LUNDS UNIVERSITET
Ekonomihögskolan

Institutionen för informatik

Persondata på Facebook

Användarnas ställningstagande till insamling av personlig data av Facebook

Kandidatuppsats 15 hp, kurs SYSK02 i informationssystem

Författare: André Johnsson
Ellinor Larsson
Amanda Liljekvist

Handledare: Odd Steen

Examinatorer: Magnus Wärja
Benjamin Weaver

Persondata på Facebook: Användarnas ställningstagande till insamling av personlig data av Facebook

Författare: André Johnsson, Ellinor Larsson och Amanda Liljekvist

Utgivare: Inst. för informatik, Ekonomihögskolan, Lund universitet

Dokumenttyp: Kandidatuppsats

Antal sidor: 54

Nyckelord: Facebook, persondata, informationsinsamling

Sammanfattning (Max. 200 ord):
Sociala medier har blivit en del av de flesta svenskars vardag, där Facebook står ut som det största sociala nätverket. Facebook används främst av yngre människor av vilka en klar majoritet använder Facebook dagligen. Facebook uppfattas som gratis, men i verkligheten betalar användarna med data som samlas in för att bland annat sälja vidare till tredje part.

Har svenskar kunskap om hur deras data samlas in och används av Facebook? Vet de om att Facebook samlar in data över huvud taget, och hur mycket kunskap besitter de i så fall om denna datainsamling? Påverkar deras kunskap hur de använder Facebook och diverse säkerhetstjänster? Vi undersökte detta genom att i en kvantitativ undersökning fråga 125 användare i åldern 20-29 år en rad frågor och utifrån svaren tog vi reda på användarnas kunskapsnivåer i ämnet. Vi ställde även frågor kring hur de använde Facebook och diverse säkerhetstjänster och jämförde sedan kunskapsnivåerna med hur de svarat att de använder Facebook och säkerhetstjänster. De flesta av användarna visade sig ha viss kunskap om Facebooks datainsamling och en viss skillnad i användandet av Facebook och säkerhetstjänster kunde härledas till användarnas kunskap i ämnet.

Innehåll

Figurer

Tabeller

1 Introduktion

Vi lever i en värld där IT blir en allt större del av våra dagliga liv. Som ett resultat av internets utveckling lämnar användare ut personlig data och genom detta riskerar de att omedvetet dela med sig av personlig data till olika företag och tjänster (Jeckmans et al., 2013). Idag använder sig företag dessutom av data mining för att samla in personlig data samt känslig information om allt som användaren gör på nätet. Syftet med detta kan vara att personalisera informationsflödet för att användaren ska få en så bra upplevelse som möjligt på internet eller i marknadsföringssyfte för att sälja diverse produkter (Toubiana et al., 2010).

Det internetanvändare oftast inte vet om är att företagen tjänar pengar på insamling av personlig data. Företagen erbjuder tjänster som utåt upplevs som gratis men i verkligheten betalar användarna med sin personliga data (Alverén, 2012). Erkin et al., (2010) skriver att det inte går att försäkra sig om att datan som säljs till tredjeparter inte missbrukas. Jeckmans et al., (2013) skriver även att det potentiella hotet mot personlig integritet ofta underskattas. Ju mer detaljerad datan är, desto större är risken att datan missbrukas om den säljs eller läcks.

Facebook är den tredje mest besökta webbsidan i Sverige idag (alexa.com, 2017) och en stor del av Sveriges befolkning använder tjänsten. När en person skapar ett konto på Facebook måste dem godkänna Facebooks användaravtal och med det godkänner de också att Facebook får samla in data om dem – på Facebook men också genom alla tredjepartssidor som är kopplade till Facebooks tjänster (Facebook.com, 2017). Eftersom få personer läser användaravtal missar de Facebooks datapolicy och saknar kunskap om vilken typ av data Facebook samlar in om dem.

Facebook använder sig utav Collaborative Filtering (CF) (Facebook.com, 2017). CF är en vanlig teknik som används inom rekommendationssystem (RS) och är enligt Ricci et al., (2011) den mest populära tekniken inom RS. Tekniken går ut på att användaren på något vis uttrycker sina preferenser genom betygsättningar och intressen på internet. CF fokuserar på att matcha ihop användare baserat på preferenser, detta för att skapa personliga rekommendationer (Terveen et al., 2010).

1.1 Problemformulering

Enligt Alverén (2012) är Sverige ett av världens mest uppkopplade länder och i med detta så finns det risker med att lämna ut personlig data. Facebook är beroende av att samla in så mycket information om sina användare som det går, just för att tjäna pengar. Tjänsten kan upplevas som gratis, men användaren betalar med sin personliga data. Allt användaren gör/lägger upp/delar med sig av på Facebook äger Facebook. Detta innebär exempelvis att om en användare raderar en bild så sparas denna hos Facebook. Det uppstår alltså risker både när användaren själv lämnar ut personlig data och när Facebook samlar in ovetande användares data. (Alverén, 2012)

Alverén (2012) diskuterar olika risker som tillkommer vid Facebooks hantering av användarnas personliga information. Det kan vara både mindre och större risker samt kort- och långsiktiga. All information samlas in och kopplas ihop för att skapa en helhetsbild av användaren. I och med att Facebook samlar in och äger all information som publiceras så upplever användarna att den personliga integriteten kränks då de inte vet vem som använder deras data eller hur den säljs vidare. Facebook är ett stort företag som dagligen utsätts för attacker och det innebär att även om Facebook hanterar den personliga datan på ett etiskt vis, finns det ingen garanti för att den inte läcks eller missbrukas av någon annan. Läckt persondata kan resultera i att användaren drabbas utav bedrägeri, identitetsstöld och utpressning. En annan viktig faktor att ta hänsyn till är filterbubblan som användarna riskerar att hamna i. Filterbubblan gör så att användaren inte får ta del av information som motsäger deras åsikter och därmed stängs de in i bubblor (Alverén, 2012).

1.2 Forskningsfråga

Vi ställer oss de här två huvudfrågorna:

- Har svenskar i åldrarna 20-29 år kunskap om hur Facebook samlar in data om dem och känner de till riskerna med hur datan hanteras, samt gör de aktivt något för att skydda sin data?

- Finns det någon skillnad i hur en användare med mer kunskap om Facebooks datainsamling skyddar sin data/inte skyddar sin data jämfört med en användare med mindre kunskap?

1.3 Syfte

Syftet med undersökningen är att ta reda på om svenskar i åldrarna 20-29 år vet i vilken utsträckning Facebook samlar in data om dem samt hur deras personliga data hanteras. Vi vill ta reda på om de känner till riskerna med Facebooks datainsamling och om de aktivt gör något för att skydda sin data. Vi vill också ta reda på om det finns skillnader i hur användare med mer kunskap om Facebooks datainsamling väljer att skydda/inte skydda sin data jämfört med användare som har mindre kunskap.

1.4 Avgränsningar

Insamling av personlig data sker överallt på internet (Alverén, 2012) och för att inte skriva generellt om all datainsamling som existerar har vi valt att avgränsa oss till den insamling som sker av Facebook. Facebook är den tredje mest besökta hemsidan i Sverige (alexa.com, 2017) och deras insamling påverkar därmed en stor del av Sveriges befolkning. Enligt 2016 års utgåva av Internetstiftelsen i Sveriges årliga undersökning *”Svenskarna och internet”* är den största användargruppen av Facebook 16-25-åringar (78% använder Facebook dagligen) följt av 26-35-åringar (70% använder Facebook dagligen) och vi har därför valt att avgränsa rampopulationen för undersökningen till svenskar i åldrarna 20-29.

2 Teoretisk bakgrund

Litteraturgenomgången tar upp artiklar och litteratur som är relevant för ämnesområdet. Här förtydligas begrepp och formuleringar. Först sammanfattas vilken litteratur det är som har använts i kapitlet om teoretisk referensram. Därefter går vi in på djupet i de olika områdena som behandlas i uppsatsen.

2.1 Teoretisk referensram

För att förtydliga har vi valt att utforma en referensram i Tabell 2.1 som berättar vilket område som tillhör vilken litteratur. Detta för att skapa en överblick över litteraturen som hjälper till under uppsatsens utformning samt att den sedan skall kunna återkopplas till enkätundersökningen.

Tabell 2.1: Teoretisk referensram

Område	Behandlar	Referenser
Facebooks hantering av persondata	(1) Hur definitionen av persondata skiljer sig för olika individer (2) Lagens definition på persondata (3) Behandlar Facebooks Data Policy och definierar allt som samlas in om Facebooks användare	(1) Alverén (2012), Såld på nätet (2) SFS (1998), Personuppgiftslag (3) facebook.com (2017), Facebook Data Policy
Facebooks hantering av andra sociala medier	(4) Beskriver att Facebook har koppling till fler än sju miljoner tredjepartssidor och samlar in information om användarnas aktiviteter även på de här sidorna (5) Hur företag bedriver marknadsföring utifrån den insamlade persondatan samt hur tredjepartsappar kan göra ändringar i användarnas sekretessinställningar (6) Behandlar Facebooks populäraste verktyg för att spåra användare	(4) Alverén (2012), Såld på nätet (5) Wang et al., (2011), Third-Party Apps on Facebook: Privacy and the Illusion of Control (6) Acar et al., (2015), Facebook Tracking Through Social Plug-ins
Personlig integritet och internet	(7) Förklarar vad begreppet integritet innebär och var det härstammar ifrån (8) Beskriver hur integritet är ett svårdefinierat begrepp som är mycket brett	(7) Warren & Bandeis (1890), The right to privacy (8) Paine et al., (2007), Internet users' percep-

	(9) Definierar en tydligare bild av var ordet integritet innebär (10) Går djupare in på hur personlig integritet och hur sociala medier går ihop samt vanliga dilemman som kan uppstå	tions of ”privacy concerns” and ”privacy actions” (9) Statens medicinsk-etiska råd (2012), Integritet (10) Bylund (2013), Personlig integritet på nätet
Risker med insamling av persondata	(11) Definierar begreppet risk samt nämner hur människor utvärderar risker (12) Nämner att användarnas oro för sin personliga integritet är ett stort problem som måste tacklas (13) Kategoriseras upp olika risker med insamling av persondata (14) Behandlar dilemmat om persondata hanteras etiskt samt att det inte finns någon garanti för att datan inte läcks eller missbrukas	(11) Cho et al. (2010), Optimistic bias about online privacy risks: Testing the moderating effects on perceived controllability and prior experience. Computers in Human Behavior (12) Featherman et al. (2010), Reducing online privacy risk to facilitate e-service adoption: the influence of perceived ease of use and corporate credibility (13) Alverén (2012), Såld på nätet (14) Toch et al. (2012), Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems. User Modeling and User-Adapted Interaction
Säkerhet	(15) Behandlar olika sätt att skydda sin persondata samt att minimera riskerna för att persondatan hamnar i fel händer eller missbrukas	(15) Xu et al. (2014), Information Security in Big Data: Privacy and Data Mining
Big Data	(16) Förklarar internets uppbyggnad och hur Big Data omfattar en komplex och enorm datamängd som ska hanteras av flera olika internettjänster. Tar även upp hur insamlingen av persondata kan komma att påverka människors privatliv	(16) Bylund (2013), Personlig integritet på nätet (17) jwork.org (2014), Small Data

	(17) Definierar begreppet Small Data (18) Definierar begreppet Micro Data (19) Förklarar att Facebook samlar in allt de kan om sina användare då det är persondata som Facebook tjänar pengar på	(18) ipums.org (2017), What are Microdata? (19) Alverén (2012), Såld på nätet
Rekommendationssystem och Big Data	(20) Beskriver rekommendationssystems uppbyggnad (21) Beskriver hur rekommendationssystem används för att erbjuda personliga rekommendationer till användarna (22) Förklarar syftet med rekommendationssystem (23) Tar upp att gratistjänster så som Facebook tjänar pengar på användarnas persondata och att persondatan säljs vidare till tredjepartssidor (24) Behandlar hur Collaborative filtering används för att matcha ihop människor baserat på preferenser (25) Behandlar vad Facebook använder sig av för rekommendationssystem och vilka fördelar de genererar för Facebook	(20) Ricci et al. (2011), Introduction to Recommender Systems Handbook, Recommender Systems Handbook (21) Erkin et al. (2010), Privacy Enhanced Recommender System (22) Jeckmans et al. (2013), Privacy in Recommender System (23) Alverén (2012), Såld på nätet (24) Terveen et al. (2001), Beyond Recommender Systems: Helping People Help Each Other (25) facebook.com (2017), Code
Mätbar kunskap	(26) Behandlar vad det innebär att ha kunskap om något samt hur kunskapen kan mätas	(26) Hunt (2003), The concept of knowledge and how to measure it

2.2 Facebooks hantering av persondata

Vad som definieras som persondata kan skilja sig åt för olika individer (Alverén, 2012). Enligt personuppgiftslagen definieras personlig information och data som all information som direkt eller indirekt kan kopplas till en fysisk person som lever (SFS, 1998:204).

För att skapa ett konto på Facebook måste användaren först godkänna Facebooks användaravtal, vilket innebär att Facebook får samla in data om användaren från deras tjänster samt från tredjepartssidor som är direkt kopplade till Facebooks tjänster. Facebook samlar in data och annan typ av information som användaren tillhandahåller när denne använder sig av Facebooks tjänster. Typ av data de samlar in är allt ifrån kön, ålder, geografisk position, språkkunskaper, arbete, utbildning, relationer, intressen, meddelanden mellan användarna etc. De

kan också se datum när användaren skapar eller lägger upp nya filer på deras hemsida. Facebook samlar även in data om hur användaren använder sig av deras tjänster, exempelvis hur ofta användaren utför olika typer av aktiviteter och vilken typ av innehåll användaren gillar och tittar på. Annan data som också samlas in är det som användare lägger upp på varandras profiler, så som bilder, meddelanden eller importerar användarens kontaktinformation. (Facebook.com, 2017)

Data från olika nätverk som användaren besöker samlas in som olika grupper och sidor som användaren är kopplad till. Här syns även hur användaren kommunicerar med andra människor. Facebook samlar också in all kontaktinformation och bilder som läggs upp. Om användaren använder sig av Facebooks tjänster för att köpa ett spel eller donera pengar så lagrar Facebook all information om transaktionen, vilket innebär att användarens kortnummer, kontonummer, verifieringsinformation och kontaktinformation sparas på deras servrar. (Facebook.com, 2017)

Facebook samlar inte enbart in persondata om vad användaren gör på hemsidan, utan också information från alla enheter användaren har installerat eller har åtkomst till tjänsten på. Exempel på data som Facebook samlar in från användarnas enheter är:

- Enhetens geografiska position, genom GPS, Bluetooth eller WiFi-signaler.
- Vilken typ av hårdvara, operativsystem som enheten har, samt enhetens inställningar, filer och mjukvara.
- Information om användarens mobiloperatör, mobilnummer, IP-nummer, webbläsare, tidszoner och språk.

Facebook har rätt till att samla in information om användaren från tredjepartssidor eller appar som är kopplade till Facebooks tjänster. Detta för att använda sig av riktad marknadsföring. Tredjepartssidorna som är kopplade till Facebook har rätt att ge ut information om användarna och deras interaktion och upplevelser på hemsidorna. Det finns även andra företag som ägs eller drivs av Facebook och även dem samlar in information om användaren. (Facebook.com, 2017)

2.3 Facebooks hantering av andra sociala medier

Facebook hanterar inte enbart persondata från egna tjänster utan de samlar även in persondata från andra sociala medier. Facebook har koppling till mer än sju miljoner tredjepartssidor (Alverén, 2012). Flertalet tredjepartsappar som har koppling till Facebook utvinner information som kan identifiera användarna och delar den informationen med olika företag som bedriver marknadsföring (Wang et al., 2011).

Enligt Acar et al. (2015) använder Facebook sig av social plug-ins för att spåra användare på internet. Social plug-ins har blivit ett oerhört populärt verktyg för hemsidors ägare då det möjliggör för deras kunder att dela och sprida vidare företagens produkter eller tjänster genom sociala medier. Facebooks mest använda plug-in är deras gilla-knapp och den är det ideala verktyget för att samla information om användarens aktiviteter på internet och spårar på detta vis

allt som användaren gör. Att spåra kan definieras som insamling av olika användares aktiviteter överallt på olika hemsidor. Den här typen av spårning som Facebook använder sig av via plug-ins brukar refereras till som tredjepartsspårning. Social plug-ins implementeras på ett sätt som tvingar användarnas webbläsare att hämta innehåll från de olika hemsidorna som besöks. Det kan vara allt från bilder till olika script. Webbläsaren länkar ihop användarnas aktiviteter och beteende med deras riktiga identiteter. (Acar et al., 2015)

Acar et al. (2015) tar även upp hur användare som inte har något Facebookkonto spåras. Alla tredjepartssidor som har en koppling till Facebook kommer att skicka cookies från de olika sidorna som besöks med information om aktiviteter men även persondata. Cookies är ett vanligt verktyg som används av olika hemsidor för att samla in smulor/spår av användarens aktiviteter på internet. Detta innebär att varje gång hemsidan skickar en förfrågan till webbservern så är all information synlig för servern, syftet med cookies är att göra flödet personligt samt att skapa användarprofiler och hantera användarnas sessioner (Bylund, 2013). Informationen som samlas in genom cookies skickas sedan vidare till Facebook. Detta är möjligt då det finns många icke-krypterade kopplingar som tillåter spårning av användare som inte använder Facebook.

Även om en användare har ställt in särskilda säkerhetsinställningar på Facebook så kan de köras över av somliga tredjepartsappar. Det innebär att informationen som en gång var privat kan upphävas och bli synlig utan att användarna har blivit notifierade att några ändringar har skett i deras säkerhetsinställningar. När en användare lägger till en app kan de inte begränsa apparnas åtkomst till deras profiler (Wang et al., 2011). Ett exempel på detta som Wang et al. (2011) tar upp är om en användare har lagt till en kalender-app. Användarens ursprungliga inställningar på Facebook gjorde att alla information om personen var anonym för alla utom användaren själv. Dock försvann filtret när användaren laddade ner appen eftersom den begärde åtkomst till all personlig information – utan att användaren själv hade vetskap om det. Det gör att den personliga informationen blir synlig för alla. (Wang et al., 2011)

2.4 Personlig integritet och internet

Enligt Warren & Bandeis (1890) är integritet ett begrepp som ursprungligen härstammar från psykologin. Integritet kan sägas vara rätten att få vara ifred. Integritet är ett väldigt svårdefinierat begrepp eftersom det är oerhört omfattande (Paine et al., 2007).

Statens medicinsk-etiska råd (2012) skriver att begreppet integritet är knutet till värdighet och menar på att människan har rätt till ett egenvärde. Personens värderingar, trosuppfattningar och åsikter får inte lov att kränkas. Vidare säger de att just personlig integritet betyder okränkthet, eller rätt att inte bli kränkt. Det handlar om att ha rätten att behålla en del information för sig själv och att själv få bestämma över vilken personlig information som lämnas ut.

Bylund (2013) behandlar personlig integritet i sin bok ”Personlig integritet på nätet”. Författaren skriver att vi alla vill bli bedömda utifrån den information som är relevant för en situation. En förutsättning för att kunna upprätthålla relationer med människor är möjligheten att kontrollera hur mycket tillgång andra har till din personliga information. Som människor vill vi kunna styra vem som har tillgång till vår persondata. Bylund (2013) tars som exempel på per-

sonlig integritet upp lanseringen av mobiltelefonen under 1980-talet. Många människor vägrade då att använda telefonen eftersom det uppfattades som en kränkning att kunna bli nådd var som helst och när som helst.

Idag spelar sociala medier en allt större roll i svenskarnas vardag och med detta kommer också en oro att de äventyrar den personliga integriteten. Bylund (2013) säger att det ofta diskuteras om personlig integritet är på väg att utarmas på grund av sociala medier men det är inte sant – det fungerar bara annorlunda idag jämfört med förr. Förändringen handlar om vad användaren delar med sig av, till vilka personer och på vilket sätt. Sociala mediers intåg innebär inte att alla delar med sig av allt till alla. Dock säker Bylund (2013) vidare att det finns stora brister i transparens och starka intressen av tjänsteleverantörer att få tillgång till personlig information av kommersiella skäl, vilket av vissa kan anses äventyra den personliga integriteten. Informationen som utnyttjas handlar om den information som skapas när en person använder internet, på sociala medier är det främst informationen som användaren själv har valt att lägga upp. Informationen kan bestå av exempelvis bilder eller uppdateringar i ett nyhetsflöde. Leverantörerna samlar också in information om det som användaren gillar och kommenterar, vem som finns i användarens nätverk samt produkter, föreningar och företag som användaren tycker om. Kunskapen om användaren används sedan för att sälja riktad marknadsföring och data till andra företag (Bylund, 2013).

2.5 Risker med insamling av persondata

Som ett resultat av teknikens framfart har nya kommunikationsteknologier skapats som möjliggör ständig uppkoppling mot världen. Detta innebär flera fördelar för samhället men den personliga integriteten riskerar att inskränkas. Risk kan definieras som en sak som tvingar eller utgör en fara för människor eller det vi värderar (Cho et al., 2010). Riskerna kan delas upp i två olika dimensioner, dessa dimensioner handlar om de personliga och samhällsmässiga nivåerna. Risker värderas utifrån hur stor sannolikheten är att en risk inträffar och om risken skulle inträffa, hur stor är den negativa effekten. (Cho et al., 2010)

Enligt Cho et al. (2010) tror inte många användare att de själva kommer råka ut för några risker då de har skapat en bild av att deras persondata inte är lika känslig som andras. Således anser många användare att deras persondata är tillräckligt skyddad och att de ej behöver vidta några säkerhetsåtgärder. Dock försvinner den bilden av säker persondata om användaren råkar ut för ett intrång. Användaren blir då mer försiktig med vilken typ av information de lämnar ut och försöker i större utsträckning att skydda sin data. (Cho et al., 2010)

Användarnas oro över den personliga integriteten är ett av de viktigaste problemen som finns i vårt tekniksamhälle och därför är det viktigt att tackla det problemet (Featherman et al., 2010). Vi kommer att nämna de vanligaste riskerna som existerar vid insamling av persondata. Alverén (2012) diskuterar olika förekommande risker med Facebooks hantering av persondata och författaren har valt att dela in riskerna i olika kategorier för att konkretisera och strukturera upp dem. Kategorierna som tas upp är:

- Pinsamheter och personens anseende kan skadas vid intrång
- Anställning och karriär kan påverkas

- Privatliv kränks
- Användaren blir utsatt för statlig övervakning
- Användaren blir drabbad av kriminell verksamhet
- Användaren riskerar att hamna i en filterbubbla

Alverén (2012) nämner både lång- och kortsiktiga risker inom de olika kategorierna som kan drabba Facebooks användare. En risk är att privata konversationer läcks och det kan vara mycket pinsamt och förnedrande. Det finns även användare som råkat ut för att en tredjeparts-sida olovligen tagit dennes profilbild och använt den i olika marknadsföringssyften. Användarna kan råka ut för identitetsstöld, bedrägeri och utpressning. En annan risk som många användare ofta inte tänker på är att de riskerar att hamna i filterbubblan. Vi har tidigare nämnt att det innebär att Facebook bestämmer vad som är viktigt för varje specifik användare utifrån vad användaren har för preferenser på internet. Användaren förlorar då möjligheten att påverka informationen som denne har tillgång till, vilket betyder att viktiga nyheter och information ibland undanhålls då Facebook inte prioriterar det. (Alverén, 2012)

Företag utsätts konstant för attacker av hackers och speciellt när det gäller ett väldigt stort företag, som Facebook. Ibland lyckas attackerna och viktig information läcks (Featherman et al., 2010). Facebook kan hantera användarnas persondata på etiskt vis men det finns ingen garanti för att persondatan inte läcks och missbrukas. En annan risk med Facebook är att omgivningen snabbt kan identifiera personlig information om olika användare utan att de känner varandra. Om det finns provat information om användaren så blir det ett integritetsproblem. (Toch et al., 2012)

2.6 Säkerhet

Xu et al., (2014) delar i deras artikel om informationssäkerhet upp datautvinningsprocessen i fyra olika roller: data provider, data collector, data miner och decision maker. ”Data provider” (hädanefter benämnd som användaren) är personen vars data samlas in av ”data collector” (datainsamlaren) som i sin tur tillhandahåller datan till ”data miner” där en utvinningsalgoritm försöker extrahera användbar data. Slutligen får ”decision maker” (beslutsfattaren) avgöra om den extraherade datan är användbar samt även på vilket sätt den är användbar. Xu et al., (2014) redogör för hur samtliga av rollerna kan vidta åtgärder för att skydda känslig data, men eftersom uppsatsen berör användaren, dvs en enskild eprson som äger en relativt liten mängd data (enligt Xu et al., (2014) känt som mikrodata) redogörs här enbart för hur användaren kan göra för att skydda sin data.

Xu et al. (2014) hävdar att det inte är fel att dela med sig av sin data. Det finns många fördelar med att företag samlar in data i form av tjänster som kan anpassas till den enskilda användaren. Användarens oro är när företag använder datan till okända eller för oetiska ändamål och att användaren då förlorar kontroll över sin data. Vad användaren kan göra för att skydda datan är att försöka återfå kontrollen. Xu et al., (2014) målar upp tre situationer för hur det är möjligt att återta kontrollen:

1. Om användaren anser viss data vara känslig kan denne neka att ge datan till datainsamlaren. Här är ”access-control”-åtgärder effektiva. Dessa åtgärder innebär helt enkelt att kontrollera datainsamlarens åtkomst till användarens data.

2. Användaren inser att persondatan har värde och är villig att utbyta informationen mot vissa förmåner som personligt anpassade tjänster eller pengar. Användaren behöver vara medveten om hur denne ska förhandla med datainsamlaren för att maximera den egna vinsten i utbytet.

3. Om användaren varken kan förhindra datainsamlaren från att få åtkomst till datan eller förhandla om persondatan har användaren ett tredje alternativ; att förvränga datan så att datainsamlaren inte på ett enkelt sätt kan urskilja användarens riktiga data.

En användare ”ger ut” datan till datainsamlaren på ett aktivt eller passivt sätt. Med aktivt menas att användaren frivilligt deltar i undersökningar eller registrerar sig på hemsidor. Med passivt menas att datan som genereras genom vanliga aktiviteter, så som att logga in på Facebook, gilla inlägg, lägga upp bilder etc. samlas in av datainsamlaren. När datan aktivt lämnas ut kan användaren själv välja att inte fylla i information som denne anser vara för privat. Om den personliga datan ändå samlas in finns det vissa metoder att ta till. (Xu et al., 2014)

Enligt Xu et al. (2014) går det att anta att användaren är orolig för att dennes handlingar på nätet skall användas av datainsamlaren på ett vis att det kränker den personliga integriteten. Användaren kan då försöka radera spåren av användningen genom att radera cookies, historik eller genom att använda vissa säkerhetsmetoder. Många av dessa metoder/funktioner är skapade för webbläsaren för att förenkla för användaren. Xu et al., (2014) skriver att nuvarande säkerhetsfunktioner kan kategoriseras i tre typer:

1. Funktioner för att minimera spårning. Kunskapen om att personlig data kan extraheras från användare gör att företag vill spåra användarna på nätet. Detta går att minska genom att använda en funktion som för att datainsamlarna inte kan använda cookies. Populära funktioner är Disconnect, Do Not Track Me och Ghostery. En stor teknik som används för att minimera spårning kallas för Do Not Track (DNT). Denna innebär att användare kan utesluta sig från spårning från hemsidor som de inte besöker. Detta signaleras genom att användaren använder sig av ett http header-namn kallat DNT, vilket betyder att användaren inte vill bli spårad. DNT är inte bara en teknik utan ett ramverk för hur företag ska agera på signalen om att användaren inte vill bli spårad.

2. Reklam- och skriptblockerare. Den här typen av webbläsarfunktion blockerar alltså reklam på hemsidor samt motverkar att skript skickar data till en okänd tredjepart. Exempel på dessa är AdBlock, Plus, NoScript och FlashBlock.

3. Krypteringsfunktioner. Det går att privatisera kommunikation som sker på nätet mellan två användare och det gör att en tredje part inte får tillgång till informationen. MailClock och TorChat är exempel på funktioner som krypterar e-mails och meddelanden som skickas via olika sociala medier. En användare kan också kryptera alla sina spår på nätet genom att använda VPN (Virtual Private Network).

I konjunktion med alla funktioner och metoder som nämns ovan bör användaren också alltid använda anti-virusprogram för att skydda datan mot intrång. Användaren kan alltså minimera

att datainsamlaren får tag i personlig data, men det finns ingen garanti för att ingen data alls samlas in. (Xu et al., 2014)

Förutom att använda funktioner som minimerar spårning kan användaren också ge ut falsk information för att i princip lura datainsamlaren. Xu et al., (2014) tar upp tre metoder för detta.

1. Genom att använda en så kallad ”sockpuppet” för att maskera de verkliga aktiviteterna på internet. En ”sockpuppet” är en falsk online-identitet som användaren gömmer sig bakom, alltså låtsas användaren vara någon annan. Genom att använda flera ”sockpuppets” kan datan som skapas av användarens aktiviteter anses tillhöra flera olika individer. Ett resultat av detta är att användarens verkliga aktiviteter blir okända för datainsamlaren och den personliga datan förblir skyddad.

2. Genom att använda en falsk identitet och skapa falsk information, nästan som en klon av sig själv. Klonidentiteten använder nätet på ett sätt som är ganska olikt den riktiga användarens mönster. När en datainsamlare sedan skall använda datan göms den riktiga datan under den massiva mängd data som klonen automatiskt har skapat för att lura datainsamlaren.

3. Genom att använda säkerhetsfunktioner som maskerar användarens identitet. När användaren registrerar sig på hemsidor eller köper något på nätet ombeds denne oftast att fylla i personlig information, så som e-mailadress, kreditkortsnummer, telefonnummer etc. En webbläsarfunktion som kallas för MaskMe kan hjälpa användaren att skapa och upprätthålla alias (eller en fasad) av personlig information. Användarna kan använda aliasen när de fyller i information och det gör att hemsidorna inte kan få tag på den riktiga informationen då den finns gömd i webbläsarfunktionen.

Något som Xu et al., (2014) också tar upp är att genom att i princip helt skydda sin data förstörs upplevelsen på internet till viss del. Det finns fördelar med att dela med sig av personlig information och dessa inkluderar exempelvis personliga produktrekommendationer som inte hade funnits om inte demografi eller historiken på internet hade sparats.

Förändringar sker hela tiden inom säkerhet på internet eftersom det är ett mycket omdebatterat ämne och en av dessa förändrningar som skett på 2010-talet är hanteringen av cookies. Bylund (2013) tar upp att sedan 2011 säger lagen om elektronisk kommunikation att de som besöker en hemsida som använder cookies måste informeras om att cookies används samt vad de används för. Användaren ska också ge samtycke till detta. Lagen efterlevs till viss del av stora webbplatser men effekten den har på användarens kunskap om att minimera insamlingen av data samt hur mycket data som samlas in är tveksam.

2.7 Big Data

Enligt Bylund (2013) så kommer allt användaren gör på internet att lagras, detta är ett resultat av hur internet är uppbyggt. Merparten av aktiviteterna på internet lagras utan användarens vetskap och informationen som lagras passerar i allmänhet ett flertal datorer på sin väg och det är sällan en enda aktör som kontrollerar informationsprocessen. Datan lagras i dataloggar hos olika tjänste- och internetleverantörer vilket innebär att all kommunikation användare emellan kommer att lagras på olika ställen. Ett exempel på att information lagras är att även

om vi raderar en bild eller annan information från Facebook så kommer den att finnas kvar, aktörerna väljer bara att dölja den för att användaren skall tro att de har raderat informationen. (Bylund, 2013)

Vi har tidigare tagit upp säkerhetsåtgärder som användaren kan ta till för att skydda sin data, detta genom att fokusera på hur användaren kan göra för att skydda sin data i Big Data-sammanhang. Därför är det viktigt att förstå vad Big Data är och hur det kan användas. Det finns olika typer av data, Big data, Small data och Micro data. Big data omfattar de stora informationsmängder som finns på internet samt vilka möjligheter de kan medföra i framtiden (Bylund, 2013). Big data är en term för en enorm och komplex datamängd som kan hanteras av olika internettjänster. Svårigheten med en sådan komplex datamängd är infångandet, lagringen, analysen, sökningen, delningen och överföringen av information samt personlig integritet. Big data refereras oftast till att förutse användarens beteendemönster eller metoder för att extrahera värdefull information från den insamlade datan (Boyd, 2011).

Small data är data som vi människor kan förstå. Våra hjärnor kan inte hantera allt för stor mängd data samtidigt, därför måste den skalas ner till mindre, mer visuella objekt för att skapa förståelse kring en stor mängd data. Exempelvis används relationer eller tabeller för att få en helhetsbild över en viss data som är greppbar för oss människor (jwork.org, 2014). Micro data utgör all individuell information som samlas in om den enskilde användaren. Micro data är mycket flexibel, den behöver inte vara beroende av redan publicerad statistik utan den genererar egen statistik om användarna själva (ipums.org, 2017).

Det finns stora svårigheter med att uppfatta vilken information som användaren lämnar efter sig och vem det är som får tillgång till den personliga datan, samt även hur insamlingen kan tänkas påverka användarens liv. Cookies är ett vanligt verktyg som används och enligt Bylund (2013) är det företagens skapande av användarprofiler som utgör det största hotet mot personlig integritet. Facebook bygger upp användarprofiler utifrån den insamlade datan om användarnas aktiviteter på internet. De bygger upp användarnas flöde med reklam, nyheter, förslag på aktiviteter i närheten samt förslag på personer som användarna kanske känner. På så vis kan Facebook även estimera vilka politiska åsikter och ekonomiska lägen användaren har. (Bylund, 2013)

Bylund (2013) menar att det finns både för- och nackdelar med användarprofilerna som bland annat Facebook skapar. En del människor uppskattar att få en mer personlig upplevelse på internet med ett flöde som är riktat efter användarens egna preferenser då det är bekvämt och användbart på många sätt. Andra hävdar motsatsen, de upplever den riktade reklamen och användarprofilerna som obehagliga då de anser sig övervakade av Facebook.

Idag finns det en stor konflikt mellan leverantörernas intresse och användarvänligheten på sociala medier (Bylund, 2013). Facebooks affärsmodell är kundanpassad reklam som grundar sig i den personliga datan som samlas in och annonserna betalas av annonsörer som vill synas på Facebooks plattform. Persondatan som Facebook innehar omfattar majoriteten av Facebooks värde (Alverén, 2012). Detta innebär att det blir en bättre affär för tjänsteleverantörerna ju mer detaljerad den personliga informationen är. Svårigheten med detta är att begränsa persondatans spridning till flera aktörer samtidigt som det är så företagen tjänar pengar idag (Bylund, 2013).

2.7.1 Rekommendationssystem och Big Data

En del av hur Big data används idag är genom rekommendationssystem. I undersökningen tar vi upp hur personlig data används för att generera personliga rekommendationer till användarna och hur detta är möjligt. Facebook använder sig av rekommendationssystem och därför är det viktigt att definiera vad det är och hur det används.

Rekommendationssystem (RS) består av olika mjukvaruverktyg och tekniker som tillhandahåller olika förslag på olika rekommendationer som kan vara av intresse för användaren. RS är alltså ett informationsfiltreringssystem som försöker förutse olika preferenser och bedömningar (Ricci et al., 2011). RS används i stor utsträckning hos applikationer för att ge användaren en personlig upplevelse. Genom att erbjuda användarna personliga rekommendationer attraherar företagen fler människor till att ansluta sig till tjänsterna (Erkin et al., 2010). Anledningen till varför RS används i så stor utsträckning beror på att det finns en enorm mängd data ute på internet som behöver sorteras och filtreras så att användaren enklare kan finna relevant data (Jeckmans et al., 2013).

Erkin et al. (2010) nämner att istället för att generera generella rekommendationer till användaren kan systemet generera personliga tjänster och teknikerna som används är självklart enormt beroende av informationen som samlas in om användarna. Ju mer information som finns om användaren, desto mer exakta blir rekommendationerna. Data som RS samlar in kan komma både direkt från användaren eller så kan RS observera användarens mönster på internet. Den stora nackdelen med att få mer exakta rekommendationer är att den personliga integriteten riskerar att hotas då den kan missbrukas. Många användare tänker inte på att en stor mängd företag som använder RS tjänar pengar på insamlingen av personlig data. Utåt upplevs många tjänster som företagen erbjuder som gratis, men i verkligheten betalar användarna alltså med sin personliga data (Alverén, 2012).

Det finns olika typer av RS som används av olika företag. Det finns Collaborative filtering, Content-based, Demographic, Knowledge-based, Context-aware, Ensemble, Social och Hybrid system (Jeckmans et al., 2013). Som vi nämnt tidigare i inledningen så använder sig Facebook av Collaborative filtering (CF) för att matcha ihop olika användare baserat på preferenser för att skapa direkta personliga rekommendationer (Terveen et al., 2001). För Facebooks del innebär CF att användarna får hjälp att upptäcka nyheter, sidor, människor, grupper, spel och olika events som anses relevanta för användaren (Code Facebook.com, 2017).

2.8 Mätbar kunskap

Frågeställningarna grundar sig i kunskap för att ta reda på om användarna vet om att Facebook samlar in data om dem. Detta kräver förståelse för vad kunskap är samt hur det är möjligt att mäta den.

Att ha kunskap om någonting innebär enligt Hunt (2003) att det man tror är korrekt och motiverat. Vidare säger han att människors kunskap påverkar deras säkerhet, effektivitet och tillfredsställelse i väldigt stor utsträckning när det kommer till hur personliga eller organisatoriska mål sätts upp. En individs beteende och prestationer beror både av kunskap som har lärts in och av erfarenhet, samt även genom kroppsliga funktioner (se, känna, muskelminne etc.).

För att definiera om kunskap finns hos en individ, måste kunskapen förutom att erhållas också kommas ihåg och användas. Om kunskapen inte används har lärandet misslyckats.

Hunt (2003) diskuterar även hur kunskap kan mätas och skriver att de flesta test som används idag för att mäta kunskap om ett visst ämne är opartiska, alltså skall inte utformarens personliga värderingar finnas med. Flervalsfrågor är vanligt förekommande för att ge en tydligare bild av om respondenten faktiskt har kunskap. Om respondenten svarar fel på en given fråga anses denne vara oinformerad. Detta anseende är dock fel enligt Hunt (2003). Det kan vara så att respondenten är felinformerad – vilket är mycket värre än att vara oinformerad. På grund av det är det viktigt att ha en distinktion mellan ett korrekt och icke-korrekt svar, således att ge respondenten en chans att svara ”Jag vet inte/Jag är inte helt säker/Jag är ganska säker” på frågan eller påståendet.

3. Metod

Arbetet inleddes med teoretisk efterforskning för att förvärva relevant kunskap till frågeställningarna. De kanaler för kunskap som användes var främst Google Scholar där vi kunde hitta tidigare forskning om bland annat datautvinning, rekommendationssystem och risker med datautvinning i allmänhet. Vi har även använt oss av den information Facebook själva publicerat genom deras hemsida samt statistik från Internetstiftelsen i Sverige. Vetenskapliga publikationer på hemsidor har också använts för att definiera en del begrepp.

3.1 Metodval

Eftersom syftet med undersökningen är att ta reda på om svenskar i åldrarna 20-29 år har kunskap om hur Facebook samlar in data om dem samt hur datan hanteras krävdes det att data samlades in från en större mängd personer – därför valde vi att genomföra en kvantitativ studie i form av en enkät.

Då vi ämnar jämföra hur personer med god kunskap om Facebooks datainsamling väljer att skydda/inte skydda sin data jämfört med en person med sämre kunskap måste kunskapen kunna mätas. Eftersom kunskap är ett abstrakt begrepp har vi valt att definiera i teorigenomgången hur kunskapen kan mätas och vi valde att ställa tio indikativa frågor om Facebooks datautvinning och kategoriserade personerna i fem olika kategorier: Mycket Dålig Kunskap, Dålig Kunskap, Viss Kunskap, Bra Kunskap och Mycket Bra Kunskap. Vi valde att kategorisera på det viset för att indikatorfrågorna kunde ge rätta svar i ett spann mellan 0-10 och att dela upp kunskapsgrupperna i ett omfång om 2 poäng per kunskapsgrupp kändes rimligt då 1 poäng genererar för många kunskapsgrupper och kunde potentiellt innehålla alldeles för många kunskapsnivåer.

Enkätstudien grundar sig alltså i en kvantitativ metod som innehöll en liten mängd förhandsdefinierad information och urvalet bestämdes före datainsamlingen genomfördes. Syftet med enkätundersökningen var att beskriva, förklara och påvisa samband för hur mycket kunskap användarna har. Slutsatsen blir därav generell och återanvändbar då resultaten kan dras till en större population (Svensson, 2015).

3.2 Utformning av enkätfrågor

Vi utformade frågorna utifrån Jacobsens (2002) förklaringar av olika faser. Enkäten är en tvärsnittsstudie vilket innebär att verkligheten studeras vid endast en tidpunkt. Tvärsnittsstudier är enligt Jacobsen (2002) förmodligen den vanligaste typen av undersökning då undersökaren slipper vänta särskilt lång tid på att datan skall kunna samlas in. På det sättet kunde vi ta reda på om svenskar i åldrarna 20-29 år vet om att Facebook samlar in data om dem och även om de aktivt gör något för att skydda sin data om de har kunskap om riskerna som finns. Vi kan också jämfört användare som har mer kunskap om Facebooks datainsamling skyddar/inte skyddar sin data jämfört med de användare med mindre kunskap. Tvärsnittsstudien tillhandahöll inte kunskap eller gav användarna någon möjlighet till utveckling av kunskap, utan undersökte helt enkelt exakt hur det är just nu och hur stor kunskap de har.

Eftersom enkätfrågorna grundar sig i användarnas kunskap valde vi att skapa en blandning av tjugo frågor och påstående som bestod av flervalssvar för att minska risken för felinformation (Hunt, 2003). Svarsalternativen bestod till viss del av en skala med fem olika svarsalternativ; ”Ja, helt säker/Ja, ganska säker/Vet ej/Nej, ganska säker/Nej, helt säker”. Genom att ge användaren en chans att inte vara helt säker på att svaret som ges är korrekt är det också möjligt att mäta olika nivåer av kunskap, nivåer som inte hade kunnat uppmätas med enbart ”Ja/Nej”-frågor.

Frågorna/påståendena är objektivt utformade då det enligt Jacobsen (2002) är mycket viktigt att ha konkreta frågor/påståenden vid undersökningar för att undvika misstolkningar. Det är också viktigt att reflektera kring felkällor som kan påverka resultaten av enkäten. Några av de vanligaste felkällorna är att frågorna inte är tydligt formulerade, att respondenten inte representerar hela rampopulationen som skall undersökas, att urvalsgruppen inte var tillräckligt stor, att det kan finnas brister i hur insamlingen av statistiken har gått tillväga samt hur den insamlade datan har bearbetats (Weisberg, 2005).

3.3 Urval av respondenter

Eftersom den största användargruppen av Facebook är 16-25-åringar följt av 26-35-råingar var det relevant att välja en tvärsnittsgrupp på användare mellan 20-29 år då dessa fortfarande till stor del använder Facebook i det dagliga livet. Detta jämfört med yngre användare som ofta använder andra sociala medier, så som Snapchat och Instagram istället (Internetstiftelsen, 2016).

Vi ansåg att det var intressant att undersöka en relativt snäv åldersgrupp och deras kunskap om Facebook för att ta reda på hur mycket kunskapen skiljde sig bland användare som till viss del befinner sig i samma stadie av livet.

3.4 Datainsamling

Enkäten skapades i gratistjänsten Google Forms och distribuerades över Facebook och delades med Facebookvänner, i grupper, samt även till utomstående då enkäten delades vidare av andra användare. Detta var möjligt då enkäten var offentlig och möjlig att dela av andra. Enkäten kom upp i Facebookanvändares flöde ett flertal gånger då den gillades och kommenterades av andra användare. Detta gjorde att genomströmningen blev större eftersom fler användare fick se den, exempelvis våra Facebookvänners andra vänner. Enkäten var aktiv i en (1) vecka och därefter låstes den. Genomströmningen hade förmodligen varit ännu lite större om enkäten varit aktiv längre än så.

Totalt samlades det in 128 svar men bortfallet blev tre stycken då dessa inte var inom den definierade åldersgruppen. Möjligheten att kunna se enskilda svar gjorde det enkelt att bortse från dessa användares svar. 125 svar på enkäten kunde därför analyseras för att sedan hjälpa till att besvara frågeställningarna.

Gratisverktyget Google Forms skapade automatiskt cirkeldiagram som delade upp svaren och under tiden enkäten var aktiv gick vi ofta in för att kontrollera antalet svar och även titta på enskilda svar för att bilda någon sorts uppfattning innan det verkliga analysarbetet började.

3.5 Kvalité av undersökningen

Det är oerhört viktigt att fastställa att undersökningen uppnådde hög undersökningskvalité för att få tillräckligt med data att analysera. Detta för att få ett tillförlitligt resultat som kunde hjälpa oss att besvara frågeställningarna. Vid utformningen av enkäten valde vi som sagt att följa Jacobsens (2002) struktur. De faktorerna som påverkade hur pålitliga och goda slutsatserna blev är: reliabilitet, validitet, källkritik och etik.

3.5.1 Reliabilitet

Vi valde att fokusera på kvalité i arbetsprocessen för att uppnå god reliabilitet. Vi ville skapa en trovärdig och tillförlitlig undersökning och som ett resultat av detta utformade vi enkätfrågorna på ett sätt som enbart genererar relevant data (Jacobsen, 2002). Vi behandlade resultaten av enkäten kort efter att enkäten stängdes för att inte låta det gå för lång tid mellan insamlingen av datan och analyseringen (Jacobsen, 2002).

3.5.2 Validitet

Enligt Jacobsen (2002) är det oerhört viktigt att validera sina källor under arbetets gång. Därför har vi under hela processen validerat och kontrollerat att källorna vi arbetat med var relevanta och tillförlitliga för undersökningen och faktiskt mätte det vi ville. Detta baserades på källans förmåga att ge oss sann information om det vi ville ta reda på i undersökningen. Validiteten ökades genom kontinuerlig kontroll gentemot annan teori och empiri (Jacobsen, 2002).

Enkäten som delades med användarna var, som tidigare nämnt uppbyggda av Google Forms där länken gav tillgång till att besvara enkäten. En nackdel med detta var att enkäten inte krävde inloggning vilket innebar att den kunde spridas felaktigt och på så vis påverka validiteten av resultatet. Vi anser dock att validiteten inte har påverkats då vi antar att användarna har besvarat enkäten utifrån bästa möjliga förmåga och har svarat sanningsenligt. Svaren var anonyma och således fanns det ingen anledning för användarna att ljuga. De hade, enligt oss, ingen yttre påverkan som skulle ge ett missvisande resultat. Redigeringslänken delades med alla uppsatsförfattare via e-mail så att alla kunde se specifika svar. Vi tror inte att länken har spridits till fel människor såvida ingen hackade våra e-mails, men det är mycket osannolikt.

3.5.3 Källkritik

Litteraturen som använts till teorin består av relativt nya artiklar och böcker. Litteraturen grundade sig för det mesta i etablerad vetenskap och var publicerade från olika universitet vilket gör att de kan anses tillförlitliga. Vi aktade oss för att plocka information från sociala medier som inte var verifierade då informationen inte nödvändigtvis var korrekt.

Det fanns vissa motsägelser i litteraturen som vi fick ta hänsyn till och hantera under arbetets gång. Det finns exempelvis många olika definitioner på personlig integritet och var det är. Det *är* ett svårdefinierat begrepp som beror av situationen, och därför fick vi läsa flera olika artiklar för att skapa en helhet att utgå ifrån. En del litteratur utgick endast från en samhällelig och mer generell beskrivning av personlig integritet och vi var mer intresserade av en IT-relaterad definition. På grund av motsägelserna var det därför svårt att ibland se en del av artiklarnas relevans för teorin.

Under tiden enkäten var aktiv mottog vi en del frågor av vad vissa begrepp betydde och vi försökte att inte förklara för mycket då hela poängen med enkätundersökningen var att ta reda på och mäta kunskap hos användarna.

3.5.4 Etik

Enligt Jacobsen (2002) är det viktigt att vara tydlig med syftet när man utför en undersökning, detta för att uppnå god etik. Därför skrev vi en inledande text när enkäten delades där vi förklarade vilken målgrupp vi riktade oss mot samt kortfattat vad enkäten handlade om. Det var också viktigt att användarna inte bara blev informerade utan att de även förstod frågorna i enkäten så resultatet blev tillförlitligt. Alla svar som samlades in var som sagt anonyma för att skydda användarnas personliga integritet (Jacobsen, 2002).

4. Resultat

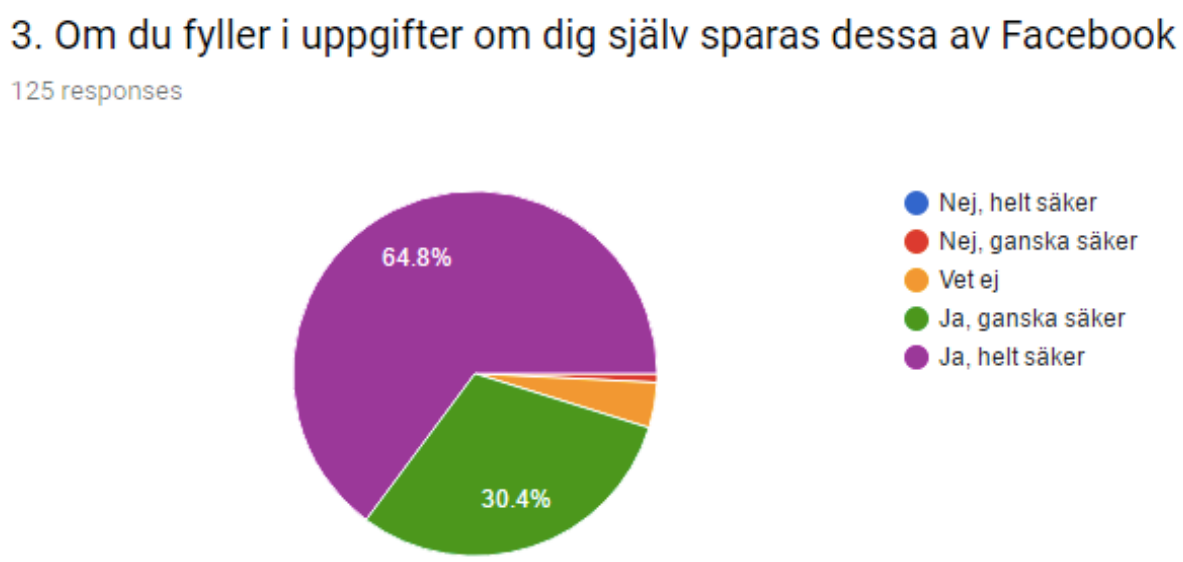
Som tidigare nämnt i metodavsnittet ställdes indikatorfrågor för att ta reda på användarnas kunskapsnivå i ämnet. När svaren från enkäten hade samlats in togs kunskapsnivån fram genom att utdela poäng enligt följande Tabell 4.

Tabell 4.1: Kunskapsnivå

Svar	**Poäng**
Korrekt, helt säker	1
Korrekt, ganska säker	0,5
Vet ej	0
Inkorrekt, ganska säker	-0,5
Inkorrekt, helt säker	-1

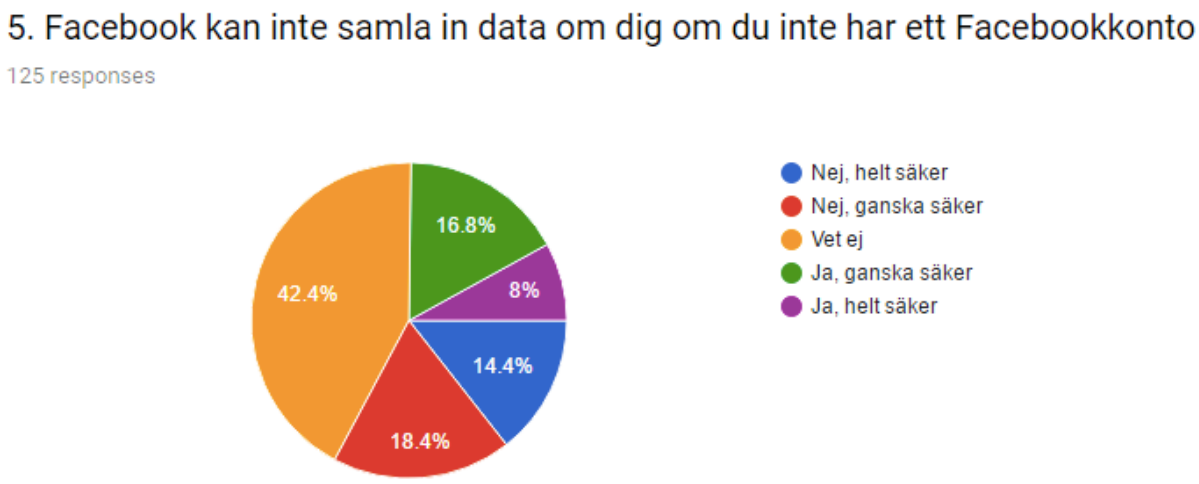
Människor agerar utifrån att den kunskap de besitter och tror är korrekt, och ju säkrare en person är på sin kunskap, desto mer säkert är det att ett beslut kommer att ske. Beslutet sker då även snabbare. Om en person besitter felaktig kunskap är det troligt att besluten och agerandet utifrån den kunskapen också är felaktigt och dåligt (Hunt, 2003). Med detta i åtanke valde vi att ge minuspoäng för inkorrekta svar eftersom det anses bättre att vara oinformerad än att vara felaktigt informerad.

Så här besvarade användarna indikatorfrågorna:

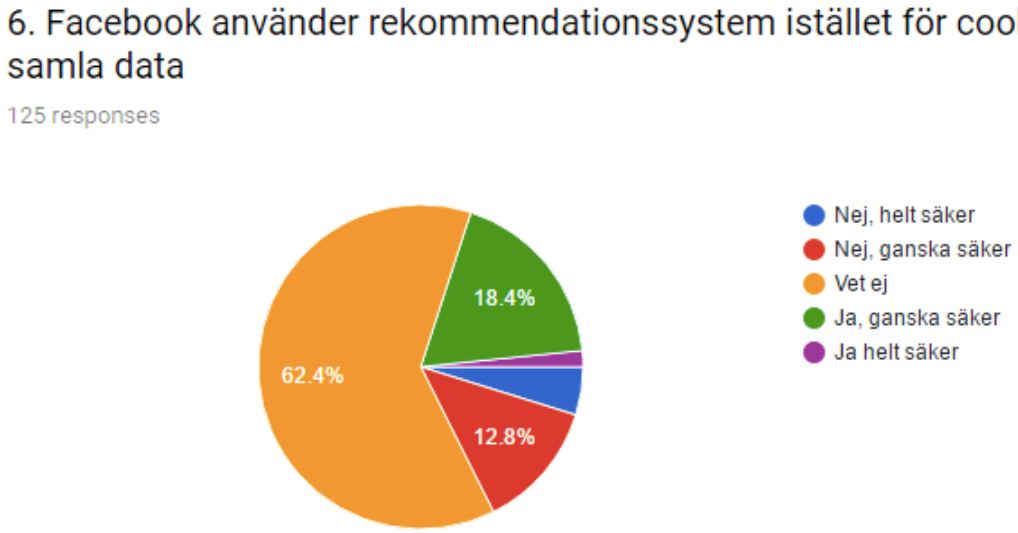


Figur 1. Indikatorfråga 1

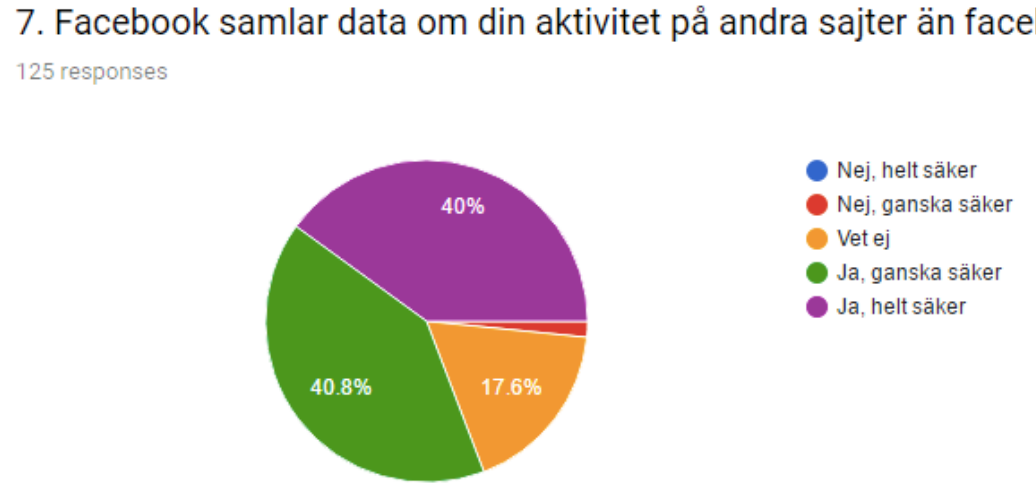
Figur 2. Indikatorfråga 2



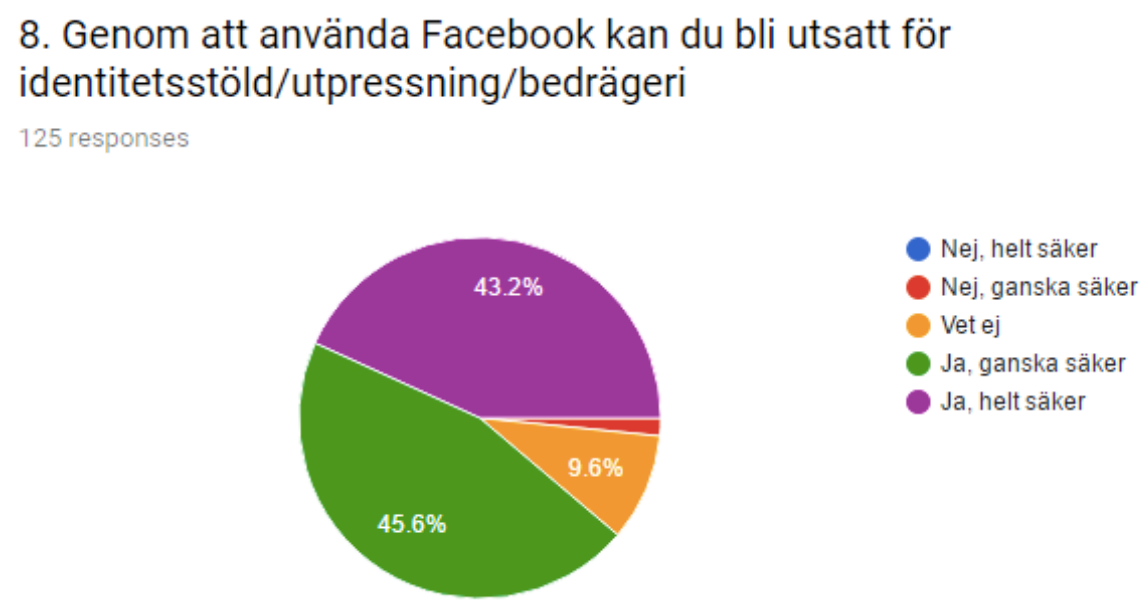
Figur 3. Indikatorfråga 3



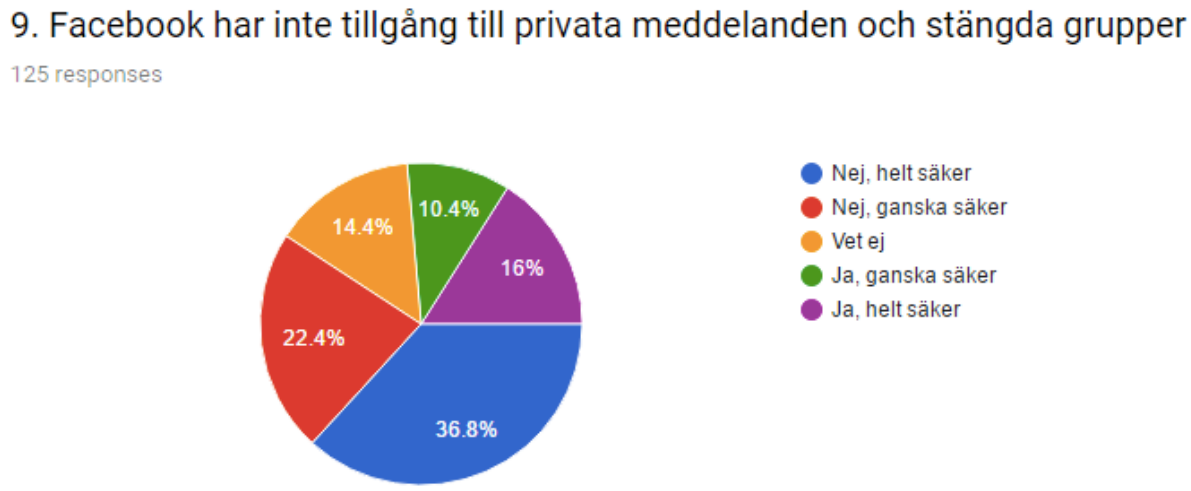
Figur 4. Indikatorfråga 4



Figur 5. Indikatorfråga 5



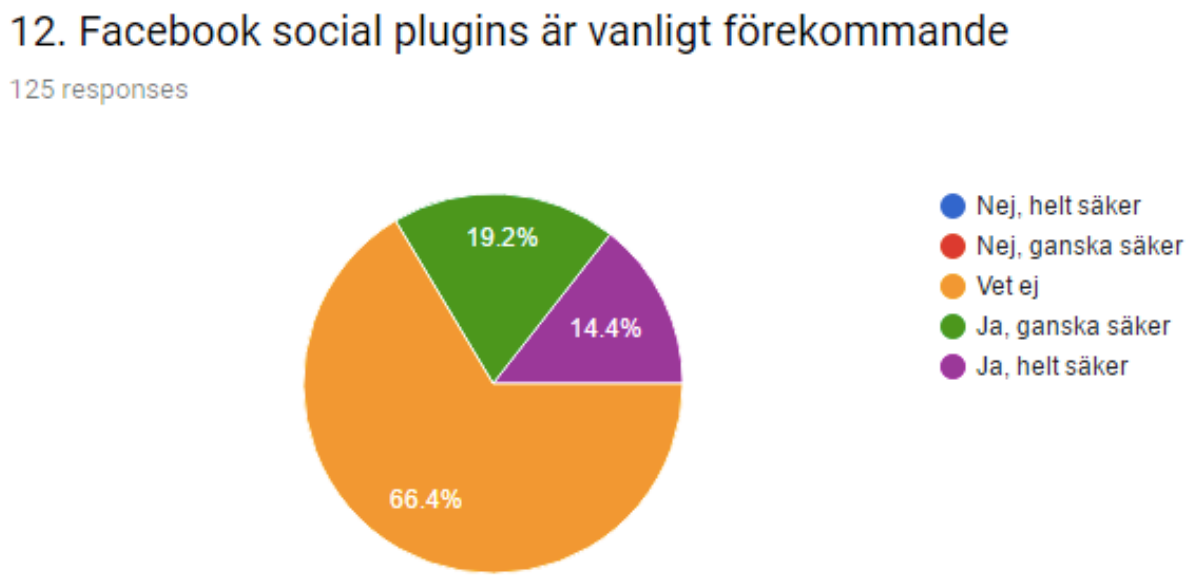
Figur 6. Indikatorfråga 6



Figur 7. Indikatorfråga 7

Figur 8. Indikatorfråga 8

Figur 9. Indikatorfråga 9

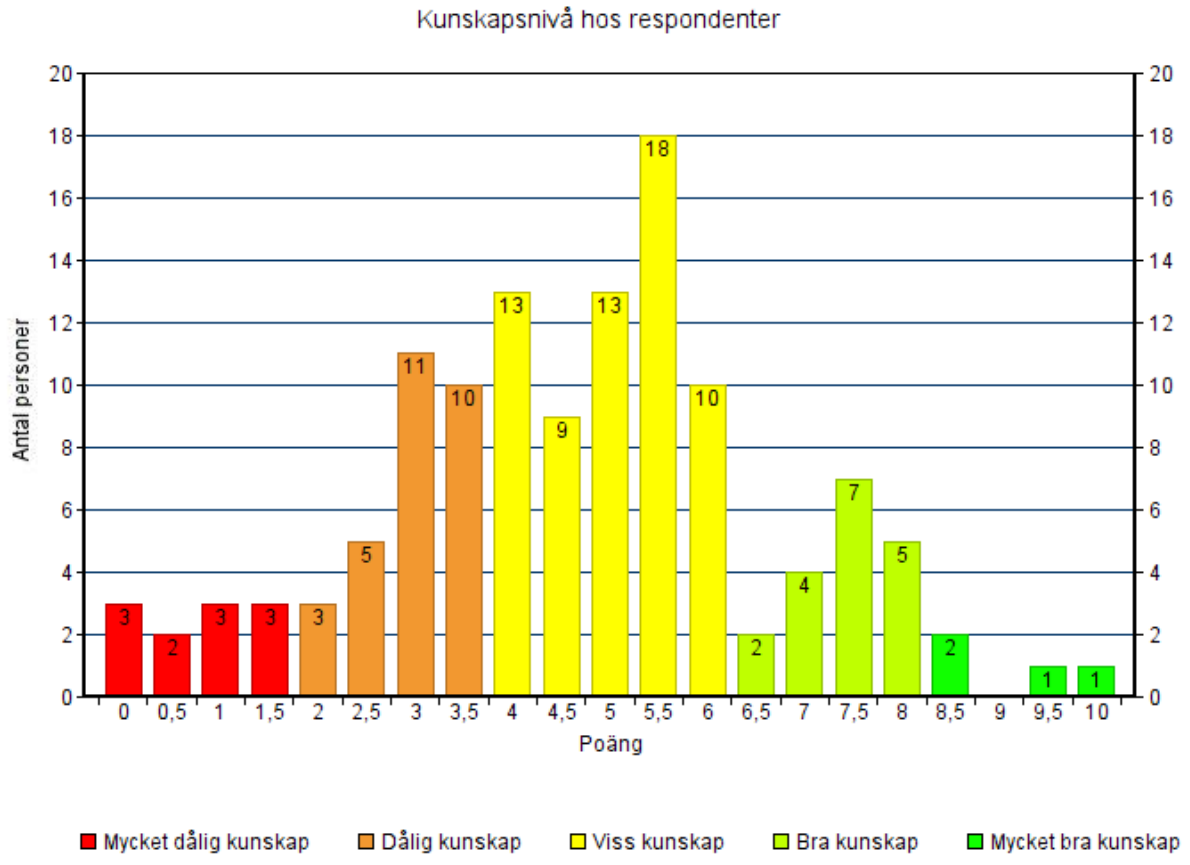


Figur 10. Indikatorfråga 10

Respondenterna delades sedan in i kunskapsgrupper utifrån indikatorfrågorna enligt följande tabell:

Tabell 4.2: Kunskapsgrupper

Kunskapsgrupp	**Poäng**
Mycket Dålig Kunskap (MDK)	0-1,5
Dålig Kunskap (DK)	2-3,5
Viss Kunskap (VK)	4-6
Bra Kunskap (BK)	6,5-8
Mycket Bra Kunskap (MBK)	8,5-10



Figur 11. Kunskapsnivå hos användare

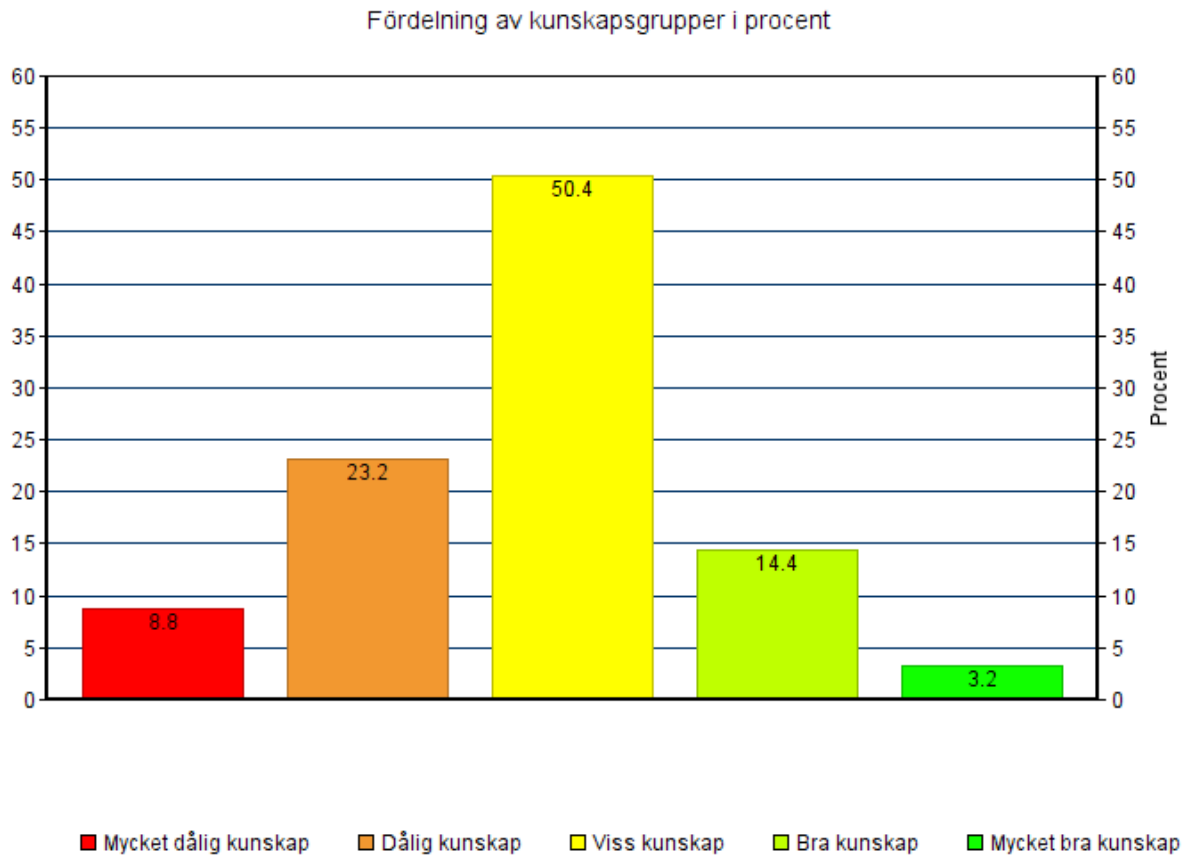
I figur 11 kan poängfördelningen för användarna ses. Vi kan se en topp mellan 3-6 (DK till VK) där de allra flesta har hamnat. Vi kan även se en mindre topp på 7,5 poäng i BK. Antalet personer på den extrema änden till vänster om toppen (MDK) är många fler än antalet personer på den högra änden av toppen (MBK). Genomsnittet är 4,652 poäng medan medianen är nästan en hel poäng högre på 5,5 poäng.

Det totala antalet personer i varje kunskapsgrupp är som följer:

Tabell 4.3: Antal personer i kunskapsgrupperna

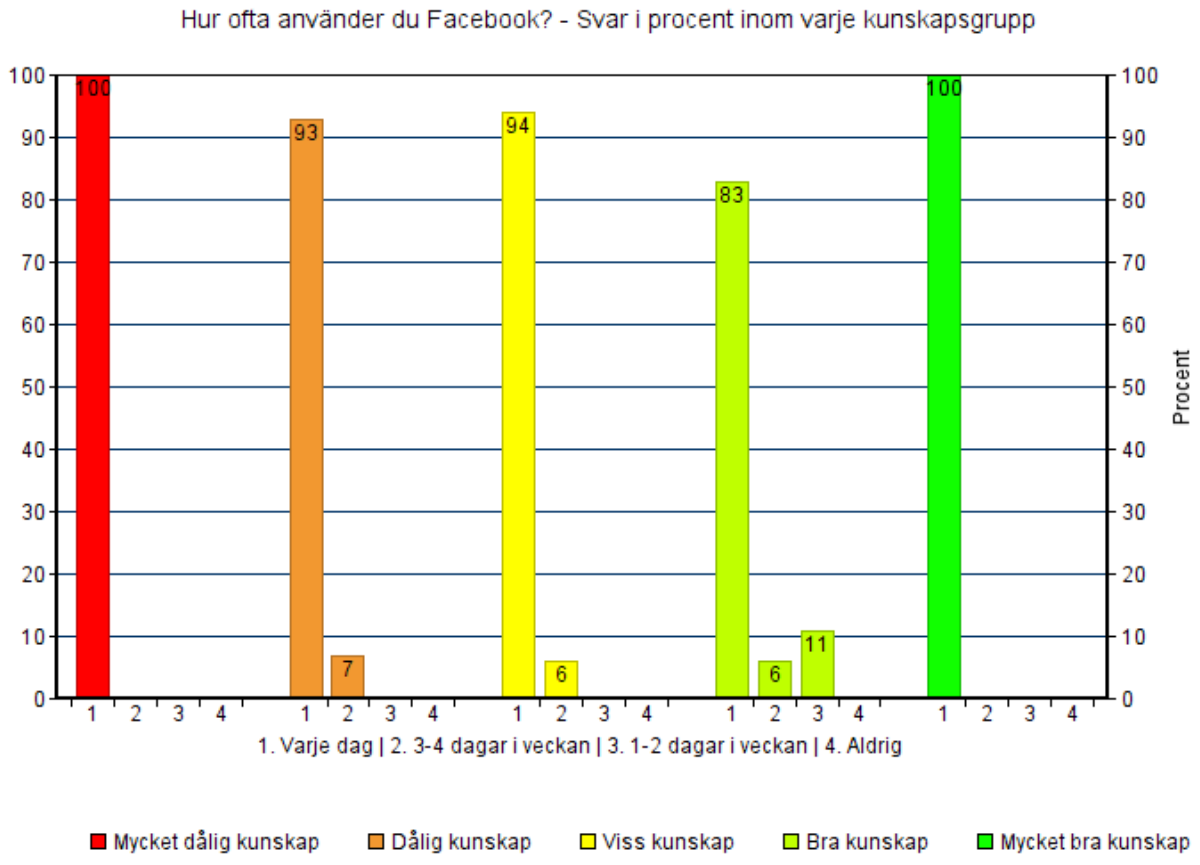
Kunskapsgrupp	Antal
Mycket Dålig Kunskap (MDK)	11
Dålig Kunskap (DK)	29
Viss Kunskap (VK)	63
Bra Kunskap (BK)	18
Mycket Bra Kunskap (MBK)	4

Tyvärr var det endast fyra personer som uppnådde MBK, vilket inte är ett bra statistiskt underlag. Gruppen MDK må ha nästan 300% fler användare än gruppen MBK, men även de har ett väldigt dåligt statistiskt underlag med endast elva användare. De övriga gruppernas användare ser något bättre ut med 29, 63 och 18 användare.



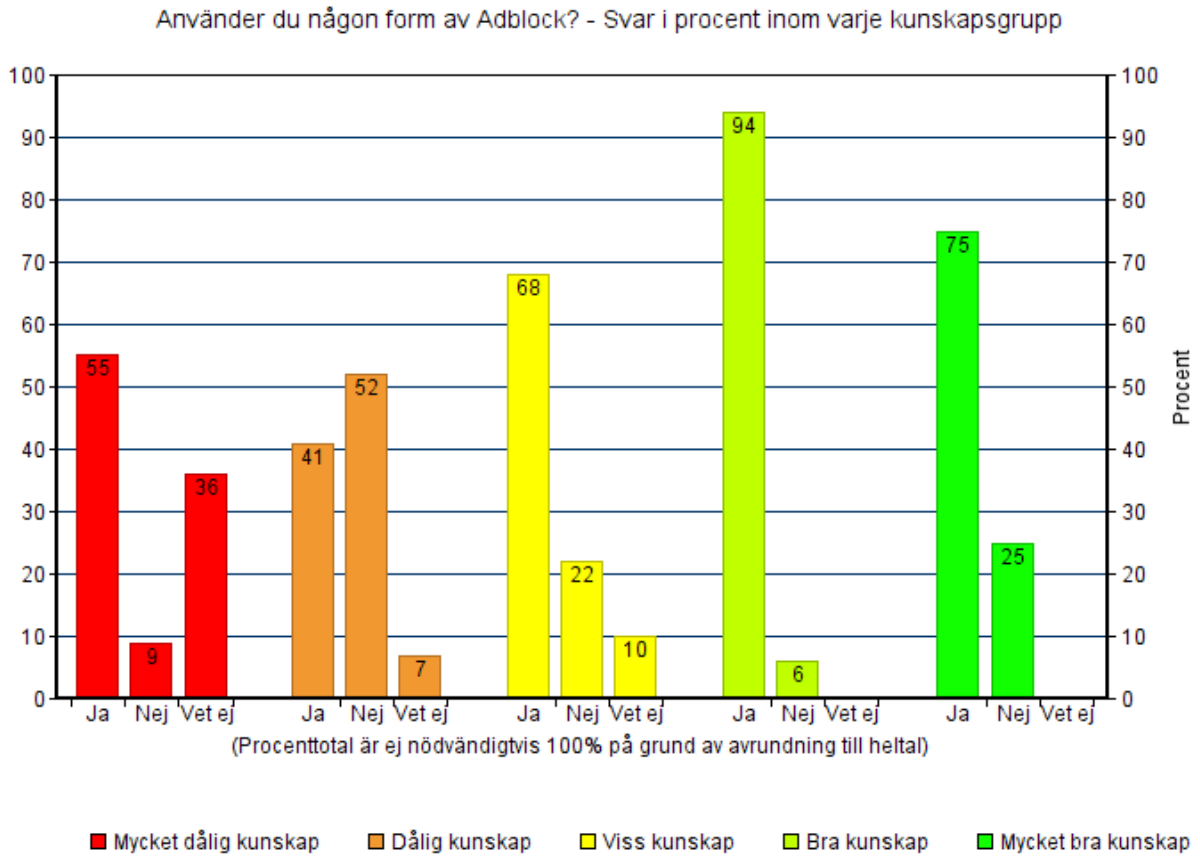
Figur 12. Fördelning av kunskapsgrupper i procent

I figur 12 kan vi se varje kunskapsgrupps fördelning i procent. En liten majoritet på 50,4% har VK, följt av 23,2% med DK, 14,4% med BK, 8,8% med MDK och slutligen 3,2% med MBK.



Figur 13. Kunskapsgruppernas frekvens i användande av Facebook i procent

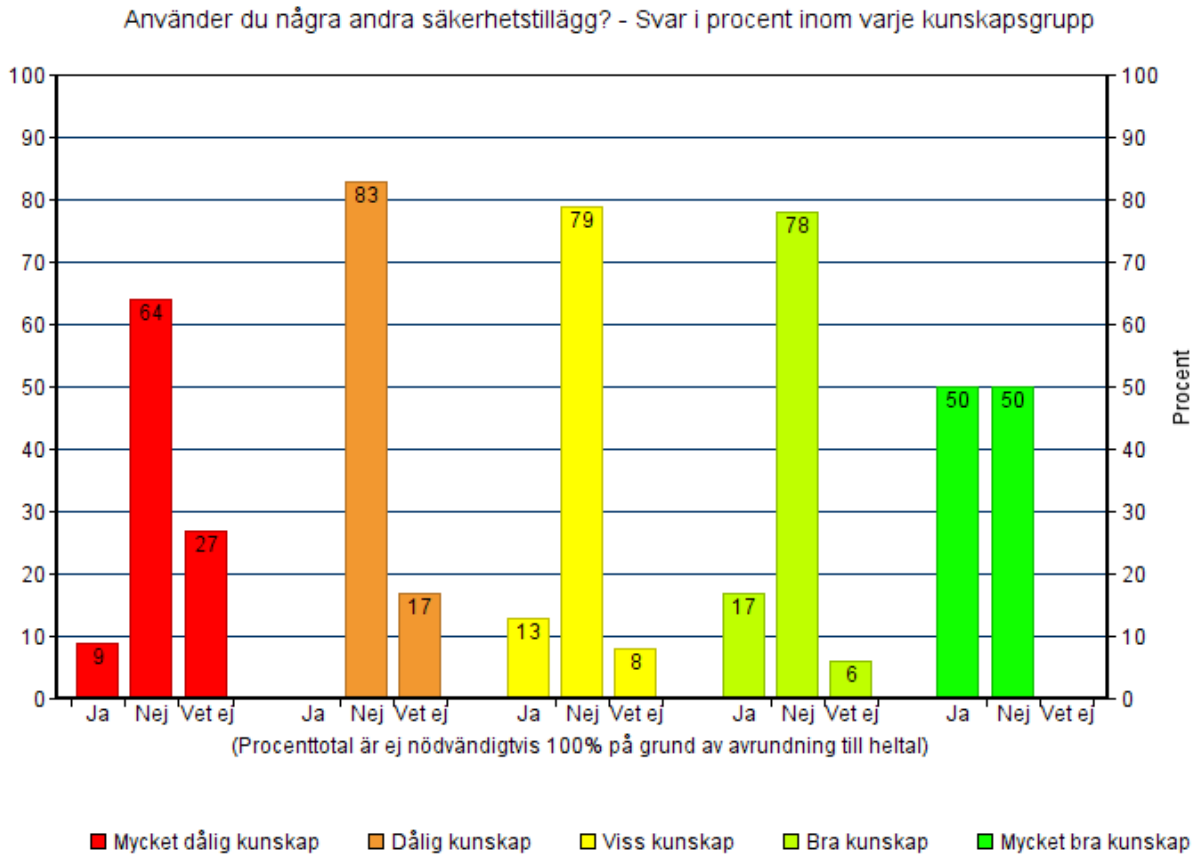
I figur 13 ser vi hur ofta de olika kunskapsgrupperna använder Facebook. Samtliga med MDK och MBK svarade att de använder Facebook varje dag. Användarna med DK och VK var nästan identiska med 93 respektive 94% som använder Facebook varje dag och 7% respektive 6% använder Facebook 3-4 dagar i veckan. Bland de i BK gick användandet varje dag ner till 83%, användandet 3-4 dagar i veckan såg likadant ut som för DK och VK på 6% och användandet 1-2 dagar i veckan steg till hela 11%. Ingen i grupperna svarade att de aldrig använde Facebook och det är bra då enkätundersökningen riktade sig till Facebookanvändare.



Figur 14. Svar i procent för om kunskapsgrupperna använder någon form av AdBlock

I figur 14 ser vi hur många procent av användarna i varje kunskapsgrupp det är som använder någon form av AdBlocker i sina webbläsare. Vi kan se att de med MDK använder AdBlocker i högre grad än de med DK (55% respektive 41%), men vi kan också se att de med MDK inte vet om de använder AdBlockers i mycket högre grad än övriga kunskapsgrupper.

Användandet av AdBlockers ser ut att öka i samband med ökad kunskapsnivå från DK till BK (41% - 68% - 94%), men mängden användare som inte vet om de använder en AdBlocker håller sig på ungefär samma nivå. Användandet av AdBlocker sjunker till 75% för gruppen med MBK, men ingen i den gruppen svarade ”Vet ej”.

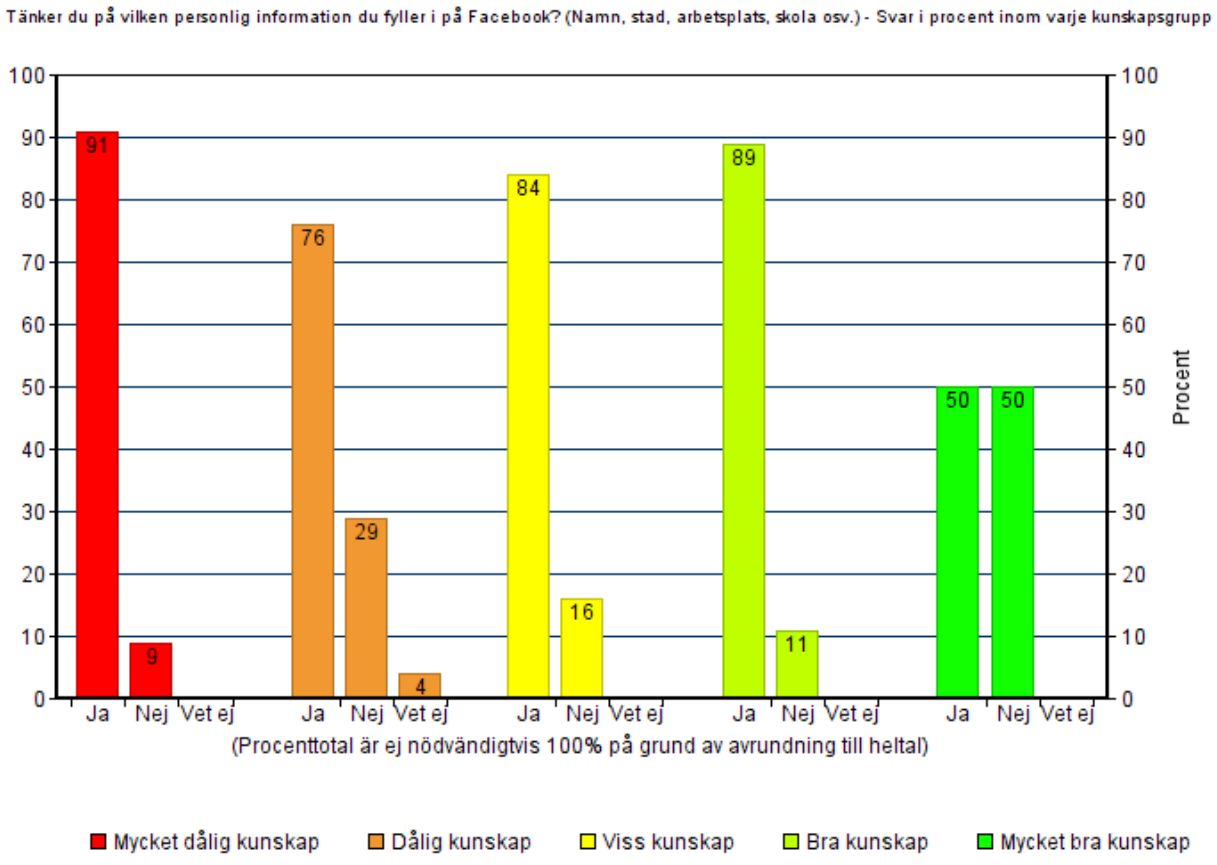


Figur 15. Svar i procent för om kunskapsgrupperna använder andra säkerhetstillägg

I figur 15 kan vi se hur de olika grupperna använder andra säkerhetstillägg i sina webbläsare och likt figur 4 kan vi se att användarna med MDK använder säkerhetstillägg i högre grad än användarna med DK. Även här är användarna med MDK överrepresenterade när det gäller att svara ”Vet ej”.

Ingen i gruppen DK svarade att de använder andra säkerhetstillägg, men jämfört med MDK hade de som inte vet om de använder andra säkerhetstillägg sjunkit till 17%. De som svarat att de använder andra säkerhetstillägg ökar för grupperna VK och BK (13% - 17%) samtidigt som de som svarat ”Vet ej” sjunker till 8% respektive 6%. Användarna som svarade ”Nej” sjunker marginellt från DK till BK (83% - 79% - 78%).

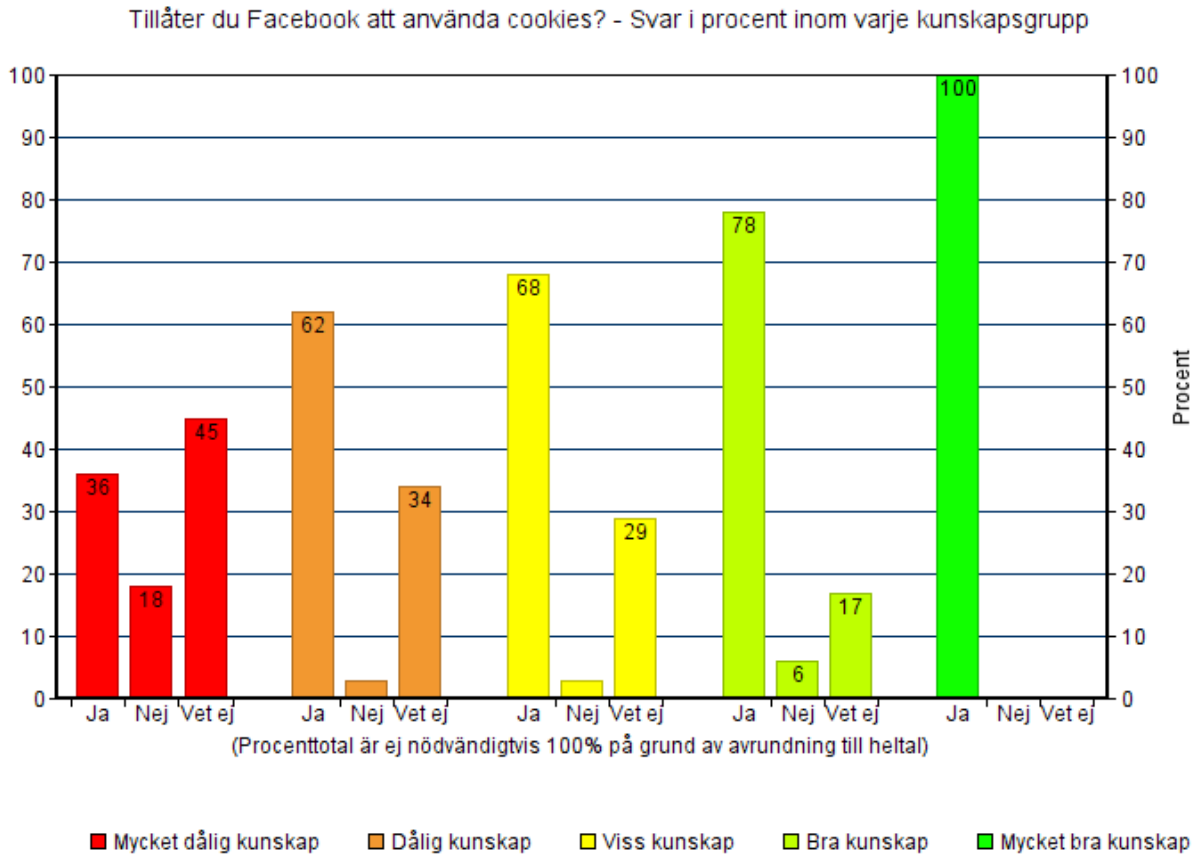
Gruppen med MBK urskiljer sig iden där 50% har svarat ”Ja”, 50% har svarat ”Nej” och ingen svarade ”Vet ej”.



Figur 16. Svar i procent för kunskapsgrupperna om de tänker på vilken personlig information de fyller i på Facebook

I figur 16 ser vi hur extremerna med MDK och MBK sticker ut igen. MDK har högst andel personer som svarat ”Ja” av alla grupper med 91% och i kontrast har MBK lägst andel personer där 50% har svarat ”Ja” och 50% har svarat ”Nej”.

Bland DK-BK ser siffrorna ut ungefär som de kan förväntas igen, där andelen som svarat ”Ja” ökar i takt med kunskapsnivån (76% - 84% - 89%). Något som däremot sticker ut, om inte särskilt mycket, är att DK har fyra användare som svarat ”Vet ej”, vilket ingen annan grupp har.

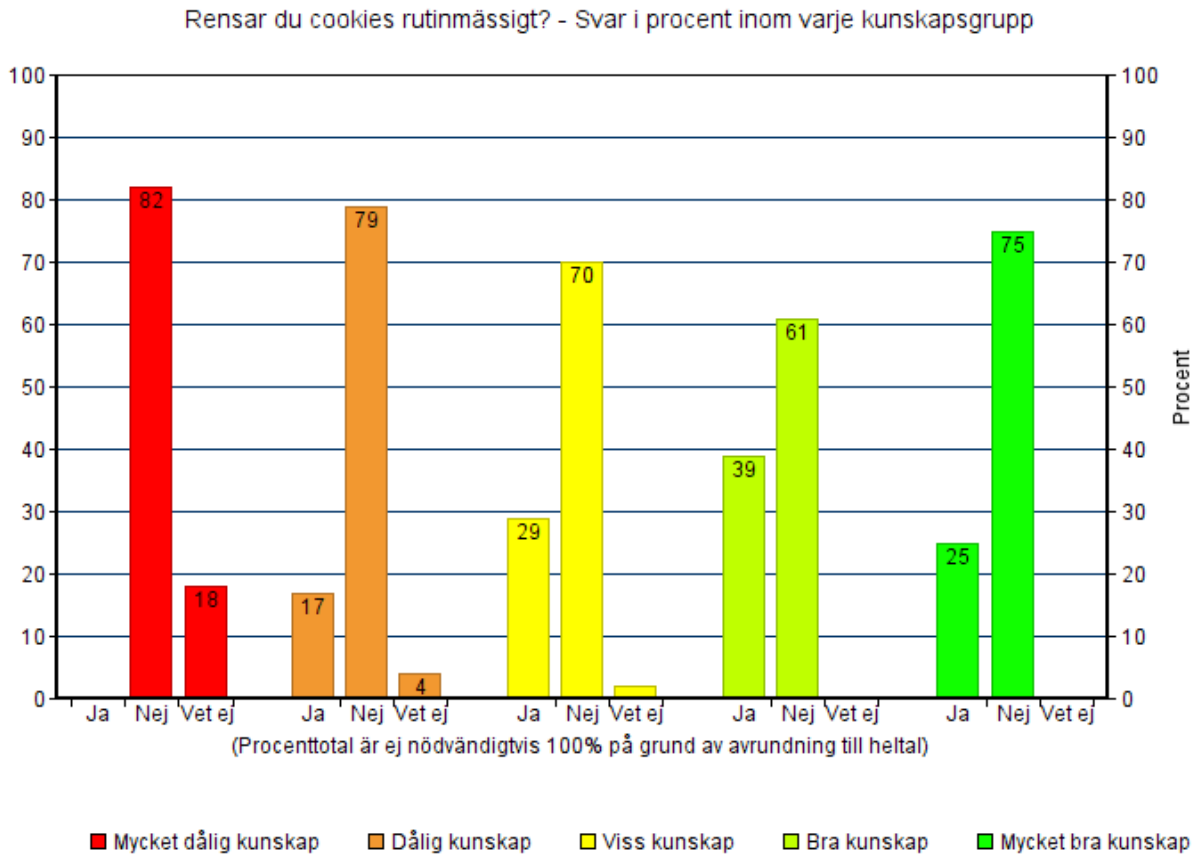


Figur 17. Svar i procent på om kunskapsgrupperna tillåter Facebook att använda cookies

MDK följer den tidigare trenden och sticker ut med högst antal ”Nej”-svar (18%) tillsammans med lägst antal ”Ja”-svar (36%) och högst antal ”Vet ej”-svar (45%). Även MBK sticker ut igen med 100% ”Ja”-svar.

Som tidigare fortsätter de tre mellangrupperna att visa mer förväntade svar jämfört med de två yttergrupperna. DK och VK har samma mängd som inte tillåter cookies (3%) och skillnaden ligger enbart i att färre användare svarat ”Vet ej” i VK (34% - 29%).

BK har snäppet högre andel som svarat att de inte accepterar cookies (6%) och enbart 17% har svarat ”Vet ej”. Tillåtandet av cookies går upp i takt med att kunskapsnivån ökar.

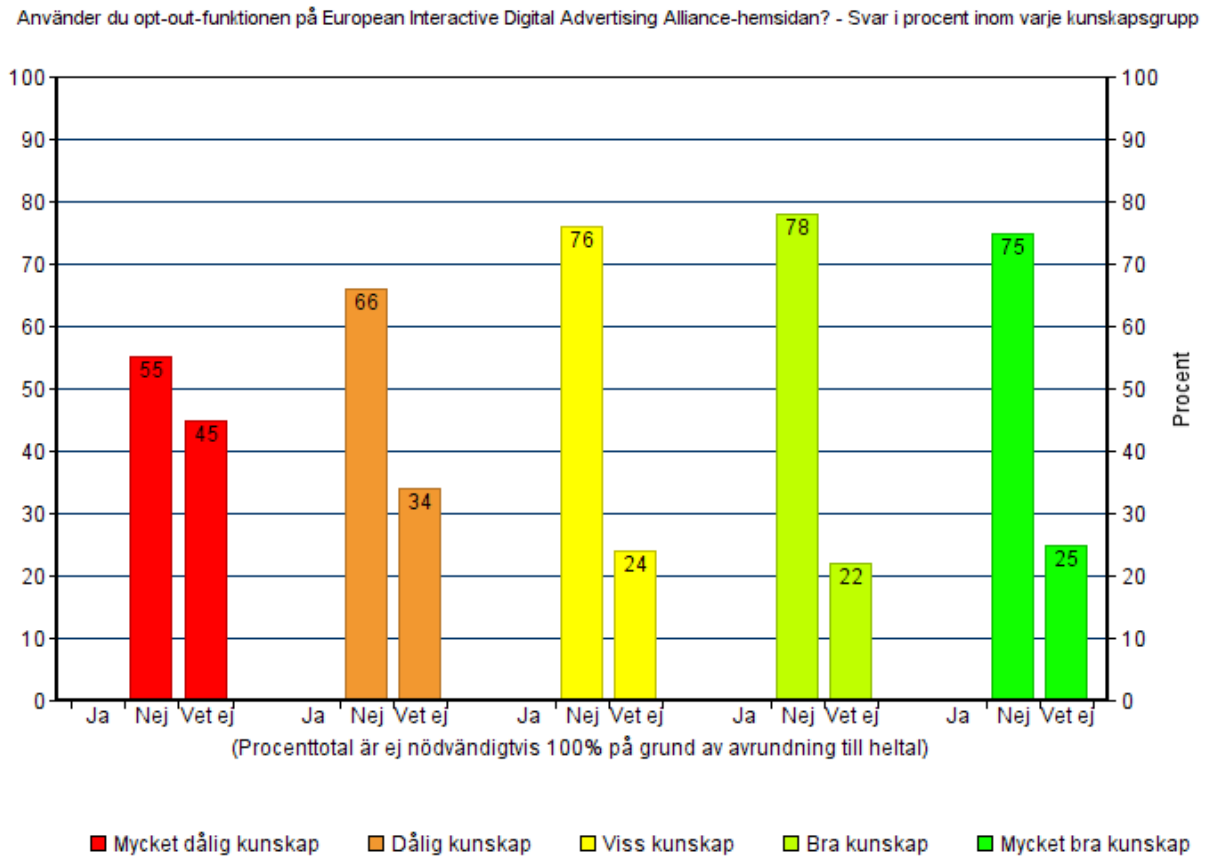


Figur 18. Svar i procent på om kunskapsgrupperna rensar cookies rutinmässigt

MDK bryter här sin tidigare trend med överraskande svar och svarar istället ungefär som förväntat. ”Ja”-svaren går upp från 0% i MDK till 17% i DK, 29% i VK och 39% i BK. ”Nej”-svaren faller från 82% i MDK till 79% i DK, 70% i VK och 61% i BK. ”Vet ej”-svaren faller från 18% i MDK till 4% i DK, 2% i VK och 0% i BK.

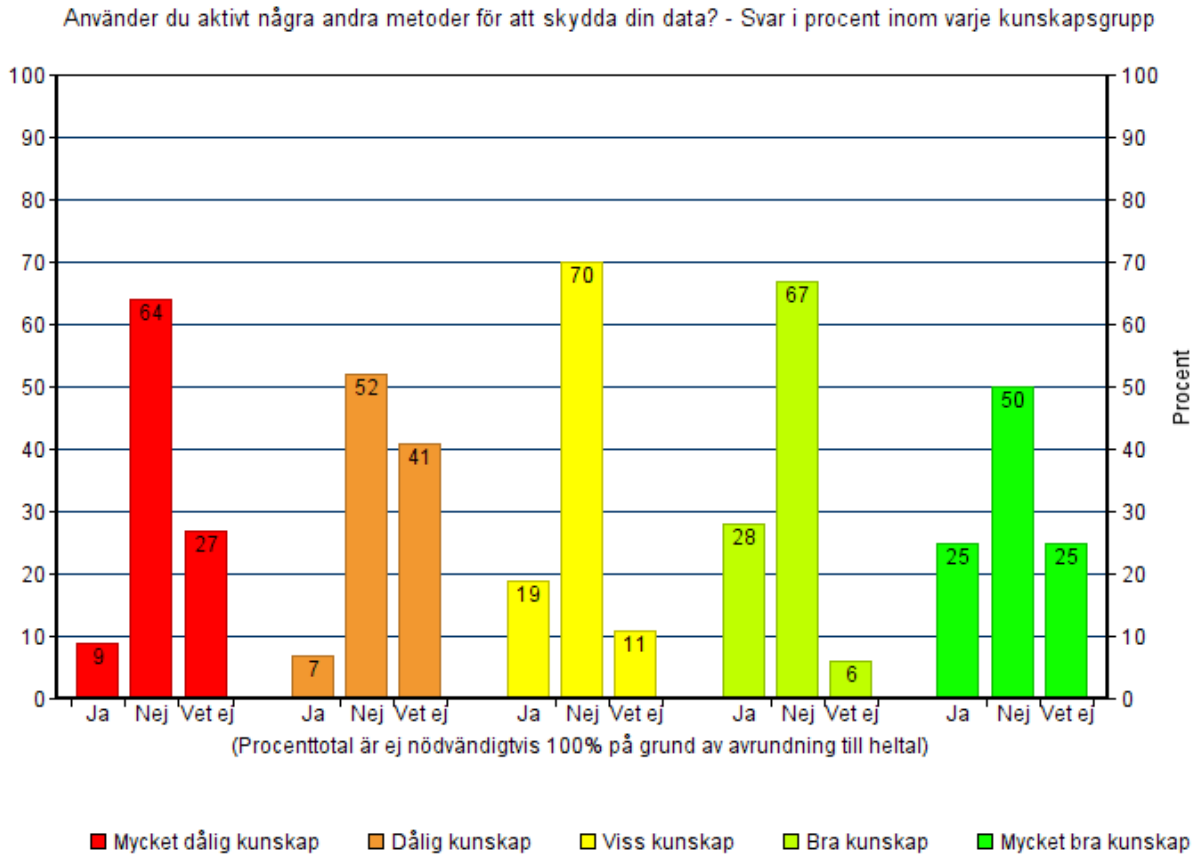
MBK fortsätter att sticka ut och följer inte alls kurvorna för de andra gruppernas svar. 25% svarar att de rensar cookies och 75% svarade att de inte rensar cookies. Ingen svarade att de inte visste.

Överlag ser ökad kunskap up att öka benägenheten att rutinmässigt rensa sina cookies, men även blad användare med bra kunskap är det en majoritet som inte gör det.



Figur 19. Svar i procent om kunskapsgrupperna använder opt-out-funktionen på EIDAA-hemsidan

Opt-out-funktionen är för att slippa bli spårad i syfte att kunna ge riktad reklam av de anslutna företagen. Ingen användare svarade att de använde funktionen och den egentligen skillnaden är att ju mer kunskap användarna har, desto mer vet de om att de inte använder funktionen.

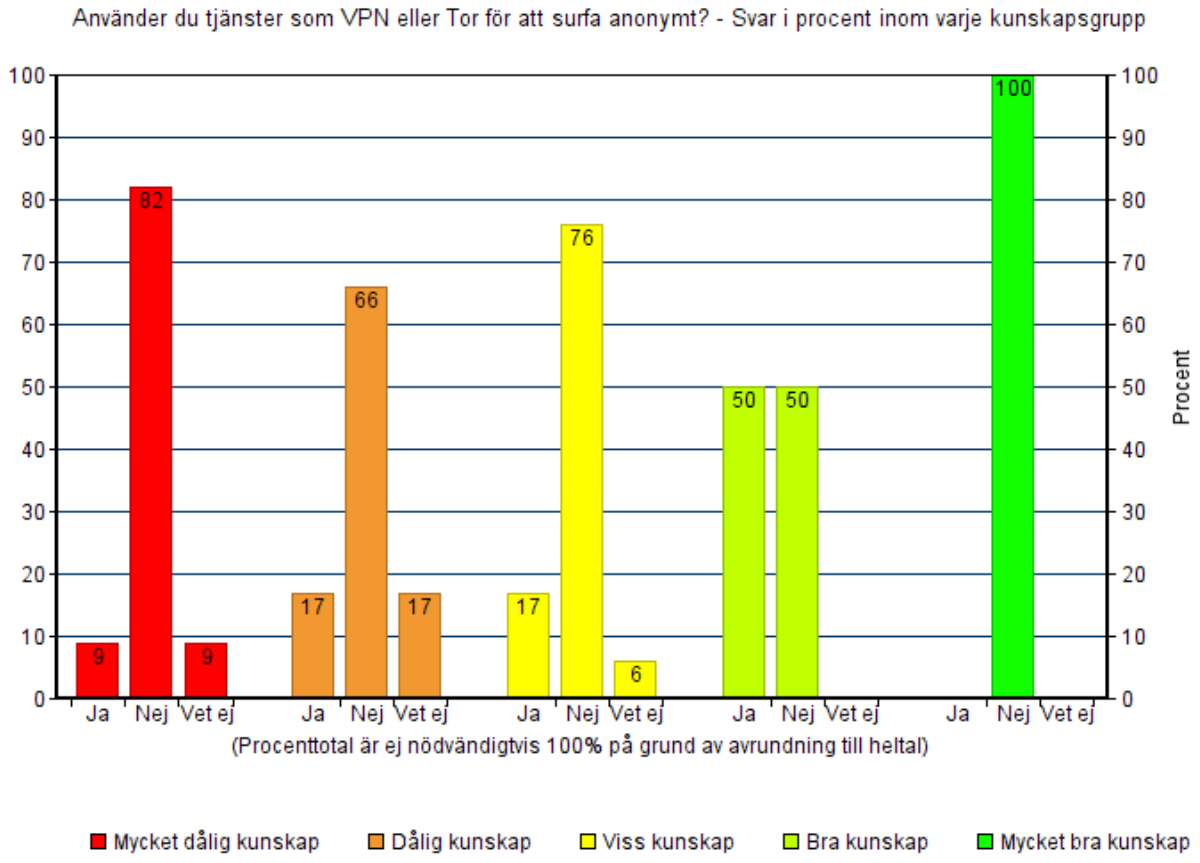


Figur 20. Svar i procent på om kunskapsgrupperna använder några andra metoder för att skydda sin data

MDK och DK går tillbaka till att ha ombytta roller jämfört med våra förväntningar. MDK ser ut att ha större kunskap kring om de använder andra metoder jämfört med DK (27% respektive 41% för ”Vet ej”). Fler användare i MDK har svarat ”Ja” jämfört med DK (9% respektive 7%). MDK har däremot också betydligt fler ”Nej”-svar än DK (64% respektive 52%).

VK och BK håller sig inom förväntningarna för ”Ja”-svaren där båda ökar relativt MDK och VK (19% för VK och 28% för BK). För ”Vet ej”-svaren minskar båda relativt MDK och DK (11% för VK och 6% för BK). Däremot har även ”Nej”-svaren ökat rejält i takt med att ”Vet ej”-svaren har minskat, där VK står högst i diagrammet med 70% tätt följt av BK med 67%.

MBK fortsätter att skilja sig från mängden med 25% ”Ja”, 50% ”Nej” och 25% ”Vet ej”.



Figur 21. Svar i procent för om kunskapsgrupperna använder VPN eller Tor för att surfa anonymt

MDK har lägst andel användare som använder VPN eller Tor med 9%, medan både DK och VK nästan fördubblar den siffran med sina 17%. Intressant är dock att MDK och DK har omvända roller när det kommer till användarna som svarade ”Vet ej”, där enbart 9% av de med MDK svarade att de inte visste, samtidigt som hela 17% av de med DK svarade att de inte visste.

Gruppen BK urskiljer sig rejält med hela 50% som svarat ”Ja” och enbart 50% som svarat ”Nej” (jämfört med 82% ”Nej”-svar för MDK, 66% ”Nej”-svar för DK, 76% ”Nej”-svar för VK och 100% ”Nej”-svar för MBK).

Även här ser det ut som att högre kunskap korrelerar med en ökad användning av tjänster som VPN och Tor, med undantag för gruppen med MBK.

5. Analys och diskussion

I analys- och diskussionsavsnittet tar vi upp områden som är relevanta för att kunna diskutera kunskapen kring Facebooks hantering av persondata. De olika kunskapsgrupperna kommer att diskuteras och jämföras. Även procentantal i cirkeldiagrammen kommer att jämföras för att ta reda på om användarna generellt har mer eller mindre kunskap när det gäller vissa frågor. Vi tar också upp och diskuterar felkällor.

5.1 Kunskap

Huvudresultaten visar att mycket få av användarna (4 stycken) har Mycket Bra Kunskap (MBK) om Facebooks hantering av personlig data. 11 stycken har Mycket Dålig Kunskap (MDK) och majoriteten (63 stycken) har Viss Kunskap (VK). Resultaten är spännande när man jämför cirkeldiagrammen med de individuella svaren. Vid första anblick på cirkeldiagrammen ser det ut som användarna har ganska bra kunskap överlag. 40 stycken av användarna hade MDK eller DK och 85 stycken hade VK, BK eller MBK.

Om vi sedan ser till hur kunskapsgrupperna har svarat är det intressant att jämföra grupperna MDK och MBK. På de flesta frågor som kunde besvaras med ”Ja”, ”Nej” eller ”Vet ej” har de med MBK väldigt sällan svarat ”Vet ej” utan bara ”Ja” eller ”Nej”. Det är mer spridningen på svaren hos de med MDK. Indikatorfrågorna gjorde det enklare för oss att mäta kunskapen hos den valda rampopulationen och detta visar på att de med MBK naturligtvis i större utsträckning var säkra på sina svar och även svarade korrekt. Det fanns ingen tvekan helt enkelt och det påvisar bra kunskap.

Vi kan konstatera att majoriteten av användarna hade VK kring ämnet, vilket inte var helt oväntat då det representerar medelvärdet. En förklaring till att väldigt få användare hade MBK är att användaren måste vara mycket insatt i Facebooks hantering av data samt allmänna säkerhetsrisker på nätet. Vi förväntade oss att de flesta av respondenterna skulle ha VK kring ämnet och så blev som sagt även fallet. Vi kan utifrån kunskapsfördelningen och användningsfrågorna (Del 2 i bilaga 1, frågor om säkerhet) anta att den stora majoriteten inte förstår innebörden av att låta sin samlas in av företag och att en mindre del av användarna inte verkar förstå att de betalar med sin data överhuvudtaget.

5.1.1 Indikatorfrågor

Indikatorfrågorna visar att en av de frågorna som användarna hade mycket dålig kunskap om var frågan: ”Facebook använder rekommendationssystem istället för cookies för att samla data”. 62,4% svarade ”Vet ej” vilket visar att majoriteten inte har någon kunskap om vad rekommendationssystem är eller hur dem används. Även frågan: ”Facebook plug-ins är vanligt förekommande” har 66,4% svarat ”Vet ej”. Även här är den indikator på att användarna är oinformerade om att detta verktyg är det som Facebook använder för att samla in information om sina användare.

59,2% av användarna var felinformerade när det gäller frågan ”Facebook har inte tillgång till privata meddelanden och stängda grupper”. Användarna trodde att det som var benämnt som

privat på Facebook (Meddelanden kallas för PM = Privat Meddelande på Facebook) faktiskt var just det – privat, men så är inte fallet. Inget på Facebook är privat för Facebook eftersom de äger all information och sparar allt i sina servrar.

Användarna hade dock bra kunskap om sex stycken av indikatorfrågorna.

- Annonserna på Facebook är slumpmässigt utvalda: 97,6% svarade rätt
- Om du fyller i uppgifter om dig själv sparas dessa av Facebook: 95,2% svarade rätt
- Genom att använda Facebook kan du bli utsatt för identitetsstöld/utpressning/bedrägeri: 88,9% svarade rätt
- När du tar bort en bild på Facebook är den helt borttagen: 86,5% svarade rätt
- Facebook samlar data om din aktivitet på andra sajter än facebook.com: 80,8% svarade rätt
- Användarna bestämmer själva över vad som visas i sina flöden på Facebook: 72% svarade rätt

Detta visar att de flesta användare har kunskap om sådant som kan anses självklart att veta. Vi kan anta att i princip ingen användare är totalt oinformerad när det gäller hantering av persondata och alla besitter någon form av kunskap, om än lite. Indikatorfrågorna ovan är sådana som många användare säkert har funderat över själva eller läst artiklar om då hantering av persondata är ett mycket omdiskuterat ämne.

5.1.2 Personlig integritet

Resultaten är mycket intressanta om man ser till hur benägna användarna är att skydda sin personliga data. I enkätundersökningen fanns inga direkta frågor som relaterade till personlig integritet men vissa svar kan tolkas i det sammanhanget.

Personlig integritet kan här anses vara möjligheten att kontrollera hur mycket tillgång andra har till din personliga information (Bylund, 2013). Påstående 4 på enkätundersökningen som säger att ”När du tar bort en bild på Facebook är den helt borttagen” har 44.4% svarat ”Nej, ganska säker” och 42,1% svarat ”Nej, helt säker”. Detta visar att 86,5% vet om att en bild aldrig försvinner från Facebook, oavsett om den raderas. Påstående 9 blir intressant om man ställer den emot påstående 4 och jämför resultatet, den säger att ”Facebook har inte tillgång till privata meddelanden och stängda grupper”. Här är svaren lite mer blandade, även om en majoritet (37,3%) har svarat att de är helt säkra på att Facebook inte har tillgång till dessa. Nästa grupp är de användare som är ganska säkra på att Facebook inte har tillgång till privata meddelanden och stängda grupper, nämligen 22,2%. Endast 26,2% av alla respondenter svarade att de var helt säkra eller ganska säkra på att Facebook har tillgång till dessa.

Svaren i påståendena 4 och 9 är motsägelsefulla i den meningen att varför tror man att Facebook sparar bilder men inte har tillgång till/sparar privata meddelanden eller stängda grupper? Användarna tycks lita på att Facebook bevarar deras personliga integritet och håller per-

sonlig data privat. Var går gränsen för vad som anses kränka den personliga integriteten? Användarna som svarat anser till stor del att bilder är okej att lagra, men inte meddelanden eller information som delas i stängda grupper. De tror dock fel, Facebook lagrar all information som användaren lägger upp/delar/skriver i sina servrar.
Påstående 8, ”Genom att använda Facebook kan du bli utsatt för identitetsstöld/utpressning/bedrägeri” har en klar majoritet, 88,9% svarat att de är helt säkra eller ganska säkra på att det kan hända. Påstående 15, ”Tänker du på vilken personlig information du fyller i på Facebook? (Namn, stad, arbetsplats, skola osv)” har majoriteten svarat ”Ja”. Detta visar på hög personlig integritet när det gäller vad användarna väljer att lägga upp om sig själva och även på försiktighet då de har kunskap om möjliga risker.

Svaren visar på att användarna håller personlig integritet förhållandevis högt men också på en del motsägelser i hur mycket de väljer att lita på Facebook och Facebooks hantering av datan. Samtidigt som de säger att Facebook håller viss data privat (meddelanden och grupper) så anser de att det finns en risk att bli utsatt för identitetsstöld/utpressning/bedrägeri och att de i så fall minskar den risken genom att tänka på vad de publicerar för information. Bör användarna inte då också fundera på vad de skriver i meddelanden till varandra och i grupper? De flesta användare gör inte det och ett svar på detta är att användarna väljer att lita på Facebook utifrån den aspekten att viss personlig information hålls privat under förutsättningarna att ingen ”oinbjuden” egentligen har tillgång till informationen. Användarna hade förmodligen svarat annorlunda på påstående 9 om de vetat om att Facebook har tillgång till privata meddelanden och stängda grupper också.

Resultaten som relateras till personlig integritet är värdefulla då det alltid är meningsfullt att undersöka hur förhållandet till personlig integritet ändras över tid. Bylund (2013) säger att förr i tiden var den personliga integriteten mindre benägen att kränkas då teknik inte fanns/användes på samma sätt som det görs idag. Personlig integritet blir enklare att kränka ju mer information vi delar med oss av på nätet och det problemet fanns inte förr. Människor sägs värna lika mycket om personlig integritet nu som förr men resultaten i enkätundersökningen säger annat.

5.1.3 Risker

Resultatet visade en oerhört stor kontrast mellan grupperna MDK och MBK i frågan om de reflekterade kring vilken typ av information de publicerar på Facebook. Gruppen MDK hade högst andel användare som svarade att de tänker på vilken typ av information som de själva väljer att publicera på Facebook. Användarna med MBK hade lägst antal användare som reflekterade kring vad de publicerar. Slutsatsen som kan dras från detta är att personerna med MDK inte har någon kunskap om hur Facebook hanterar deras data och hur den säljs vidare. Därför är de mycket mer försiktiga med vad de väljer att publicera då de inte har insikt i Facebooks rutiner och villkor. De användare i gruppen MBK vet i större utsträckning hur Facebook hanterar deras persondata och är inte särskilt oroliga över riskerna. Precis som Cho et al. (2010) tar upp så har många användare en bild av att deras egna persondata inte är värdefull och på så vis inte lika känslig som andras. Många användare anser därför att deras data är tillräckligt skyddad och att de inte behöver vidta några extra åtgärder. Dock är det viktigt att ta hänsyn till att endast 4 användare hade MBK och endast 11 användare hade MDK, vilket är ett extremt litet urval. Då användarna var så få i varje grupp finns risken att resultaten är skeva.

Cirka hälften av användarna hamnade i gruppen VK och endast 18 användare hade BK om Facebooks hantering av persondata. På frågan ”Facebook kan inte samla in data om dig om du inte har ett Facebookkonto” blev det väldigt spridda svar. Endast 7,9% svarade att de var helt säkra på att Facebook kan samla in data om användaren även utan ett Facebookkonto, vilket är korrekt. Högsta andelen användare (42%) svarade ”Vet ej”. Det visar att användarna till viss del är oinformerade om hur Facebook hanterar persondata utanför den egna tjänsten.

Resultatet på frågan ”Facebook samlar data om din aktivitet på andra sajter än facebook.com” hade dock en större andel av användare svarat ”Ja, helt säker” eller ”Ja, ganska säker”. Detta visar på att användarna tror att om de har ett Facebookkonto så kan Facebook samla information om dem även på andra sajter. Det tycks finnas en förståelse för att Facebook är kopplat till andra hemsidor och tjänster och att där finns information att hämta. Emellertid kan vi konstatera att de flesta användare inte vet om att Facebook har möjlighet att samla information om dig även om du inte har ett Facebookkonto, det räcker att använda andra tjänster som är kopplade till Facebook.

Majoriteten av användarna som svarade på frågan ”Facebook social plug-ins är vanligt förekommande” svarade ”Vet ej”. Social plug-ins är ett oerhört populärt verktyg som Facebook använder sig av för att spåra sina användare på internet. Den mest använda plug-in är ”gillaknappen” som är ett utmärkt verktyg för att samla in information om sina användares aktiviteter på nätet. Social plug-ins tvingar användarnas webbläsare att hämta och spara innehåll från de besökta hemsidorna (Acar et al., 2015). Användarna har alltså en viss kunskap om hur Facebook hanterar deras persondata och vilka risker som finns med det, men kunskapen kan anses väldigt begränsad då majoriteten inte förstår hur Facebook spårar sina användare och att inget som finns på Facebook är privat, därför förstår de inte *alla möjliga* risker som existerar heller.

5.1.4 Säkerhet

Resultatet visade att användare med mer kunskap om området gör ett aktivt val att skydda sin personliga data i högre utsträckning jämfört med de användare som har sämre kunskap. Det finns en genomgående trend både när det gäller att använda någon form av AdBlocker eller andra säkerhetstillägg i webbläsaren. Resultatet visar också att användare med MDK in vet om de skyddar sin data eller inte. Användarna med MBK använde någon form av AdBlocker men användandet av säkerhetstillägg sjönk sedan när vi tittade på användarna med BK. Det här kan tolkas som att de med MBK känner till riskerna och förstår vad olika säkerhetstillägg är för något och vad de gör.

Facebooks kan spåra sina användare genom cookies. Gruppen med MDK hade den högsta andelen användare som svarade ”Vet ej” på frågan om de tillåter Facebook att använda sig av cookies, vilket inte är förvånande. Grupperna VK och BK hade lika stor mängd användare som svarade att de tillåter cookies, men VK hade fler användare som svarade ”Vet ej” jämfört med BK. Vi kan alltså konstatera att med ökad kunskapsnivå gör användarna ett aktivt val att tillåta cookies. När det gäller frågan om användarna rutinmässigt rensade cookies fanns det inte en enda användare i gruppen med MDK som svarade ”Ja”. Mängden användare som svarat ”Ja” ökar i takt med kunskapsnivån och är som högst i gruppen med BK, där 39% rutinmässigt rensade cookies. I gruppen med MBK sjunker andelen ”Ja” till 25%, men det är mycket få användare i den gruppen så det är svårt att säga att det inte är tillräckligt bra. Ingen i grupperna med BK eller MBK svarade ”Vet ej”. Användarna med mer kunskap har således

aktivt valt att tillåta/inte tillåta cookies samt att rensa/inte rensa cookies rutinmässigt. Precis som Xu et al. (2014) tar upp så väljer en del användare att inte skydda sin data fullt ut då det kan förstöra upplevelsen på internet till viss del. Fördelarna med att dela med sig av persondata är att användarna får personliga rekommendationer på internet och dessa personliga rekommendationer hade inte existerat om inte Facebook kunde spåra användarnas historik samt demografi på internet.

5.2 Felkällor

Som vi nämnde i metodavsnittet är några av de vanligaste felkällorna att frågorna inte är tydligt formulerade, användarna representerar inte hela rampopulationen som skall undersökas, urvalsgruppen var inte tillräckligt stor, det kan finnas brister i hur insamlingen av statistiken har gått tillväga samt hur den insamlade datan har bearbetats.

Samtliga av dessa felkällor är befintliga i vår undersökning och har troligtvis påverkar resultatet. Ett tydligt exempel är en användare som lämnade feedback på att framförallt en av våra frågor kunde misstolkas. Användaren menade att indikatorfråga 5, ”Facebook kan inte samla in data om dig om du inte har ett Facebookkonto” kan besvaras med både ja och nej och fortfarande vara sann; ”Ja, Facebook kan inte samla in data om dig om du inte har ett Facebookkonto” och ”Nej, Facebook kan inte samla in data om dig om du inte har ett Facebookkonto” är båda korrekta svar. Detta kan ha påverkat resultatet då folk med kunskap kanske har feltolkat frågan på grund av otydlighet och svarat vad vi tolkar som ett felaktigt svar. Otydligheten kan potentiellt finns hos andra frågor också, eftersom de besvarades med ”Ja” eller ”Nej” istället för ”Sant” eller ”Falskt”. Otydligheten hade således eliminerats om ”Sant” eller ”Falskt” hade använts som svar istället.

Då vi främst distribuerade enkäten på Facebook till vänner, studiekamrater och studentlivskamrater finns det en stor risk att användarna inte representerar rampopulationen 20-29-åringar i Sverige. De flesta av våra vänner är antagligen i åldern 20-25 år snarare än 20-29 år och eftersom vi – författarna – själva studerar Systemvetenskap finns det många i vår närhet som också gör det. Det här riskerar att ge skeva resultat för rampopulationen. Urvalsgruppens storlek påverkar också resultatet. För väldigt säkra resultat i en enkätundersökning bör man ha omkring 1000 respondenter, vilket betyder att 125 svar inte ger jättesäkra resultat. När användarna sedan delas in i olika kunskapsgrupper blir urvalet ännu mindre och svaren därmed mer osäkra. Det märks tydligt i stapeldiagrammen där gruppen med MBK enbart har 4 användare.

Tittar vi på hur vi har samlat in datan använde vi oss av en enkät med indikatorfrågor som tillsammans är tänkta att indikera hur bra kunskap en användare har. Men är frågorna vi valt verkligen indikativa för kunskap? Vi tror det, men det behöver inte nödvändigtvis vara så. Om fallet inte är så är undersökningen baserad på icke-indikativ data. Vi behöver också fråga om vi har bearbetat datan korrekt. Användarna är indelade med poänggivning där Korrekt svar, helt säker gav +1 poäng, Korrekt svar, ganska säker gav +0,5 poäng, Vet ej gav 0 poäng, Inkorrekt svar, ganska säker gav -0,5 poäng och Inkorrekt svar, helt säker gav -1 poäng. Finns det några problem med poänggivningen? Möjligtvis, att ge minuspoäng för felaktiga svar behöver inte vara rimligt för att mäta användarnas kunskap.

För att sammanfatta finns det en mängd felkällor som kan ha förvrängt resultaten från sanningen. Tydligare ställda frågor, större urval, mer slumpmässigt utvalt urval och noggrannare

insamling och bearbetning av datan hade kunnat generera ännu bättre svar, och detta är något att sträva efter vid en ny undersökning.

6. Slutsats

I det här avslutande avsnittet kommer vi att besvara forskningsfrågorna utifrån enkätundersökningen vi genomfört med stöd i teorin.

6.1 Har svenskar i åldrarna 20-29 år kunskap om hur Facebook samlar in data om dem och känner de till riskerna med hur datan hanteras, samt gör de aktivt något för att skydda sin data?

Att döma av enkätundersökningen är svaret att svenskar i åldrarna 20-29 år har *viss* kunskap om att Facebook samlar in data om dem samt även hur Facebook gör det. De känner också till viss del om riskerna med hur datan hanteras och en del av användarna använder olika säkerhetstillägg för att skydda sin data. Självklart skiftar nivån på kunskapen beroende på användaren själv och egentligen var respondenterna för få för att få fram ett fullt tillförlitligt resultat, i synnerhet för gruppen med MBK som enbart hade 4 respondenter. De andra grupperna kan anses lite mer tillförlitliga då respondenterna där var fler. Vi kan även konstatera att de med MDK hade högst andel användare som svarade ”Vet ej” på indikatorfrågorna, vilket innebär att användarna i den gruppen troligen är utsatta för risker i högre utsträckning jämfört med de andra kunskapsgrupperna.

6.2 Finns det någon skillnad i hur en användare med mer kunskap om Facebooks datainsamling skyddar sin data/inte skyddar sin data jämfört med en användare med mindre kunskap?

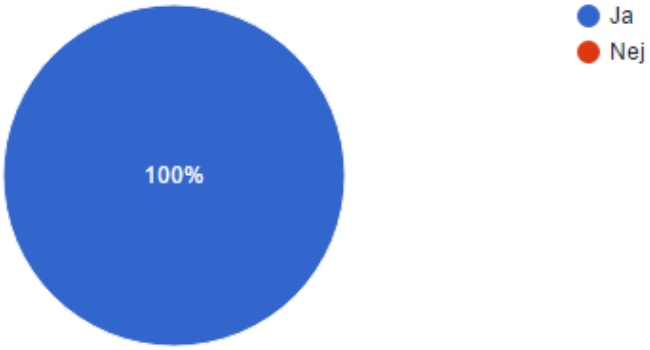
Ja, det finns en skillnad i hur en användare med mer kunskap om Facebooks datainsamling väljer att skydda/inte skydda sig jämfört med användarna med mindre kunskap. Vi kan se att användandet av säkerhetstillägg och andra typer av säkerhetsmetoder ökar i takt med att kunskapen ökar. Vi kan konstatera att kunskapsgrupperna med mindre kunskap inte alls vet om de använder några säkerhetstillägg eller metoder. Vår slutsats är att Facebooks datainsamling påverkar användarens sätt att använda Facebook, men att det krävs mycket kunskap om ämnet för att användarna aktivt ska välja att skydda sin data med hjälp av olika säkerhetstillägg.

Bilaga 1: Svar från enkätundersökning

Del 1

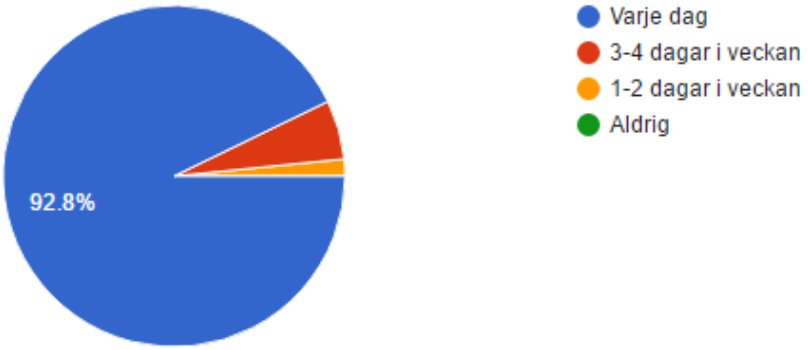
1. Är du mellan 20-29 år gammal?

125 responses



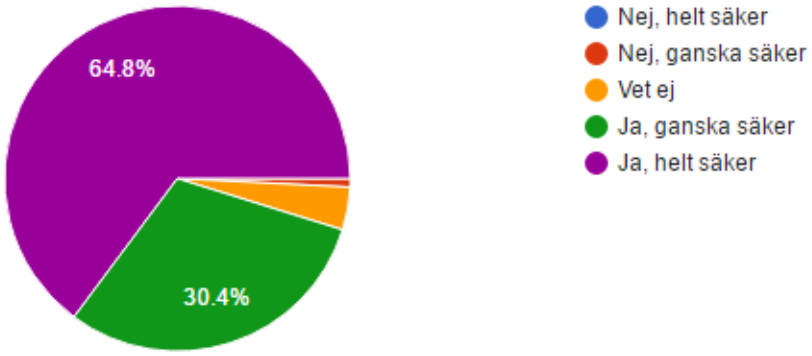
2. Hur ofta använder du Facebook?

125 responses



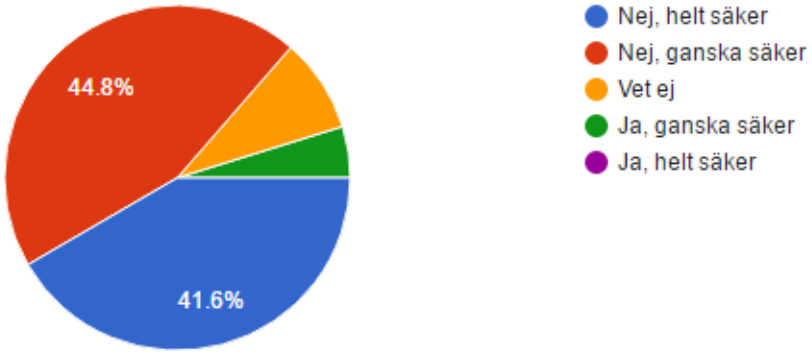
3. Om du fyller i uppgifter om dig själv sparas dessa av Facebook

125 responses



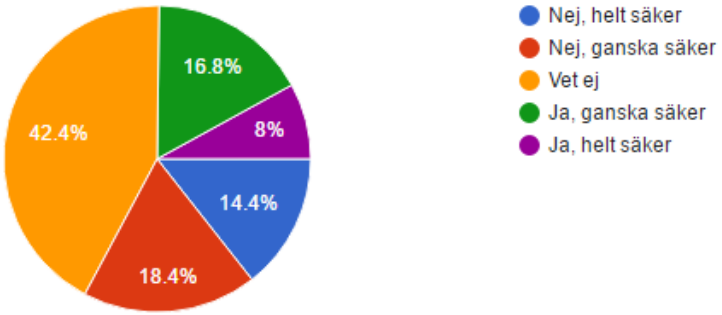
4. När du tar bort en bild på Facebook är den helt borttagen

125 responses



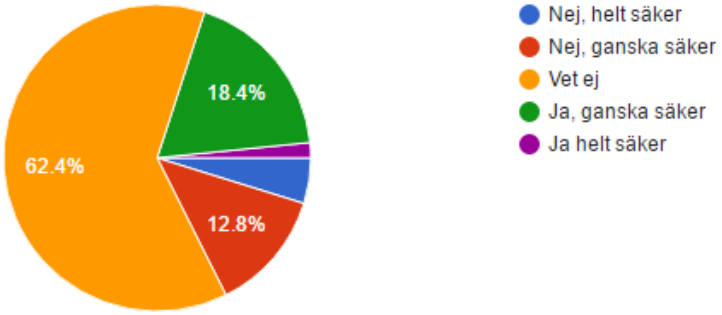
5. Facebook kan inte samla in data om dig om du inte har ett Facebookkonto

125 responses



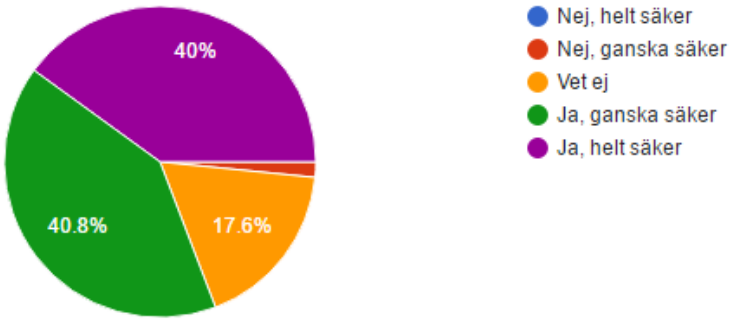
6. Facebook använder rekommendationssystem istället för cookies för att samla data

125 responses



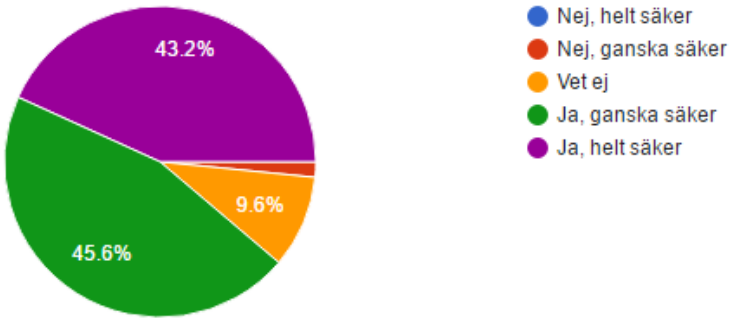
7. Facebook samlar data om din aktivitet på andra sajter än facebook.com

125 responses



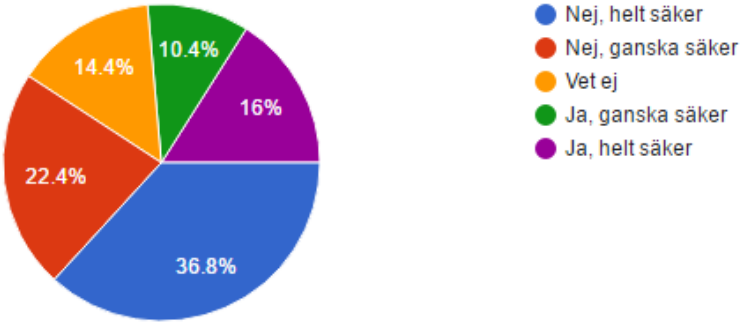
8. Genom att använda Facebook kan du bli utsatt för identitetsstöld/utpressning/bedrägeri

125 responses



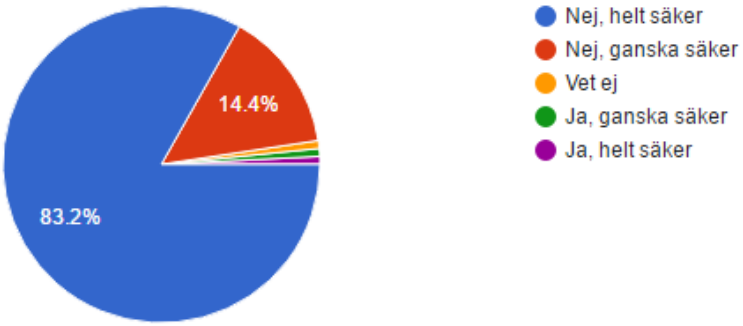
9. Facebook har inte tillgång till privata meddelanden och stängda grupper

125 responses



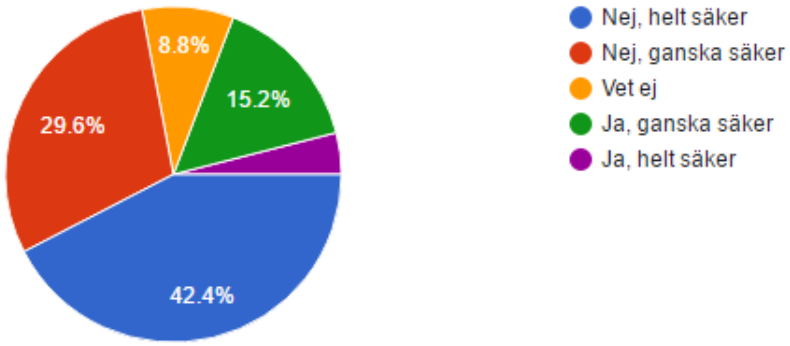
10. Annonserna på Facebook är slumpmässigt utvalda

125 responses



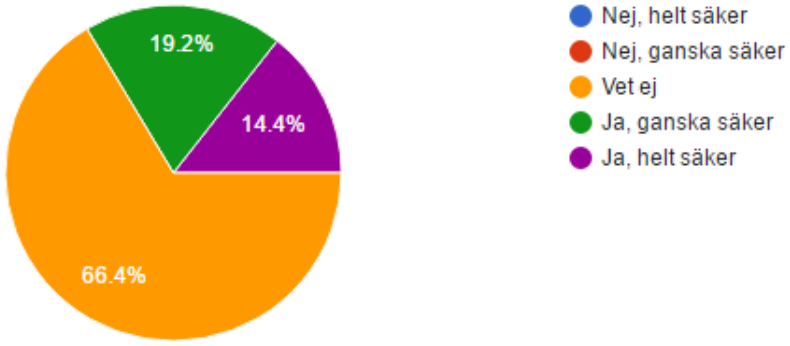
11. Användarna bestämmer själva över vad som visas i sina flöden på Facebook

125 responses



12. Facebook social plugins är vanligt förekommande

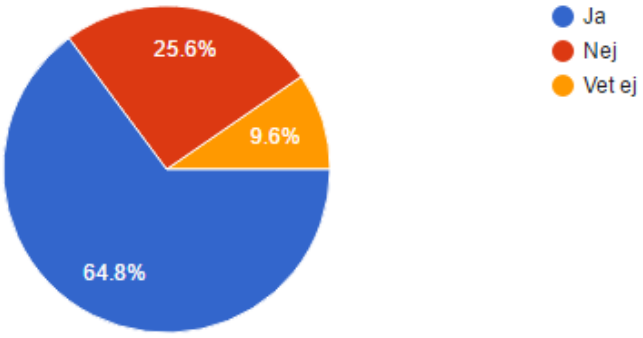
125 responses



Del 2

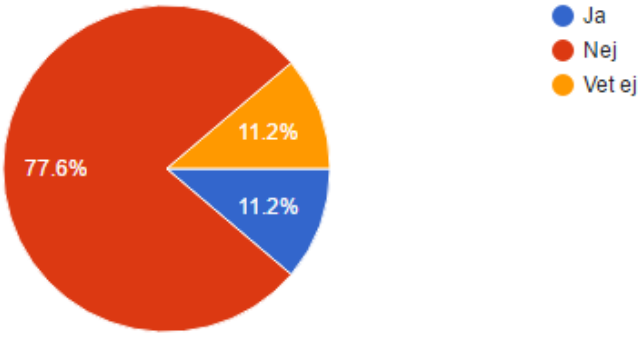
13. Använder du någon form av adblock?

125 responses



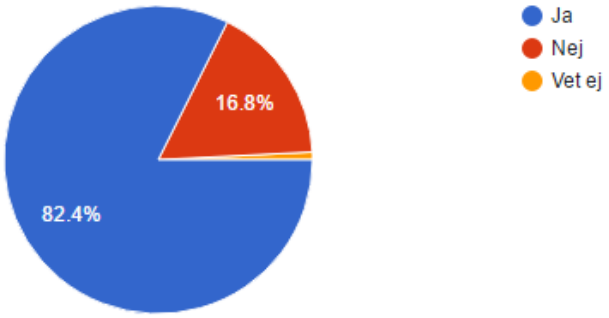
14. Använder du några andra säkerhetstillägg i din webbläsare? (T.ex. Facebook Disconnect)

125 responses



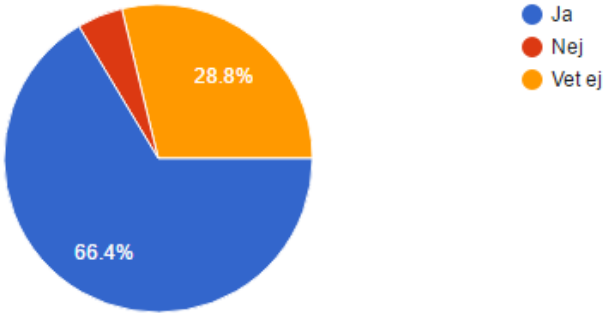
15. Tänker du på vilken personlig information du fyller i på Facebook? (Namn, stad, arbetsplats, skola osv.)

125 responses



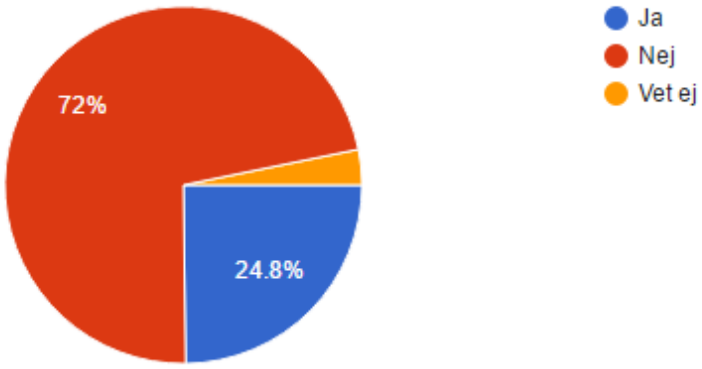
16. Tillåter du Facebook att använda cookies?

125 responses



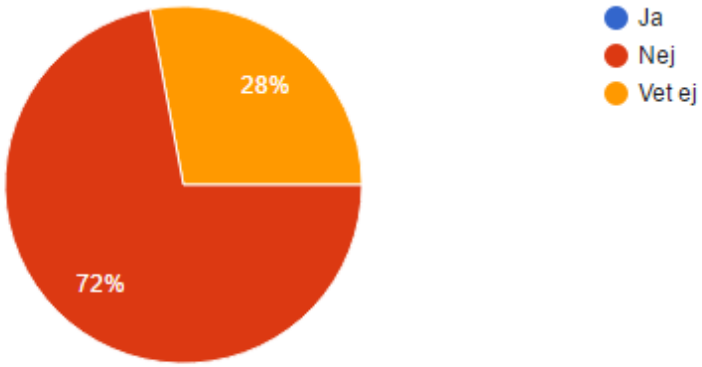
17. Rensar du cookies rutinmässigt?

125 responses



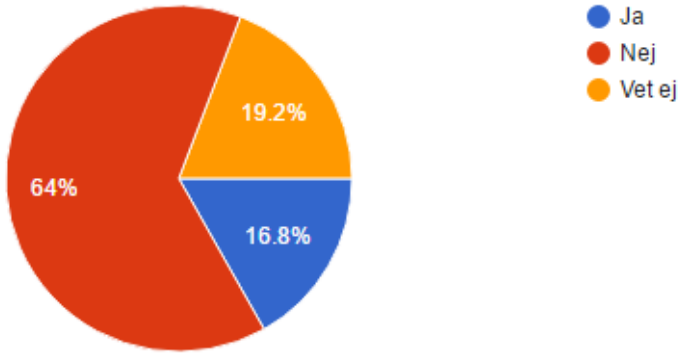
18. Använder du opt-out-funktionen på European Interactive Digital Advertising Alliance-hemsidan?

125 responses



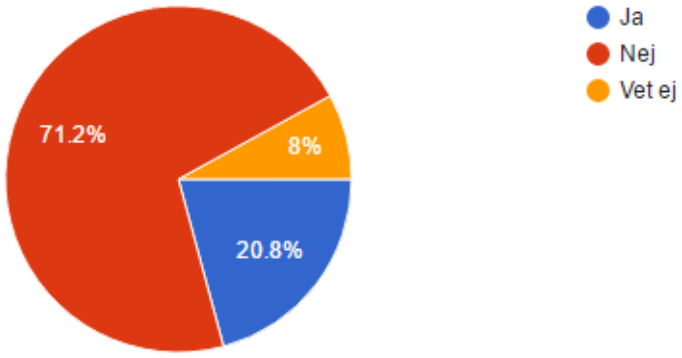
19. Använder du aktivt några andra metoder för att skydda din data?

125 responses



20. Använder du tjänster som VPN eller Tor för att surfa anonymt?

125 responses



Referenser

Acar, G., et al. (2015) Facebook Tracking Through Social Plug-ins: Technical report prepared for the Belgian Privacy Commission

Alexa Top Sites in Sweden. Hämtas 28 mars 2017 från: http://www.alexa.com/topsites/countries/SE

Alverén, F. (2012) Såld på nätet: Falun: ScandBook AB

Boyd, D., Crawford, K. (2011) Six Provocations for Big Data. Social Science Research Network: A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society

Bylund, M. (2013) Personlig integritet på nätet. ScandBooks AB, 2013. 1:a upplagan

Code Facebook. Hämtas 12 april 2017 från: https://code.facebook.com/posts/861999383875667/recommending-items-to-more-than-a-billion-people

Cho, H., Lee, J., Chung. S, (2010) Optimistic bias about online privacy risks: Testing the moderating effects of perceived controllability and prior experience. Computers in Human Behavior, ISSN 0747-5632, 2010, Volume 26, Issue 5

Erkin, Z., et al. (2010) Privacy Enhanced Recommender System. In: 31st Symposium on Information Theory in the Benelux, WIC, Rotterdam, the Netherlands (pp. 34-42)

Facebook Data Policy. Hämtad 28 mars 2017 från: https://www.facebook.com/policy.php

Featherman, M. S., Miayzaki, A. D., Sprott D. E. (2010): Reducing online privacy risk to facilitate e-service adoption: the influence of perceived ease of use and corporate credibility. The Journal of Services Marketing, ISSN 0887-6045, 03/2010, Volume 24, Issue 3.
Hunt, D. (2003) The concept of knowledge and how to measure it: Journal of Intellectual Capital

IPUMS International. What are Microdata? Hämtad 3 maj 2017 från : https://international.ipums.org/international-action/faq#ques7

Jacobsen, D. I. 2002. Vad, hur Varför? Om metodval i företagsekonomi och andra samhällsvetenskapliga ämnen. Studentlitteratur, Lund. ISBN: 9789144040967, 503 s

Jeckmans, A., et al. (2013) Privacy in Recommender System: Springer London.

Xu, L. et al. (2014) Information Security in Big Data: Privacy and Data Mining: IEEE

Ricci, F., Rokach, L., Shapira, B.(2011) Introduction to Recommender Systems Handbook, Recommender Systems Handbook: Springer.

SFS (1998:204) Personuppgiftslag (1998:204) Stockholm: Justitiedepartementet L6 (Elektronisk) Tillgänglig: <http://www.riksdagen.se/sv/dokument-lagar/dokument/svensk-forfattningssamling/personuppgiftslag-1998204_sfs-1998-204/> (2017-03-28)

Small Data. Hämtad 3 maj 2017 från: http://jwork.org/main/node/18

STATENS MEDICINSK-ETISKA RÅD. Integritet. Hämtad 19 april 2017 från: http://www.smer.se/etik/integritet/

Svenskarna och internet 2016. Undersökning om svenskarnas internetvanor. Hämtad 20 april 2017 från: https://www.iis.se/docs/Svenskarna_och_internet_2016.pdf

Svensson, P. (2015) Kvalitativ och kvantitativ undersökningsmetodik. Hämtad 19 april 2017 från: https://student.portal.chalmers.se/sv/chalmersstudier/programinformation/maskinteknik/kandidatarbete/Documents/20150225%20Vetenskapsmetodik%20f%C3%B6rel%202%20PS.pdf

Terveen, L., et al. (2001) Beyond Recommender Systems: Helping People Help Each Other: Addison-Wesley.

Toubiana, V., et al. (2010): Adnostic: Privacy Preserving Targeted Advertising: New York University

Toch, E., Wang., Y., Cranor, L. (2012): Personalization and privacy: a survey of privacy risks and remedies in personalization-based systems. User Modeling and User-Adapted Interaction, ISSN 0924-1868, 03/2012, Volume 22, Issue 1-2.

Paine, C., Reips, U., Stieger, S., Joinson, A., Buchanan, T. (2007): Internet users' perceptions of 'privacy concerns' and 'privacy actions'. International Journal of Human - Computer Studies, ISSN 1071-5819, 2007, Volume 65, Issue 6

Wang, N., et al. (2011): Third-Party Apps on Facebook: Privacy and the Illusion of Control: The Pennsylvania State University

Warren, S., Brandeis, L.D. (1890): The right to privacy. Harvard Law Review

Wesiberg, H. (2005) The Total Survey Error Approach: The University of Chicago Press