökningen påstod att deras företag samlar in data direkt från individer (genom t.ex deras kundkonto) och att 42% samlar in data från andra företag genom kommersiella partnerskap eller avtal om data-delning (Cooper et al., 2015). 33% hävdar att de samlar in data från uppkopplade enheter och köper data från tredjepartsleverantörer som samlar in, delar och säljer data (Cooper et al., 2015).

Larsson och Ledendal (2017) säger att det inte går att förneka att det ligger ett konkret ekonomiskt värde i personuppgifter, utan snarare tvärt om så är personuppgifter en värdekälla som har blivit ledande inom den digitala ekonomin. En stor anledning till att persondata samlas in av företag är för att ta reda på mer information om konsumenter, och på så sätt ha en aktiv möjlighet att influera köpprocesser i både nutid samt i framtiden (Larsson & Ledendal, 2017). Det ger också möjligheter att rikta reklam, sälja reklam och öka tillväxt på sikt för företag (Larsson & Ledendal, 2017).

2.3 Datainsamling på sociala medier

Deltagande i sociala medier har vuxit kraftigt de senaste åren (Bylund, 2013). Det finns stora variationer i hur olika individer använder sig av sociala medier, där en del användare delar med sig av allt som händer i sina liv frivilligt medan andra är mer återhållsamma och privata (Bylund, 2013). Vidare beskriver Bylund (2013) en förändring i betydelsen av personlig integritet på internet har uppkommit till följd av den ökade användningen av sociala medier och att denna förändring i första hand handlar om vad vi delar med oss av, på vilket sätt och till vilka.

Personlig information på internet kan vara frivilligt angiven, exempelvis genom att användaren fyller i information i ett formulär eller går med i en tjänst som t.ex ett socialt nätverk, eller så samlas data in indirekt utan användarens vetskap med hjälp av olika metoder (Bujlow et al., 2015). Information och data som används med syftet att skapa användarprofiler på webben

utgörs till största delen av sidoinformation om användarens aktiviteter som skapas när användaren besöker olika webbsidor på internet (Bylund, 2013).

Bylund (2013) uttrycker dock att för sociala medier så handlar det i första hand om information som användare själv valt att sprida, som exempelvis uppdateringar i ett nyhetsflöde eller liknande frivilligt angiven information. Många användare förstår att information som de själva väljer att sprida och dela på sociala nätverk finns tillgänglig för andra (Tuunainen et al., 2009). Men när data samlas in i ett "osynligt" nätverk av insamlingsmetoder och flera olika aktörer blir det svårare att överblicka vilken data som delas och vilka som får tillgång till datan (Christl, 2017).

Utöver den frivilligt angivna informationen så samlar leverantörer av tjänster som sociala medier in data om användarens beteende, aktiviteter och intressen vilket ger möjligheter att dra slutsatser om olika individers intressen, vanor och umgänge (Bylund, 2013). Data som samlas in kan vara av teknisk natur, som vilken webbläsare, operativsystem eller hårdvara användaren brukar för tillfället men datan innefattar ofta mer känslig information som geografisk plats, användarens preferenser och webbhistorik (Bujlow et al., 2015).

Facebook är, som tidigare nämnt, det största sociala nätverket med cirka 2.3 miljarder användare (Statista, 2019b). För att Facebook ska kunna fungera och kontinuerligt uppdatera och anpassa sin struktur och sitt utbud av tjänster krävs det att information om användarna och deras beteende samlas in och analyseras (Facebook, 2018).

Facebook samlar in data om exempelvis innehållet användaren skapar och delar, kommunikationen mellan användare och geografisk plats där en fil skapades eller ett foto togs (Facebook, 2018). Denna information analyseras automatiskt av systemet med syftet att kunna leverera personligt anpassade tjänster och innehåll till användaren samt förslag på grupper, event eller ämnen som är av intresse för användaren (Facebook, 2018).

Ett forskningsteam kallat Share Lab gjorde en omfattande undersökning om den datainsamling som sker på Facebook för att belysa och kartlägga alla inputs och outputs i systemet och vilka steg detta "osynliga" system tar (Joler & Petrovski, 2016). Joler och Petrovskis (2016) metodik baseras på olika undersökningsverktyg som appliceras på offentligt tillgängliga resurser, som exempelvis Facebooks Data Policy och cookies eller pixel-teknologi hos tredjepartsaktörer.

Enligt forskningen från Joler och Petrovski (2016) kan Facebook på olika sätt att extrahera data från användares beteende och aktiviteter både inom och utanför Facebooks domän kategoriseras i två huvudkategorier:

- Aktiviteter och beteende

Varje Like användaren väljer att göra, varje steg eller klick på hemsidan, varenda foto som laddas upp och varje event användaren deltar i registreras och lagras av Facebook i deras databas (Joler & Petrovski, 2016). Denna kategori består av alla interaktioner mellan användare och systemet, innehåll användaren skapar eller laddar upp och vilka sidor inom Facebook som användaren besöker - i princip allt en användare väljer att göra på Facebook, vilket således gör att denna information representerar användarens preferenser och interaktioner dynamiskt i realtid (Joler & Petrovski, 2016).

- Frivilligt delgiven information

Den andra kategorin består av all den information som användaren frivilligt matat in och valt att dela om sig själv, som namn, adress och telefonnummer. Denna information skiljer sig från den andra kategorin då den är statisk, sällan uppdateras och baseras helt och hållet på vad användaren själv väljer att delge, vilket innebär att det finns risk att informationen är felaktig eller missledande (Joler & Petrovski, 2016).

Facebook samlar in data från användarnas datorer, telefoner och andra enheter samt enheter på samma nätverk för att kunna göra sin tjänst så effektiv som möjligt (Facebook, 2018). Data som samlas in inkluderar enhetens attribut, signaler som Bluetooth och närliggande Wi-Fi access points, inställningar, nätverksanslutningar, cookiedata samt operationer som utförs på enheten (Facebook, 2018). Datainsamling om dessa operationer inbegriper exempelvis om ett fönster på enheten är maximerat eller minimerat och användarens musrörelser på skärmen, med syftet att bedöma om användaren är en robot eller en människa (Senate Commerce, 2018).

Genom att kombinera denna data med ytterligare information som exempelvis användarens operativsystem, IP-adress och webbläsare samt data insamlad via cookies och Facebook pixel lyckas Facebook utläsa och förutspå en mängd olika beteendemönster (Joler & Petrovski, 2016).

Användare, inklusive avancerade användare, har en tendens att frånsä betydelser av att läsa användarvillkoren när de installerar en applikation på sina mobila enheter vilket innebär att, i många fall, applikationen får tillgång till mer information än vad som krävs för att den ska fungera (Joler & Petrovski, 2016).

När en användare installerar Facebook och Facebook Messenger på sin mobila enhet ger hen tillåtelse till applikationen att komma åt en omfattande och mångsidig mängd data som finns på den mobila enheten (Joler & Petrovski, 2016). Data som blir tillgänglig för applikationen innefattar olika typer av inkräktande information som exempelvis enhetens unika identifieringskod, exakta geografiska position, identiteten på alla kontakter som är inlagda i enheten, innehåll i SMS-meddelanden och samtalshistorik (Joler & Petrovski, 2016). Utöver detta får applikationen även tillåtelse att spela in ljud, hämta information om användarens uppkoppling och ladda ned filer utan att varna användaren (Joler & Petrovski, 2016).

2.3.1 Cookies

Cookies är små delar av data som placeras i webbläsarens lagringsutrymme av webbservern (Bujlow et al., 2015). När en användare besöker en hemsida för första gången så kan en unik identifieringskod lagras i cookie-filen på användarens dator vilket innebär att hemsidan kan känna igen användaren vid återbesök genom att göra identifieringskoden åtkomlig om och om igen (Christl & Spiekermann, 2016).

Det finns flera typer av cookies, som exempelvis "session cookies" vilka upphör att existera när sessionen är över, det vill säga när webbläsaren stängs ned till skillnad från "persistent cookies" som kan lagras i flera timmar, dagar eller år hos användaren (Bujlow et al, 2015).

Cookies ämnar att spåra användarens interaktioner på olika hemsidor och spara information som användaren matat in eller för att autentisera användaren vid behov (Bujlow et al, 2015). Cookies placeras på användarens dator på två olika sätt: antingen av ett JavaScript som använder sig av ett API-call, eller genom HTTP responser som innehåller en Set-Cookie header (Roesner et al., 2012). Cookies blir automatiskt bifogade när en HTTP-förfrågning utförs till domänen som cookien tillhör med hjälp av cookie headers vilket gör att olika tjänster kan ta del av cookien (Bujlow et al., 2015). De kan också göras tillgängliga genom att ett JavaScript API uttryckligen gör en förfrågan om att få tillgång, för att sedan skickas vidare till servern (Bujlow et al., 2015).

2.3.2 *Självzensurering*

Självzensurering är att hindra sig själv från tala, som visar sig tydligt inom sociala medier där användare kan frukta att spamma andra användare i sitt nätverk med ointressant eller onödigt innehåll, eller för att upprätthålla sin självbild över flera sociala kontexter samtidigt (Das & Kramer, 2013).

Självzensurering inom sociala medier tillåter användare att skriva, forma och uttrycka sina tankar innan de delar dem med världen, med möjligheten att sedan radera det, något som inte riktigt går i ett samtal mellan två människor, och som skapar ett nytt fenomen som Das och Kramer (2013) kallar för "sista minuten självzensurering". Detta kan, enligt Das och Kramer (2013) både hjälpa och stjälpa, då användare kan antingen lyckas eller misslyckas med att uppnå ett slags potentiellt socialt värde när de inte delar och sociala medierna förlorar värde då det inte har skapats något innehåll.

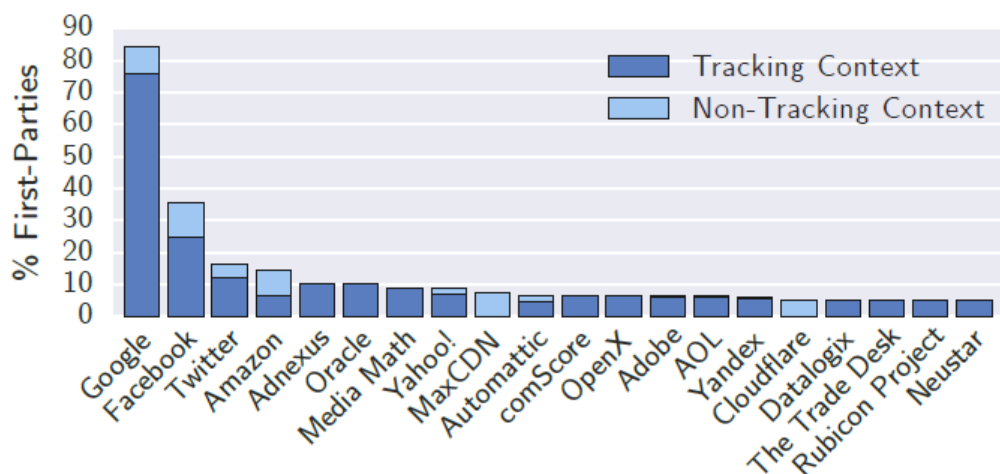
I en studie av Das och Kramer (2013) som baserades på inlägg och kommentarer skrivna av användare på Facebook som sedan självzensurerats, så sparades inte dessa inlägg och kommentarer i sin helhet, utan skickades tillbaka som ett anonymiserat binärt värde att det hade skapats innehåll som sedan inte skickats vidare. Att spara text som användare skriver är dock ingenting ovanligt på andra webbsidor enligt Goldbeck (2013) som påpekar att Googles e-mailtjänst Gmail sparar användares skrivna e-mails även om användaren stänger ned fönstret. Men medan Google gör detta för att hjälpa användaren, så förväntar sig inte användare på Facebook att deras oskickade inlägg sparas, och användaren tjänar inte heller någonting på det (Goldbeck, 2013).

All data som skickades tillbaka till Facebook var anonymiserad, trots det har det skapat värde för Facebook då Das och Kramer (2013) i sin studie fått många resultat gällande självzensurering. De kom bland annat fram till att män självzensurerar mer än kvinnor, inlägg självzensureras mer än kommentarer och att de som har mer kontroll över vilka som ser ens inlägg genom begränsningsverktyg självzensurerar mindre (Das & Kramer, 2013).

2.3.3 *Spårningsmetoder*

När användare surfar på nätet och interagerar med olika webbsidor så observeras de av både webbsidorna som samlar in data direkt men även av tredjepartsaktörer som utnyttjar olika spårningsmetoder vilka ofta är dolda i bakgrunden utan att användaren kan se dem

(Englehardt & Narayanan, 2016). Tredjepartsaktörer kan samla in användares webbhistorik genom att kombinera cookies med andra spårningsmetoder som gör det möjligt för dem att identifiera användaren och ta del av information om användarens aktivitet som t.ex vilken webbsida användaren besöker just nu (Englehardt & Narayanan, 2016).



Figur 2.1: Organisationer med flest aktiva tredjepartsaktörer vid en analys av en miljon populära webbsidor (Englehardt & Narayanan, 2016, s.8).

En av dessa spårningsmetoder är en spårningspixel, som ofta är en minimal transparent bild, och när en person besöker webbsidan så hämtas bilden från en server och sedan registrerar att användaren har använt webbsidan (Bylund, 2013). Ännu ett sätt att spåra användare är via så kallad “fingerprint tracking” som inträffar när en användare besöker en webbsida, så sänder dess uppkopplade enhet väldigt specifik information om sitt system och inställningar till webbsidan (Englehardt & Narayanan, 2016). Det går även att blanda olika tekniker som automatiserat lyckas samla in spårdata som har blivit raderad, och dessa kallas “zombiekaka” eller “superkaka” (Bylund, 2013). Data som samlas in via dessa spårningsmetoder kan sedan användas till bland annat riktad reklam, prisdiskriminering, för att bedöma kreditvärdighet samt att bedöma hälsotillstånd (Bujlow et al., 2015).

Facebook spårar användares musrörelser, och företaget håller även koll på om användaren har igång tjänsten i förgrunden eller bakgrunden, avslöjade deras VD Mark Zuckerberg (Senate Commerce, 2018). Anledningen till denna spårning är att det hjälper till att avgöra om användaren är en människa eller om det är en robot (Senate Commerce, 2018). Det avslöjades även att företaget via sina partnerskap och sina affärsverktyg tillhandahåller information om användare utanför Facebook, där det inte spelar någon roll om en användare är medlem på Facebook eller inte (Senate Commerce, 2018). Information som de tillhandahåller är exempelvis användares enhetsinformation, vilka köp som görs och vilka webbsidor som användare besöker (Senate Commerce, 2018).

Facebook har även utvecklat ett eget analysverktyg kallat Facebook-pixeln, som är en liten remsa JavaScript-kod som gör det möjligt att övervaka och spåra besökare på en webbsida som har implementerat pixel-teknologin (Facebook Developers, 2019). Pixeln sitter på olika webbsidor och är ihopkopplad med Facebook Ads, som gör det möjligt för företag som

annonserar att spåra och se vilka handlingar som användare gör som ett resultat av annonseringen (Tikno, 2017). Detta kan sedan länkas ihop med Facebook och användas för marknadsföring på Facebook genom att rikta personlig reklam till användare om produkter, som exempelvis en besökare tidigare varit i kontakt med på deras webbsida och hjälper ägaren av Facebook-pixeln att mäta effektiviteten på sina annonser till följd av deras Facebook-annonser (Facebook Business, 2019a). Beroende på vilken sida och vart pixeln placeras så går det att mäta och logga olika aktiviteter, exempelvis vid en kassa-sida som kan mäta köp eller vid en registreringssida som kan mäta registrering (Gordon et al., 2018). Detta gör att både företaget som har annonserat och Facebook kan spåra användare, och det går även om användaren inte har klickat på själva reklamen, utan bara besökt sidan (Gordon et al., 2018).

Ett annat analysverktyg som Facebook har utvecklat är Facebook Analytics, som bygger på data från över två miljarder människor i Facebooks Community (Facebook Analytics, 2018). Facebook Analytics arbetar i samverkan med Facebook-pixeln och sträcker sig över flera kanaler, vilket innebär att de företag som nyttjar tjänsten kan spåra användares hela aktivitetsförlopp och se deras resa från exempelvis Facebook-sida, till webbsida och sedan till företagets app (Heitzman, 2017). Om en användare har besökt en produkt genom exempelvis en app men har använt sig utav en webbsida på dator för att sedan köpa den, så sparas alla dessa aktiviteter som användaren utför i dessa kanaler (Heitzman, 2017).

Fram till för några år sedan kunde marknadsförare som använde sig av Facebook och Google endast inrikta sig på olika individers profiler som tagits fram helt och hållet baserat på deras beteende online (Christl, 2017). År 2012 började Facebook tillåta företag att ladda upp listor över mailadresser och telefonnummer från deras databaser till Facebooks plattform, mer specifikt till Facebook Custom Audiences (Facebook Business, 2019b).

Trots att dessa mailadresser och telefonnummer konverteras till pseudonym kod vid överföringen så har detta gett Facebook möjligheten att göra en direkt koppling mellan företagets kunder och kundernas Facebook-konto (Christl, 2017). Med hjälp av detta kan även företag på ett systematiskt sätt koppla samman sin egna data om sina kunder med data som Facebook tillhandahåller (Christl, 2017). Företag kan ta del av väldigt specifik information om kunders beteende, som deras besök på hemsidor eller olika typer av transaktioner i realtid och omedelbart vidarebefordra detta till Facebook - som i sin tur direkt kan anpassa sin tjänst baserat på informationen för att exempelvis visa personligt anpassad reklam (Christl, 2017).

2.3.4 Sociala plug-ins

Åtskilliga hemsidor har implementerat Facebook Like-knappen för att låta medlemmar på Facebook dela med sig av sina intressen vilket har gjort att knappen blivit ett viktigt affärsverktyg för företag som levererar och distribuerar digital innehåll (Roosendaal, 2011). I en artikel från 2011 framgår det att Like-knappen också fungerar som ett verktyg för att placera cookies på användares datorer, vare sig de använder sig av knappen eller ej - användaren behöver inte ens vara medlem på Facebook för att data om dem ska kunna samlas in då aktiviteten hos icke-medlemmar kan sparas i ett separat dataset som sedan kopplas till individen när hen går med i tjänsten (Roosendaal, 2011).

En studie utförd på University of Cambridge visar att det är möjligt att förutspå attribut som etnicitet, religiös och politisk syn, relationsstatus, kön, sexuell läggning samt personlig

konsumtion av alkohol och cigaretter endast genom att analysera Facebook Likes (Kosinski et al., 2013).

Analysen baserades på data från 58,466 användare från USA som deltog i en enkätundersökning och frivilligt lämnade demografisk information via en applikation kallad "my-Personality" (Kosinski et al., 2013). Denna applikation bestod av ett frågeformulär om användarens personlighet och analyserade även användarens "likes" på Facebook. Med hjälp av denna applikation lyckades forskarna förutspå känsliga personliga attribut på ett träffsäkert sätt, genom att utgå från cirka 170 "likes" per Facebook-användare (Kosinski et al., 2013).

Förutsägelser om individers attribut från digitala uppgifter kan ha negativa implikationer eftersom de på ett enkelt sätt kan appliceras på en stor mängd olika människor utan att dessa människor har lämnat samtycke eller ens har vetskap om att deras information används i detta syfte (Kosinski et al., 2013).

2.3.5 Hur informerade är användare om datainsamling

Nowak och Phelps (1997) identifierar och kategoriserar hur informerade konsumenter är om datainsamling som antingen fullt informerad om insamlingen och att datan kan användas, att konsumenten känner till insamlingen men inte att datan används eller att konsumenten är helt oinformerad om både insamling och användningen av data.

Hotet mot konsumentens integritet beskrivs som minimalt för fullt informerade konsumenter men att det blir allt större och har fler negativa konsekvenser i fallet där konsumenten är oinformerad om vilken data som samlas in och vad den används till för syfte (Graeff & Harmon, 2002). Konsumenter uppfattar att marknadsförare och företag som använder sig av personlig data inte lägger stor vikt vid konsumentens integritet och uppskattar generellt inte att för mycket av deras personliga data samlas in (Graeff & Harmon, 2002). Däremot anser många konsumenter att, trots att de tvingas dela med sig av för mycket personlig data, de på detta sätt kan få bättre anpassade tjänster eller mer relevant reklam och nöjer sig med detta som compensation för att dela sin data (Phelps et al., 2000).

Vissa konsumenter bekymrar sig över att återförsäljare säljer vidare och tjänar pengar på information om deras köpmönster till andra marknadsförare (Graeff & Harmon, 2002). Många konsumenter upplever en ökad brist på kontroll över sin persondata och har en begränsad insyn i hur datan används, vilket i sin tur gör att många blir oroade över att tredjepartsaktörer kan få tag i deras data, och möjligheten att deras data kan nyttjas på ett sätt än vad det tidigare var avsett för (Kshetri, 2014). Hur problematiskt datainsamling om konsumentens personliga information och köpmönster uppfattas beror till stor del på vilken sorts data som samlas in, hur den kommer användas och vilken sorts compensation som erbjuds (Graeff & Harmon, 2002).

2.4 Teoretiskt resultat

I denna uppsats kommer vi bedöma hur informerade användare är om datainsamling utifrån de tre nivåer som Nowak och Phelps (1997) beskriver. Nowak och Phelps (1997) baserar

dessa tre nivåer på sitt arbete om hur informerade konsumenter är om datainsamling och nivåerna ser ut på följande sätt:

- Fullt informerad - användaren är informerad om att data samlas in och att den utnyttjas
- Halvt informerad - användaren är informerad om att data samlas in men inte att den utnyttjas
- Oinformerad - användaren känner varken till att data samlas in eller att den utnyttjas

Genom att kombinera dessa tre nivåer av hur informerade användare är om datainsamling (Nowak & Phelps, 1997) med Richter och Kochs (2008) definition av basfunktioner i sociala medier och att data samlas in (Christl & Spiekermann, 2016; Joler & Petrovski, 2016; Bujlow et al., 2015) samt att data används (Bylund, 2013; Larsson & Ledendal, 2017), så kan vi utvärdera till vilken grad användare är informerade om datainsamling på sociala medier.

För att en användare ska klassificeras som fullt informerad (Nowak & Phelps, 1997) så ska användaren känna till att det samlas in data på sociala medier (Richter & Koch, 2008; Christl & Spiekermann, 2016; Joler & Petrovski, 2016; Bujlow et al., 2015) samt att datan utnyttjas (Bylund, 2013; Larsson & Ledendal, 2017). Richter och Kochs (2008) definition av basfunktionerna inom sociala medier utgår ifrån att användarna kan utföra olika handlingar, så som att skriva i ett sökfält eller direkt och indirekt informationsutbyte, som i sin tur samlar in data (Christl & Spiekermann, 2016; Joler & Petrovski, 2016; Bujlow et al., 2015). Bylund (2013) och Larsson och Ledendal (2017) berättar även konkret att denna data kan säljas vidare som beteendedata och som riktad marknadsföring.

En användare som klassificeras som halvt informerad (Nowak & Phelps, 1997) ska känna till att datainsamling på sociala medier äger rum (Richter & Kochs, 2008; Christl & Spiekermann, 2016; Joler & Petrovski, 2016; Bujlow et al., 2015) men behöver inte nödvändigtvis vara informerad om att datan utnyttjas.

För att en användare skall klassificeras som oinformerad utefter Nowak och Phelps (1997) tre nivåer av till vilken grad användare är informerade så skall användaren varken känna till att datan samlas in eller används.

2.5 Litteraturtabell

I kapitel 2.5 så har det skapats en litteraturtabell för att ge en överblick över litteraturgenomgången samt förtydliga vilken litteratur som tillhör vilket område. Litteraturtabellen skall även gå att återkopplas till enkätundersökningen och beskrivs som Tabell 2.1.

Tabell 2.1: Litteraturtabell

Område	Behandlar	Referenser
Sociala medier	Definition av sociala medier Sociala mediernas användningsområde från olika perspektiv	Merriam Webster (2019) Statista (2019a) Statista (2019b) Statista (2019c) Boyd & Ellison (2007)

		Richter & Koch (2008) Bylund (2013) Whiting & Williams (2013) Asur & Huberman (2010) Baird & Parasnis (2011) Facebook (2016) Bechmann (2014) Appelgren & Leckner (2016)
Beskrivning av data	Definition av data Olika typer av data och hur dessa typer kan grupperas	Zinz (2007) Larsson & Ledendal (2017) Datainspektionen (u.å.) Christl & Spiekermann (2016) Busby et al. (2012) Cooper et al. (2016)
Datainsamling på sociala medier	Förklarar direkt och indirekt datainsamling på sociala medier med fokus på Facebook Datainsamling om aktiviteter och beteende	Bylund (2013) Bujlow et al. (2015) Tuunainen et al. (2009) Christl (2017) Statista (2019b) Facebook Data Policy (2018) Joler & Petrovski (2016) Senate Commerce (2018)
Cookies	Behandlar dataextrahering med hjälp av cookies	Christl & Spiekermann (2016) Bujlow et al. (2015) Roesner et al. (2012)
Själv censurering	Behandlar fenomenet själv censurering på sociala medier	Das & Kramer (2013) Goldberg (2013)
Spårningsmetoder	Beskriver olika spårningsmetoder Tar upp att musrörelser & förgrund/bakgrund loggas Tar upp pixel-teknologi Tar upp analytics-funktioner Facebooks & företags gemensamma spårning av användare	Englehardt & Narayanan (2016) Bylund (2013) Bujlow et al. (2015) Senate Commerce (2018) Facebook Developers (2019) Tikno (2017) Facebook Business (2019a) Gordon et al. (2018) Facebook Analytics (2019) Heitzman (2017) Christl (2017) Facebook Business (2019b)
Sociala plug-ins	Sociala plug-ins som verktyg för datainsamling Förutsägelser om känsliga privata attribut	Roosendaal (2011) Kosinski et al. (2013)

Hur informerade är användare om datainsamling på sociala medier	Behandlar till vilken grad användare är informerade om datainsamling samt problematik kring datainsamling	Nowak & Phelps (1997) Graeff & Harmon (2002) Phelps et al. (2000) Kshetri (2014)
---	---	---

3 Metod

3.1 Kvantitativ metod

Vi kommer att göra en kvantitativ analys av området baserat på en enkätundersökning. Då syftet med uppsatsen är att ta reda på hur informerade användare är angående datainsamling på sociala medier, så anser vi att vi skall söka så många svar som möjligt för att få en så bred uppfattning som möjligt. Detta stärks av Jacobsen (2002) som anser att en undersökning som skall få många svar är bäst lämpad att utföra med den kvantitativa metoden. Eftersom att den information som författarna ämnar att samla in också är definierad i förväg så lämpar sig den kvantitativa metoden väl, påpekar Jacobsen (2002). Tidigare forskning om vilken vetenskap som användare har om datainsamling på sociala medier, så som Tuunainen et al. (2009) använder sig även av enkätundersökning, vilket ytterligare påvisar att detta är en metod som går att använda inom området.

Denna enkät kommer att läggas ut och spridas på Facebook, och kommer att delas av författarna i deras flöden, med möjlighet för andra användare att dela vidare enkäten för att få en större spridning. Enligt Jacobsen (2002) är en av syftena till en kvantitativ undersökning just att få en stor spridning och många svar, vilket vi då hoppas uppfylla. Anledningen till att enkäten endast kommer att delas över Facebook, är att frågorna är riktade till facebookanvändare och att det då riktas till rätt urval. Hade enkäten spridits till personer som inte är facebookanvändare eller medlemmar på sociala medier så skulle dessa svar räknas med som alla andra, men författarna anser att de svar som skulle lämnas inte skulle vara relevanta för frågeställningen.

Enkäten har därför utformats så att de som deltar i enkäten kommer att få svara på frågor och påståenden, där varje svarsalternativ är kopplat till en av de olika graderna av hur informerade användare är baserat på Nowaks och Phelps (1997) skala. Skalan är i tre olika nivåer av hur informerade användare är, alltså är alltid ett svarsalternativ kopplat till fullt informerad, ett svarsalternativ till halvt informerad och ett svarsalternativ till oinformerad. Då svarsalternativet för oinformerad ansågs vara för bestämt av författarna, det vill säga att en enkättagare kanske inte kände sig bekväm eller säker på att svara på det, så lades även ett fjärde svarsalternativ till, som genomgående hade namnet "Vet inte.". Jacobsen (2002) påpekar att detta kan vara bra då det inte är en större poäng att tvinga deltagarna att ta ställning i frågor som de inte har någon kunskap om. Detta indikerar även att den som svarar inte vet, och kopplas alltid till nivån oinformerad.

Att göra en kvantitativ enkät standardiserad är enligt Jacobsen (2002) viktig, då det tvingar in enkättagarna i de på förhand definierade kategorierna, och alla deltagare behandlas

likadant. Undantagen i detta är fråga ett och två, som båda är till för att kategorisera målgrupp, och för att få en överblick över respondenterna.

3.2 Enkätutformning

Då Nowak och Phelps (1997) nivåer handlar om hur informerade användare är kring datainsamling, så är syftet med enkäten att ta reda på hur informerade enkätdeltagarna är utefter dessa nivåer. Jacobsen (2002) berättar att det inte är bestämt om det är bäst med frågor eller påståenden och utefter detta så har mest påståenden använts i denna enkät, då författarna ansåg att detta passade in bäst i hur svarsalternativen blev utformade och kopplade mot Nowak och Phelps (1997) teori. Författarna valde att ha liknande svarsalternativ på de flesta frågor för att skapa minimal förvirring för uppgiftslämnarna, men någonting som istället då kan uppkomma är en så kallad "ja eller nej-effekt", där uppgiftslämnare kommer in i en rytm och väljer svar efter vana (Jacobsen, 2002). Denna effekt försökte författarna minimera genom att på en fråga ändra riktningen, och ansåg att det räckte sett till antalet frågor i enkäten. Jacobsen (2002) som säger att en kvantitativ ansats måste vara standardiserad, något som författarna försökte uppnå med hjälp av standardiserade svarsalternativ. I början av enkäten så fick deltagarna exempel på vad data kan användas till, så att de lättare skulle förstå vad det innebar i svarsalternativet.

Enkäten skapades i Google Forms, vilken även hjälper till med grafisk statistik över enkäten, något som var lockande för författarna. Enkäten började med att ställa "ofarliga" frågor till deltagarna, om deras ålder och hur ofta de använder Facebook, dels för att få en överblick över vårt urval, kategorisera målgrupp och för att se skillnader i hur informerade användare är beroende på användningsfrekvens. Jacobsen (2002) påpekar att denna sortens frågor är bra att inleda med för att undvika misstänksamhet som sedan kan resultera i att de inte fyller i formuläret. Författarna utformade påståendena mot syftet för att hitta frågor som ansågs relevanta för att uppfylla syftet. Detta gjordes med koppling från litteraturgenomgången. När författarna utformat påståendena så testades de mot en liten grupp för att se om det fungerade, en metod som Jacobsen (2002) anser är bra att använda för att se om enkäten fungerar. Enkäten togs emot väl av den lilla gruppen men författarna insåg att några små justeringar behövde göras för att göra det tydligare, som sedan verkställdes. I tabell 3.1 visas enkäten i sin helhet med de olika frågorna/påståendena med frågenummer, vilket område som frågan tillhör samt en tillhörande motivering till frågan. I tabell 3.2 visas svarsalternativ för fråga 3-6, 8-10. I tabell 3.3 visas svarsalternativ till fråga 7.

Tabell 3.1: Enkätutformning

Frågenummer:	Område	Fråga/påstående	Motivering
1.	Personlig	Hur gammal är du? (Textsvar)	För att få en överblick hur urvalet ser ut.
2.	Sociala medier	Hur ofta använder du Facebook? (1-5)	Kategorisering av målgrupp. För att se om användningsfrekvens påverkar hur informerad användaren är.

3.	Datainsamling / Sociala plug-ins	Det krävs att du har ett Facebook-konto för att Facebook ska kunna samla in data om dig. (1-4)	För att se om enkättagarna är informerade om att Facebooks datainsamling även kan ske utanför Facebook.
4.	Datainsamling	Facebook kan samla in data om dig när du surfar på andra webbsidor än facebook.com. (1-4)	För att se om enkättagarna är informerade om att Facebooks datainsamling även kan ske utanför Facebook.
5.	Datainsamling	Facebook kan samla in data om andra uppkopplade enheter på samma nätverk som du är uppkopplad på.(1-4)	För att se om användaren känner till att Facebook samlar data om enheter som är uppkopplade på samma nätverk.
6.	Cookies / Sociala plug-ins	En webbsida som använder sig av Facebook Like-knappen kan skicka data om din aktivitet till Facebook utan att du använder Facebook Like-knappen. (1-4)	För att bedöma om användaren förstår att cookies kan placeras i webbläsaren utan att Like-knappen nyttjas.
7.	Självcensurering	Kommentarer och inlägg som du skriver på Facebook men raderar innan de är skickade, är helt borttagna. (1-4)	För att se om enkättagarna vet att inlägg och kommentarer som skrivs på Facebook inte skickas fortfarande kan sparas.
8.	Spårningsmetoder	Facebook kan spåra dina musrörelser på skärmen när du använder Facebook via din dator. (1-4)	Är användaren informerad om denna spårningsmetod?
9.	Spårningsmetoder	Facebook kan bedöma om du har Facebook igång i bakgrunden eller förgrunden (minimerad/maximerad). (1-4)	Är användaren informerad om denna spårningsmetod?
10.	Spårningsmetoder	Företag kan spåra dig på internet via ditt Facebook-konto när du är utloggad från tjänsten. (1-4)	Känner användaren till att Facebook spårar deras aktivitet även när de är utloggade?

Tabell 3.2: Tabell över värden i enkäten

Fullt informerad	Ja, och datan används
Halvt informerad	Ja, men datan används inte
Oinformerad	Nej, data samlas inte in och används inte.
Oinformerad	Vet inte.

Tabell 3.3: Tabell över värden i enkäten

Fullt informerad	Nej, de sparas och datan används.
Halvt informerad	Nej, de sparas men datan används inte.
Oinformerad	Ja, datan är helt borttagen.

Oinformerad	Vet inte.
-------------	-----------

3.3 Urval och spridning av enkät

Då syftet med uppsatsen är att ta reda på hur informerade användare på sociala medier är om olika datainsamlingsmetoder som sker, så valde författarna att sprida enkäten på det största sociala nätverket Facebook. Författarna började med att skicka ut den första versionen av enkäten till en liten grupp bekanta för att testa den och se om den fungerade, något som Jacobsen (2002) påpekar är bra för att se om någonting behöver ändras innan enkäten används i verkligheten. Sedan så spred författarna enkäten i sina personliga Facebook-flöden och nådde ut till sina nätverk, med uppmaning om att det var fritt att dela vidare och sprida enkäten ytterligare. Till sist så skickades även enkäten med uppmaning om att göra den i chattgrupper på Facebook. Detta var ett tillvägagångssätt som författarna ansåg lämpligt inom den tidsram som är satt på uppsatsen.

Anledningen till att enkätens frågor var riktade mot Facebook i synnerhet, och endast delades på Facebook är för att författarna anser att Facebook är representativt för sociala medier. Ett skäl till detta är att Facebook är störst inom sociala medier (Statista, 2019b). Författarna anser även att Facebook uppfyller de sex basfunktioner som är karaktäriserande för sociala medier som Richter och Koch (2008) beskriver. Boyd och Ellison (2007) påstår att den tekniska aspekten inte skiljer sig markant mellan olika sociala medier, utan endast de kulturella delarna. Men då Facebook är störst inom sociala medier (Statista, 2019b) så anser vi att de kulturella skillnaderna som Boyd och Ellison (2007) nämner är minimala, så är det rimligt att anta att de flesta kulturer är representerade på Facebook.

Ännu en anledning till att enkäten inte spreds på andra sociala medier är att frågorna/påståendena är riktade mot just Facebook, och att ingen av författarna har ett annat konto på någon annan sociala medier-plattform och författarna bedömer då att det inte skulle bli någon större spridning, eller nå ut till en annan publik än den som redan nås på Facebook. Författarna ansåg vidare att det inte skulle hinnas med att bygga upp profiler och skaffa ett nätverk på andra sociala medier-plattformar på tiden som uppsatsen skrevs. Huvuddelen av de som svarade på enkäten är förmodligen runt samma ålder som författarna, 25-30 år, då författarnas nätverk till största del består av människor i samma ålder. Däremot så kommer det säkert vara blandade åldrar, då båda författarna har spridda åldrar i sina nätverk.

För att få en hög svarsfrekvens på alla frågor, så bedömde författarna att varje fråga skulle vara obligatorisk. Jacobsen (2002) påpekar att ett urval inte bör underskrida 100 enheter som en tumregel, då det både gör det svårt att analysera informationen och felmarginalen blir stor. Författarna hade detta i åtanke och hoppades att få så många svar som möjligt, men valde att stänga enkätundersökningen vid 101 svar, detta för att lyckas studera resultatet inom den givna tidsramen.

3.4 Analysmetod

Frågorna som enkättagarna kan svara på är skapad med svarsalternativ utefter Nowak och Phelps (1997) nivåer om hur informerade användare är om datainsamling. För att ha möjlighet att bedöma vilken nivå på skalan som en användare tillhör, så operationaliserades begrepp i ett tidigare skede, vilket i sin tur innebar att författarna kan göra en förhandskategorisering och sedan kunna standardisera informationen i sifferform. Författarna är ute efter att få en effektiv statistisk analys för att kunna utröna resultat, något som operationalisering och förhandskategorisering möjliggör enligt Jacobsen (2002).

Då enkäten är gjord i Google Forms, så autogenereras en grafisk sammanfattning i diagram, vilken kommer att användas för att presentera resultaten, och är enligt Jacobsen (2002) ett vanligt tillvägagångssätt att presentera insamlad empiri. Författarna kommer sedan att ha möjligheten att utläsa och kunna presentera väldigt precist över hur många som har svarat på de olika svarsalternativen och därmed kunna utröna hur många som tillhör respektive nivå av informerad om de olika datainsamlingsmetoderna.

3.5 Validitet och Reliabilitet

Syftet med en undersökning är att samla in empiri och för att denna empiri ska vara användbar krävs det att den är giltig och relevant samt tillförlitlig och trovärdig (Jacobsen, 2002). För att empirin ska vara relevant krävs det att vi mäter det vi faktiskt önskar att mäta och att datan som samlas in också har relevans i andra sammanhang (Jacobsen, 2002).

För att undvika att ställa ledande frågor och göra en undersökning som genererar relevant data är det viktigt att utforma frågeformuläret på ett korrekt sätt och se till att frågorna ställs på ett neutralt sätt utan att ett förutbestämt ställningstagande har gjorts (Jacobsen, 2002). Enligt Jacobsen (2002) är det även viktigt att se till att enkäten inte innehåller frågor där deltagaren uppfattar att det finns ett rätt eller fel svar, vilket är något vi har försökt uppfylla genom att ställa neutrala frågor som inte ger deltagaren möjlighet att misstänka att vi är ute efter ett specifikt svar. Vi har även med svarsalternativet "Vet inte" för att undvika att deltagarna väljer något av de andra alternativen trots att de inte vet svaret på frågan eller påståendet. Vi behandlade svarsresultaten på enkätundersökningen omedelbart efter att undersökningen avslutades för att undvika att det gick för lång tid mellan insamlingen och analysen av datan.

En kvantitativ undersökning kräver enligt Jacobsen (2002) både en god intern och extern giltighet. Intern giltighet innebär att det som vi ämnar att undersöka faktiskt undersöks, vilket vi har försökt att göra genom konkreta påståenden som gör det lätt för deltagarna att förstå vad vi är ute efter. För att uppnå en god extern giltighet krävs det att den insamlade datan kan appliceras i andra liknande sammanhang och att resultatet går att generalisera i någorlunda hög utsträckning (Jacobsen, 2002). Med utgångspunkt i litteraturtabellen (Tabell 2.1) har en hög validitet försökt uppnås genom att utforma påståendena på ett sätt som gör det möjligt att mäta det som är giltigt och relevant för uppsatsen. För att lyckas skapa en studie som är trovärdig och tillförlitlig skriver Jacobsen (2002) att den måste vara genomförd på ett trovärdig

sätt, undersökningen ska gå att lita på och den ska inte innehålla några mätfel. En av påståendena i enkäten hade annorlunda svarsalternativ jämfört med de andra påståendena (se Tabell 3.2 och Tabell 3.3) men dessa svarsalternativ kopplades till teorin på ett annat sätt jämfört med svarsalternativen i de andra frågorna för att undvika mätfel.

3.6 Etik

Vår empiri samlas in genom en enkätundersökning som har utförts på ett sätt som behåller tillförlitligheten av undersökningen genom att ta hänsyn till de tre grundförutsättningar som Jacobsen (2002) nämner att en undersökning bör innehålla och dessa grundförutsättningar är: informerat samtycke, krav på privatliv och krav att bli korrekt återgiven.

Informerat samtycke innebär enligt Jacobsen (2002) att respondenterna frivilligt deltar i undersökningen och att de är medvetna om möjliga risker och vinster deras deltagande kan innebära. I enkätens början så beskrivs syftet med enkäten så att deltagarna var medvetna om varför enkätundersökningen gjordes, och kommer därmed kunna bedöma riskerna och vinsterna i sitt eget deltagande. Enkätundersökningen är anonym då enskilda svar ej går att koppla till en specifik respondent, vilket även styrks då det är fullt möjligt för en respondent att göra enkäten vid vilket tillfälle som önskas, diskret över internet och inte i en fysisk miljö som skulle kunna medföra en viss påtryckning.

Enligt Jacobsen (2002) så innebär krav på privatliv hur stor risk det finns att respondenterna kan bli identifierad utifrån datan som samlas in, och även om hur privat och känslig den insamlade datan är. Eftersom vi inte är intresserade av någon personlig information förutom ålder bedöms dessa potentiella risker för respondenterna som små. Då frågorna handlar om hur informerade användare är om datainsamling, så anser vi att detta inte heller är någon speciell känslig eller privat data som samlas in. Undantaget som författarna ser till detta skulle vara om någon enkättagare inte skulle vilja visa sig som oinformerad eller om enkättagarna anser det vara personligt då de förmodligen nås av enkäten via sin privata facebook-sida. Jacobsen (2002) påtalar att risken för identifikation av en deltagare blir lägre om det är få kvalitativa svarsalternativ, vilket enkätundersökningen undgår genom att använda sig utav standardiserade svarsalternativ.

Krav på att bli korrekt återgiven är enligt Jacobsen (2002) att datan som samlas in skall stanna inom sammanhanget och inte modifieras så att den fortfarande representerar åsikterna som respondenterna har. I författarnas studie så blir diagram och tabeller sammanställda och presenteras för varje fråga som har ställts så att det blir enkelt att se hur många respondenter som har svarat på vilken fråga.

3.7 Felkällor i metod

Vid utförandet av en enkätundersökning finns det risk att stöta på olika felkällor som kan leda till att svaren kan bli irrelevanta för undersökningen (Weisberg, 2005). En av de vanligaste felkällorna är att respondenterna inte representerar den population som enkätundersökningen ämnar att granska (Weisberg, 2005). Denna felkälla har försökts undvikas genom att sprida

enkäten på det sätt som beskrivs i 3.3 vilket innebär att enkäten förhoppningsvis träffat rätt deltagare och kan således ses som representativ för rampopulationen.

Då enkäten inte krävde att deltagarna loggade in eller på något sätt kopplades till individen som valt att svara var det möjligt för en deltagare att fylla i enkäten flera gånger. Det är svårt att säkerställa att detta inte skett utan att ställa krav på inloggning, men då det inte finns någon anledning för deltagarna att sabotera undersökningen genom att medvetet ge felaktiga svar eller svara flera gånger så valdes alternativet som saknade krav på inloggning, eftersom detta alternativ gjorde enkäten mer lättillgänglig vilket också innebär att undersökningen kunde få en högre svarsfrekvens.

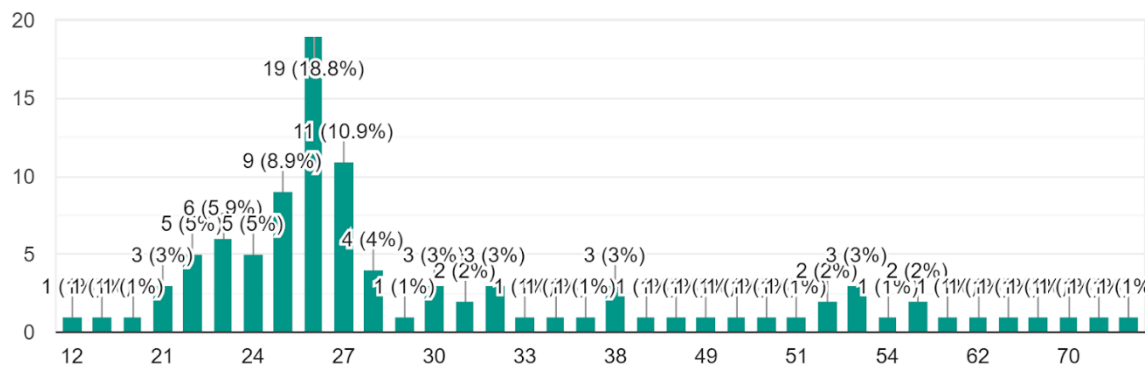
En ytterligare felkälla som kan uppkomma är olika typer av mätfel, som kan bero på att frågorna eller påståendena inte är tillräckligt tydligt utformade vilket öppnar upp för respondenterna att göra egna tolkningar och således kan svaren på undersökningsfrågorna förlora betydelse och giltighet (Weisberg, 2005). För att undvika denna felkälla utformades frågorna på ett standardiserat sätt där varje svarsalternativ hade en klar koppling till den teori som används vid analysen av empirin (Tabell 3.2, Tabell 3.3).

4 Resultat

I detta kapitel så kommer författarna att presentera de svar som de har tillhandahållit i enkätundersökningen. Enkättagarnas svar kommer att presenteras i diagram för att på så sätt göra det lätt att utläsa resultaten. Denna enkät var tillgänglig att svara på i fem dagar och den totala svarmängden var 101. Varje fråga i enkäten var obligatorisk, vilket har gjort att det är lika många svar på varje fråga. För att det skall vara lätt att utläsa resultatet, så är viss text inte medtagen i diagrammet, som exempelvis alla olika åldrar eller vissa procentsatser. I vissa fall så förklaras dessa procentsatser under det presenterade diagrammet, i andra fall så hänvisar vi till den fullständiga enkäten i appendix.

1. Hur gammal är du?

101 responses

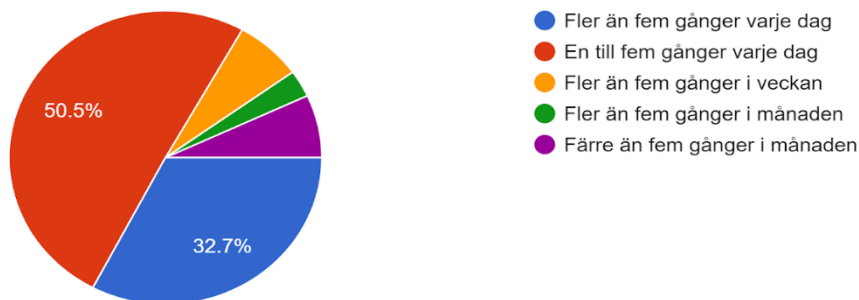


Figur 4.1: Svar på enkätfråga 1

Figur 4.1 illustrerar åldersfördelningen hos de personer som svarat på enkätundersökningen. Det går att utläsa att respondenterna befinner sig i ett åldersspann mellan 12-77 år, där 26-åringar representerar den största gruppen på 18,8% av respondenterna, följt av 27-åringar med 10,9% och 25-åringar med 8,9%.

2. Hur ofta använder du Facebook?

101 responses



Figur 4.2: Svar på enkätfråga 2

Figur 4.2 illustrerar användningsfrekvensen av Facebook hos respondenterna som har deltagit i enkätundersökningen. 32,7% av respondenterna använder Facebook fler än fem gånger varje dag och 50,5% använder Facebook en till fem gånger dagligen. 6,9% använder Facebook fler än fem gånger i veckan, 3% fler än fem gånger i månaden och till sist är det 6,9% som använder Facebook färre än fem gånger i månaden.

3. Facebook kan samla in data om dig utan att du har ett Facebook-konto.

101 responses



Figur 4.3: Svar på enkätfråga 3

Figur 4.3 illustrerar respondenternas svar på fråga 3, "Facebook kan samla in data om dig utan att du har ett Facebook-konto". 49,5% svarade "Ja, och datan används.", 3% svarade "Ja, men datan används inte.", 14,9% svarade "Nej, data samlas inte in och används inte." samt 32,7% svarade "Vet inte."

4. Facebook kan samla in data om dig när du surfar på andra webbsidor än facebook.com.

101 responses



Figur 4.4: Svar på enkätfråga 4

Figur 4.4 illustrerar respondenternas svar på fråga 4, "Facebook kan samla in data om dig när du surfar på andra webbsidor än facebook.com.". 63,4% svarade "Ja, och datan används.", 2% svarade "Ja, men datan används inte.", 6,9% svarade "Nej, data samlas inte in och används inte" och 27,7% svarade "Vet inte."

5. Facebook kan samla in data om andra uppkopplade enheter på samma nätverk som du är uppkopplad på.

101 responses

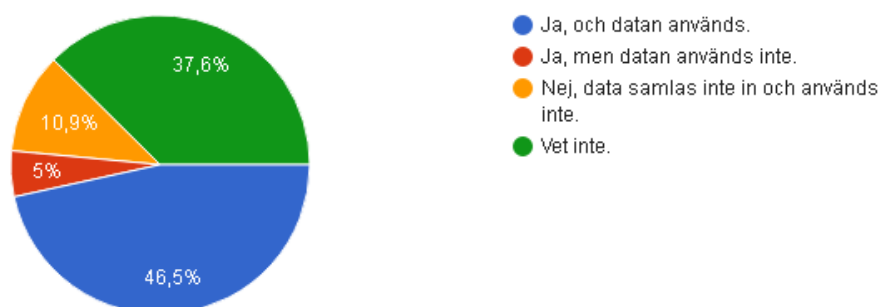


Figur 4.5: Svar på enkätfråga 5

Figur 4.5 illustrerar respondenternas svar på fråga 5, "Facebook kan samla in data om andra uppkopplade enheter på samma nätverk som du är uppkopplad på.". 37,6% svarade "Ja, och datan används.", 1% svarade "Ja, men datan används inte.", 14,9% svarade "Nej, data samlas inte in och används inte." och 46,5% svarade "Vet inte."

6. En webbsida som använder sig av Facebook Like-knappen kan skicka data om din aktivitet till Facebook utan att du använder Facebook Like-knappen.

101 svar

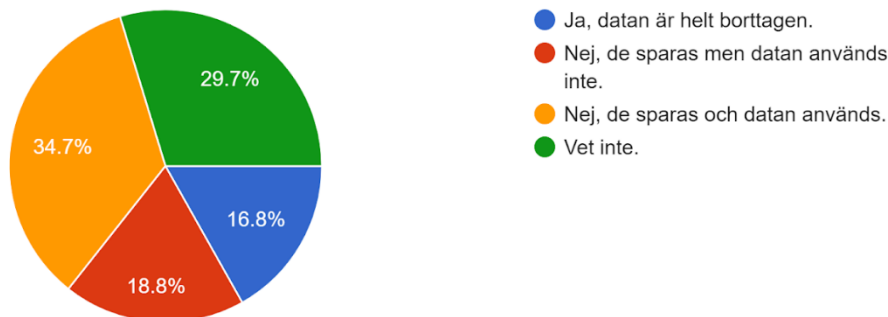


Figur 4.6: Svar på enkätfråga 6

Figur 4.6 illustrerar respondenternas svar på fråga 6, "En webbsida som använder sig av Facebook Like-knappen kan skicka data om din aktivitet till Facebook utan att du använder Facebook Like-knappen.". 46,5% svarade "Ja, och datan används.", 5% svarade "Ja, men datan används inte.", 10,9% svarade "Nej, data samlas inte in och används inte." och 37,6% svarade "Vet inte."

7. Kommentarer och inlägg som du skriver på Facebook men raderar innan de är skickade, är helt borttagna.

101 responses



Figur 4.7: Svar på enkätfråga 7

Figur 4.7 illustrerar respondenternas svar på fråga 7, “Kommentarer och inlägg som du skriver på Facebook men raderar innan de är skickade, är helt borttagna.”. 16,8% svarade “Ja, datan är helt borttagen.”, 18,8% svarade “Nej, de sparas men datan används inte.”, 34,7% svarade “Nej, de sparas och datan används.” och 29,7% svarade “Vet inte.”.

8. Facebook kan spåra dina musrörelser på skärmen när du använder Facebook via din dator.

101 responses



Figur 4.8: Svar på enkätfråga 8

Figur 4.8 illustrerar respondenternas svar på fråga 8, “Facebook kan spåra dina musrörelser på skärmen när du använder facebook via din dator.”. 34,7% svarade “Ja, och datan används.”, 7,9% svarade “Ja, men datan används inte.”, 15,8% svarade “Nej, data samlas inte in och används inte.” och 41,6% svarade “Vet inte.”.

9. Facebook kan bedöma om du har Facebook igång i bakgrunden eller förgrunden (minimerad/maximerad).

101 responses



Figur 4.9: Svar på enkätfråga 9

Figur 4.9 illustrerar respondenternas svar på fråga 9, "Facebook kan bedöma om du har Facebook igång i bakgrunden eller förgrunden (minimerad/maximerad)". 58,4% svarade "Ja, och datan används.", 12,9% svarade "Ja, men datan används inte.", 2% svarade "Nej, data samlas inte in och används inte." och 26,7% svarade "Vet inte."

10. Företag kan spåra dig på internet via ditt Facebook-konto när du är utloggad från tjänsten.

101 responses



Figur 4.10: Svar på enkätfråga 10

Figur 4.10 illustrerar respondenternas svar på fråga 10, "Företag kan spåra dig på internet via ditt Facebook-konto när du är utloggad från tjänsten". 43,6 % svarade "Ja, och datan används.", 5% svarade "Ja, men datan används inte.", 11,9 % svarade "Nej, data samlas inte in och används inte." och 39,6% svarade "Vet inte."

5 Diskussion

5.1 Användningsfrekvens

Åldersfördelningen hos de personer som har svarat på enkätundersökningen var relativt spridd (Figur 4.1). Respondenterna befinner sig i ett åldersspann mellan 12 år och 77 år men en stor del av respondenterna (38,6%) är mellan 25-27 år vilket är enligt våra förväntningar, då enkäten spreds i författarnas personliga nätverk, och majoriteten av de som är i författarnas nätverk befinner sig mellan dessa åldrar.

Utifrån resultatet på fråga två (Figur 4.2) om användningsfrekvens går det att se att en tydlig majoritet av respondenterna använder Facebook en till fem gånger varje dag (50,5%) eller fler än fem gånger varje dag (32,7%). Endast 16,8% av respondenterna valde något av de tre andra svarsalternativen om användningsfrekvens vilket gör att underlaget för att jämföra dessa tre grupper är ganska bristande och en generalisering kan bli svår.

Vid en jämförelse mellan de högst representerade grupperna och hur deras svar ter sig över resten av enkätundersökningen går det att utläsa relativt små skillnader i de flesta av frågorna, vilket kan tyda på att användningsfrekvensen inte påverkar hur informerad en användare är. De frågor som står ut är framförallt fråga sju (Figur 4.7) där 48,5% av de som använder Facebook fler än fem gånger om dagen var bekanta med att kommentarer och inlägg som skrivs men raderas innan de skickas inte blir borttagna helt och hållet, jämfört med gruppen som använder Facebook en till fem gånger dagligen där endast 27,5% av respondenterna kände till att datan inte tas bort helt. Det går även att utläsa en markant skillnad i resultatet från fråga åtta (Figur 4.8) där 48,5% av de som använder Facebook fem gånger eller fler dagligen uppgav att Facebook kan spåra musrörelser på skärmen medan för de som använder Facebook en till fem gånger dagligen var det bara 23,5% som hade kännedom om denna spårningsmetod.

5.2 Datainsamling

5.2.1 *Datainsamling på sociala medier*

Resultatet på fråga tre (Figur 4.3) visade att 49,5% av respondenterna kände till att Facebook kan samla in data om individer utan att de är medlemmar på Facebook, vilket innebär att dessa respondenter kan klassificeras som fullt informerade om denna datainsamlingsmetod. Roosendaal (2011) beskriver att data om individer som inte ännu har gått med i tjänsten kan samlas in och sparas i separat dataset som sedan kan kopplas till individen när hen går med i tjänsten. Från resultatet kan vi utläsa att det var 14,9% av respondenterna som istället trodde att data varken kunde samlas in eller användas och 32,7% som inte var säkra på om insamling av data även pågår trots att ett Facebook-konto saknas. Endast 3% av respondenterna valde svarsalternativet som innebar att data samlas in men att den inte utnyttjas.

När en användare skapar ett Facebook-konto tilldelas användaren ett unikt identifikationsnummer genom att en cookie placeras hos användaren och sparas för att underlätta vid autentisering av framtida besök men även för att användarens aktivitet ska kunna spåras (Roosendaal, 2011). Att ungefär hälften av respondenterna kan klassificeras som fullt informerade

innebär inte nödvändigtvis att det känner till det tekniska tillvägagångssättet för insamling av deras persondata men resultatet kan betyda att många respondenter förstår att Facebook är kapabla av att samla in och spåra användare på sätt som inte framgår särskilt tydligt.

Svarsresultaten från fråga fyra (Figur 4.4) indikerar att en majoritet av respondenterna är informerade om att Facebook har möjlighet att samla in data om deras beteende och aktiviteter från andra webbsidor än facebook.com. Det var 63,4% av respondenterna som valde svarsalternativet att data samlas in och används vilket kan vara intressant att jämföra med de 46,5% som valde samma alternativ på fråga sex (Figur 4.6) eftersom ett vanligt förekommande sätt som Facebook samlar in data från andra webbsidor är via sociala plug-ins, främst genom en inbäddad Facebook Like-knapp (Roosendaal, 2011). Resultatet pekar på att respondenterna är informerade om att Facebooks datainsamling sträcker sig utanför Facebooks domän men att det kan vara svårt att ta del av exakt hur datainsamlingen går till, vilket kan innebära att även om en användare klassificeras som fullt informerad så betyder inte det att användaren tvunget delar med sig av mindre data än en oinformerad användare.

Genom att granska resultatet från fråga fem (Figur 4.5) går det att utläsa att en stor mängd respondenter var osäkra på om Facebook kan samla in data om andra uppkopplade enheter på samma nätverk. 46,5% svarade att de inte visste och 14,9% ansåg att det inte är möjligt för Facebook att samla in data från andra enheter på samma nätverk. Det framgår i Facebooks datapolitik att denna sorts data samlas in och används till olika syften, men utifrån resultatet skulle en anledning till att en majoritet av användarna inte kände till detta kunna vara att få användare läser igenom avtalet innan de lämnar sitt samtycke. Bechman (2014) uttrycker en oro över att användare ombes lämna sitt samtycke på internet via ett klick och det går att diskutera huruvida ett enkelt klick för att godkänna ett avtal kan likställas med den formalitet som ett godkännande ofta innebär när det görs ansikte mot ansikte. Detta överensstämmer med författarnas åsikter och författarna anser att detta kan vara en starkt bidragande faktor till varför 61,4% av användarna i denna fråga kan klassificeras som oinformerade i enlighet med Nowak och Phelps (1997) nivåer.

5.2.2 Sociala plug-ins, cookies och själv censurering

Om vi tittar på resultatet från fråga sex (Figur 4.6) så var det 46,5% som hade svarat att data samlas in och används, vilket gör att de klassas som fullt informerade. På samma fråga var det 48,5% som utifrån sina svar klassas som oinformerade, och de resterande 5% klassas som halvt informerade. Vi kan tydligt se att resultatet är jämnt mellan de som är fullt informerade och oinformerade, och det enda som egentligen sticker ut är att endast 5% klassas som halvt informerade.

Roosendaal (2011) förklarar att webbsidor som har valt att implementera Facebook Like-knappen kan samla in data om de som använder sig utav de webbsidorna genom cookies, och att det även fungerar om en besökare inte använder knappen. För att minska eventuella missförstånd vad denna knapp är för något, så bifogades en bild med ett exempel på hur den kan se ut på en webbsida. Det kan vara så att många tror att en webbsida som de besökt inte alls har haft en koppling till Facebook, och även fastän en social plug-in som Facebook Like-knappen visar att det finns en koppling, så kan det kanske inte uppfattas så av alla som besöker sidan och så kan det vara mycket möjligt att de tror att knappen behövs användas för att data skall samlas in.

Detta kan vara en anledning varför den största delen respondenter utifrån sina svar har klassats som oinformerade. Det är även två svarsalternativ kopplade till klassificeringen oinformerad, där den ena tyder på att respondenten inte vet, och är därmed oinformerad, och den andra där respondenten anser att data inte samlas in eller används, vilket innebär att den är felinformerad, vilket i sin tur innebär oinformerad om hur det egentligen går till. Detta kan också ha varit en bidragande faktor till att en sådan stor andel blir klassificerade som oinformerade. Det var nästan lika många som klassades som fullt informerade, och en anledning till detta kan vara att de själva kunnat härleda att om sociala plug-ins från Facebook finns på en webbsida så kan de även ha med ett finger i spelet och därmed ta del av trafiken som sker på webbsidan.

När vi tittar på resultatet i fråga sju (Figur 4.7) så ser vi att 34,7% har svarat att kommentarer och inlägg som skrivs sparas och används, vilket gör att de klassas som fullt informerade. Det var 29,7% som svarade att de inte visste, vilket klassar dem som oinformerade och 16,8% som tror att datan blev helt borttagen, vilket även klassar dem som oinformerade. 18,8% av respondenterna klassas som halvt informerade då de trodde att datan sparas men inte används.

Något som kan ha påverkat svaren till den här frågan är att den är ställd åt andra hållet, där svarsalternativen är kopplade mot nivåerna av informerade fast åt andra hållet än de tidigare frågorna. En anledning till att den största delen klassas som oinformerade, 46,5%, kan vara det att en kan tänka att det som inte har skickats eller publicerats, inte heller har nått Facebooks servrar, och därmed så skall datan inte kunna gå att lagra, och då absolut inte att använda. Detta är likt det som Goldberg (2013) påstår, att användare på Facebook inte förväntar sig att oskickade inlägg sparas.

Men, så är det inte, då enligt Das och Kramer (2013) så skickas denna data till Facebook, som sedan har använt datan till att profilera användare och se vilka som väljer att inte publicera det som de har skrivit, det vill säga själv censurerat. En anledning till att de som är fullt informerade visste om att denna data sparas och används, är att de skulle kunna ha haft tankarna i att detta är förekommande på andra webbsidor, så som Goldberg (2013) påpekar att Google gör i sin e-mailtjänst Gmail, och därför anse att detta även kan förekomma på sociala medier. Det kan även vara så att de tänker att om data samlas in, så används det, för annars skulle det inte finnas någon egentlig anledning till att samla in datan.

Att veta att text som skrivs men inte skickas även kan sparas, likt i Gmail, går även att applicera på de 18,8% som klassades som halvt informerade (Goldberg, 2013). Skillnaden är att de även svarade att datan inte används till något, vilket kan innebära att det kan vara svårt att förstå vad sådan data skall användas till i detta sammanhang. De som klassades som halvt informerade i denna fråga får ändå anses som relativt många. Detta på grund av att det är den frågan med flest som har svarat att de tror att datan sparas men inte används, och att de uppgår till nästan en femtedel av alla som har svarat på frågan.

När vi ser på resultatet från fråga sex och sju (Figur 4.6, Figur 4.7) så går det att utläsa att i båda fallen så var ungefär hälften av respondenterna inte varse om att data kan samlas in på dessa sätt, och inte heller att det används, vilket innebär att de uppfyller det som Nowak och Phelps (1997) beskriver som oinformerad om datainsamling.

5.2.3 Spårningsmetoder

Enligt resultaten på fråga åtta (Figur 4.8) där frågan var om Facebook kan spåra respondentens musrörelser på skärmen när den använder Facebook på datorn, så ser vi att en majoritet med 41,6% har svarat att de inte vet om det händer, som tillsammans med de 15,8% som anser att data inte samlas in på detta sätt eller används, utgör en större majoritet på 57,4% som sedan får klassas som oinformerade enligt Nowak och Phelps (1997) skala eftersom denna data samlas in (Senate Commerce, 2018). Denna data används sedan till att avgöra om det är en människa eller en bot som navigerar på webbsidan (Senate Commerce, 2018).

Att det är en majoritet som inte tror att detta sker ser författarna flera anledningar till. En anledning skulle kunna vara att detta uppdagades i utfrågningen med Facebooks VD Mark Zuckerberg för mindre än ett år sedan, vilket kan ses som relativt nytt och därmed inte har fått en sådan stor spridning (Senate Commerce, 2018). En annat skäl kan vara att runt denna tidpunkten så pratades det mycket om Cambridge Analytica-skandalen, vilket kan ha gjort att detta hamnade i skymundan och inte fick så stor uppmärksamhet, så att det inte nådde så många användare (Badshah, 2018).

En annan orsak kan vara att frågan explicit frågar om hur det kan samlas in när en användare använder en dator. Då smartphones och tablets är det sätt som flest använder Facebook genom, så är det tänkbart att musrörelser kan kännas främmande för vissa om de inte använder sig utav en stationär eller bärbar dator när de är inne på Facebook (Statista, 2019c). Ytterligare en anledning till att så få svarade att datan sparas och används i detta fall kan vara det att få läser igenom användaravtal (Appelgren & Leckner, 2016).

En anledning som däremot talar för att de 34,7% som visste om att denna data samlas in och används, kan också vara kopplat till Cambridge Analytica-skandalen (Badshah, 2018), då det finns en möjlighet att de blev mer informerade om i vilken utsträckning data kan samlas in, eller att de blev intresserade att ta reda på mer och följde utvecklingen i Facebooks hantering av data som senare uppkom i förhöret (Senate Commerce, 2018).

När vi tittar på resultatet i fråga nio (Figur 4.9) så kan vi utläsa att den största majoriteten som svarat vet om att Facebook samlar in data och kan bedöma om en har tjänsten igång i bakgrunden eller i förgrunden. Detta skiljer sig från tidigare frågan om spårningsmetoder, fråga åtta (Figur 4.8), där majoriteten ansåg att den tekniken inte används till att samla in data. Vår försvaren på dessa spårningsmetoder skiljer sig vet inte författarna, då båda metoderna uppkom i Zuckerbergs utfrågning av senaten (Senate Commerce, 2018). Ett skäl skulle kunna vara att de som svarat helt enkelt tänker sig att detta (fråga nio) är en spårningsmetod som är lättare att utföra. Ytterligare en anledning skulle kunna vara att denna fråga riktar sig till användning från flera olika sorters enheter och inte är fäst vid att användaren måste använda Facebook via sin dator som i fråga åtta (Figur 4.8), och eftersom smartphones och tablets är de enheter som flest besöker Facebook genom, så är det inte otänkbart att detta påverkar resultatet (Statista, 2019c).

I fråga nio (Figur 4.9) så går det att se att 12,9% har svarat att de tror att data samlas in men inte används. Detta är förhållandevis en stor mängd respondenter som har svarat detta svar sett till de andra frågorna, och där skulle en anledning kunna vara att respondenterna helt enkelt inte ser vad data som samlas in av den här typen skulle ha för användning. Något som klart går att se är i alla fall att fler som svarat på enkäten är informerade om spårningsmetoden

i fråga nio (Figur 4.9) snarare än den i fråga åtta (Figur 4.8), och räknas in som fullt informerade angående den datainsamlingsmetoden enligt Nowak och Phelps (1997).

Resultatet från fråga tio (Figur 4.10) påvisar att de flesta respondenterna, 43,6% har svarat att företag kan spåra dem på internet via ens Facebook-konto även när en är utloggad från tjänsten. Däremot, om man slår ihop antalet som har svarat "Vet inte", 39,6%, och att datan inte samlas in och inte används, 11,9%, så ser vi att de tillsammans blir 51,5%, och när de då är kopplade till nivåerna av Nowak och Phelps (1997) så innebär det att majoriteten klassas som oinformerad, men att det inte är en stor skillnad mellan andelen fullt informerade och oinformerade.

Att en stor andel är informerade om denna spårningsmetod kan bero på att de har sett personlig reklam på Facebook som är baserat på webbsidor som de tidigare har besökt, något som är möjligt genom Facebook Pixel-teknologin (Facebook Business, 2019a). Att den största andelen klassas som oinformerad om den här spårningsmetoden kan vara att respondenterna kan tänka sig att när man är utloggad från en tjänst, så får den inte längre reda på vad en gör. Det kan även vara så att de har klickat sig igenom ett användaravtal som beskriver detta, men inte läst det (Appelgren & Leckner, 2016). Endast 5% svarade att de tror att datan samlas in men inte används, och en anledning till att det inte var fler som valde detta svar skulle kunna bero på att det är lätt att tro att om data samlas in så används det.

5.2.4 *Diskussion om teorins begränsningar*

Teorin som används för att bedöma till vilken grad användare är informerade om datainsamling utgår från Nowak och Phelps arbete kring datainsamling från 1997. Det har skett stora förändringar sedan 1997 när det gäller datainsamling till följd av bland annat att teknologin har utvecklats och genom uppkomsten av sociala medier. Däremot innebär detta inte att teorin inte är applicerbar när det gäller datainsamling idag, då teorin beskriver olika grader av hur informerad en användare är oberoende av hur själva datainsamlingen går till.

Däremot så anser författarna att teorin från Nowak och Phelps (1997) hade kunnat utvecklats och möjligtvis fått fler grader eller annorlunda kriterier, framförallt när det kommer till de användare som klassificeras som halvt informerade. Halvt informerad innebär att användaren vet om att datan samlas in men inte att den används, vilket författarna anser blir en aning motsägelsefullt; varför skulle data samlas in om det inte skulle användas till någonting? I och med att data har blivit ett centralt begrepp och en stor handelsvara (Larsson & Ledendal, 2017) och med bland annat Cambridge Analytica-skandalen (Badshah, 2018) så är det säkerligen många som har hört och lärt sig vad data kan användas till, och att data ofta samlas in med ett ändamål. Med detta, och när enkätundersökningen har klara svar kopplade till dessa olika nivåer, så blir det, som författarna märkte, förmodligen inte många som kommer att svara att data samlas in men inte används till något. Detta påvisar en brist i teorin, eller möjligtvis hur den används av författarna i detta fallet.

Ännu en sak som författarna märkte när det kopplades klara svarsalternativ mot dessa olika nivåer av Nowak och Phelps (1997) är att för att en användare skall klassas som oinformerad, så skall de varken veta om att data samlas in eller används. Hade det då endast funnits ett alternativ som säger att data inte samlas in eller används, så skulle respondenterna bli tvungna att svara på ett svarsalternativ som de kanske inte står bakom. Genom att då lägga till möjligheten att svara att en inte vet, så fick respondenterna ett alternativ som tog bort detta problem. Det skulle dock gå att säga att ett nytt problem skapades, då två svarsalternativ blir kopplade

till oinformerad, och att respondenterna då får det lättare att klassas som oinformerad. Författarna tyckte dock att det tidigare var ett större problem, och valde att köra på det andra alternativet.

Det går ändå att dra en slutsats utifrån Nowak och Phelps (1997) teori, även om dessa begränsningar som diskuterats kan ha påverkat hur svaren blev och vad som resulterade utav det. Det går dock att se till vilken grad som användare är informerade om dessa olika sätt som deras data samlas in på.

6 Slutsats

Uppsatsen syfte har varit att ta reda på till vilken grad användare på sociala medier är informerade om hur deras data samlas in på sociala medier. För att detta skall vara möjligt att besvara, så har författarna haft följande forskningsfråga som utgångspunkt:

Hur informerade är användare på sociala medier om i vilken utsträckning deras data samlas in?

För att ta reda på detta, så utfördes en enkätundersökning där deltagarna fick ge svar på påståenden om olika datainsamlingsmetoder som sker på sociala medier.

Det är svårt att utröna om användningsfrekvensen på sociala medier påverkar hur informerade användare är om datainsamlingen, men det finns vissa indikationer på att frekvensen har en inverkan. I tre av åtta datainsamlingsmetoder så klassificerades en majoritet av användare som fullt informerade, och i fem av åtta datainsamlingsmetoder så klassificerades en majoritet av användare som oinformerade. Detta visar att användare är mer informerade om vissa datainsamlingsmetoder än andra, men att användare överlag är oinformerade om majoriteten av de datainsamlingsmetoder som utreds i uppsatsen.

Författarna anser att det behövs undersökningar som är mer omfattande, både i form av fler datainsamlingsmetoder samt större urval för att styrka slutsatserna som har framkommit i denna uppsats. Då användandet av sociala medier ökar, så lär även datainsamlingen öka, och författarna tycker att det hade varit spännande att se om användarna blir mer informerade i och med detta. Författarna ser även att det är möjligt att användarna blir mindre informerade på grund av detta, då det säkerligen kommer att uppkomma nya metoder att samla in data på.

Appendix 1 - Enkät

Enkätundersökning om datainsamling på sociala medier

Hej! Vi är två studenter på det Systemvetenskapliga kandidatprogrammet vid Lunds Universitet som skriver vår kandidatuppsats, som fokuserar på hur informerade användare är om datainsamling på sociala medier.

Svara på frågorna/påståendena utefter vad du tror stämmer. När vi skriver att 'datan används' betyder det att datan används till olika syften som exempelvis för att skräddarsy annonser, förbättra användarupplevelsen och annat.

Tack på förhand från Benjamin Idzikowski och Victor Jennekvist!

***Obligatorisk**

1. Hur gammal är du? *

Ditt svar

2. Hur ofta använder du Facebook? *

- Fler än fem gånger varje dag
- En till fem gånger varje dag
- Fler än fem gånger i veckan
- Fler än fem gånger i månaden
- Färre än fem gånger i månaden

3. Facebook kan samla in data om dig utan att du har ett Facebook-konto. *

- Ja, och datan används.
- Ja, men datan används inte.
- Nej, data samlas inte in och används inte.
- Vet inte.

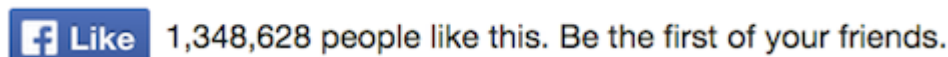
4. Facebook kan samla in data om dig när du surfar på andra webbsidor än [facebook.com](https://www.facebook.com). *

- Ja, och datan används.
- Ja, men datan används inte.
- Nej, data samlas inte in och används inte.
- Vet inte.

5. Facebook kan samla in data om andra uppkopplade enheter på samma nätverk som du är uppkopplad på. *

- Ja, och datan används.
- Ja, men datan används inte.
- Nej, data samlas inte in och används inte.
- Vet inte.

6. En webbsida som använder sig av Facebook Like-knappen kan skicka data om din aktivitet till Facebook utan att du använder Facebook Like-knappen. *



- Ja, och datan används.
- Ja, men datan används inte.
- Nej, data samlas inte in och används inte.
- Vet inte.

7. Kommentarer och inlägg som du skriver på Facebook men raderar innan de är skickade, är helt borttagna. *

- Ja, datan är helt borttagen.
- Nej, de sparas men datan används inte.
- Nej, de sparas och datan används.
- Vet inte.

8. Facebook kan spåra dina musrörelser på skärmen när du använder Facebook via din dator. *

- Ja, och datan används.
- Ja, men datan används inte.
- Nej, data samlas inte in och används inte.
- Vet inte.

9. Facebook kan bedöma om du har Facebook igång i bakgrunden eller förgrunden (minimerad/maximerad). *

- Ja, och datan används.
- Ja, men datan används inte.
- Nej, data samlas inte in och används inte.
- Vet inte.

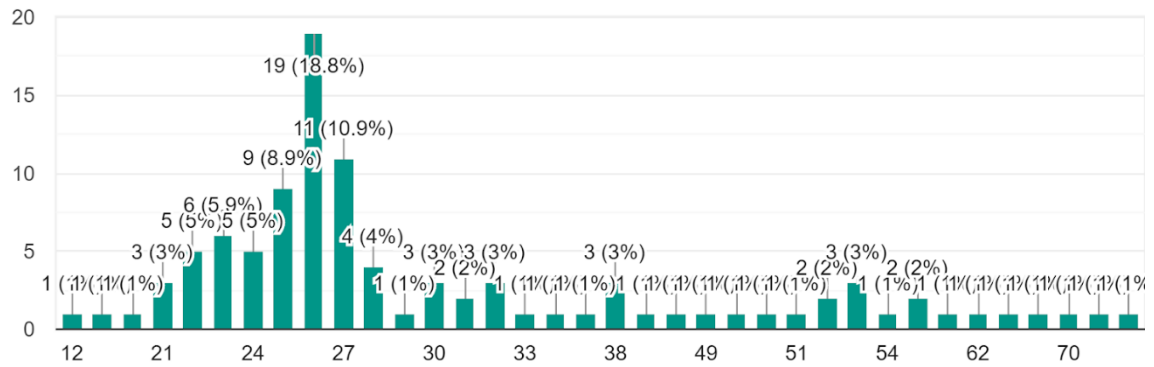
10. Företag kan spåra dig på internet via ditt Facebook-konto när du är utloggad från tjänsten. *

- Ja, och datan används.
- Ja, men datan används inte.
- Nej, data samlas inte in och används inte.
- Vet inte.

Appendix 2 – Enkät svar

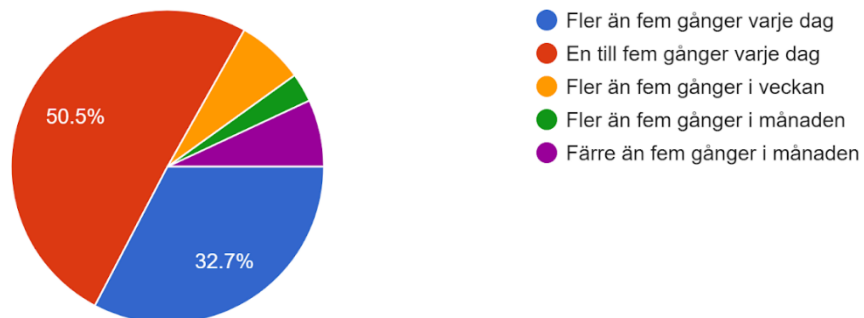
1. Hur gammal är du?

101 responses



2. Hur ofta använder du Facebook?

101 responses



3. Facebook kan samla in data om dig utan att du har ett Facebook-konto.

101 responses



4. Facebook kan samla in data om dig när du surfar på andra webbsidor än facebook.com.

101 responses



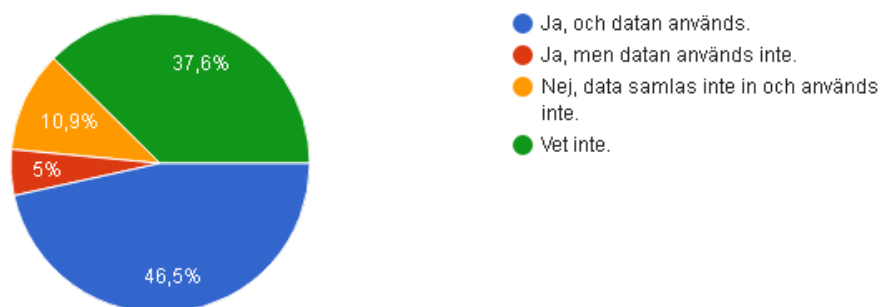
5. Facebook kan samla in data om andra uppkopplade enheter på samma nätverk som du är uppkopplad på.

101 responses



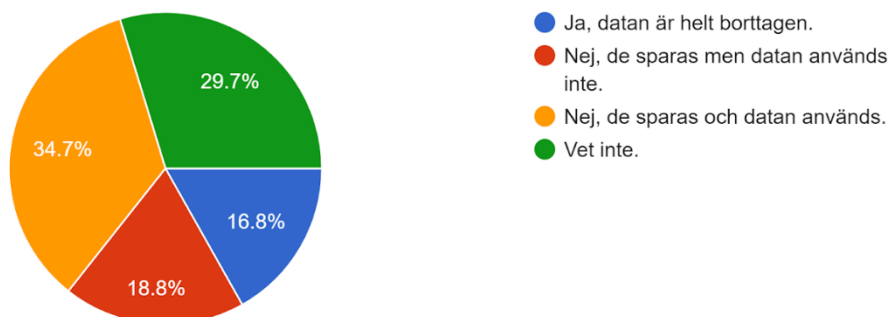
6. En webbsida som använder sig av Facebook Like-knappen kan skicka data om din aktivitet till Facebook utan att du använder Facebook Like-knappen.

101 svar



7. Kommentarer och inlägg som du skriver på Facebook men raderar innan de är skickade, är helt borttagna.

101 responses



8. Facebook kan spåra dina musrörelser på skärmen när du använder Facebook via din dator.

101 responses



9. Facebook kan bedöma om du har Facebook igång i bakgrunden eller förgrunden (minimerad/maximerad).

101 responses



10. Företag kan spåra dig på internet via ditt Facebook-konto när du är utloggad från tjänsten.

101 responses



Referenser

Appelgren, E. & Leckner, S. (2016). "Att dela eller inte dela - användarnas inställning till insamling av personlig data", i Ohlsson, Oscarsson och Solevid (red.) Ekvilibrium, SOM-institutets forskarantologi nummer 67. Göteborg: SOM-institutet, Göteborgs universitet. Tillgänglig online:

https://som.gu.se/digitalAssets/1577/1577248_att-dela-eller-inte-dela.pdf

Asur, S. & Huberman, B.A. (2010). *Predicting the Future With Social Media*. International Conference on Web Intelligence and Intelligent Agent Technology. Tillgänglig online:

<https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=5616710&tag=1>

Badshah, N. (2018). *Facebook to contact the 87 million users affected by data breach*, The Guardian, 8 April. Tillgänglig online:

<https://www.theguardian.com/technology/2018/apr/08/facebook-to-contact-the-87-million-users-affected-by-data-breach>

Baird, C.H. & Parasnis, G. (2011). *From social media to social customer relationship management*. VOL. 39 NO. 5 2011, pp. 30-37, Q Emerald Group Publishing Limited. Tillgänglig online:

<http://heuristic.kaist.ac.kr/cylee/xpolicy/TermProject/15/4.%20social%20media.pdf>

Bechmann, A. (2014). *Non-Informed Consent Cultures: Privacy Policies and App Contracts on Facebook*. Journal of Media Business Studies, 11:1, 21-38. Tillgänglig online:

<https://www.tandfonline.com/doi/pdf/10.1080/16522354.2014.11073574?needAccess=true>

Boyd, D.M. & Ellison, N.B. (2007). *Social Network Sites: Definition, History and Scholarship*. Journal of Computer-Mediated Communication. Volume 13, Issue 1, Pages 210-230.

Tillgänglig online: <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1083-6101.2007.00393.x?scrollTo=references>

Bujlow, T., Carela-Español, V., Solé-Pareta, J., & Barlet-Ros, P. (2015). *A survey on Web Tracking: Mechanisms, Implications, and Defenses*. Under review, July 2015. Tillgänglig online:

<http://arxiv.org/abs/1507.07872>

Busby, E., Hammoud, T., Rose, J., & Prashad, R. (2012). *The evolution of online-user data*.

The Boston Consulting Group, 2012. Tillgänglig online: <https://www.bcg.com/publications/2012/marketing-technology-evolution-of-online-user-data.aspx>

Bylund, M. (2013). *Personlig integritet på nätet*. Fores. Tillgänglig online: https://fores.se/wp-content/uploads/2013/11/FORES-Personligintegritet_web_enkelsid.pdf

Cadwalladr, C. & Graham-Harrison, E. (2018). *Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach*, The Guardian, 17 Mars. Tillgänglig online: <https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election>

Cegłowski, M. (2016): The Moral Economy of Tech. Text version of remarks, SASE conference, Berkeley, June 26, 2016. Tillgänglig online: http://idlewords.com/talks/sase_panel.htm

Christl, W. (2017). *Corporate surveillance in everyday life*. Cracked labs. Tillgänglig online: https://crackedlabs.org/dl/CrackedLabs_Christl_CorporateSurveillance.pdf

Christl, W. & Spiekermann, S. (2016). *Networks of Control. A Report on Corporate Surveillance, Digital Tracking, Big Data*. Tillgänglig online: https://crackedlabs.org/dl/Christl_Spiekermann_Networks_Of_Control.pdf

Cooper, T. & LaSalle, R. (2015). *Guarding and growing personal data value*. Accenture Institute of High Performance. Tillgänglig online: https://www.accenture.com/_acnmedia/PDF-4/Accenture-Guarding-and-Growing-Personal-Data-Value-POV-Low-Res.pdf

Das, S. & Kramer, A. (2013). *Self-censorship on Facebook*. Association for the Advancement of Artificial Intelligence. Tillgänglig online: <https://research.fb.com/wp-content/uploads/2016/11/self-censorship-on-facebook.pdf>

Datainspektionen. [u.å]. *Vad är egentligen en personuppgift?* Datainspektionen. Tillgänglig online: <https://www.datainspektionen.se/vagledning/en-introduktion-till-dataskyddsförordningen/vad-ar-en-personuppgift/> (Hämtad 2019-05-15).

Englehardt, S. & Narayanan, A. (2016). *Online Tracking: A 1-million-site Measurement and Analysis*. Princeton University, Princeton, NJ, USA, Tech Rep. Tillgänglig online: http://randomwalker.info/publications/OpenWPM_1_million_site_tracking_measurement.pdf

Equifax-Harris. (1996). *Consumer Privacy Survey*. Atlanta, GA: Equifax Corporate Marketing Department.

EU-kommissionen. (2014). Pressmeddelande: "Dataskyddsdagen 2014: Viviane Reding efterlyser en ny överenskommelse om dataskydd för Europa". 28 Januari. Tillgänglig online: http://europa.eu/rapid/press-release_IP-14-70_sv.htm

Facebook Analytics. (2018). *Personbaserade analyser för en flerkanalvärld*. Facebook. Tillgänglig online: <https://analytics.facebook.com/> (Hämtad 2019-04-21).

Facebook Business. (2019a). *Om Facebook-pixeln*. Facebook. Tillgänglig online: <https://sv-se.facebook.com/business/help/742478679120153> (Hämtad 2019-04-20).

Facebook Business. (2019b). *Om anpassade målgrupper från kundlistor*. Facebook. Tillgänglig online: <https://www.facebook.com/business/help/341425252616329> (Hämtad 2019-04-21).

Facebook Developers. (2019). *Facebook Pixel*. Facebook. Tillgänglig online: <https://developers.facebook.com/docs/facebook-pixel/> (Hämtad 2019-04-20).

Facebook. (2018). *Datapolicy*. Tillgänglig online: <https://www.facebook.com/policy.php>

Facebook. (2016). *Three Million Business Stories. What's yours?*. Facebook. Tillgänglig online: <https://www.facebook.com/business/news/3-million-advertisers> (Hämtad 2019-04-12)

Goldbeck, J. (2013). *On second thought... Facebook wants to know why you didn't publish that status update you started writing*. Slate, 13 December. Tillgänglig online: <https://slate.com/technology/2013/12/facebook-self-censorship-what-happens-to-the-posts-you-dont-publish.html>

Gordon, B.R., Zettelmeyer, F., Bhargava, N., & Chapsky, D. (2018). *A Comparison of Approaches to Advertising Measurement: Evidence from Big Field Experiments at Facebook*. Marketing Science, Volume 38, Issue 2, Pages 193-364. Tillgänglig online: https://www.kellogg.northwestern.edu/faculty/gordon_b/files/fb_comparison.pdf

Graeff, T. R. & Harmon, S. (2002). *Collecting and Using Personal Data: Consumers' Awareness and Concerns*. Journal of Consumer Marketing, vol. 19, no. 4. Tillgänglig online: <https://www.emeraldinsight.com/doi/pdfplus/10.1108/07363760210433627>

Heitzman, A. (2017). *Facebook Analytics: A Beginner's Guide*. Search Engine Journal, 22 December, Tillgänglig online: <https://www.searchenginejournal.com/facebook-analytics-beginners-guide/228627/#close>

Jacobsen, D.I. (2002). *Vad, Hur Och Varför: Om Metodval I Företagsekonomi Och Andra Samhällsvetenskapliga Ämnen: Studentlitteratur.*

Joler, V. & Petrovski, A. (2016). *Immaterial Labour and Data Harvesting, Facebook Algorithmic Factory*. Share Lab. Tillgänglig online: <https://labs.rs/en/facebook-algorithmic-factory-immaterial-labour-and-data-harvesting/>

Kanter, J. (2018). *Facebook is tracking you in ways you never knew - here's the crazy amount of data it sucks up*. Business Insider, 12 Juni. Tillgänglig online: <https://nordic.businessinsider.com/facebook-reveals-all-the-way-it-tracks-user-behaviour-2018-6>

Kosinski, M., Stillwell, D., & Graepelb, T. (2013): *Private Traits and Attributes Are Predictable from Digital Records of Human Behavior*. PNAS, March 2013. Tillgänglig online: <http://www.pnas.org/content/110/15/5802>

Kshetri, N. (2014). *Big data's impact on privacy, security and consumer welfare*. Telecommunications Policy, 38(11). Tillgänglig online: https://libres.uncg.edu/ir/uncg/f/N_Kshetri_Big_2014.pdf

Larsson, S (2018). *Den kvantifierade konsumenten: Om behovet av tillit och transparens på datadrivna marknader*. Fores. Policy Brief 2018:1. Tillgänglig online: <http://fores.se/wp-content/uploads/2018/02/Policy-Brief-2018-1-%E2%80%93-Den-kvantifierade-konsumenten.pdf>

Larsson, S. & Ledendal, J. (2017). *Personuppgifter som betalningsmedel*. Karlstad: Konsumentverket. 2017:4. Tillgänglig online: <https://www.konsumentverket.se/globalassets/publikationer/produkter-och-tjanster/gemensamt/rapport-2017-4-personuppgifter-som-betalmedel-konsumentverket.pdf>

Nowak, G. & Phelps, J. (1997) *Direct marketing and the use of individual-level consumer information: Determining how and when "privacy" matter*, Journal of Direct Marketing. Volume 11 Number 4.

Pasquale, F. (2015). *The Black Box Society: The secret algorithms that control money and information*, Harvard University Press. Tillgänglig online: <http://rale.english.ucsb.edu/wp-content/Engl800/Pasquale-blackbox.pdf>

Pew. (2014). *Public Perceptions of Privacy and Security in the Post-Snowden Era*. Pew Research Center. Tillgänglig online: <https://www.pewinternet.org/2014/11/12/public-privacy-perceptions/>

Phelps, J., Nowak, G., & Ferrell, E. (2000). *Privacy concerns and consumer willingness to provide personal information*. *Journal of Public Policy & Marketing*, 19(1), 27-41.

Richter, A. & Koch, M. (2008) *Functions of Social Networking Services*. Cooperation Systems Center Munich. Tillgänglig online: <https://dl.eusset.eu/bitstream/20.500.12015/2802/1/00467.pdf>

Roesner, F., Kohno, T., & Wetherall, D. (2012). *Detecting and defending against third-party tracking on the web*. Proceedings of the 9th USENIX Conference on Networked Systems Design and Implementation (NSDI2012). University of Washington. Tillgänglig online: <https://www.usenix.org/system/files/conference/nsdi12/nsdi12-final17.pdf>

Roosendaal, A. (2011). *Facebook tracks and traces everyone: Like this!* Tilburg Law School Legal Studies Research Paper Series No. 03/2011. Tillgänglig online: <https://mikz.net/log/plaatjes/facebook.pdf>

Social Media. [u.å]. *Merriam Webster*. Tillgänglig online: <https://www.merriam-webster.com/dictionary/social%20media> (Hämtad 2019-04-20).

Statista. (2019a). *Number of social media users worldwide from 2010 to 2021 (in billions)*. Statista. Tillgänglig online: <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/> (Hämtad 2019-04-25).

Statista. (2019b). *Most popular social networks worldwide as of April 2019*. Statista. Tillgänglig online: <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/> (Hämtad 2019-04-25)

Statista. (2019c). *Device usage of Facebook users worldwide as of January 2019*. Statista. Tillgänglig online: <https://www.statista.com/statistics/377808/distribution-of-facebook-users-by-device/> (Hämtad 2019-04-25).

Tikno. (2017). *Exploring the acceptance for pixel technology implementation in Facebook ads among advertisers in Indonesia*. The 9th International Conference on Advances in Information Technology (IAIT-2017), Pages 96-107. Tillgänglig online: <https://www.knepublishing.com/index.php/Kne-Social/article/view/1399/3213>

Tuunainen, V. K., Pitkänen, O., & Hovi, M. (2009). *Users' awareness of privacy on online social networking sites - Case Facebook*. Bled 2009 Proceedings. 42. Tillgänglig online: <https://pdfs.semanticscholar.org/9b83/3ca55abd01f842c764804706750743c67115.pdf>

US Senate. Committee on Commerce, Science and Transportation. (2018). *Facebook, Social Media Privacy, and the Use and Abuse of Data*. Washington, D.C. 20510-6125. Tillgänglig online:

https://www.commerce.senate.gov/public/_cache/files/9d8e069d-2670-4530-bcdc-d3a63a8831c4/7C8DE61421D13E86FC6855CC2EA7AEA7.senate-commerce-committee-combined-qfrs-06.11.2018.pdf

Weisberg, H. (2005). *The Total Survey Error Approach*. The University of Chicago Press. Tillgänglig online:

https://www.academia.edu/31859666/Weisberg_2005_TheTotal_Survey_Error_Approach

Whiting, A. & Williams, D. (2013). *Why people use social media: a uses and gratifications approach*. *Qualitative Market Research: An International Journal*, Vol. 16, No. 4, P.362-369. Tillgänglig online:

https://www.researchgate.net/publication/237566776_Why_people_use_social_media_A_uses_and_gratifications_approach

Wong, J.C. (2018). *Mark Zuckerberg faces tough questions in two-day congressional testimony - as it happened*. *The Guardian*, 4 November. Tillgänglig Online:

<https://www.theguardian.com/technology/live/2018/apr/11/mark-zuckerberg-testimony-live-updates-house-congress-cambridge-analytica>

Zins, Chaim (2007). *Conceptual Approaches for Defining Data, Information and Knowledge*. *Journal of the American Society for Information Science and Technology*. Tillgänglig online:

http://www.success.co.il/is/zins_definitions_dik.pdf