PRIVACY PARADOX OP ONLINE PLATFORMEN

een kwantitatief onderzoek

Jessy Bosman
11056045
24-6-2018

Begeleider: dr. D. Heinhuis
2e Examinator: ir. A.M. Stolwijk
Bachelorscriptie Informatiekunde
Faculteit der Natuurwetenschappen, Wiskunde en Informatica
Universiteit van Amsterdam

Abstract

Dagelijks wordt er veel informatie op het internet gedeeld. Een deel hiervan is persoonlijke informatie. Deze informatie kan gebruikt worden om mensen te identificeren en te categoriseren, waardoor persoonlijke advertenties kunnen worden gemaakt. Gebruikers hechten een grote waarde aan privacy, maar dit blijkt niet uit hun gedrag online. Deze tegenstelling wordt het privacy paradox genoemd. In dit onderzoek wordt er gekeken welke factoren dit privacy paradox veroorzaken onder studenten. De literatuur schrijft de oorzaken toe aan vertrouwen, onwetendheid en uitruilbaarheid van data tegen voordelen. Echter is er verdeling over de mening of het privacy paradox wel of niet bestaat. Dit onderzoek vindt ondersteuning voor het bestaan van een privacy paradox. De volgende factoren zijn onderzocht: invloed van vertrouwen op het delen van data, de invloed van service of korting op het delen van data, een incompleetheid aan kennis over data en invloed van verschillende soorten data op het delen van data. Deze factoren worden vergeleken met de oorspronkelijke intentie om privacy te bewaren. Voor alle factoren wordt ondersteuning gevonden. Er wordt geconcludeerd dat alhoewel privacy belangrijk wordt gevonden, er niet wordt gekozen voor privacy wanneer er genoeg voordelen tegenover het inleveren van privacy staan. Een mogelijke verklaring hiervoor is dat privacy niet wordt gezien als een algemeen recht, maar een verhandelbaar goed, dat tegen de juiste prijs te koop is.

Inhoudsopgave

1. Achtergrond

Het internet is niet meer weg te denken uit onze dagelijkse bezigheden. Sociale netwerken kunnen worden onderhouden door middel van sociale media, zoals Facebook. Zoekmachines als Google maken het doorzoeken van het grote aantal bestanden op het internet mogelijk. Webshops als Amazon en Bol maken het mogelijk om verschillende producten eenvoudig online te bestellen en te laten thuisbezorgen. Wat veel gebruikers echter niet beseffen is dat alles wat ze op het internet doen privacygevoelige data genereert. Dit kan gekoppeld worden aan de data die al over de persoon bekend is, waardoor er een goed beeld kan worden verkregen van wie de gebruiker is en waar deze zich dagelijks mee bezighoudt. Maar de vraag is welke data er wordt opgeslagen en waar deze voor wordt gebruikt. Hieronder volgen een aantal voorbeelden (figuur 1):

Figuur 1: Overzicht bedrijven Personal Data
Sources: Facebook, 2018a; 2018b; Hunter et al., 2009; Privacy.google.com, 2018; Policies.google.com, 2018; Snap.com, 2018; Xu et al., 2016; Young, 2014.

Wat beseft moet worden is dat al onze persoonlijke gegevens opgeslagen en gekoppeld kunnen worden aan elkaar. Hieraan kunnen alle handelingen die worden uitgevoerd, bijna elke muisklik, worden toegevoegd. Deze informatie kan gebruikt worden om een exact profiel van een gebruiker te maken. Dit kan worden gebruikt voor advertenties en analyse. In het geval dat deze informatie uitlekt door bijvoorbeeld een hack kan dit nog meer gevolgen hebben. Een van de uiterste 'gevaren' is het principe van social sorting. Social sorting is het indelen en categoriseren van mensen met behulp van de data die over ze beschikbaar is (Stoddart, 2014). Een voorbeeld van een toepassing van social sorting is de '*no flight list*', een lijst die vliegen naar de Verenigde Staten verbiedt. Deze lijst wordt gebaseerd op de beschikbare persoonlijke data van de desbetreffende persoon, zoals bijvoorbeeld bezochte landen (Martijn & Tokmetzis, 2018).

2. Probleemstelling

Het probleem dat hier ontstaat is dat er veel meer verschillende soorten data bestaat en wordt bijgehouden dan er wordt beseft. Onder het mom van "Ik heb niks te verbergen" (Martijn & Tokmetzis, 2018, p.17) wordt er onzorgvuldig omgegaan met persoonlijke data en het delen hiervan

(Debatin et al., 2009), zonder de gevolgen hiervan in te zien. Privacy op zich is belangrijk voor mensen, maar in afwegingen, zoals privacy of veiligheid, wordt er niet voor privacy gekozen (Martijn & Tokmetzis, 2018). Zo beschrijft Acquisti (2010) dat alhoewel mensen privacy waarderen, ze niet willen betalen voor privacy, en bereid zijn persoonlijke data te leveren in ruil voor bijvoorbeeld kortingen of andere voordelen. Xie et al. (2006) beschrijven dit gedrag als het '*risk-benefit*' perspectief. Persoonlijke informatie wordt afgestaan als de voordelen opwegen tegen de nadelen. Mensen zijn bereid persoonlijke informatie te geven voor economische of sociale voordelen. In deze overweging om informatie te geven speelt vertrouwen ook een grote rol (Norberg et al., 2007). Persoonlijke data wordt te makkelijk weggegeven door gebruikers van het internet en dit kan tot intimidatie, stalking of misbruik van (gestolen) data leiden (Debatin et al., 2009). Enerzijds wordt er bezorgdheid geuit over privacy, maar anderzijds blijkt dit niet uit het online gedrag van mensen. Dit fenomeen staat bekend als het **privacy paradox** (Barnes, 2006).

Het probleem dat ontstaat bij dit paradox is dat gebruikers denken privacybewust te zijn, maar dit in de praktijk niet tot uiting komt. Er ontstaat een waardenpluralisme, er moeten morele afwegingen worden gemaakt waarbij conflicterende waarden tegen elkaar moeten worden afgewogen. Voorbeelden hiervan zijn privacy tegenover veiligheid, of privacy tegenover korting. Het is het één óf het ander, er moet een keuze worden gemaakt tussen wél of niet privacy voorop stellen en het tweede is vaak het geval. Een mogelijke oorzaak hiervan is dat veel gebruikers zich niet bewust zijn welke data precies van ze wordt verzameld en op welke manieren. Hierdoor is men zich niet goed bewust welke gevolgen het weggeven van privacygevoelige gegevens heeft. Een mogelijk gevolg is dat het manipulatie van mensen kan veroorzaken, bijvoorbeeld door gebruikers naar bepaalde keuzes te sturen. Dit wordt gebaseerd op de data die over hen beschikbaar is, zoals gebeurt bij gericht adverteren.

Om het bestaan van het privacy paradox te ondersteunen dan wel af te vallen wordt er in dit onderzoek gekeken naar de intentie en het gedrag omtrent privacybewustzijn op online platformen, met als doelgroep studenten. Er wordt gehypothetiseerd dat hoewel privacy belangrijk wordt gevonden, het toch het onderspit delft als privacy wordt afgewogen tegen andere waarden, zoals service.

3. Hoofdvraag

Hieruit wordt de volgende hoofdvraag geformuleerd:
"In hoeverre is er sprake van een privacy paradox op online platformen?"

Om dit te kunnen beantwoorden moet er allereerst gekeken worden naar de oorzaken die dit paradox mogelijk veroorzaken. Dit leidt tot de vraag:
Qtk: "Welke theoretische verklaringen ondersteunen of ontkrachten het privacy paradox?"

In de praktische en academische relevantie wordt verkennend onderzocht hoe privacy zich verhoudt in een online omgeving waar privacy slechts gering bestaat. In het theoretisch kader wordt nader

onderzocht welke theorieën er zijn die het bestaan van het privacy paradox kunnen verklaren of ontkrachten, om zo een beeld te krijgen over hoe bepaalde afwegingen met betrekking tot privacy worden gemaakt.

4. Praktische en academische relevantie

Koppelen van data

Om te achterhalen hoe privacy bestaat in een online omgeving wordt er gekeken naar welke data er over gebruikers bestaat. Hiervoor maakt Acquisti (2004a, p.180) onderscheid tussen online en offline identities in E-commerce. "The online identity might carry information about an individual's tastes, her evaluation of a certain good, her browsing behavior, her purchase history, etc.: On the other side, the offline identity represents the actual identity of an individual, as revealed by identifiers such as credit card numbers and social security numbers". Deze data is echter soms ook aan elkaar te koppelen, waardoor er heel veel persoonlijke data bekend is over een persoon. Er wordt beschreven dat ook derde partijen hier soms toegang tot kunnen hebben. Het gevaar dat hierbij ontstaat is dat de data gebruikt kan worden om gebruikers te sturen en te beïnvloeden.

Gedrag op sociale netwerken

Deze persoonlijke data kan worden verkregen op bijvoorbeeld sociale platformen. In het artikel van Debatin et al. (2009) wordt beschreven dat personen die gebruik maken van Facebook zich wel bewust zijn van de privacy issues, maar desondanks persoonlijke informatie grootschalig delen met het medium. Gebruikers geloven dat privacyschending anderen wel kan overkomen maar henzelf niet. Het onderzoek schrijft deze denkwijze toe aan "a combination of highgratification, usage patterns, and a psychological mechanism similar to third-person effect". De voordelen van verbonden zijn met een sociaal netwerk wegen op tegen de privacygevoeligheid. Govani en Pashley (2005) bevestigen dat gebruikers gewillig veel informatie geven aan Facebook en suggereren dat dit mogelijk komt door teveel vertrouwen dat de data veilig is. In het onderzoek wordt gekeken naar wat er voor soort informatie wordt gegeven en ook wordt er gekeken naar waarom er wordt deelgenomen aan Facebook. Zo valt op dat meer dan 80% hun E-mail, geboortedag en interesses opgeeft. De voornaamste reden om gebruik te maken van het platform is om zo in contact te blijven met vrienden.

Online Social Footprint

Het gevolg van het verkrijgen van deze data is dat hieruit *Online behavioral advertising* kan ontstaan, het online volgen van gebruikers om zo persoonlijke voorkeuren te kunnen achterhalen (Toubiana et al., 2010). Dit heeft als gevolg dat er gericht persoonlijke advertenties gemaakt kunnen worden. Dit is een inbreuk op privacy en er kan misbruik van worden gemaakt (Juels, 2001). Het onderzoek van Irani et al. (2011) beschrijft dat door de combinatie van verschillende online sociale netwerken een *Online Social Footprint* gemaakt kan worden. Dit combineert de informatie die op de verschillende platformen beschikbaar is om zo een groter geheel aan informatie over een persoon te krijgen. Deze informatie kan gebruikt worden voor bijvoorbeeld fraude.

Samengevat gaan mensen vrij nonchalant om met data. De voordelen van het verbonden zijn en het gemak van online winkelen wegen op tegen het vrijgeven van persoonlijke data. Ook is niet iedereen zich ervan bewust in hoeverre die data terug te leiden is naar de persoon zelf.

5. Theoretisch Kader

De voordelen van het gebruik van het internet en sociale media zorgen ervoor dat privacy op een tweede plek komt te staan, vaak zonder dat dit gerealiseerd wordt. Dit leidt tot een scheve verhouding, het privacy paradox. Het begrip privacy paradox wordt beschreven in het literatuuronderzoek van Kokolakis (2017) als "the dichotomy between privacy attitude and privacy behaviour", een verschil in houding ten opzichte van privacy en het waargenomen gedrag omtrent privacy. Privacy is in dit geval het controleren welke gegevens worden verzameld en opgeslagen en hoe deze worden gebruikt, informatie privacy genoemd (Kokolakis, 2017). In de wetenschap heerst er grote twijfel over het bestaan van het privacy paradox. Er zijn theorieën gevormd en ondersteuningen gevonden om het bestaan te verklaren of in twijfel te trekken. Deze worden hieronder uiteengezet en hieruit worden hypotheses opgesteld.

5.1 Ondersteunen van het privacy paradox

Onmiddellijke bevrediging

Acquisti (2004b) beschrijft mogelijke redenen om dit effect te verklaren in E-commerce. Het onderzoek beschrijft dat mensen neigen naar onmiddellijke bevrediging. Dit wil zeggen dat de voorkeur uitgaat naar meteen een beloning krijgen en daardoor toekomstige nadelen worden genegeerd. Ook worden de risico's van privacy gezien als onrealistisch, mensen verwachten niet dat het ze overkomt.

Rationaliteit

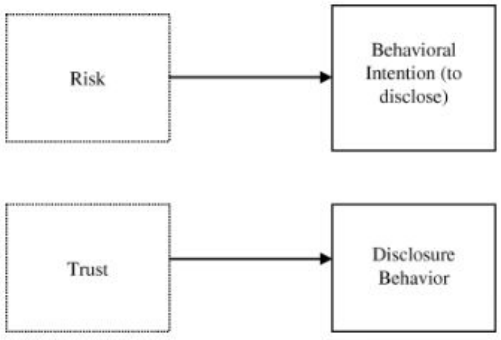
In een vervolgonderzoek van Acquisti & Grossklags (2005) wordt er verder gekeken naar de rationaliteit achter het privacy paradox. Allereerst beschrijven ze *'bounded rationality'*, de keuze wordt beïnvloed door incomplete informatie, denkvermogen en tijdspanne waarin de keuze gemaakt moet worden. Zo laat het onderzoek zien dat van de respondenten die privacy zeer belangrijk vinden 41% de privacy voorwaarden niet of nauwelijks leest. Dit duidt op een incompleetheid van informatie bij het maken van privacy keuzes.

Waarde van data

Hierop aansluitend is er ook onderzoek gedaan naar de geldelijke waarde van privacy. Zo is uit het onderzoek van Carrascal et al. (2013) gebleken dat gebruikers bereid zijn om voor een relatief lage prijs data te verkopen (bijvoorbeeld hun browsegeschiedenis voor gemiddeld 7 euro). Hieruit kan worden opgemaakt dat mensen slechts een kleine waarde hechten aan hun privacy.

Risico & vertrouwen

Norberg et al. (2007) verklaren het verschil tussen intentie en gedrag omtrent privacy doordat ze door verschillende oorzaken gestuurd worden. Zo wordt de intentie om informatie te delen bepaald door het ingeschatte risico, terwijl het gedrag zelf wordt beïnvloed door vertrouwen in de ontvangende partij (Figuur 2). Doordat risico en vertrouwen geen invloed op elkaar hebben, leiden deze tot inconsistenties in intentie en gedrag. Dit zou het privacy paradox kunnen verklaren.



Figuur 2: Risk - Trust op behavior

Source: The privacy paradox: Personal information disclosure intentions versus behaviors. by Norberg, P. A., Horne, D. R., & Horne, D. A. (2007).

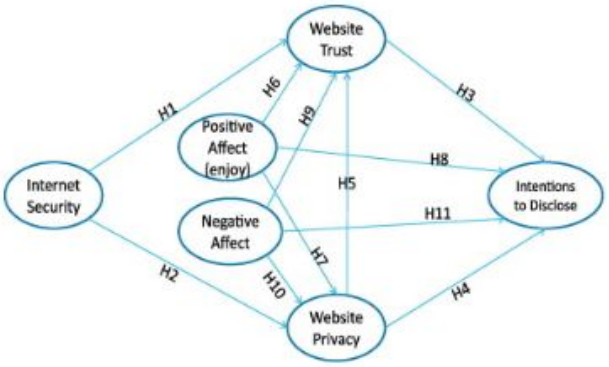
5.2 Ontkrachten van het privacy paradox

Jongeren gewend aan online omgeving

Ook zijn er onderzoeken die als uitkomst hebben dat het privacy paradox niet voorkomt. Blank et al. (2014) hebben onderzoek gedaan naar het privacy paradox tussen leeftijdsgroepen, waarbij wordt gekeken naar demografische gegevens, comfortabelheid om persoonlijke informatie te geven (telefoonnummer, e-mail, postcode, geboortedatum en naam) en in hoeverre ze vinden dat dit een inbreuk is op privacy. Uit het onderzoek blijkt dat vooral bij jongeren de intentie en het gedrag omtrent online privacy overeenkomt. De verklaring die hiervoor wordt gegeven is dat jongeren zo gewend zijn aan het gebruik van internet in het dagelijkse sociale leven dat persoonlijke informatie vrijgeven 'normaal' is.

Vertrouwen & ervaringen

In het onderzoek van Wakefield (2013) wordt ook een overeenkomst gevonden in intentie en gedrag in plaats van een tegenstelling. Het onderzoek test gedrag van internetgebruikers op twee commerciële websites. Het volgende model (figuur 3) werd getest. Hierin wordt de intentie om informatie vrij te geven bepaald uit vertrouwen en privacy van een website en of er sprake is van een positieve of negatieve ervaring. Dit in tegenstelling tot het model (figuur 2) van Norberg et al. (2007), waarin risico de intentie bepaalt in tegenstelling tot vertrouwen.



Figuur 3: Variabelen intention to disclose

Source: The influence of user affect in online information disclosure. By Wakefield, R. (2013). The Journal of Strategic Information Systems, 22(2), 157-174.

Verschillende soorten data

Het is mogelijk dat het wel of niet delen van data afhankelijk is van hoe persoonlijk de data is. Het soort data kan het gedrag om data te delen beïnvloeden. Taddicken (2014) beschrijft dat er persoonlijke data (zoals naam en e-mail) en gevoelige persoonlijke data (zoals medische gegevens) bestaat. Het zou mogelijk kunnen zijn dat hierin een verschil in deelgedrag in bestaat. Zo is het mogelijk dat medische data minder snel gedeeld wordt dan een naam, omdat deze data gevoeliger ligt voor de persoon zelf.

5.3 Hypotheses

Samengevat worden er een aantal verklaringen in de literatuur gevonden voor het privacy paradox, waaruit de volgende hypotheses kunnen worden gesteld:

H1:	*"Er bestaat een incompleetheid aan informatie, gebruikers zijn niet bekend met welke data er wordt bijgehouden en waar deze voor wordt gebruikt."*
H2:	*"Er wordt weinig (monetaire) waarde gehecht aan privacy, data is ruilbaar voor service en sociale interactie of korting."*
H3:	*"Vertrouwen zorgt ervoor dat er desondanks de intentie toch data wordt gedeeld."*
H4:	*"Verschillende niveaus in persoonlijkheid van data hebben invloed op het gedrag."*

6. Methode

Om te onderzoeken of er sprake is van een privacy paradox wordt er gebruikgemaakt van een questionnaire. Het onderzoek richt zich op studenten en kijkt naar online omgevingen. Om het privacy paradox te onderzoeken moet de intentie om privacy te waarborgen vergeleken worden met het daadwerkelijke gedrag in het waarborgen van persoonlijke data.

Hypotheses onderzoeken

Allereerst wordt er een aantal algemene vragen opgesteld die bedoeld zijn om de privacy intentie te meten. Er worden vragen gesteld over wat gebruikers denken over het beschermen van hun data, zoals '*Ik vind het belangrijk dat mijn privacy gewaarborgd blijft*' en '*Ik ben verantwoordelijk om mijn eigen data privé te houden*', om zo een algemene intentie te meten.

Om hypothese 2 te beantwoorden over het effect van beloningen op data worden er stellingen gemaakt over de inwisselbaarheid van data. Dit kijkt hoe vaak gebruikers bereid zijn om persoonlijke data in te ruilen om hiervan te profiteren, op een schaal 5 puntsschaal van nooit tot altijd. Voorbeelden van vragen zijn '*Ik ben bereid meer data op te geven als de service daardoor gratis is*' en '*Ik ben bereid om mijn e-mail op te geven in ruil voor korting*'.

Vertrouwen in een bedrijf of service kan er mogelijk voor zorgen dat de oorspronkelijke intentie wordt genegeerd (hypothese 3). Om dit te onderzoeken worden er vragen gesteld als '*Ik geef echte gegevens op als ik het betreffende bedrijf vertrouw*'.

Om de vraag te beantwoorden of datasoort deelgedrag beïnvloedt (hypothese 4) wordt er een lijst gegeven met verschillende soorten data, zoals 'naam' of 'medische gegevens'. De frequentie wordt gemeten op een 5 puntsschaal van nooit tot altijd. Op deze manier kan er worden onderzocht of bepaalde soorten data sneller of minder snel wordt gedeeld. Zo beschrijft Taddicken (2014) dat er persoonlijke data (zoals naam en e-mail) en gevoelige data (zoals medische gegevens) bestaat. Hier kan een verschil in deelgedrag in bestaan.

Om gebrek aan kennis over data te meten (hypothese 1) wordt als laatste aan de respondenten gevraagd of zij kunnen aanwijzen welke data populaire online platformen (in dit onderzoek Google, Facebook en Snapchat) verzamelen zoals staat beschreven in de privacyvoorwaarden. Dit wordt uitgevoerd met behulp van een meerkeuze raster. Dit kan laten zien of gebruikers op de hoogte zijn van wat er wordt verzameld en verder ondersteunen of de intentie om privacy te beschermen leidt tot het beter kennen van de privacyvoorwaarden van de bedrijven.

Procedure

De volgorde van de vragen wordt willekeurig bepaald om een bias in antwoord volgorde te voorkomen. Voor een volledig overzicht van de enquête zie de link van bijlage 1. Voorafgaand wordt er een introductietekst gegeven om een beeld te schetsen over het onderwerp; het gebruik van persoonlijke data. Hierbij wordt niet beschreven dat het onderzoek gericht is op het onderzoeken van privacy intentie en gedrag, om een bias te voorkomen. Voor het afnemen van de vragenlijst wordt Google Forms gebruikt, een online omgeving voor Questionnaires. Na het invullen van de vragenlijst wordt de respondent bedankt voor zijn of haar tijd.

Er wordt gebruikgemaakt van convenience sampling om respondenten te selecteren, waarbij mensen worden gekozen die makkelijk te bereiken zijn (Burns & Burns, 2008), in plaats van een willekeurig sample.

7. Resultaten

Descriptives

De gegevens uit de vragenlijst worden gebruikt om statistisch onderzoek uit te kunnen voeren. Allereerst wordt er gekeken naar de algemene informatie van de respondenten. In totaal is de enquête beantwoord door 30 respondenten *(N = 30)*, waarvan 14 mannelijk en 16 vrouwelijk, met een minimale leeftijd van 18 en een maximale leeftijd van 25 jaar *(M = 20.57, SD = 1.813)*. Hiervan hebben 20 respondenten van een WO/Universitaire opleiding genoten *(N = 20)*. Voor een volledig overzicht van opleidingsniveaus zie tabel 1.

Opleidingsniveau

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	HAVO	1	3,3	3,3	3,3
	HBO	2	6,7	6,7	10,0
	MBO	2	6,7	6,7	16,7
	VWO	5	16,7	16,7	33,3
	WO/Universitair	20	66,7	66,7	100,0
	Total	30	100,0	100,0	

tabel 1: frequentieverdeling opleidingsniveau

Algemene privacy intentie

Van het totaal aantal respondenten geeft 93.3% geeft aan dat privacy belangrijk *(53.3%)* tot zeer belangrijk *(40%)* is. Daarnaast is 6.3% hier neutraal over en vindt niemand het onbelangrijk (tabel 2). De mediaan van de vraag of het belangrijk is dat privacy gewaarborgd blijft is 4.00 *(Median = 4.00, SD = 0.606; tabel 2)*. Respondenten geven met een mediaan van 4.00 aan dat ze verantwoordelijk zijn voor hun eigen data *(Median = 4.00, SD = 0.923; tabel 3)*.

Ik vind het belangrijk dat mijn privacy gewaarborgd blijft.

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	3	2	6,7	6,7	6,7
	4	16	53,3	53,3	60,0
	5	12	40,0	40,0	100,0
	Total	30	100,0	100,0	

Tabel 2: Frequentieverdeling "Ik vind het belangrijk dat mijn privacy gewaarborgd blijft"

Statistics

		Ik vind het belangrijk dat mijn privacy gewaarborgd blijft.	Ik ben verantwoordelijk om mijn eigen data privé te houden.	Vrienden zijn verantwoordelijk voor het privé houden van mijn data.	Bedrijven zijn verantwoordelijk voor het privé houden van mijn data.	Ik ben bang dat mijn data misbruikt wordt.	Ik ben bang dat mijn data gestolen wordt.
N	Valid	30	30	30	30	30	30
	Missing	0	0	0	0	0	0
Mean		4,33	3,90	3,13	4,13	3,00	3,00
Median		4,00	4,00	4,00	4,00	3,00	3,00
Std. Deviation		,606	,923	1,137	,681	1,083	,947
Variance		,368	,852	1,292	,464	1,172	,897
Minimum		3	2	1	2	1	1
Maximum		5	5	5	5	5	5

Tabel 3: Statistics vragen over privacy

Q2: "Er wordt weinig (monetaire) waarde gehecht aan privacy, data is ruilbaar voor service en sociale interactie of korting."

Er wordt gehypothetiseerd dat service of korting ten gevolge heeft dat privacy wordt opgegeven, ondanks dat privacy belangrijk is. Om dit te onderzoeken wordt er gebruikgemaakt van een dependant ANOVA test. De schaal van de variabele *"Ik vind privacy belangrijk"* wordt omgekeerd. Hierdoor kan privacy tegenover service/korting worden gezet, waarbij een lage score privacy voorop stelt en een hoge score privacy achterstelt.

Mauchly's test geeft aan dat de aanname voor sphericity niet verworpen wordt *($\chi^2 = 12.282$, $p = .585$)*. De nulhypothese dat er geen verschil is tussen intentie in privacy en gedrag bij service/korting kan verworpen worden *($F(5) = 16.134$, $p = <.001$, $\eta p^2 = .357$; Tabel 4)*.

Tests of Within-Subjects Effects

Measure: privacygedrag

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
voordeel	Sphericity Assumed	53,978	5	10,796	16,134	,000	,357	80,670	1,000
	Greenhouse-Geisser	53,978	4,205	12,835	16,134	,000	,357	67,850	1,000
	Huynh-Feldt	53,978	5,000	10,796	16,134	,000	,357	80,670	1,000
	Lower-bound	53,978	1,000	53,978	16,134	,000	,357	16,134	,973
Error(voordeel)	Sphericity Assumed	97,022	145	,669					
	Greenhouse-Geisser	97,022	121,957	,796					
	Huynh-Feldt	97,022	145,000	,669					
	Lower-bound	97,022	29,000	3,346					

a. Computed using alpha = ,05

Tabel 4: One-way ANOVA with Repeated Measures

Het gemiddelde privacybelang *(M = 1.667, SD = 0.111; tabel 5)* wijkt significant af van de gemiddelden van service/korting. Zo geeft de variabele "e-mail opgeven in ruil voor korting" *(M = 3.133, SD = 0.196; tabel 5)* aan dat er geregeld e-mailadressen worden opgegeven voor korting, wat niet correspondeert met het belang van privacy.

1. voordeel

Measure: privacygedrag

voordeel		Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	Ik vind privacy belangrijk (reversed)	1,667	,111	1,440	1,893
2	Persoonlijke informatie opgeven aan vertrouwde bedrijven	3,167	,173	2,812	3,521
3	E-mail opgeven in ruil voor korting	3,133	,196	2,732	3,534
4	Meer data opgeven voor gratis service	2,567	,164	2,231	2,902
5	Voordelen sociaal netwerk wegen op tegen privacyverlies	3,100	,205	2,680	3,520
6	Ik geef mijn gegevens op om korting te ontvangen	2,300	,180	1,931	2,669

Tabel 5: Estimated Marginal Means privacybelang tegenover service of korting

H4:"Verschillende niveaus in persoonlijkheid van data hebben invloed op het gedrag."

Om te onderzoeken of er een verschil is in het deelgedrag tussen verschillende soorten data worden de verschillende soorten data uiteengezet met een dependant ANOVA test. Mauchly's test geeft aan dat de aanname voor sphericity verworpen moet worden *($\chi^2 = 131.201$, $p < .001$)*. De nulhypothese dat er geen verschil is tussen de soorten met betrekking tot deelgedrag kan verworpen worden *($F(6.814) = 31.038$, $p = <.001$, $\eta p^2 = .517$; Tabel 6)*.

Tests of Within-Subjects Effects

Measure: Deelgedrag

Source		Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared	Noncent. Parameter	Observed Power[a]
gedeelde_data	Sphericity Assumed	227,390	12	18,949	31,038	,000	,517	372,461	1,000
	Greenhouse-Geisser	227,390	6,814	33,373	31,038	,000	,517	211,481	1,000
	Huynh-Feldt	227,390	9,123	24,925	31,038	,000	,517	283,162	1,000
	Lower-bound	227,390	1,000	227,390	31,038	,000	,517	31,038	1,000
Error(gedeelde_data)	Sphericity Assumed	212,456	348	,611					
	Greenhouse-Geisser	212,456	197,593	1,075					
	Huynh-Feldt	212,456	264,566	,803					
	Lower-bound	212,456	29,000	7,326					

a. Computed using alpha = ,05

Tabel 6: One-way ANOVA with Repeated Measures

Zo worden 'medische gegevens' het minst vaak gedeeld *(M = 1.067, SD = .046; tabel 7)* en 'naam' het meest *(M = 3.567, SD = .213; tabel 7)*. 'Medische gegevens' zijn een vorm van gevoelige persoonlijke data, terwijl 'naam' een voorbeeld is van minder gevoelige data.

DataType

Measure: Deelgedrag

DataType		Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	Medische Gegevens	1,067	,046	,972	1,161
2	Relatiestatus	1,667	,138	1,384	1,950
3	Studie / Werk	2,533	,184	2,157	2,910
4	Wat ik leuk vind	3,067	,203	2,651	3,482
5	Politieke voorkeur	1,300	,128	1,038	1,562
6	Adres	1,600	,149	1,296	1,904
7	E-mail	2,700	,254	2,180	3,220
8	Telefoonnummer	1,667	,154	1,351	1,982
9	Foto's / Videos	2,400	,156	2,081	2,719
10	Geboortedatum	2,733	,203	2,318	3,149
11	Locatie	1,867	,150	1,561	2,173
12	ID-kaart nummer	1,067	,046	,972	1,161
13	Naam	3,567	,213	3,132	4,002

Tabel 7: Estimated Marginal Means Data Soorten

H3: "Vertrouwen zorgt ervoor dat er desondanks de intentie toch data wordt gedeeld."

Om te onderzoeken of vertrouwen in een bedrijf invloed heeft op het delen van informatie wordt er gebruikgemaakt van spearman's rho correlatiecoëfficiënt. Er wordt een significante, gemiddeld sterke positieve correlatie gevonden tussen hoe vaak er wordt gedeeld op platformen en het vertrouwen in een website als vrienden *($r_s = .417$, $p = .022$; tabel 8)* of de persoonlijke omgeving *($r_s = .438$, $p = .015$; tabel 8)* hierin vertrouwen. Het delen van informatie leidt tot minder privacy, waardoor vertrouwen privacy negatief beïnvloedt.

Correlations

			Hoe vaak deelt u op online platformen?	Ik vertrouw een bedrijf eerder als mijn vrienden het vertrouwen.	Ik vertrouw een bedrijf eerder als mijn omgeving het vertrouwd.
Spearman's rho	Hoe vaak deelt u op online platformen?	Correlation Coefficient	1,000	,417*	,438*
		Sig. (2-tailed)	.	,022	,015
		N	30	30	30

tabel 8: Spearman Correlatie frequentie delen en vertrouwen

H1: "Er bestaat een incompleetheid aan informatie, gebruikers zijn niet bekend met welke data er wordt bijgehouden en waar deze voor wordt gebruikt."

Om een beeld te krijgen van de kennis van respondenten over welke gegevens worden verzameld is gevraagd aan te kruisen welke gegevens worden verzameld door de bedrijven Google, Facebook en Snapchat. Dit zijn bekende online platformen die veelal gebruikt worden. Hierbij wordt verondersteld wordt dat deze informatie als bekend kan worden beschouwd in het geval dat privacy belangrijk is.

De verschillende gegevens worden allereerst apart van elkaar beschouwd. One-sample T-test wordt gebruikt met een testwaarde van 3, waarbij wordt uitgegaan dat de antwoorden juist worden

aangekruist. Hier worden significante afwijkingen gevonden tussen de variabelen en de veronderstelde kennis. De nulhypothese “er is geen verschil tussen daadwerkelijke kennis en veronderstelde kennis” kan verworpen worden. Zo is de kennis voor gezichtsherkenning *(95% CI, -.84 tot -.23; tabel 9)* significant lager dan de veronderstelde kennis van score 3 *(t(29) = -6.196, p < .001; tabel 9)*.

One-Sample Test

	Test Value = 3					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
Telefoongegevens	-3,565	29	,001	-,533	-,84	-,23
code_voor_gezichtsherkenning	-6,196	29	,000	-1,300	-1,73	-,87
Browsegeschiedenis	-7,549	29	,000	-,867	-1,10	-,63
Clickgeschiedenis	-4,817	29	,000	-,667	-,95	-,38
Apparaatgegevens	-4,000	29	,000	-,533	-,81	-,26
Locatie	-2,796	29	,009	-,367	-,63	-,10
Verzonden_berichten	-4,428	29	,000	-,733	-1,07	-,39
IPadres	-4,011	29	,000	-,567	-,86	-,28
Locatiegeschiedenis	-3,638	28	,001	-,379	-,59	-,17

tabel 9: One-Sample T-Test data verzameld door bedrijven

Daarnaast wordt er gekeken of dit verschil kan worden verklaard door een verschil tussen kennis over de dataverzameling van de verschillende bedrijven. Om dit te onderzoeken wordt een totaalscore voor ieder bedrijf gemaakt door de antwoorden van iedere soort data op te tellen en vervolgens een One-Sample T-Test uitgevoerd. De maximale totaalscore heeft een waarde van 9, het totaal van het aantal verschillende soorten data. De kennis over de dataverzameling van Google *(t(29) = -4.469, p < .001; tabel 10)*, Facebook *(t(29) = -4.889, p < .001; tabel 10)* en Snapchat *(t(29) = -5.943, p < .001; tabel 10)* zijn significant lager dan de score van 9. Hierbij wijkt Snapchat het meest af, met een gemiddelde van 5,40 *(M = 5.40, SD = 3.318; tabel 11)*.

One-Sample Test

	Test Value = 9					
	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
totaal Google	-4,469	29	,000	-1,100	-1,60	-,60
totaal Facebook	-4,889	29	,000	-1,233	-1,75	-,72
totaal Snapchat	-5,943	29	,000	-3,600	-4,84	-2,36

Tabel 10: One-Sample T-Test dataverzameling van bedrijven, gegroepeerd

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
totaal Google	30	7,90	1,348	,246
totaal Facebook	30	7,77	1,382	,252
totaal Snapchat	30	5,40	3,318	,606

Tabel 11: Sample Statistics kennis tussen bedrijven

8. Conclusie

Om te onderzoeken in hoeverre er sprake is van een privacy paradox is onderzocht hoe het belang van privacy zich verhoudt als er moet worden gekozen voor privacy of een andere waarde (korting/service en vertrouwen) en mogelijke verklaringen hiervan (data soort en kennis). De resultaten laten ondersteuning zien voor alle gestelde hypotheses (figuur 4).

Figuur 4: Ondersteunde hypotheses

Privacy of korting

Uit de resultaten blijkt dat alhoewel privacy belangrijk wordt gevonden, deze mogelijk toch wordt opgegeven wanneer er een beloning tegenover staat. Persoonlijke data is dus inwisselbaar wanneer de beloning meer waarde heeft dan het behouden van privacy. Dit ondersteunt de resultaten uit het onderzoek van Carrascal et al. (2013), waarin wordt beschreven dat data tegen een relatief lage prijs wordt verkocht.

Privacy of vertrouwen

Uit de resultaten kan worden geconcludeerd dat het behouden van privacy in de vorm van persoonlijke data minder belangrijk wordt naarmate de voordelen van het delen van data toenemen. Vertrouwen in bedrijven en voldoende voordelen zorgen er dus voor dat er meer data wordt gedeeld, ongeacht de oorspronkelijke intentie om data persoonlijk te houden. Het model van Norberg et al. (2007) (figuur 2; p.7) suggereert dat het privacy paradox ontstaat omdat risico de privacy intentie beïnvloed en vertrouwen het daadwerkelijke gedrag. Dit onderzoek ondersteunt dat vertrouwen invloed heeft op het delen van privacy ongeacht de oorspronkelijke intentie. Dit spreekt het model van Wakefield (2013) (figuur 3; p.8) tegen, waarin vertrouwen juist de oorspronkelijke intentie zou beïnvloeden en daarom het bestaan van een privacy paradox zou tegenspreken.

Soorten data

De privacy afwegingen worden beïnvloed door het soort data. Zo worden data als medische gegevens minder snel gedeeld dan bijvoorbeeld namen. Medische gegevens zijn gevoelige gegevens (Taddicken, 2014). Dit kan verklaren waarom deze minder snel gedeeld worden. Hierdoor wordt de drempel voor

het delen van data verhoogd. Mogelijk wordt deze data nog steeds gedeeld als de beloning maar groot genoeg is.

Kennis over privacy

De hoge gewilligheid om persoonlijke data in te leveren kan mogelijk worden verklaard door onwetendheid over hoeveel persoonlijke data kan worden bijgehouden en waar deze voor gebruikt kan worden. Zo kan data gemakkelijk aan elkaar gekoppeld worden om zo profielen van mensen te maken (Acquisti, 2004a). Uit de resultaten blijkt dat gebruikers niet goed bekend zijn met welke data over ze kan worden verzameld, terwijl dit wel verwacht zou worden aangezien privacy belangrijk wordt gevonden. Dit kan mogelijkerwijs verklaard worden door het vertrouwen in bedrijven (sectie '*Privacy en vertrouwen'; p. 15*), waar vertrouwen leidt tot acceptatie zonder precies te weten welke gegevens worden verzameld. Het onderzoek van Acquisti en Grossklags (2005) ondersteunt dat incomplete informatie kan leiden tot niet goed onderbouwde keuzes.

Het privacy paradox

Met deze informatie kan de hoofdvraag beantwoord worden:
"In hoeverre is er sprake van een privacy paradox op online platformen?"
Wanneer er sprake is van een privacy paradox is er een verschil in houding en gedrag ten opzichte van privacy. Privacy wordt belangrijk gevonden, maar in de praktijk worden er toch persoonlijke gegevens gedeeld. Ook het onderzoek laat zien dat privacy belangrijk wordt gevonden, maar dit in de praktijk niet persé wordt nageleefd. Vertrouwen in online platformen en onwetendheid van de gevaren en gevolgen van het delen van data leiden tot een onvermogen om goede overwogen keuzes te maken. Privacy wordt niet gezien als een algemeen recht, maar een verhandelbaar goed, dat tegen de juiste prijs te koop is.

9. Discussie

Steekproef

Een goede afspiegeling van de samenleving krijgen uit een steekproef is vrij lastig te realiseren. In het onderzoek is gebruikgemaakt van convenience sampling, het verkrijgen van respondenten die makkelijk te bereiken zijn. Dit heeft als gevolg dat er (mogelijk) geen goede afspiegeling wordt gemaakt van de samenleving. Zo heeft van de respondenten 73.4% (tabel 1, p.9) een HBO of WO/Universitaire opleiding genoten, terwijl dit in Nederland lager ligt op 30.1% (Onderwijsincijfers.nl, 2018).

Ook de sample size (N = 30) is niet groot, waardoor mogelijke significante effecten in de samenleving niet kunnen worden gemeten. Een mogelijke oorzaak hiervan is dat rond deze tijd veel onderzoeken worden uitgevoerd, waardoor veel respondenten nodig zijn en deze daardoor niet veel questionnaires willen invullen.

Methode

Wegens de beperkte tijd voor het onderzoek is er gekozen om een kwantitatieve analyse uit te voeren met behulp van een questionnaire. Met meer tijd was een mixed methods onderzoek, de combinatie

van kwantitatieve en kwalitatieve analyse, mogelijk een betere keuze geweest. Op deze manier is het mogelijk te achterhalen waar de echte problemen binnen de doelgroep zitten en welke verklaringen hier door de respondenten zelf voor worden gegeven. Zo kunnen de mogelijke afwegingen in privacy gedrag verklaard worden.

Questionnaire

De onderzoeken uit de theorie maakten geen gebruik van erkende modellen om privacy intentie en gedrag te toetsen. De questionnaire uit dit onderzoek is grotendeels gebaseerd op vragen uit eerdere onderzoeken, waardoor deze mogelijk niet optimaal zijn. Ook maakt dit onderzoek onderscheid tussen verschillende platformen (Google, Facebook en Snapchat), terwijl deze in de enquêtes uit de theorie vaak samen worden genomen als 'online platformen', waardoor de vragen uit de questionnaire iets kunnen afwijken.

10. Future research

Uit dit onderzoek blijkt dat persoonlijke data een waarde heeft. In toekomstig onderzoek zou het interessant kunnen zijn om de waarde van persoonlijke data zo goed mogelijk vast te stellen. Hierbij kan bijvoorbeeld gekeken worden naar welke afwegingen worden gemaakt, hoe verschillende soorten data hier invloed op hebben en de psychologische denkwijze die erachter zit. Dit zou een indicatie kunnen geven van het belang van privacy en hoe privacy zich verhoudt tegenover andere normen en waarden, zoals veiligheid.

Een ander onderzoek zou kunnen testen of betere voorlichting over persoonlijke data ervoor kan zorgen dat persoonlijke data privé blijft of meer waard wordt. Hier kan uit worden opgemaakt welke rol persoonlijke data in het persoonlijke leven gaat spelen wanneer er bekend is hoeveel informatie kan worden gehaald uit een klein beetje data.

Referenties

1. Acquisti A. (2004a). Privacy and Security of Personal Information. In: Camp L.J., Lewis S. (eds) Economics of Information Security. Advances in Information Security, vol 12. Springer, Springer, Boston, MA.

2. Acquisti, A. (2004b). Privacy in electronic commerce and the economics of immediate gratification. In Proceedings of the 5th ACM conference on Electronic commerce (pp. 21-29). ACM. doi> 10.1145/988772.988777

3. Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. IEEE Security & Privacy, 3(1), 26-33.

4. Acquisti, A. (2010). The economics of personal data and the economics of privacy. 35-40. http://repository.cmu.edu/heinzworks/332/

5. Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. First Monday, 11(9). http://journals.uic.edu/ojs/index.php/fm/article/view/1394/1312

6. Blank, G., Bolsover, G., & Dubois, E. (2014). A new privacy paradox. In Annual Meeting of the American Sociological Association. San Francisco, CA. Retrieved https://www.oxfordmartin.ox.ac.uk/downloads/A%20New%20Privacy%20Paradox%20April%202014.pdf (Vol. 202014).

7. Burns, R. & Burns, R. (2008). Business research methods and statistics using SPSS (1st ed.). Los Angeles: SAGE.

8. Carrascal, J. P., Riederer, C., Erramilli, V., Cherubini, M., & de Oliveira, R. (2013). Your browsing behavior for a big mac: Economics of personal information online. In Proceedings of the 22nd international conference on World Wide Web (pp. 189-200). ACM.

9. Debatin, B. , Lovejoy, J. P., Horn, A. and Hughes, B. N. (2009), Facebook and Online Privacy: Attitudes, Behaviors, and Unintended Consequences. Journal of Computer-Mediated Communication, 15: 83-108. doi:10.1111/j.1083-6101.2009.01494.x https://onlinelibrary.wiley.com/doi/full/10.1111/j.1083-6101.2009.01494.x

10. Facebook. (2018a). Accessing Your Facebook Data | Facebook Help Centre | Facebook. [online] Available at: https://www.facebook.com/help/405183566203254 [Accessed 12 Apr. 2018].

11. Facebook. (2018b). Data Policy. [online] Available at: https://www.facebook.com/about/privacy [Accessed 12 Apr. 2018].

12. Govani, T., & Pashley, H. (2005). Student awareness of the privacy implications when using Facebook. unpublished paper presented at the "Privacy Poster Fair" at the Carnegie Mellon University School of Library and Information Science, 9, 1-17.

13. Hunter, R., De Lotto, R. J., Frank, A., Gassmann, B., Hallawell, A., Heiser, J., & Taylor, D. (2009). What does Google know. *Report G00158124. Stamford, CT: Gartner Research*.

14. Irani, D., Webb, S., Li, K., & Pu, C. (2011). Modeling unintended personal-information leakage from multiple online social networks. *IEEE Internet Computing*, *15*(3), 13-19. http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5696719

15. Juels, A. (2001). Targeted advertising... and privacy too. In Cryptographers' Track at the RSA Conference (pp. 408-424). Springer, Berlin, Heidelberg.

16. Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. Computers & Security, 64, 122-134.

17. Martijn, M. & Tokmetzis, D. (2018). Je hebt wél iets te verbergen. 5th ed. de Correspondent, pp.145-160. ISBN 9789082821611

18. Nl-nl.facebook.com. (2018). Helpcentrum. [online] Available at: https://nl-nl.facebook.com/business/help/742478679120153?helpref=page_content [Accessed 13 Apr. 2018].

19. Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of Consumer Affairs*, *41*(1), 100-126. https://onlinelibrary.wiley.com/doi/full/10.1111/j.1745-6606.2006.00070.x

20. Onderwijsincijfers.nl. (2018). Hoogst behaald onderwijsniveau. [online] Available at: https://www.onderwijsincijfers.nl/kengetallen/sectoroverstijgend/nederlands-onderwijsstelsel/hoogst-behaalde-onderwijsniveau [Accessed 16 Jun. 2018].

21. Policies.google.com. (2018). Privacybeleid. [online] Available at: https://policies.google.com/privacy?hl=nl [Accessed 11 Apr. 2018].

22. Privacy.google.com. (2018). Google Privacy | Waarom gegevensbescherming belangrijk is. [online] Available at: https://privacy.google.com/intl/nl/your-data.html [Accessed 11 Apr. 2018].

23. Snap.com. (2018). Privacy Center – Snap Inc.. [online] Available at: https://www.snap.com/en-US/privacy/privacy-policy/ [Accessed 14 Apr. 2018].

24. Stoddart, E. (2014). (In) visibility Before Privacy: A Theological Ethics of Surveillance as Social Sorting. Studies in Christian Ethics, 27(1), 33-49.

25. Taddicken, M. (2014). The 'privacy paradox' in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of Computer-Mediated Communication*, *19*(2), 248-273.

26. Toubiana, V., Narayanan, A., Boneh, D., Nissenbaum, H., & Barocas, S. (2010). Adnostic: Privacy preserving targeted advertising.

27. Wakefield, R. (2013). The influence of user affect in online information disclosure. The Journal of Strategic Information Systems, 22(2), 157-174.

28. Xie, E., Teo, H. H., & Wan, W. (2006). Volunteering personal information on the internet: Effects of reputation, privacy notices, and rewards on online consumer behavior. *Marketing letters*, *17*(1), 61-74. https://link.springer.com/content/pdf/10.1007/s11002-006-4147-1.pdf

29. Xu, B., Chang, P., Welker, C. L., Bazarova, N. N., & Cosley, D. (2016, February). Automatic archiving versus default deletion: what Snapchat tells us about ephemerality in design. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing* (pp. 1662-1675). ACM.

30. Young, D. (2014). Now You See It, Now You Don't... Or Do You?: Snapchat's Deceptive Promotion Of Vanishing Messages Violates Federal Trade Commission Regulations, 30 J. Marshall J. Info. Tech. & Privacy L. 827 (2014). The John Marshall Journal of Information Technology & Privacy Law, 30(4), 6.Boston, MA

Bijlage

bijlage 1: Link Questionnaire
https://goo.gl/forms/CFa2OC5K0aflPhaF3