Trusting organisations online

Contextual integrity in public privacy concern

Study:	Erasmus University Rotterdam – Erasmus School of History, Culture and Communication – Department of Media & Communication – Digitalisation, Surveillance & Societies
Name:	Casper Kerklaan
Student number:	510837
Course title:	Master thesis
Course number:	CM5000
Supervisor:	dr. Jorge Pereira Campos
Date:	June 23, 2022
Word count:	16061 words

Many thanks to: dr. Jorge Pereira Campos & dr. Joao Fernando Ferreira Goncalves

---

Key words: Privacy concern, Contextual Integrity, Privacy Knowledge, Organisational Trust, Repeated Measures ANOVA

Abstract

The introduction of the General Data Protection Regulation (GDPR) in the European Union has brought important changes in how personal information gets distributed and handled online, by individuals and organisations. Applying the theory of contextual integrity, holding that to approach data subjects' privacy we ought to look towards the flow of information in specific contexts, this study tests whether the collection and use of different types and amounts of personal data, collected by distinct public, private and hybrid organisations, result in significant differences in privacy concern among Dutch citizen-consumers. Moreover, individual factors have been considered, given how organisational trust was expected to negatively affect privacy concern, with less trust leading to more individual worry. Individual privacy knowledge, the extent to which people know how to handle and secure their data, was thought to strengthen this effect.

To test these expectations, this study made use of an already existing dataset, created from a survey that was filled out by 510 Dutch participants, 311 of whom were included in the analyses of this study. To determine the role of contextual integrity in public privacy concern, respondents' answers on approximately 6912 unique vignette combinations were used, portraying hypothetical scenarios in which individuals shared data to different organisations, asking to what extent they felt worried about the sharing of their personal information in each circumstance. These vignettes were analysed using repeated measures ANOVA, also known as one-way within-subjects ANOVA, allowing for the testing of contextual differences in privacy concern depending on the organisation collecting the data, the amount of data collected, and the type of data collected. Additionally, a hierarchical regression analysis was conducted to account for individual factors influencing privacy concern, using survey responses outside the scope of the vignette study. This analysis could explain some of the variation in organisational trust, privacy knowledge, and privacy concern among individuals. These factors have been controlled for using age, levels of education, and frequency of smartphone use.

The findings revealed that organisational trust was negatively related to privacy concern, however, privacy knowledge did not strengthen this relationship as a moderating variable, though it was still found to be a separate significant negative predictor of privacy concern instead. Significant differences were found in privacy concern depending on the organisation that collected the data, though this finding was only limited to some organisations, and varied scarcely depending on the type or amount of data. This study has provided new insights regarding public privacy concern in a context where strict privacy policies, courtesy of the GDPR, dictate how organisations can and should handle the personal data of the individual. The increased knowledge that the EU's GDPR has provided its citizen-consumers has shown to have changed people's perceptions on their rights to privacy and their concerns over organisations' collection of their personal data.

Table of contents

Introduction

Nowadays, when people go online to visit a website, they are often confronted with a message or a pop-up requesting them to check or adjust their privacy settings: so-called cookie banners (Nocera, 2022). These messages pop up with the intent to make the customer visiting a website aware of how their personal information is being gathered by the provider of the webservice, and it offers them a tool to adjust their preferences on what data they share about themselves. Service providers on the other hand want to gather and process the data they collect of their customers for purposes like targeted advertising, improving customer experience on the site, and monitoring their behaviour on the site in general (Kulyk et al., 2018), all tools to ensure that the customer optimally experience their website's functionalities, and to keep them interested in its contents for as long as possible.

But associated with web providers' intent to collect data from their respective visitors, is a risk of privacy breaches for internet users, where personal information is either unlawfully collected by the providers (Cullen & Reilly, 2008), or the data are maliciously obtained by third parties that were not supposed to gain these data (Choi et al., 2018), this can be either the consequence of lack of security or data protection by the service provider that originally collected the data, or an oversight caused by human error (Leering et al., 2022; Luo et al., 2011).

To prevent this from happening, the European Union used, while they are placed at risk of monetary penalisation under the circumstance that they violate these individuals' rights to privacy. This law is better known as the General Data Protection Regulation (GDPR), and its jurisdiction apply to all organisations that handle data pertaining to European Union citizens (Wolford, 2020). This regulation recognises the right of the individual to have all personal information about themselves removed if they wish that to be the case (Tsesis, 2014), a perspective that contrasts that of the privacy regulations in other nations, such as the US, where online platforms' desires to build better profiles of citizen-consumers are prioritised over these rights (Adomavicius & Tuzhilin, 2001; Myers, 2014).

With the regulation setting a limit to the amount of sensitive data that organisations can collect about citizen-consumers, for organisations, both public and private, this can come at a financial cost. Public institutions want to collect data from internet users to decrease such financial costs (Rogge et al., 2017), while private corporations want to use it to make financial

profits instead (Winter & Davidson, 2019), and if neither can collect the information they want, rather than need, they risk losing money because of it.

The risks for the individual when disclosing their personal data, and the benefits that organisations can obtain by collecting such information, makes for a dichotomy that can lead to an increase in privacy concern among citizen-consumers (Bandyopadhyay, 2012). This originates from a lowered trust in the organisations that want to use their personal information (Cullen & Reilly, 2008). The implementation of the GDPR in 2018 has made it so that organisations can only collect the amount of personal data that they need from their online customers, for example to make the website that a user visits run without issue (Adomavicius & Tuzhilin, 2001), or to confirm an individual's identity in case of an online purchase on a web shop. This process of limiting the amount and type of data collected to only what is necessary, is called data minimisation (Zarsky, 2017), and it reveals that the GDPR uses an approach based on contextual integrity (Nissenbaum, 2004). This is a standard that views each context in which personal information is requested in a different light, only highlighting the data required within the given context or scenario. For each actor involved, for each data type, or the amount of data shared and collected, differences in people's perceptions of data privacy can be found, for example in how willing data subjects are to share this data with interested parties (Roeber et al., 2015).

The focus of this study has been placed on the Netherlands, for two reasons. First, given the Netherlands' status as an EU member state (European Union, n.d.), Dutch citizen-consumers and organisations are subject to the GDPR's privacy law (Wolford, 2020). Second, and what makes the Netherlands different from other member states, is the hybrid role of its healthcare. Where health insurance companies have become privatised since a law change in 2006 (Maarse et al., 2016), there still exist public healthcare institutions as well (Toebes & San Giorgi, 2015). This divide is not visible in the case of for example police or local government, both of which are exclusively public institutions in the Netherlands, or social media platforms like Twitter and Facebook, which are private companies that intend to make profit by appealing to visitors online (Myers, 2014; Tsesis, 2014).

This study expands upon the theory of contextual integrity by using it to explore whether differences can be found in people's privacy concerns depending on the hypothetical scenarios which they were provided, through answering vignette surveys. The research question asks: *How*

are Dutch citizens' concerns about personal data usage affected by their perceptions of privacy and their trust in public and private organisations, following the implementation of the GDPR in the European Union?

With the GDPR having been implemented in the European Union as recently as 2018 (Goddard, 2017), this study offers Roeber and colleagues (2015) studied users' willingness to share personal information before the implementation of the GDPR. Research on privacy concern is similar, given that the addition of these new regulations can affect people's views on data security and their rights to privacy (Custers et al., 2018; Goddard, 2017; Trepte et al., 2015). Consequently, this can shed new light on how one's worry over the control of their own data can be affected once they become more knowledgeable about the potential repercussions that come with sharing their personal information (Cullen & Reilly, 2008).

Recent research no longer seems interested in understanding the concept of privacy and people's perceptions thereof, but rather, it seeks to explain how people interact with their own rights to privacy, and their concerns pertaining to it. In a similar fashion, this study seeks to understand how organisational trust, or the loss thereof, affects citizens' online interactions with e-commerce as well as e-government, and what lies in between in the form of as hybrid Dutch health care. The negative consequences of a loss of trust in institutions online are what marks down the importance to obtain a greater understanding of the conditions under which institutional trust improves or worsens when citizens interact with e-government and e-commerce.

In this study, a literature review was conducted to define and elaborate on the most relevant concepts included in the research question, including individual privacy, personal information, organisational trust, as well as privacy knowledge and concern, and how these all come together through the theory of contextual integrity. Following that, in the methodology, two methods of analyses have been worked out, one to explain the role that individual factors play in public privacy concern, with the other working out how contextual factors play part in it, the findings of both types of analyses having been written down in the results section. Finally, the hypotheses that were tested receive elaboration in the conclusion, followed by some mentions of research limitations and proposals for future research.

Literature review

Privacy in an information age

As with most definitional debates, when the term privacy first became popularised in the age of information technology as a social issue among researchers in the social sciences, it was not introduced with a single universal definition (Margulis, 2003). At the beginning of this era, Westin (1967) and Altman (1975) already provided different theories on what privacy encompasses. Both theories did however agree on the fact that privacy was something that the individual could control for themselves. Knowing this, problems can arise when individuals' control over their own privacy, by the consequence of outside parties' and actors' interferences, is lost. For this reason, instead of looking for a universal meaning of the concept, one must instead focus on how information privacy is understood, interpreted, and acted upon by individuals and groups falling victim to such outside influences (Margulis, 2003; Steeves, 2009).

The violation of people's privacy has been a long-standing topic of concern (Steeves, 2009). This concern has been further fuelled by the recent developments of digital technologies, like the popularisation of the internet at the end of the 20th century, as well as the rise of Web 2.0 social media only two decades later (Allen, 2012). This sudden surge in popularity of social media like Facebook and Twitter has left these same platforms unable to properly manage their users' personal information and their privacy in general (Kumar & Kumar, 2010). This inability bore risk of data leakage, but a lack of legislation enacted by government institutions worldwide also meant that companies could misuse and mismanage personal information for commercial purposes (Bandyopadhyay, 2012).

Consequently, new measures have increasingly been put in place by regional, national, and transnational agencies to ensure people's personal information remains protected and secure, such as the GDPR in the European Union and the Health Information Interoperability and Accountability Act (HIPAA) in the United States (Appari et al., 2009; Solove, 2016). Contrastingly, the Freedom of Information Act in the US serves as an example showing how government institutions are required by law to disclose information to the public when it is asked of them (Halstuk & Chamberlin, 2006). Privacy laws on different societal levels can thereby be seen as serving the needs of a single entity (i.e., the individual). By demanding of institutions that they release undisclosed information regarding their policies and codes of conduct, citizens can also uncover in what ways the government handles their personal data, and whether they are

doing so ethically and fairly. The implementation of the GDPR has thus increased individual agency among data subjects, given their growing knowledge over their privacy rights (Goddard, 2017; Strycharz et al., 2020). Consequently, citizens have become critical over both the legality of different organisations' data collection strategies, as well as the effectiveness of the legislation that aims to resolve these issues. The increased awareness of the data subject on part of their personal information, creates a new power balance, with actors collecting personal data as the opposing force. For the public and private organisations in the Netherlands, this means performing a balancing act involving transparency and accountability of organisation policy to maintain and increase citizen and consumer trust and abiding by GDPR rules (Goddard, 2017; Janssen & Van den Hoven, 2015), while at the same time upholding confidentiality for the purpose of protecting sensitive information, or instead to make profits from it (Winter & Davidson, 2019).

Institutional and corporate data collection

The technological developments that arose in conjunction with the information age have led to organisations adopting new approaches to gain, or retain, a competitive advantage over company rivals (Prescott, 2014). Such new approaches also involve the collection and leveraging of private data of individuals. For organisations in the European Union however, because of the recent introduction of the GDPR, it has become more difficult than for their American counterparts to compete in the global economy. This is because in the latter nation, companies can still adopt strategies that allow them to skirt GDPR regulation on personal data collection (Voss & Houser, 2019), not to mention collect and use data from American citizen-consumers with comparatively less legal resistance.

Moreover, American law surrounding data privacy does not always favour the individual, and instead shows a keenness towards the corporations and desires for data collection instead, a position which contradicts that of the GDPR (Myers, 2014). In the case of social media companies for example, EU citizens have "the right to be forgotten" (Tsesis, 2014), whereas in the US, the content that social media users have shared on the platforms, including potential personal information, can remain on the platform forever, given how US law does not demand of social media networks to remove it. The power balance between the public and the organisations thus does not exist in countries where permanent data retention is prioritised. Data subjects

falling under such legislations maintain comparatively less personal agency than do citizens in EU member states, a visible cultural difference between these regions (Myers, 2014).

In more recent years however, Grimmelmann (2020) saw developments in US law that have begun favouring the side of the citizen-consumers. California and its California Consumer Privacy Act (CCPA), for one, has taken an approach akin to the EU's GDPR in the protection of personal information. Despite this, most US states still favour the first amendment right for the freedom of speech and press over individual privacy, still treating the balance between privacy and security as a duality, albeit unjustly, according to Dawes (2014) and Ghioshray (2006). The complete disappearance of individual security will never lead to the complete liberty of the individual, meaning that the dichotomy between the two concepts is untrue. The presence of this duality in American culture can make it difficult for EU organisations to involve themselves with actors based in the US (Myers, 2014), as the latter group would need to prioritise GDPR regulations over their own principles to continue trading information with the former group (Voss & Houser, 2019).

Private companies can obtain a competitive advantage over other organisations by using sensitive personal information to create increasingly accurate and effective profiles of their consumers, which allows them to better target these individuals in several different ways. First, companies can use the acquired information for the personalisation of online services – and for social media platforms user-generated content as well – towards customers (Adomavicius & Tuzhilin, 2001; Vesanen & Raulas, 2006). Customers from streaming platforms like YouTube and Netflix can then be recommended videos and movies that suit their tastes more, based on browsing patterns of not only themselves, but also other customers "similar to you" (Netflix, n.d.; YouTube, n.d.). Such personalisation allows for companies to better target their customers' needs while interacting or making transactions online. With better consumer profiling and personalisation, the products and content that users are recommended can become more representative of their own interests (Gauch et al., 2007). Website visitors can come to see this as beneficial, and it is viewed as especially helpful on the long term, as more becomes known of the user, allowing for more accuracy in their recommendations (Telzrow & Kobsa, 2004). The personalisation of online web services can add to customers' *flow experience* (Wang et al., 2019), making activity on these platforms more immersive and engaging, though at the same time making them more susceptible to privacy violations, as a false sense of security can lead to

a disregard to its risks. To visit specific parts of a website, users may for example have to allow the use of more cookies, opening more of their personal data to the service provider.

When having gathered personal data, private organisations can opt to further sell their acquired data to other, third-party firms and actors (Hoofnagle, 2003; Schroeder, 2016). Organisations that make profit off data using this method are also called data brokers. The role of these data brokers has become increasingly important in the information age as the amount of data required for companies to uphold an advantage over competitors rises. Rather than directly having personal information transferred over from the individual to the organisation that requests it – the most legally sound method of data collection – the latter group can then instead opt to obtain it through these third parties (Conger et al., 2013).

The collection of sensitive data through multiple disparate data brokers are one way to allow for the analysis and subsequent combination of databases, creating new data altogether (Gates & Matthews, 2014; Schroeder, 2016) The sale of personal data is not exclusive to the private sector either. Hoofnagle (2003) mentioned the role of data brokerage in the purchase and downloading of databases by public institutions like police forces. Commercial data brokers take the interest of law enforcement into account by reorganising the data they have acquired in such a way, that it is easier for law enforcement agencies to leverage and analyse them.

Public institutions tend to collect data not for financial gain, but rather to decrease financial costs (Rogge et al., 2017). The bureaucratic procedures of public institutions are often time and cost-intensive, but by reducing the amount of physical interaction required between government and citizens through the de-materialisation of the bureaucratic process, in essence delegating administrative operation from the town hall to the computer, costs for personnel can for one be lowered, though at the cost of citizens-consumer privacy.

The third-party involvement in buying new data for more advanced analysis is but one strategy that public and private organisations alike adopt in the data collection procedure. Data mining is another technique commonly used by organisations to make customer profiles, for one to improve personalisation when web services are being used (Adomavicius & Tuzhilin, 2001), for another to improve public safety by providing an algorithmic tool through which personal information from different sources can be linked to uncover previously unseen patterns of deviant behaviour (Fienberg, 2006; Miller, 2014; Zarsky, 2011). Given this knowledge, “data mining provides its users with answers to questions they did not know to ask” (Zarsky, 2002, p.

6). This description exemplifies how the grouping of certain types of data can allow for the creation of new data entirely (Gates & Matthews, 2014), where before, without application of data mining algorithms, such linkages had been overlooked.

Strategies such as data brokerage and data mining are often used to circumvent the legal requirement of receiving user consent for the collection, analysis, and sharing of personal information, resulting in several legal and ethical issues on individual privacy that can affect citizen-consumer trust in corporations and institutions. One way to ascertain the protection of people's privacy, even when organisations collect data about them, is through the anonymisation of their data (Carvalho et al., 2020). Laws regarding data security become more lenient once data are anonymised (Telzrow & Kobsa, 2004), yet this does not always guarantee that the data are untraceable. On the contrary, the involvement of third parties can increase the risk of re-identification of internet users, rendering the solution of anonymisation moot (Büchi et al., 2019; Porter, 2008). Because third-party data brokers can collect personal information from multiple sources, they also become capable of recombining it, and thus matching separate parts of data regarding the identity of a single individual.

Moreover, anonymity is not equally effective in every context; on the one hand, it can be applied properly when data are gathered pertaining to citizen-consumers' activities in making online purchases or interacting with e-government. On the other hand, on social networks, where personal information may have already been shared to a more public audience – though not intended for an audience as large as the platform itself, or data brokers that purchase personal information from it – the security of personal information is lower because of this public sharing (Martin & Shilton, 2016). Collectors of data from social media argue that data that are shared on these platforms are already made public anyway. However, the role of this information changes once it is displaced for uses other than its original, and this can lead to issues on consent with the individuals whose data have been collected (Politou et al., 2018; Zimmer, 2010). Additionally, data shared on social media platforms are among the highest in confidentiality (Roeber et al., 2015), meaning that people are very unwilling to provide any information shared through these platforms to interested parties, compared to many other types of personal data.

The second issue that can arise because of third-party involvement, is the subsequent increased risk of fourth-party involvement (Conger et al., 2013). The threat to individual privacy can grow when, with the addition of data brokers third parties in the collection and analysis of

personal information, malicious actors are also given more options to obtain the data in less than legally sound ways. This is under the assumption that the inclusion of data brokerage is not already illegal, given how organisations must disclose it to their online customers or clients when their personal information is further being distributed (Caudill & Murphy, 2000). According to Tsesis (2014), setting up a time limit for third party data retention could already help decrease the risk of data theft from malicious fourth-party actors and data brokers alike.

For data mining, the accumulation of data from multiple different sources can, like how third parties might accidentally come to uncover it, lead to the re-identification of individuals online (Moorosi & Marivate, 2015). From a collection of different datasets, acquired from different actors, entirely new data can be created that can further profile the individuals whose personal data are subject to this process of analysis (Gates & Matthews, 2014). By that point, an issue arises that asks, to whom does this new data belong, for this individual could have provided data to different firms, deliberately keeping certain types of data apart, while never having been aware of, or having consented to, its use in the creation of new personal information that can be attributed to them. Someone could have chosen to share sensitive health information to their health insurance provider, while at the same time shopping for clothes online and providing the webstore information about larger clothes sizes and selectively purchasing cheaper products to wear. These two data points are then collected by data brokers and sold to a health agency, where conclusions could be drawn connecting health issues to the individual being overweight or having a low income.

Citizen-consumer online behaviour

The effects of the legal and ethical issues pertaining to the indirect strategies for data collection can be found in behavioural changes among citizen-consumers when interacting with organisations online. For the individual, the risk of a data breach on the part of large corporations and institutions means a careful assessment of what sensitive data one wants to share online, and to whom. This can serve as a problem however, since research has determined that the human factor is the weakest link in the assurance and maintenance of online data security, whether it be because of non-compliant behaviour on the part of the internet users – the citizen-consumers – themselves (Leering et al., 2022), or because of a general lack of knowledge in how certain aspects of data security technology work, paired with the consequent susceptibility to malicious

actors who intend to exploit such human error on part of either the consumer or webservice provider (Luo et al., 2011; West et al., 2009). Even so much as a corporate employee clicking on the wrong link in a spam email can already lead to a computer system being compromised and a subsequent data breach involving users' sensitive data. Cullen and Reilly (2008) have argued that it is not the involvement of malicious actors that can most direly affect citizen trust in e-government, but rather, the perceived incompetence of government employees and their interactions with digital technologies is to blame.

Herein, then, lies another dilemma. First, the institutional or corporate service provider wants to learn as much about the citizen-consumer as possible, as more data about the individual can mean a better understanding of their browsing patterns, and it can help improve users' experience of visiting a website, as well as the latter's functionality, leading consumers to stick around longer or visit more often (Kulyk et al., 2018). Data are for example collected using cookies, small files of text on which personal information is stored, which a web browser then remembers once a user revisits the website on which they were first activated. On social media, algorithms are used to increase user engagement on the platforms (Saura et al., 2020; Zakon, 2020), for example by recommending or forwarding user-generated content that is shocking or controversial and that sparks discussion among commenters. To do this, collecting personal information is a requirement, but users are still left powerless, knowing that they can still be recommended content algorithmically based on other people with similar browsing patterns as them, even without providing their own personal information (Sarwar et al., 2000).

At the same time, on the opposite side of the system, people are apprehensive about sharing personal information to the actors that rely so much on it to keep users engaged online and to offer them positive experiences using their services, in part because they do not trust of service providers that they will use their data for the things that they have explained to use it for (Beldad et al., 2011). Such behaviour changes however, once organisations make more transparent how and for what reason they are going to use this information. In essence, for users it is no longer important why they must share their data, as soon as corporations and institutions make efforts to become more transparent (Dinev & Hart, 2006; Leszczynski, 2015; Wu et al., 2012).

To prevent privacy violations, people adopt risk avoidance behaviours as they make use of the different webservices from corporations and public institutions alike. Such behaviours

were found to be driven by one's own conscious choice (Yao & Linz, 2008), in as much as people are themselves confident in controlling their own actions rationally when protecting their personal data online. These conclusions, built on Ajzen's (2011) theory of planned behaviour, need to be scrutinised however, knowing that, despite individuals' perceptions of their own behavioural control, these findings can turn out to be less accurate for observed behaviour (Armitage & Conner, 2001). To clarify, the way people perceive their own actions does not always match how they are seen to be acting by others. Human beings can instead be seen as irrational actors (Ajzen, 2011), whose actions – and thereby also the choices they make – though seen as rationally justified by the individual who undertakes them, can be influenced by stress factors such as the threat of data leakage from online organisational platforms (Boerman et al., 2021; Marett et al., 2011). In such scenarios too, it is their feelings of perceived threat, and their self-perceived confidence in protecting themselves that motivates them to handle their own data more securely (Boerman et al., 2021). If such a mindset is not present, people are instead more likely to give up on protecting sensitive information about themselves (Agozie & Kaya, 2021; Zhang et al., 2020).

Dinev and Hart (2006) argue that complete online privacy is practically unattainable, and that for this reason, people must be selective about what personal information they are willing to provide and to whom. The researchers based their findings on the so-called privacy calculus (Culnan & Armstrong, 1999; Dinev et al., 2006), that argues that a balance exists between factors that either inhibit or drive interactions and transactions between web users and web services. For example, when the trust that people feel toward an organisation and its online platform is higher than the personal concerns they have with their own online privacy, they will feel more inclined to fulfil a transaction. When the opposite is the case, such inclination is not present. Again however, this privacy calculus falsely assumes that people make decisions based on well thought out rationale (Moloney & Potì, 2013). The reality is that, though people can still critically reflect on what decision to make on privacy disclosure when the information that they are presented involves aspects more directly associated with the direct consequences of sharing personal information at hand – think of individual concerns about privacy or the potential rewards people can receive such as improved experience on online platforms or better website functionality – any information outside of this scope but still closely associated with privacy risk can lead to heuristic and irrational data protection decision-making (Wang et al., 2019).

Consequently, internet users can choose to adopt several disclosure avoidance strategies to minimise the amount of personal information they have to share with online organisations (Leering et al., 2022; White, 2004). First, they can choose to avoid using online transactions and interactions altogether and purchase their wares in physical locations, rather than virtual ones (Cho et al., 2009). Second, they can decide to opt-out of disclosing their personal information, an option that has in recent years become a requirement for online platforms to offer to their customers because of the Europe's GDPR, for example through cookie-management (Irwin, 2022; Kulyk et al., 2018). Third, users can be more proactive in protecting their data, for example by installing using privacy enhancing software on their computers or smartphones, like cookie-blockers. A fourth strategy that people can and do use involves the more malicious act of offering a platform false information about oneself, diverting organisations from one's actual personal data (Zhou et al., 2022).

Sometimes though, when regulations that have been put in place by data protection authorities turn out to be unhelpful in solving issues surrounding breaches of personal data, this can lead to feelings of discouragement among citizens (Agozie & Kaya, 2021; Zhang et al., 2020), and a consequent lack of intent from these data subjects to protect their data using any disclosure avoidance strategies. This problem can lead to internet users developing a mindset that no longer sees the selection of personal preferences as a useful tool to protect one's personal information, and thus they become more likely to just skip over these notices as they pop up on their screen (Kulyk et al., 2018). Accordingly, people no longer remain vigilant online about when and where to provide services with sensitive data about themselves. This weariness people experience when attempting to secure their personal information works counterproductive for the implementation of privacy policies that are supposed to further protect them from data breaches and the malicious collection of information by unethical data brokers or criminal hackers (Conger et al., 2013; Ziar et al., 2019). People who show more concern about their privacy being violated, then end up acting then act less in accordance with their intent of keeping their privacy secure. Such consumer behaviour can be regulated by the flow of information. How much they perceive this flow of information to be correct, can have an influence on their willingness to disclose personal information to interested parties. Contextual integrity is an adequate framework to approach this.

Contextual integrity

For web services created by different organisations, different public perceptions on privacy can apply. Whether an individual's privacy is perceived to be violated by these organisations is determined by what Nissenbaum (2004) has come to call "contextual integrity". It is upheld when no sensitive information is collected or analysed in a context where it should not have been. If this becomes the case however, contextual integrity is breached.

The author argued, through this framework, that for any one difference between situations in which privacy issues may arise, actors involved have a unique response in handling the privacy or security of data subjects and their personal information. With the introduction of the GDPR in Europe, Dutch citizen-consumers were presented with a situation involving more agency and autonomy on the monitoring and sharing of data, whereby the security of data subjects' personal information became their own responsibility, rather than an information resource controlled by public and private organisations (Strycharz et al., 2020; Zarsky, 2017), leading to novel responses on part of their privacy and data security (Presthus & Sørum, 2019).

Additionally, organisations were confronted with new regulations that relate to two types of informational norms explained in the contextual integrity framework (Nissenbaum, 2004). These are the two cornerstones that determine whether the contextual integrity of personal information is breached or not. First, there are the norms of appropriateness, that determine what types of information can be shared in what contexts. A patient can expect of their doctor to collect and use their health records to monitor their well-being. However, a patient would feel less inclined to share their social media activity with their doctor, as they would deem it less appropriate to disclose such personal information in the context of health care.

In the case of citizen-consumers accepting data disclosure online, Chin and colleagues (2012) found that a willingness to perform privacy-related tasks online was largely dependent on the kinds of applications that citizen-consumers were working with and their knowledge thereof, as they feared that the use of certain types of applications could negatively affect users' privacy and online security. Besides uncovering the most prevalent weakness of most, if not all data security systems, Leering and colleagues (2022) also revealed how certain situational factors can affect the behaviour of citizen-consumers when interacting with online webservices. One such factors asks just how sensitive the personal information is, that an online platforms ask for. This varies for each datatype, and it varies to everyone, as explained by the cultural differences on

data security behaviour pertaining to similar forms of personal information (Cho et al., 2009; LaBrie et al., 2017; Wu et al., 2012).

Second, there are the norms of distribution. Where norms of appropriateness ask *what* types of personal information are shared with interested parties, the norms of distribution determine *to whom* sending these data is appropriate. This accounts firstly for the distribution of data between the data subject and the organisation collecting it, but also between such organisations and third-parties, whose presence in the transactions is not always known to the individual to whose data is being distributed (Hoofnagle, 2003; Schroeder, 2016).

To resolve this issue, the GDPR has put in place a regulation that demands of organisations to only collect the minimum amount of data required for it to be able to achieve its outwardly expressed goal. This is called data minimisation (Zarsky, 2017). EU law has made it so that any goals organisations have that require the collection and use of personal information, are made clear to the individual whose data are requested, and that their consent has been given in return (Goddard, 2017; Kulyk et al., 2018). Moreover, no additional information is collected and analysed from the individuals that have provided it. Thus, given its norms of appropriateness and distribution, the framework of contextual integrity brings about the obligation of corporations and public institutions to only collect data that is relevant to the organisation at hand, and nothing more (Nissenbaum, 2004).

Roeber and colleagues (2015) have found significant contextual differences in public willingness to share depending on differing data types, organisations to whom these data are sent, methods of data gathering, and the organisations' reasons for gathering these data. Although the authors have uncovered these differences, their article does not further draw on how they have come to exist, or why data sharing preferences in these contexts are in that order specifically. Moreover, the date of the GDPR's implementation in the European Union in 2018 exceeds the date that Roeber and colleagues' (2015) article was written and published. Since then, the GDPR has addressed many privacy challenges in the EU, primarily on the importance of individual consent (Goddard, 2017), and the public has become more informed about, firstly, what their data are being used for (Kulyk et al., 2018), and secondly, what this could imply for their personal privacy when going online and how much they know about it (Presthus & Sørum, 2019). This is not to say that all challenges pertaining to online privacy have dissipated, however. For example, EU member states – including the Netherlands, the country whose

citizens this present paper focuses on – can still decide for themselves, within a certain, though unestablished scope, to what extent they apply the GDPR's regulations on their own citizens (Custers et al., 2018; Goddard, 2017), with laws regarding children's privacy and member states' different ages of consent being one such grey areas.

Privacy knowledge and concerns

The legal and ethical issues that arise because of corporate and institutional data collection, as well as the at times questionable behaviour of internet users and government or corporate workers alike (Luo et al., 2011; West et al., 2009), have given cause to a rise in privacy concerns among citizen-consumers (Bandyopadhyay, 2012). This worry is linked with the degree to which individuals are aware of the problems that can arise in the maintenance of data security online.

Bandyopadhyay (2012) found that the extent to which individuals felt concerned about their own privacy online could be affected by, firstly, to what extent they saw themselves capable of controlling the disclosure of their own private data, and secondly, how vulnerable they felt regarding risks of personal information disclosure they had no control over. Once online web services have collected data, they become able to determine themselves how this data is further processed. They can make third parties involved, which are interested in collecting data, at times illegally, with the purpose of learning more about internet users' browsing patterns, interests, and general online behaviour (Choi et al., 2018; Ziar et al., 2019). Such parties can then use these data to offer web users targeted advertisements on products that they might be interested in. One example of such a scandal involved Facebook and Cambridge Analytica (Confessore, 2018), where the social media platform collected personal data from users without their consent, and then sold them to Cambridge Analytica, who in turn used these data to run politically inclined advertisements during the 2016 US election campaigns.

People's knowledge in data security and online privacy can be measured by the extent to which they are internet literate and possess social awareness on digital privacy issues that organisations, internet users, and other stakeholders face when interacting with each other on the web (Bandyopadhyay, 2012; Dinev & Hart, 2005). Social awareness can be expanded upon using the works of Park (2011) and Trepte and colleagues (2015). The former listed three different dimensions – the latter group of authors expanding upon it with two more – which people could be informed on regarding digital privacy. Testing the first three dimensions of

familiarity with privacy technology, knowledge of organisational application of these technologies, and general understanding of policies pertaining to individual privacy, revealed that online privacy behaviour could be accurately predicted based on their privacy knowledge using these factors. With the introduction of the GDPR directive in the EU, and the subsequent improvement on public knowledge surrounding online privacy (Goddard, 2017), Trepte and colleagues' (2015) addition of two more dimensions on users' knowledge of European directives and user strategies on privacy control has also borne relevance in countries like the Netherlands. The argument for the use of different dimensions in measuring privacy knowledge was that it cannot be considered a unidimensional concept, but rather an umbrella term pertaining to the public ability to recognise different aspects of online data security and privacy.

Studies have shown that privacy knowledge can have an impact on potential data subjects' concerns pertaining to their personal information, as well as their confidence in solving these issues (Bandyopadhyay, 2012). This has been evidenced by the risk of privacy fatigue among internet users who felt that their efforts or the efforts of others to protect their private data ended up having no effect (Agozie & Kaya, 2021; Choi et al., 2018).

Worries can exist over how sensitive data are acquired and used, while people are at the same time unaware of the possibility that their personal information is already being used (Graeff & Harmon, 2002). A lack of transparency from institutions and corporation on their data collection and analysis methods can lead to an information imbalance, while providing users with knowledge can lower individuals' feelings of cynicism about data privacy (Agozie & Kaya, 2021). At the same time, the provision of information about data use policies on the part of organisations or privacy advocates must not be too selective or undetailed, as users' abilities to make properly informed decisions can be affected by what and how much knowledge they can obtain about it (Knijnenburg et al., 2017; Wang et al., 2019). Disclaimers on data collection that are not detailed enough will not help to resolve the information inequality between organisations and data subjects. Cookie disclaimers on websites showcase this by offering web users the option to save their privacy preferences and informing them on their ability to opt out (Kulyk et al., 2018), while upon visiting a website, these optional cookies can already automatically be turned on by the service provider. This means that the data subject user still needs to manually deselect these options before their actual personal preferences can be saved.

At the same time, making privacy policies too lengthy and complicated can also negatively impact public privacy knowledge, as individuals can decide to skip over the contents and end up not reading the terms and conditions in the disclaimers at all (Rao et al., 2016). An organisation's privacy regulations end up becoming less transparent because data subjects avoid opening the black box that describes the strategies online platforms adopt to collect personal data. It is a matter of public perception that determines how willing individuals are to disclose their personal information to webservices online. The more robust and transparent privacy regulations are, the more trust people are willing to put in the institutions that have implemented these policies (Nati, 2018; Wirtz et al., 2007). Based on this evidence, the following hypothesis is proposed:

H1: There is a negative relationship between citizens' trust in institutions and corporations, and their privacy concerns.

Organisational trust

In this research paper, the different organisations that exist as stakeholders in the privacy debate involve both public institutions, and private corporations. Both groups have different reasons for which they collect, analyse, and sometimes redistribute personal information from individuals who are active online, the former intent on improving citizens' security both online and in real life (Zarsky, 2011), and the latter aiming to gain an advantage over competitors in their trade (Prescott, 2014). Consequently, the premise of contextual integrity is made visible in a comparison between private and public organisations, and what different privacy expectations people have for each of them (Cullen & Reilly, 2008). For example, interaction with e-government is seen as being more reliable and data secure compared to e-commerce.

For both public and private organisations count, if people know more about the details of their privacy policies and regulations, and the organisations are transparent about how personal information is handled, citizens show less privacy concern (Carter & McBride, 2009; Wirtz et al., 2007; Zukowski & Brown, 2007). Additionally, organisations becoming transparent in how personal information is gathered and used can positively affect how much data subjects can learn, and thus know, about how to protect their own sensitive data (Morey et al., 2016). Companies and institutions being open about how internet users' data are collected and used can

better inform the individual about how secure their data is, and what they can possibly do to further secure it, for example by personalising their data sharing preferences through informed consent (Kulyk et al., 2018). Subsequently, by getting rid of the obscurity or complicatedness of data policies, citizen-consumers can become more trusting towards these organisations (Janic et al., 2013), and data subjects' confidence in their perceived knowledge on data privacy and security can increase. These findings exemplify the role of privacy knowledge as a moderating factor between organisational trust and privacy concern, leading to the creation of the second hypothesis.

> *H2: Individual privacy knowledge and one's confidence therein strengthens the negative relationship between organisational trust and privacy concern.*

Trust and concern in context

A recurring theme in privacy as contextual integrity is the divisive role that the cooperation between public and private organisations plays (Winter & Davidson, 2019). Preferably, social media platforms are not given access to an individual's personal health information. People would also prefer that the shops they order from online are not shared information about their criminal record if they have one. For each context, individuals have different concerns about what information is provided to what actor, for how long, and for which causes. Generally, citizens are more trusting towards public institutions when it comes to sharing personal information, in part because it is at times mandatory for people to provide these institutions with these data (Cullen & Reilly, 2008). For private organisations, the opposite is true, as such companies are allowed to ask for personal information of an individual, but said individual is not mandated to provide this to them. Based on these findings, a third hypothesis is formulated as follows:

> *H3a: Dutch citizen-consumers are more concerned when their personal information is shared with private organisations than they are when it is shared with public organisations.*

In the Dutch context, the role of the health care system is unique. In 2006, the Dutch government opted for a privatisation of its health insurance system, with the intent of increasing competition in health care (Toebes & San Giorgi, 2014). As Dutch healthcare organisations adapted to this new situation, such competition was found in the form of innovation, whereby institutional actors make use of resources from both public and private origins, to allow for health care arrangements that are more hybrid in their construction (Tuohy, 2012). The process of privatisation was found to have a downside as well however, as ten years on from its initiation, this change had negatively affected Dutch citizen-consumer trust in health insurance organisations (Maarse et al., 2016). Moreover, Appari and colleagues (2009) have revealed how in private healthcare facilities, patients' personal information can be more at risk than in public facilities, as professionals from private facilities are generally less confident in their abilities to protect sensitive data. Their behavioural intent to protect sensitive data played a significant role in this perception of self-efficacy. That the intent for taking active or proactive action in applying data security measures appeared lower in private healthcare facilities, can in turn also bring about increased risk of data breaches (Herath & Rao, 2009). Thus, the shift from health care institutions from public to private gives it an ambiguous role in the relationship between institutional trust and public privacy concern. Hence the following hypothesis:

> *H3b: Dutch citizen-consumers are more concerned when personal information is shared with hybridised institutions like health care than they are when it is shared with public institutions, though their concerns are lower when compared to other private organisations.*

As Zarsky (2002) pointed out, with some methods of data collection, data brokers and other actors can come to learn more about their data subjects than they initially expected or intended to learn. Despite efforts from internet users to anonymise their personal information, methods like data mining – where this collection of excess data is especially prevalent – can bear risk to their privacy, as multiple data sources can be linked with other records, as well as be combined, in turn creating new data relating to the individuals in question (Büchi et al., 2019; Buratović et al., 2012). Solutions have been on the rise ever since this issue became apparent, for example in the form of privacy preserving data mining through methods like k-anonymisation (Aldeen et al.,

2015), where a minimum amount of data subjects can be analysed at once, meaning that research on personal information can never be focused on, or associated to a single individual, or too small group of them (Buratović et al., 2012). Despite this, such solutions are not always implemented by organisations, because it can bear risk of financial loss in the reselling of the data (Fienberg, 2006; Porter, 2008). After all, the more data can be sold, the more profits can be made.

The reselling of personal information requires the involvement of third-party actors (Hoofnagle, 2003; Schroeder, 2016), and the risks associated with data mining and the linking of large amounts of personal information can lead to the collection and use of new data the creation of which the data subject has never consented to (Gates & Matthews, 2014). This knowledge could determine that regardless of what actor collects the data, there are significant contextual differences in privacy concerns among Dutch-citizen-consumers, based on how much personal information is collected. As such, the following hypothesis was formulated:

> *H4: Dutch citizen-consumers are more concerned for their privacy when more of their personal information is collected by public, hybrid, and private organisations.*

The linkage of different datasets to create new data requires the use of different types of information (Gates & Matthews, 2014), like how an individual's self-provided address, which is personal information, and a government-documented estimate of housing prices in that address's area can lead an agency that collects the data to create new personal information in the form of an estimate of the individual's monthly income. This risk illustrates why individuals can become wary about sharing their personal information online (Beldad et al. 2011), and it can explain why people are more willing to share certain types of data over others (Roeber et al. (2015). After all, combining the data on an individual's age with the housing price estimate in a specific area will not allow the agency to estimate one's monthly income, for the agency does not know whether the individual lives in that area or not, only how old they are.

As such, how worried individuals become about the provision of different types of data in different contexts, can vary as well. Martin and Shilton (2016) have already revealed that privacy expectations among mobile phone users were not met to an equal degree when their data was being tracked as compared to when they were shown targeted advertisements. They also found

that collecting key word information could meet privacy expectations in one scenario while not doing so in the other, revealing contextual differences. Moreover, where keyword information was deemed to have met such privacy expectations in the scenario of targeted advertising, for other types of information, like contact information or location data, mobile app users found that privacy expectations were not met (Martin & Shilton, 2016). Thus, the final hypothesis reads as follows:

> *H5: There are significant contextual differences in privacy concerns among Dutch citizen-consumers, depending on what type of personal data is collected.*

Methodology

Data collection

For this project, an existing dataset was used with data obtained by researchers from the Erasmus University Rotterdam, who created a survey in conjunction with researchers from the University of Maryland and the University of Wisconsin-Milwaukee. The researchers who collected the data received ethics approval from the ethics board at their University. In total, 510 respondents filled out the questionnaire from which the data were gathered. The requirements for filling out the questionnaire were that the respondent had to be at least eighteen years old, residing in the Netherlands. Finally, they had to be in possession of a smartphone and make use of social media platforms reasonably often, because of the role of social media apps and platforms in public privacy concerns as private organisations (Fatima et al., 2019). All survey responses were answered in May 2019, in the span of a single week, ranging from May the 8th until May the 15th. All questionnaire answers were processed in a way so that respondents remained anonymous.

The dataset from this survey was chosen for the proposed study because it made use of vignette questions. Vignettes are short lines of texts consisting of several elements describing hypothetical scenarios, that are replaceable and interchangeable with one another, all with the purpose of letting an individual answer a more overarching question using multiple variable instances of these vignettes (Atzmüller & Steiner, 2010; Hainmueller et al., 2015). To give an example from the dataset that this study intends to make use of, one such hypothetical scenarios read as follows: “your doctor collects data about the complete history of your physical activity (via your phone statistics), to determine how healthy you are, for the purpose of preventing the

spread of illnesses." This sentence shows an actor, the doctor, gathering a specific type of data for a specific purpose. In vignette studies, a large collection of such hypothetical scenarios is created, in this case each vignette consisting of a unique combination of actors, types of data, and aims for collecting these data.

The reason for undertaking this vignette study, lies in its improvements on both the internal validity that traditional survey research has little of, and the external validity that classical experimental designs lack (Atzmüller & Steiner, 2010). The low internal validity of traditional surveys gets removed because vignette surveys bear less risk of multicollinearity, given how the variables of each vignette are unique for all respondents. This creates less of a chance of explanatory variables overlapping with each other. Similarly, the low external validity of classical experiments disappears, because vignette studies, like surveys, possess a larger sample of respondents which it can draw its answers from. Besides that, because the sample size increases in vignette studies, its internal validity improves even further as it solves the risk of full multicollinearity in exceedingly small samples like what can sometimes be found in experiments (Voss, 2005).

The survey and dataset created for this research was based on Martin and Shilton's (2016) work on contextual factors affecting privacy judgements on tracking and advertising vignettes. In their research they analysed at once both the multifactorial vignettes as well as the variability accounting for individual differences in privacy judgements. After all, if everybody was the same as one another, they would all give the same privacy judgement for the same vignette. When presented with another vignette, their answers would all still be the same, albeit changed because of the new vignette, indicating contextual differences in privacy judgements. However, not everyone the same, and thus, everyone might have a different answer to the same vignette because of that. These differences also need to be accounted for in the research, and hence control variables are to be added in the analysis.

For this reason, a separate hierarchical regression analysis has been conducted to account for individual differences between respondents and the general population that they represent. In this analysis, the contextual differences accounted for in the multifactorial vignettes have not been included, for they are analysed separately using a repeated measures analysis of variance, also known as a one-way within-subjects ANOVA. Instead, several variables pertaining to individual organisational trust, privacy knowledge and privacy concerns have been added into

three different scales in order to uncover how these relate to each other without taking contextual integrity into account.

In total, 510 respondents filled out the questionnaire between May 8th and May 15th, 2019. Out of those respondents, only those who filled out all the answers relevant to the analyses in this research were included in the survey (N = 311), meaning at least one answer for each variable used in the repeated measures ANOVA, as well as all the answers in the scale variables used for the hierarchical regression, leaving a valid response rate of 60.98%. 155 of these respondents were male, and 156 were female, and they were on average 47 years old. 124 participants (39.87%) were high educated, 122 participants (39.23%) were middle educated, and 65 participants (20.90%) were low educated.

Operationalisation

Vignette design

For this research on contextual integrity in people's privacy concerns, the multifactorial vignettes were the most relevant variables to be included in the first analysis. While filling in the survey, respondents were presented with 32 vignettes out of an estimated total of 6*3*8*6*8 = 6912 unique vignette combinations (see appendix A), where they were asked to what extent they felt concerned for their privacy when a particular actor required a specific type of information for a certain amount of time, how they would come to use it, and for what purpose they would collect it. Respondents had the possibility of answering any of these questions, showcasing that the vignette experiment used a multifactorial design (Atzmüller & Steiiner, 2010). Since each respondent was asked to answer for only 32 of these hypothetical situations, and all respondents received a unique selection, the vignette experiment derived from the dataset also used a between-subjects design.

With each vignette, respondents were presented with two questions. The first question asked whether people found the collection and use of personal information in the hypothesised scenario justified or not, with the latter asking how much such collection and use of this data worried them. For this analysis, only the latter question was considered, the reasons being that the justifiability of data usage means approximately the opposite of the sorrowfulness thereof. To use only the data on sorrow, the data first had to be cleaned. In the dataset originally exported from Qualtrics – on which the university researchers had collected the survey responses – to

SPSS, all answers to the vignettes' questions were ordered in pairs, one below the other. Each pair of answers had to be manually separated from one another before only the latter answer, on individual worry, could be used in the analysis. By reversing the scores that respondents gave to either of these questions, they could end up meaning almost the same thing, rendering the inclusion of either one with the other moot. Moreover, with the focus of this research having been placed on addressing public privacy concerns in the Netherlands, the respondents' answers to the statement, "this use of my data worries me," was deemed most applicable in the conjoint analysis.

All vignettes' answers were scored on a five-point Likert scale, asking people to what extent they agreed with this statement on the individual's feelings of concern. A score of 1 indicated that they completely disagreed with the statement, and a score of 5 signified their complete agreement with it. Higher scores thus mean an increase in people's privacy concerns in the different contexts that they were presented with through these vignettes.

Privacy concern

For the hierarchical regression analysis, two independent variables were used to uncover citizen-consumers' individual differences in the relationship between organisational trust and public privacy concerns. The last-mentioned concept was the dependent variable in this analysis. It was measured using a mean scale variable relating to the worries that respondents had about mobile phone applications and the privacy risks associated with these applications in general. The scale consisted of nine items, each holding a statement about which the respondents were asked to what extent they agreed or disagreed with them. The results were scored on a five-point Likert scale, where a score of 1 meant the individual completely disagreed with the statement, and a score of 5 meant they completely agreed with it. An example of such statements, translated from Dutch, read as follows: "*it worries me that mobile phone applications can monitor the activities on my mobile phone*". The higher a respondent scored on the scale variable, the more they agreed to the statements, and the more concerned they were about their individual privacy.

To create the scale variable on privacy concern, an exploratory factor analysis and a reliability analysis were subsequently conducted on all nine relevant items. The first analysis revealed that all items loaded onto a single dimension (see table 1), with only one item having an eigenvalue above 1, KMO = .91, X^2 (N = 311, 36), 1471.26, $p < .001$. The second analysis

showed that the resulting scale ended up being very reliable (α = .89), and though the deletion of two items could have led to a higher Cronbach's alpha, the reliability for the scale at hand was already satisfactory for an accurate depiction of citizen-consumers' individual privacy concerns.

Table 1: factor and reliability analyses for the scale variable of privacy concern (N = 311)

Item	*Privacy concern*
It concerns me that mobile apps share my personal data with other parties without my permission.	.85
When I share personal data to use mobile apps, I am worried that apps use my information for other ends as well.	.84
It worries me that apps can monitor the activities on my mobile phone.	.82
My mobile app activity causes information that I deem private to become more available to other parties than I would want.	.82
It worries me that mobile apps collect too much information about me.	.80
I have a feeling that, due to my mobile app activity, others know more about me than I feel comfortable with.	.78
It worries me that mobile apps use my personal data for other ends without informing me or asking permission from me.	.76
Because of my mobile app activity, information has been gathered about me that could harm my privacy.	.46
I believe that my mobile phone's location is being monitored at least part of the time.	.44
R^2	.55
Cronbach's α	.89

Organisational trust

The first independent variable used in the hierarchical regression was a scale variable called relating to citizen-consumers' trust in organisations. Given how this variable intended to look at the individual differences in this trust, rather than the contextual differences as presented in the multifactorial vignettes, the items included in this mean scale referred to both private corporations as well as public institutions, this factor illustrating people's organisational trust in a

more general sense. As such, for each item, a respondent was asked to what extent they agreed to a given statement, an example being: "*I trust that the social media or messaging app that I use the most does what is best for the consumer*". A distinction was made between private companies from both the Netherlands and the United States, depending on where the survey was conducted. However, in the Dutch context, organisational trust in American companies still applies, given how many social media corporations, for one, originated from the US (Myers, 2014). Furthermore, with the GDPR regulation having been implemented in the European Union, and thus the Netherlands as well, any data collection from companies and other organisations from outside Europe are still required to abide by GDPR rules, as it involves the personal data from people living inside the EU (Koch, 2020; Wolford, 2020). Hence, the items for Dutch and US-based companies were both included in the initial factor analysis.

Items in the scale were scored in the same fashion as was the case for the dependent variable on privacy concern, with participants being asked to what extent, on a five-point Likert scale, they agreed to the statements they were presented. Each item focussed on a different organisation, the scores of which were computed in a mean scale to create a general assessment of the individual's organisational trust. A score of 1 indicated complete disagreement to a statement, while a score of 5 saw the participant completely agreeing, meaning that the higher one scored on the scale, the more trusting they were towards the organisations in question.

The factor analysis' rotated component matrix revealed one dimension onto which all private and public organisations loaded by a score of .7 or higher (see table 2), KMO = .87, X^2 (N = 311, 21), 1734.60, $p < .001$. Following that, the reliability analysis with all seven items included verified this, the scale scoring as very reliable ($\alpha = .92$), its reliability not increasing if any of the items were to be deleted.

Privacy knowledge and confidence

The third and final variable created for the hierarchical regression analysis was a mean scale pertaining to how confident people were in their own knowledge of handling their personal data, what it encompasses, and how data subjects can themselves control who receives what data. This variable, though still a mean scale, adopted three items that instead of a Likert scale, made use of a slide scale ranging from 1 to 100 to determine how certain they were in their knowledge about

these aspects of personal data collection. Confidence in the individual's perceived knowledge on how to control their own data was matched a higher score on the mean scale.

Table 2: factor and reliability analyses for the scale variable of organisational trust (N = 311)

Item	*Organisational trust*
I trust that the European government does what is right.	.87
I trust that the Dutch government does what is right.	.87
I trust that Dutch companies do what is best for the customer.	.87
I trust that my local government (incl. law enforcement) does what is right.	.85
I trust that American companies do what is best for the customer.	.77
I trust that the social media or messaging app that I use does what is best for the customer.	.76
I trust that news media report news in a just manner.	.70
R^2	.67
Cronbach's α	.92

Table 3: factor and reliability analyses for the scale variable of privacy knowledge (N = 311)

Item	*Privacy knowledge*
How sure are you of your general knowledge in the different types of data that your smartphone shares with mobile applications?	.90
How sure are you that you possess the knowledge required to protect your online privacy and safety?	.89
How sure are you of your ability to determine what and how your information is shared online?	.87
R^2	.79
Cronbach's α	.87

The validity and reliability of this variable was again verified using a factor analysis and a reliability analysis respectively (see table 3), KMO = .73, X^2 (311, 3), 450.39, $p < .001$, revealing a single dimension with an eigenvalue above 1 on which all three items could be

loaded, and the resulting scale being very reliable as well ($\alpha = .87$). The reliability of the scale did not increase in case an item was deleted. Finally, to make the variable of individuals' privacy knowledge into an interactive or moderating variable, the mean scale that was initially created using the three items measuring privacy knowledge was multiplied by the independent variable of organisational trust.

Control variables

In the hierarchical regression analysis, three control variables have been used to account for and lessen the influence of confounding variables that were not being tested in the hypotheses of this research (Spector & Brannick, 2011). The inclusion of age and education level were deemed important because of the findings by Zukowski & Brown (2007), who concluded that these two factors significantly affected public privacy concerns relating to their activity on the internet.

The first of these two control variables was measured using a single question that asked the respondents how old they were, in years, on the day that they were filling in the questionnaire. The participants answers were included in the analysis as a continuous variable. Age was included as a control variable because there are considerable generational differences in how individuals approach internet usage and how frequently they are active online (Cha, 2010). Elderly people are generally more hesitant to share personal data online, given their sensitivity to any issues that can negatively impact their personal privacy (Zukowski & Brown, 2007), like how older individuals have less control over the information they share online (Park, 2011), making them more susceptible to cybercrimes (Ziar et al., 2019).

The second control variable, education level, asked what level of education respondents last completed at the time they filled out the questionnaire. This demographic variable was chosen because of what Zukowski and Brown (2007) found on it, that the higher the education was that the individual had received, the less concerned they were about their data privacy. In the survey, respondents could select one out of eight options, which were in turn placed into three distinct categories. The first category, involving four options pertaining to the completion or only partial completion of secondary levels of education, high school for example, or anything below that, was defined as the respondent being "lower educated". The second category, involving the completion of higher levels of secondary education, like senior general secondary education and pre-university education, or the completion of secondary vocational education, was defined

"middle educated". The third and final category, involving the completion of higher education studies, such as higher professional education or university, or anything higher than that, was defined as "higher educated". These three categories were placed into a single ordinal variable, which was in turn made into two dummy variables, with "lower educated" as the reference category.

The third control variable pertains to the individual's frequency of mobile phone use. Though evidence is vague about whether smartphone use frequency affects privacy concern or the other way around (Brown & Muchira, 2004; Chen & Wen, 2019; O'Neil, 2001), research has repeatedly proven that there is a negative relationship between the two factors (Yao & Zhang., 2008). Similarly, a positive relationship was found between internet usage frequency and individual privacy self-efficacy, also known as one's knowledge in privacy and security-related issues (Yao et al., 2007), and with privacy knowledge hypothesised to strengthen the negative effect of organisational trust on privacy concern, internet usage frequency can indirectly be linked to this negative effect as well. Internet usage on smartphones comes into play more prevalently with the addition of social networking sites or platforms (Salehan & Negahban, 2013; Tsetsi & Rains, 2017). It can be seen as the driving factor for smartphone addiction, and with that, increased smartphone activity among users. In the survey, participants were asked how often they used their smartphone daily. The answers were given using a slide scale ranging from 1 to 100, with the frequency of smartphone use increasing in conjunction with higher scores on the slide scale.

Data analysis

To make a distinction between the individual and the contextual factors that can affect privacy concern among internet users, two separate analyses were used to measure each. A hierarchical regression analysis was applied to the individual factors, and a repeated measures ANOVA was used to account for the contextual factors.

The hierarchical regression analysis was measured in three steps, using the scale for privacy concern as the dependent variable in all of them. In the first step, only the direct relationship between organisational trust and privacy concern was tested. Only in the second step the moderating variable of privacy knowledge was added to see, first, whether the variance of the analysis increased with the addition of this new variable, and second, whether the inclusion of

the moderating variable led to a change in strength of the direct relationship between trust and concern. Finally, in the third step the control variables were added to the regression, to account for extraneous variables that were initially not included in the analysis or the research.

Before the second step of the hierarchical regression analysis could be run properly, the interaction variable was found to multicollinear (Where VIF > 8). For this reason, all variables involved in the creation of the interactor – the scales on privacy knowledge and organisational trust, as well as the moderating variable itself – had to be centred. After this the variables were found to no longer be multicollinear.

The repeated measures ANOVA was conducted using the vignette design of the survey. Initially, the intention was to make use of a conjoint analysis on the multifactorial vignettes using an orthogonal design, given how this method would have been able to determine what contextual traits – whether it be actor, data type, or otherwise – data subjects are least and most concerned about (Hainmueller et al., 2014). However the dataset was found to be incompatible with an orthogonal design, rendering the use of conjoint analysis on the vignette survey difficult. Instead, the choice was made to use one-way within-subjects ANOVA, given the method's ability to distinguish between factors scored using the same rating system. It can thus differentiate between the different actors in the "actor" group, and the same can be done for the factors on "data type" and "data amount".

For the repeated measures ANOVA, several new variables had to be created to account for the different items that were distinguished from each other in each factor. This meant the creation of new scale variables, each of which only included vignette responses pertaining to one given item. In the case of the "actor" factor, this meant six separate variables were computed for exclusive responses about either police, local government, one's doctor, the HR department in the company the respondent hypothetically works for, a social media platform, and data brokers. These variables can be further categorised, with the first two actors being public institutions, the doctor representing the hybrid organisation of health care in the Netherlands, and finally the last three actors being considered private corporations.

For each of these variables, the answers were given on the same Likert scale as was the case with the separate vignettes in the survey itself. Given how all vignettes already measured the same thing – to what extent the use of data in any given hypothetical context worried the individual – no factor or reliability analysis had to be conducted to create these new variables.

Using the repeated measures ANOVA, it was tested whether within a factor, there were significant differences between the items in the answers given to them about the extent to which the collection and use of certain data concerned the respondents. This method of analysis allowed for the ranking of different factor items, making it possible to answer the relevant hypotheses relating to contextual integrity. The factors for purpose and end goal were not analysed, seeing how for some actors do not intend to, or do not even have the right to, achieve certain types of goals. For example, one's doctor and a company's HR department are not authorised to combat crime and terrorism like police and government are. For the collection of different data types, this is different, as one's permission to collect these data entirely depends on whether the data subject consents to it, regardless of what actor may be involved (Goddard, 2017).

Results

Hierarchical regression analysis

In table 4, descriptive statistics are shown on all variables, predictors, criterium and control, that were used to analyse the individual factors that can account for privacy concern. To test hypotheses 1 and 2, a hierarchical regression analysis was conducted in three steps. The first model tested only the direct relationship between organisational trust and privacy concern, followed by the second model in which the scale and interactive variable of privacy knowledge were included to determine whether this direct relationship was strengthened by it and finalised with the inclusion of all three control variables – age, education level, and smartphone use frequency – in the third and final model.

The three models resulting from the hierarchical regression analysis are shown in table 5. Model 1 revealed a significant relationship between the predictor of organisational trust and the criterium of privacy concern, $F (1, 302) = 16.93$, $p < .001$, $R^2 = .05$. Organisational trust was found to negatively affect privacy concern among survey participants. ($\beta = -.23$, $p < .001$).

The second model, testing the interactive variable, was found to be significant as well, $F (3, 300) = 8.75$, $p < .001$, $R^2 = .08$. Moreover, by including the relevant variables to the model, the predictive value significantly improved as well, $\Delta R^2 = .03$, $F (2, 300) = 4.47$, $p = .01$. Organisational trust remained a significant negative predictor ($\beta = -.20$, $p < .001$), yet the interactive variable of privacy knowledge on the relationship between organisational trust and

privacy concern was found not significant, as well as negative (β = -.00 p = .960). The negative coefficient indicates that the interactor would have weakened, rather than strengthened the direct relationship, had it been significant. However, the second model also reveals a significant negative direct relationship between privacy knowledge and privacy concern (β = -.17, p = .003). This means that, though not useful as a moderator, higher privacy knowledge can still lead to a decrease in privacy concern among individuals, as found in earlier academic works (Bandyopadhyay, 2012; Trepte et al., 2015).

Table 4: descriptive results (N = 311)

Variables	Range	Mean (SD)
Dependent variable		
Privacy concern	4 (1 - 5)	3.88 (0.66)
Independent variables		
Organisational trust	4 (1 - 5)	2.85 (0.91)
Privacy knowledge	100 (0 - 100)	52.80 (20.01)
Control variables		
Age	48 (18 - 66)	47.00 (13.51)
Level of education (ref. low)		
High level of education	-	0.40 (0.49)
Medium level of education	-	0.39 (0.49)
Frequency of smartphone use	100 (0 - 100)	65.72 (20.90)

The third and final model, with the control variables included, is also significant, F (7, 296) = 5.19, p < .001, R^2 = .11. However, the predictive value no longer significantly increased like was the case between models 1 and 2, ΔR^2 = .03, F (4, 296) = 2.39, p = .051. The predictors for organisational trust (β = -.18, p = .001) and privacy knowledge (β = -.16, p = .01) remained significant and both relationships with the criterium privacy concern remained negative. The interactive variable of privacy knowledge also showed no change in significance or relationship direction (β = -.02, p = .77). It remained non-significant and negatively related to the direct relationship that was tested in the first model. The results in all three models revealed that, when

controlling for age, education level, and frequency of smartphone use, hypothesis 1 can be accepted, while hypothesis 2 must be rejected. Out of the three control variables that were included in the final model, only age was found to be a significant predictor of privacy concern. This is a positive relationship ($\beta = .17$, $p = .01$), where an increase in age indicates higher privacy concern among participants.

Table 5: regression models 1, 2 and 3 (N = 311)

	Model 1	**Model 2**	**Model 3**
	b (SE)	b (SE)	b (SE)
Dependent variable			
Privacy concern	3.876	3.877	3.245
	(0.037)	(0.037)	(0.233)
Independent variables			
Organisational trust	-0.165***	-0.146***	-0,132**
	(0,040)	(0,040)	(0,040)
Privacy knowledge		-0,006**	-0,005*
		(0,002)	(0,002)
Interaction between organisational trust and privacy knowledge		-0,000	-0,001
		(0,002)	(0,002)
Control variables			
Age			0,008**
			(0,003)
Opleidingsniveau (Hoog)			0.146
			(0,099)
Opleidingsniveau (Midden)			0,143
			(0,100)
Frequency of smartphone use			0,002
			(0,002)
	$R^2 = .05$***	$\Delta R^2 = .08$*	$\Delta R^2 = .11$
	$p < 0.001$	$p = .012$	$p = .051$

* p<0.05; ** p<0.01; *** p<0.001 (two-tailed tests).

Repeated measures ANOVA

Three hypotheses were tested using one-way within-subjects ANOVA to determine whether contextual factors played a role in privacy concern. Each factor was studied separately, meaning that three distinct analyses were performed using variables of the different actors, the amount of data gathered, and the types of data collected, that were included in the vignettes.

The first analysis revealed significantly different results on privacy concern for the actors included in the test, $F(4.27, 1322.39) = 19.76$, $p < .001$, partial $\eta 2 = .06$. The post-hoc pairwise comparisons verified this, as they revealed nine relationships between actors that were statistically significant. Out of those comparisons, six were found to be relevant for verifying hypotheses 3a and 3b. The actor on police was found to be significantly different from all other actors, including those of the hybrid and private organisations. Citizen-consumers were found to be less concerned about sharing their data to the police ($M = 3.86$, $SD = 1.01$) compared to government ($M = 3.99$, $SD = 1.11$), where $p = .01$, one's doctor ($M = 4.04$, $SD = 0.98$), where $p < .001$, the HR department in the company the individual hypothetically works for ($M = 4.19$, $SD = 1.07$), where $p < .001$, social media companies ($M = 3.98$, $SD = 1.07$), where $p = .01$, and data brokers ($M = 4.10$, $SD = 1.05$), where $p < .001$.

Moreover, citizen-consumers were less significantly concerned for their own privacy when sharing data to the government compared to HR, $p < .001$. These differences reveal at least partial support for hypothesis 3a, that argues that public institutions spark less privacy concern among citizen-consumers compared to hybrid and private organisations. The comparison of the hybrid health care actor with the private corporations revealed but one significant difference, between one's own doctor and one's own company's HR department, where about the former, people had less privacy concerns, $p = .03$. This relationship as well as that between the police and one's doctor, again reveal at least partial support for hypothesis 3b.

The police-government comparison revealed a new insight that within the group of public institutions there also exist significant differences in public privacy concern, and the same rule of thumb can be applied to private corporations as well, as privacy concern directed towards HR was significantly greater compared with social media platforms, where $p < .001$, and data brokers, where $p = .01$. Finally, privacy concerns were significantly lower for social media platforms compared to data brokers, $p = .002$.

The second ANOVA, concerning the amount of data collected by interested parties, revealed no significant differences between a week's worth and a year's worth of collected data, or even the complete history of the type of data collected, $F(2, 620) = 1.27$, $p = .28$, partial $\eta 2 = .04$. As such, hypothesis 4 must be rejected, and an increased amount of data collected does not lead increased privacy concerns among citizen-consumers.

The third and final repeated measures ANOVA did find significant difference in privacy concern between data types, $F(6.47, 2006.66) = 3.96$, $p < .001$, partial $\eta 2 = .01$. The post-hoc pairwise comparisons revealed only one such difference however, between the data types of social media messages ($M = 3.96$, $SD = 1.01$) and photo and video messages ($M = 4.08$, $SD = 1.04$), both of which scored lowest and highest in mean public privacy concern respectively in the vignette survey, $p = .02$. Despite the limited finding, this result still shows partial support for hypothesis 5, given how these two data types were significantly different from each other in the perceived privacy concern.

Conclusion & Discussion

This study addresses the role that organisational trust could have on privacy concern, while considering individual differences in citizen-consumers' privacy knowledge, as well as contextual differences in what interested parties are provided with what kind of personal information, and how much thereof. Dutch citizens were chosen as the research subjects, given the position of the Netherlands as a member state of the European Union (European Union, n.d.), as well as the unique role of Dutch health care, given the privatisation of hospitals and the Dutch health insurance system as recently as 2006 (Toebes & San Giorgi, 2014), making for a new, more hybridised organisational form to be compared with public institutions and private corporations. Using both a repeated measures ANOVA to analyse a vignette survey that provided this study with contextual differences in privacy concern, and a hierarchical regression analysis to take individual factors affecting privacy concern into account, this study sought to answer the question: *How are Dutch citizens' concerns about personal data usage affected by their perceptions of privacy and their trust in public and private organisations, following the implementation of the GDPR in the European Union?*

It turns out that both contextual integrity and individual traits were important in explaining the effects that factors like organisational trust and privacy knowledge had on privacy

concern in the Netherlands. The three repeated measures ANOVA applied to the vignette responses revealed that understanding the context is crucial in the face of public privacy concern, though not always in the form that was expected before the analyses were conducted. For example, though differences were found in privacy concern among Dutch citizen-consumers depending on what kind of organisational actor hypothetically requested and retrieved their personal information, not all private corporations sparked significantly more privacy concern over public institutions or hybridised health care. Despite this, all relationships between actors pointed towards the right predictions: public institutions caused the least amount of privacy concern, followed by the hybridised Dutch health care organisations, with the most public concern having been directed towards either social media platforms, data brokers, or one's HR department respectively.

At least the direction in which privacy concern rises, low for public institutions and high high for private organisations, has found support in the first hypotheses that focussed on contextual integrity in privacy. However, when it came to the amount of data being shared, or the type of data being requested, the differences in privacy concern that were found, were very few to none. Such results from the analyses contradict the findings made by Roeber and colleagues (2015), who did find significant differences in people's willingness to share their personal information between all data types they tested. For the contextual integrity framework, these findings mean that the norms of distributions (Nissenbaum, 2004), which ask with whom personal information is being shared, are seen to exist within Dutch society after the implementation of the GDPR in the EU. On the other hand, norms of appropriateness appear to be less present, as people feel less concerned over any differences in what or how much data is shared with interested organisations. Thus the contextual integrity framework was seen to affect privacy concern among Dutch citizens, this effect was almost exclusively limited to the public's perception of the organisations collecting their personal data.

Three things can explain this disparity. First, and most relevant to this study, Roeber and colleaguess' (2015) study was conducted prior to the GDPR's implementation in the European Union in 2018 (Goddard, 2017; Voss & Houser, 2019). Thus, it could not have considered the impact of that the regulation has had on public privacy concern in the EU's member states, including the Netherlands. The disparity between the present study and the one conducted by Roeber and colleagues (2015) could indicate that the implementation of the GDPR in the

European Union has led to a more even distribution in people's privacy concern. This can be the case because the regulation made people more informed about their privacy rights (Irwin, 2022; Kulyk et al., 2018), or they were more often confronted with it whenever they went online, being presented with the choice to adjust their privacy preferences on every website they visited. The GDPR demands that data subjects are given more control over how their personal information is collected and used.

Second, given that willingness to disclose and concern over the privacy of one's personal data are two different concepts, one can conclude that an individual can still be willing to share personal information to interested parties despite their concerns over it. Privacy fatigue can play a role in this as well (Agozie & Kaya, 2021; Choi et al., 2018). Though privacy concern can be low among citizen-consumers, in case of privacy fatigue they would still provide their own sensitive data to interested parties, because they lack faith in the effectiveness of the national laws and regulations that prevent organisations from collecting data for prohibited uses (Conger et al., 2013). Alternatively, despite their concern over the data they provide to organisational actors, people find that the benefits of data disclosure outweigh the potential pitfalls that they are worried about.

Third, the data types that were analysed in the present study and the research conducted by Roeber and colleagues (2015) do not match one another. Thus, where the different types of data were found to cause significant differences in people's willingness to provide their data to organisations, the data types that were mentioned in the vignettes had no such relationships. However, location data and social media messages were two types of data that were included in both studies, and the possibility of this third explanation is made moot when finding that the relationship was significantly different in the first study, while not significant in the second.

This study has tested and verified that the causes for privacy concern are not exclusively context-based. Individual factors, such as organisational trust and age, can influence privacy concern as well, and subsequently lead to variation in privacy concern among the Dutch populace. People who are more trusting in organisations, both private and public, tend to feel less concerned about their privacy than less trusting individuals. At the same time, older people tend to express higher concern than younger people, matching the findings made by Cha (2007) and Zukowski and Brown (2010).

On the other hand, where previous research found support for the effect of control variables like education level and smartphone use frequency on privacy concern, these conclusions were not matched in this study. The introduction of the GDPR could explain why education level no longer controlled for privacy concern, as the regulation has led to a more widespread knowledge on privacy rights and data security among internet users, indiscriminate of their level of education (Kulyk et al., 2018; Presthus & Sørum, 2019). The non-significance of smartphone use frequency can be explained by the degree to which internet activity has risen among citizen-consumers in recent years.

Though one's privacy knowledge, that is, how much people know about how to control and secure their personal data, was initially thought to strengthen the negative relationship between organisational trust and privacy concern, the results revealed that no such thing was the case. The relationship between privacy knowledge and concern is more direct than that (Bandyopadhyay, 2012), and knowledge or literacy played no interactive role with regards to the first tested relationship. This finding indicates that the things that individuals learn about rights to privacy and data security are not the consequence of the GDPR's demand that organisations become more transparent about how and why they collect internet users' personal information (Goddard, 2017; Morey et al., 2016). Alternatives are that data subjects are taught these things by other actors, like advocates for individual privacy (Agozie & Kaya, 2021), or they take initiative by reading the organisational requirements of data security and transparency from the GDPR itself.

The use of repeated measures ANOVA as a research method to account for contextual integrity in the vignette study has brought about some concern for the study's validity and reliability. The vignette survey was originally created using a between-subjects design, where each respondent is given different vignette questions to fill out (Atzmüller & Steiner, 2010). However, an issue with the creation of the dataset led to a problem where the vignettes could not be analysed in in the way they were intended, for example through conjoint analysis using an orthogonal design. Instead, different scale variables were created, each referring to a different item in the three factors to which they belonged, but a multifactorial design of the vignette survey made it not possible to check for these variables' validity and reliability and were instead at once included in the one-way within-subjects ANOVA that this study used. This is recognised as a considerable limitation in the research process. Another limitation can be found in the fact

that, despite noting some possibilities explaining the differences in the results on privacy concern between this study and Roeber and colleagues' (2015), no definitive conclusions can be drawn about these differences when these data have not yet been compared to each other.

For the multiple regression analysis, a limitation was found in the use of the control variables, and the control for smartphone use frequency in particular. This variable was found not to predict privacy concern despite prior evidence that it could lead to increased concerns among internet users (Yao & Zhang, 2008). The reason for this could lie in the way the control variable was constructed prior to the distribution of the survey. The slide scale described in the operationalisation allows the survey participant to determine for themselves to what extent they are either very infrequently, or very frequently, using their smartphones, though this estimate is purely subjective based on how active these users perceive themselves to be. Two individuals who use their smartphones for an equal amount of time on a daily or weekly basis could provide the researchers with two vastly different answers on the scale, depending on how active they deem themselves to be, compared to other users. A suggestion for future research to solve this problem, would be to use answer categories with time-based estimates, for example through asking the survey participant how many hours per day or per week on average they used their smartphone.

A proposal to solve the issues with the inconclusive comparison between the current study and Roeber and colleagues' (2015) work, would be to analyse vignette surveys on privacy concern that were created and distributed among EU residents prior to the implementation of the GDPR in 2018, in the form of a longitudinal study, comparing pre-GDPR European privacy concern to its post-GDPR counterpart. However, no such vignette studies were conducted prior to 2018, making such a task in possible. Hence, the next-best solution for this problem would be to conduct a comparative study, where the privacy concerns of EU citizen-consumers subject to the GDPR regulations are compared to those of citizen-consumers in countries or regions where the GDPR's rules do not apply, and residents are thus not under its protection. The right to be forgotten that applies to member states of the European Union, for one, does not apply to US residents, where the rights of the companies collecting the data are prioritised (Myers, 2014; Tsesis, 2014), allowing for an interesting comparison in privacy concerns between nations with and without the GDPR privacy laws.

References

Adomavicius, G., & Tuzhilin, A. (2001). Using data mining methods to build customer profiles. *Computer, 34*(2), 74-82. https://doi.org/10.1109/2.901170

Agozie, D. Q., & Kaya, T. (2021). Discerning the effect of privacy information transparency on privacy fatigue in e-government. *Government Information Quarterly, 38*(4). https://doi.org/10.1016/j.giq.2021.101601

Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology & Health, 26*(9), 1113-1127. https://doi.org/10.1080/08870446.2011.613995

Aldeen, Y. A. A. S., Salleh, M., & Razzaque, M. A. (2015). A comprehensive review on privacy preserving data mining. *SpringerPlus, 4*(694). https://doi.org/10.1186/s40064-015-1481-x

Allen, M. (2012). What was Web 2.0? Versions as the dominant mode of internet history. *New Media & Society, 15*(2), 260-275. https://doi.org/10.1177/1461444812451567

Altman, I. (1975). The environment and social behavior. Monterey, CA: Brooks/Cole.

Appari, A., Johnson, M. E., & Anthony, D. L. (2009). HIPAA compliance: An institutional theory perspective. *Proceedings of the American Conference on Information Systems (AMCIS).*

Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology, 40*, 471-499.

Atzmüller, C., & Steiner, P. M. (2010). Experimental vignette studies in survey research. *Methodology, 6*(3), 128-138. https://doi.org/10.1027/1614-2241/a000014

Bandyopadhyay, S. (2012). Consumers' online privacy concerns: Causes and effects. *Innovative Marketing, 8*(3), 32-39.

Beldad, A., De Jong, M., & Steehouder, M. (2011). I trust not therefore it must be risky: Determinants of the perceived risks of disclosing personal data for e-government transactions. *Computers in Human Behavior, 27*(6), 2233-2242. https://doi.org/10.1016/j.chb.2011.07.002

Boerman, S. C., Kruikemeijer, S., & Borgesius, F. J. Z. (2021). Exploring motivations for online privacy protection behavior: Insights from panel data. *Communication Research, 48*(7), 953-977. https://doi.org/10.1177/0093650218800915

Brown, M., & Muchira, R. (2004). Investigating the relationship between internet privacy concerns and online purchase behavior. *Journal of Electronic Commerce Research, 5*(1), 62-70.

Büchi, M., Fosch-Villaronga, E., Lutz, C., Tamò-Larrieux, A., Velidi, S., & Viljoen, S. (2019). The chilling effects of algorithmic profiling: Mapping the issues. *Computer Law & Security Review, 36*. https://doi.org/10.1016/j.clsr.2019.105367

Buratović, I., Miličević, M., & Žubrinić, K. (2012). Effects of data anonymization on the data mining results. In *2012 Proceedings of the 35th International Convention MIPRO* (pp. 1965-1969). IEEE.

Carter, L., & McBride, A. (2009). Information privacy concerns and e-government: A research agenda. *Transforming Government: People, Process and Policy, 4*(1), 10-13. https://doi.org/10.1108/17506161011028777

Carvalho, A. P., Canedo, E. D., Carvalho, F. P., & Carvalho, P. H. P. (2020). Big data, anonymisation and governance to personal data protection. *The 21st Annual International Conference on Digital Government Research*, 185-195. https://doi.org/10.1145/3396956.3398253

Caudill, E. M., & Murphy, P. E. (2000). Consumer online privacy: Legal and ethical issues. *Journal of Public Policy and Marketing, 19*(1), 7-19.

Cha, J. (2010). Factors affecting the frequency and amount of social networking site use: Motivations, perceptions, and privacy concerns. *First Monday, 15*(12). https://doi.org/10.5210/fm.v15i12.2889

Chen, Y. K., & Wen, C. R. (2019). Taiwanese university students' smartphone use and the privacy paradox. *Communicar, 60*(27), 61-69. https://doi.org/10.3916/C60-2019-06

Chin, E., Felt, A. P., Sekar, V., & Wagner, D. (2012). Measuring user confidence in smartphone security and privacy. *Proceedings of the Eighth Symposium on Usable Privacy and Security, 1*, 1-16. https://doi.org/10.1145/2335356.2335358

Cho, H., Rivera-Sánchez, M., & Lin, S. S. (2009). A multinational study on online privacy: Global concerns and local responses. *New Media & Society, 11*(3), 395-416. https://doi.org/10.1177/1461444808101618

Choi, H., Park, J., & Jung, Y. (2018). The role of privacy fatigue in online consumer behavior. *Computers in Human Behavior, 81*, 42-51. https://doi.org/10.1016/j.chb.2017.12.001

Confessore, N. (2018, April 4). *Cambridge analytica and Facebook: The scandal and the fallout so far*. The New York Times. https://www-nytimes-com.eur.idm.oclc.org/2018/04/04/us/politics/cambridge-analytica-scandal-fallout.html

Conger, S., Pratt, J. H., & Loch, K. D. (2013). Personal information privacy and emerging technologies. *Info Systems Journal, 23*, 401-417. https://doi.org/10.1111/j.1365-2575.2012.00402.x

Cullen, R., & Reilly, P. (2008). Information privacy and trust in government: A citizen-based perspective from New-Zealand. *Journal of Information Technology & Politics, 4*(3), 61-80. https://doi.org/10.1080/19331680801915066

Culnan, M., J., & Armstrong, P. K. (1999). Information privacy concerns, procedural fairness, and impersonal trust: An empirical investigation. *Organization Science, 10*(1), 104-115. https://doi.org/10.1287/orsc.10.1.104

Custers, B., Dechesne, F., Sears, A. M., Tani, T., & Van der Hof, S. (2018). A comparison of data protection legislation and policies across the EU. *Computer Law & Security Review, 34*(2), 234-243. https://doi.org/10.1016/j.clsr.2017.09.001

Dawes, S. (2014). Press freedom, privacy, and the public sphere. *Journalism Studies, 15*(1), 17-32. https://doi.org/10.1080/1461670X.2013.765637

Dinev, T., Bellotto, M., Hart, P., Russo, V., Serra, I., & Colautti, C. (2006). Privacy calculus model in e-commerce – A study of Italy and the United States. *European Journal of Information Systems, 15*(4), 389-402. https://doi.org/10.1057/palgrave.ejis.3000590

Dinev, T., & Hart, P. (2005). Internet privacy concerns and social awareness as determinants of intention to transact. *International Journal of Electronic Commerce, 10*(2), 7-29. https://doi.org/10.2753/JEC1086-4415100201

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information Systems Research, 17*(1), 61-80. https://doi.org/10.1287/isre.1060.0080

European Union. (n.d.). *Netherlands*. Retrieved May 31, 2022, from https://european-union.europa.eu/principles-countries-history/country-profiles/netherlands_en

Fatima, R., Yasin, A., Liu, L., Wang, J., Afzal, W., & Yasin, A. (2019). Sharing information online rationally: An observation of user privacy concerns and awareness using serious game. *Journal of Information Security and Applications, 48*. https://doi.org/10.1016/j.jisa.2019.06.007

Fienberg, S. E. (2006). Privacy and confidentiality in an e-commerce world: Data mining, data warehousing, matching, and disclosure limitation. *Statistical science, 21*(2), 143-154. https://doi.org/10.1214/088342306000000240

Gates, C., & Matthews, P. (2014). Data is the new currency. *Proceedings of the 2014 New Security Paradigms Workshop*, 105-116. https://doi.org/10.1145/2683467.2683477

Gauch, S., Speretta, M., Chandramouli, A., & Micarelli, A. (2007). User profiles for personalized information access. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The adaptive web: Methods and strategies of web personalization* (pp. 54-89). Springer.

Ghioshray, S. (2006). Untangling the legal paradigm of indefinite detention: Security, liberty and false dichotomy in the aftermath of 9/11. *St. Thomas Law Review, 19*(2), 249-280.

Goddard, M. (2017). The EU General Data Protection Regulation (GDPR): European regulation that has a global impact. *International Journal of Market Research, 59*(6), 703-705. https://doi.org/10.2501/IJMR-2017-050

Graeff, T. R., & Harmon, S. (2002). Collecting and using personal data: Consumers' awareness and concerns. *Journal of Consumer Marketing, 19*(4), 302-318. https://doi.org/10.1108/07363760210433627

Grimmelmann, J. (2020). Law and technology: A recent renaissance in privacy law. *Communications of the ACM, 63*(9), 24-27. https://doi.org/10.1145/3411049

Hainmueller, J., Hangartner, D., & Yamamoto, T. (2015). Validating vignette and conjoint survey experiments against real-world behavior. *PNAS, 112*(8), 2395-2400. https://doi.org/10.1073/pnas.1416587112

Hainmueller, J., Hopkins, D. J., & Yamamoto, T. (2014). Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments. *Political Analysis, 22*, 1-30. https://doi.org/10.1093/pan/mpt024

Halstuk, M. E., & Chamberlin, B. F. (2006). The freedom of information act 1966-2006: A retrospective on the rise of privacy protection over the public interest in knowing what the government's up to. *Communication Law and Policy, 11*(4), 511-564. https://doi.org/10.1207/s15326926clp1104_3

Herath, T., & Rao, H. R. (2009). Protection motivation and deterrence: A framework for security policy compliance in organisations. *European Journal of Information Systems, 18*, 106-125. https://doi.org/10.1057/ejis.2009.6

Hoofnagle, C. J. (2003). Big Brother's little helpers: How ChoicePoint and other commercial data brokers collect and package your data for law enforcement. *North Carolina Journal of International Law and Commercial Regulation, 29*(4), 595-638.

Irwin, L. (2022, April 12). *How the GDPR affects cookie policies*. IT Governance. https://www.itgovernance.eu/blog/en/how-the-gdpr-affects-cookie-policies

Janic, M., Wijbenga, J. P., & Veugen, T. (2013). Transparency enhancing tools (TETs): An overview. In *2013 Third Workshop on Socio-Technical Aspects in Security and* Trust, (pp. 18-25). IEEE. https://doi.org/10.1109/STAST.2013.11

Janssen, M., & Van den Hoven, J. (2015). Big and Open Linked Data (BOLD) in government: A challenge to transparency and privacy? *Government Information Quarterly, 32*(4), 363-368. https://doi.org/10.1016/j.giq.2015.11.007

Knijnenburg, B. P., Raybourn, E., Cherry, D., Wilkinson, D., Sivakumar, S., & Sloan, H. (2017). Death to the privacy calculus? https://doi.org/10.2139/ssrn.2923806

Koch, R. (2020). *What is considered personal data under the EU GDPR?* GDPR. https://gdpr.eu/eu-gdpr-personal-data/

Kulyk, O., Hilt, A., Gerber, N., & Volkamer, M. (2018). "This website uses Cookies": Users' perceptions and reactions to the Cookie disclaimer. *3rd European Workshop on Usable Security.* https://doi.org/10.14722/eurousec.2018.23012

Kumar, A., & Kumar, P. (2010). Managing privacy of user generated information in a Web 2.0 world. *Journal of Information Privacy & Security, 6*(2), 3-16. https://doi.org/10.1080/15536548.2010.10855885

LaBrie, R. C., Steinke, G. H., Li, X., & Cazier, J. A. (2017). Big data analytics sentiment: US-China reaction to data collection by business and government. *Technological Forecasting & Social Change, 130,* 45-55. https://doi.org/10.1016/j.techfore.2017.06.029

Leering, A., Van den Wijngaert, L., & Nikou, S. (2022). More honour'd in the breach: Predicting noncompliant behaviour through individual, situational and habitual factors. *Behaviour & Information Technology, 41*(3), 519-534. https://doi.org/10.1080/0144929X.2020.1822444

Leszczynski, A. (2015). Spatial big data and anxieties of control. *Environment and Planning D: Society and Space, 33*(6), 965-984. https://doi.org/10.1177/0263775815595814

Luo, X. R., Brody, R., Seazzu, A., & Burd, S. (2011). Social engineering: The neglected human factor for information security management. *Information Resources Management Journal, 24*(3), 1-8. https://doi.org/10.4018/irmj.2011070101

Maarse, H., Jeurissen, P., & Ruwaard, D. (2016). Results of the market-oriented reform in the Netherlands: A review. *Health Economics, Policy and Law, 11*(2), 161-178. 10.1017/S1744133115000353

Marett, K., McNab, A. L., & Harris, R. B. (2011). Social networking websites and posting personal information: An evaluation of protection motivation theory. *AIS Transaction on Human-Computer Interaction, 3*(3), 170-188. https://aisel.aisnet.org/thci/vol3/iss3/2

Margulis, S. T. (2003). On the status and contribution of Westin's and Altman's theories of privacy. *Journal of Social Issues, 59*(2), 411-429.

Martin, K., Shilton, K. (2016). Putting mobile application privacy in context: An empirical study of user privacy expectations for mobile devices, *The Information Society, 32*(3), 200-216. https://doi.org/10.1080/01972243.2016.1153012

Miller, K. (2014). Total surveillance, big data, and predictive crime technology: Privacy's perfect storm. *Journal of Technology Law & Policy, 19*(1), 105-146.

Moloney, M., & Potì, V. (2013). A behavioural perspective on the privacy calculus model. https://doi.org/10.2139/ssrn.2310535

Moorosi, N., & Marivate, V. (2015). Privacy in mining crime data from social media: A South African perspective. *2015 Second International Conference on Information Security and Cyber Forensics*, 171-175. https://doi.org/10.1109/InfoSec.2015.7435524

Morey, T., Forbath, T., & Schoop, A. (2016). Customer data: Designing for tranparency and trust. *Harvard Business Review, 93*(5), 96-105.

Myers, C. (2014). Digital immortality vs. "the right to be forgotten": A comparison of U.S. and E.U. laws concerning social media privacy. *Romanian Journal of Communication and Public Relations, 16*(3), 47-60. https://doi.org/10.21018/rjcpr.2014.3.175

Nati, M. (2018). *Personal data receipts: How transparency increases consumer trust* [White paper]. Catapult Digital. https://www.digicatapult.org.uk/wp-content/uploads/2021/11/Personal_Data_Receipts_r1.5_2.pdf

Netflix. (n.d.). *How Netflix's recommendation system works.* Retrieved June 19, 2022, from https://help.netflix.com/en/node/100639

Nissenbaum, H. (2004). Privacy as contextual integrity. *Washington Law Review, 79*(1), 119-158.

Nocera, J. (2022). How Cookie Banners Backfired. *The New York Times*. https://www.nytimes.com/2022/01/29/business/dealbook/how-cookie-banners-backfired.html

O'Neil, D. (2001). Analysis of internet users' level of online privacy concerns. *Social Science Computer Review, 19*(1), 17-31.

Park, Y. J. (2011). Digital literacy and privacy behavior online. *Communication Research, 40*(2), 215-236. https://doi.org/10.1177/0093650211418338

Politou, E., Alepis, E., & Patsakis, C. (2018). Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. *Journal of Cybersecurity*, 1-20. https://doi.org/10.1093/cybsec/tyy001

Porter, C. C. (2008). De-identified data and third party data mining: the risk of re-identification of personal information. *Shidler Journal of Law, Commerce & Technology, 5*(1), 1-8.

Prescott, M. E. (2014). Big data and competitive advantage at Nielsen. *Management Decision, 52*(3), 573-601. https://doi.org/10.1108/MD-09-2013-0437

Presthus, W., & Sørum, H. (2019). Consumer perspectives on information privacy following the implementation of the GDPR. *International Journal of Information Systems and Project Management, 7*(3), 19-34. https://doi.org/10.12821/ijispm070302

Rao, A., Schaub, F., Sadeh, N., & Acquisti, A. (2016). Expecting the unexpected: Understanding mismatched privacy expectations online. *Proceedings of the Twelfth Symposium on Usable Privacy and Security*, 77-96.

Roeber, B., Rehse, O., Knorrek, R., & Thomsen, B. (2015). Personal data: How context shapes consumers' data sharing with organizations from various sectors. *Electronic Markets, 25*, 95-108. https://doi.org/10.1007/s12525-015-0183-0

Rogge, N., Agasisti, T., & De Witte, K. (2017). Big data and the measurement of public organizations' performance and efficiency: The state-of-the-art. *Public Policy and Administration, 32*(4), 263-281. https://doi.org/10.1177/0952076716687355

Salehan, M., & Negahban, A. (2013). Social networking on smartphones: When mobile phones become addictive. *Computers in Human Behavior, 29*(6), 2632-2639. https://doi.org/10.1016/j.chb.2013.07.003

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000). Analysis of recommendation algorithms for e-commerce. *Proceedings of the 2nd ACM Conference on Electronic Commerce*, 158-167. https://doi.org/10.1145/352871.352887

Saura, J. R., Palacios-Marqués, D., & Itturicha-Fernández, A. (2020). Ethical design in social media: Assessing the main performance measurements of user online behavior modification. *Journal of Business Research, 129*, 271-281. https://doi.org/10.1016/j.jbusres.2021.03.001

Schroeder, R. (2016). Big data business models: Challenges and opportunities. *Cogent Social Sciences, 2*(1). https://doi.org/10.1080/23311886.2016.1166924

Solove, D. J. (2016). A brief history of information privacy law. In K. J. Mathews (Ed.), *Proskauer on privacy: A guide to privacy and data security law in the information age.* (2nd ed., pp. 1-52). Practicing Law Institute.

Spector, P. E., & Brannick, M. T. (2011). Methodological urban legends: The misuse of statistical control variables. *Organizational Research Methods, 14*(2), 287-305. https://doi.org/10.1177/1094428110369842

Steeves, V. (2009). Reclaiming the social value of privacy. In V. Steeves, C. Lucock, & I, Keer (Eds.), *Privacy, identity and anonymity in a network world: Lessons from the identity trail* (pp. 191-208). Oxford University Press.

Strycharz, J., Ausloos, J., & Helberger, N. (2020). Data protection or data frustration? individual perceptions and attitudes towards the GDPR. *European Data Protection Law Review (EDPL), 6*(3), 407-421.

Telzrow, M., & Kobsa, A. (2004). Impacts of user privacy preferences on personalized systems. In C. Karat, J. O. Blom, & J. Karat (Eds.), *Designing personalized user experiences in e-commerce* (pp. 315-322). Kluwer Academic Publishers.

Toebes, B., & San Giorgi, M. (2014). Dutch realities: Evaluating healthcare reform in the Netherlands from a human rights perspective. In B. Toebes, R. Ferguson, M. M. Markovic, & O. Nnamuchi (Eds.), *The right to health: A multi-country study of law, policy and practice* (pp. 403-436). https://doi.org/10.1007/978-94-6265-014-5_14

Trepte, S., Teutsch, D., Masur, P. K., Eicher, C., Fischer, M., Hennhöfer, A., & Lind, F. (2015). Do people know about privacy and data protection strategies? Towards the "online privacy literacy scale" (OPLIS). In S. Gutwirth, R. Leenes, & P. De Hert (Eds.), *Reforming European Data Protection Law* (pp. 333-365). https://doi.org/10.1007/978-94-017-9385-8_14

Tsesis, A. (2014). The right to be forgotten and erasure: Privacy, data brokers, and the indefinite retention of data. *Wake Forest Law Review, 48*, 101-151.

Tsetsi, E., & Rains, S. A. (2017). Smartphone internet access and use: Extending the digital divide and usage gap. *Mobile Media & Communication, 5*(3), 239-255. https://doi.org/10.1177/2050157917708329

Tuohy, C. H. (2012). Reform and the politics of hybridization in mature health care states. *Journal of Health Politics, Policy and Law*, *37*(4), 611-632. https://doi.org/10.1215/03616878-1597448

Vesanen, J., & Raulas, M. (2006). Building bridges for personalization: A process model for marketing. *Journal of Interactive Marketing, 20*(1). https://doi.org/10.1002/dir.20052

Voss, D. S. (2005). Multicollinearity. *Encyclopedia of Social Measurement, 2*, 759-770.

Voss, W. G., & Houser, K. A. (2019). Personal data and the GDPR: Providing a competitive advantage for U.S. companies. *American Business Law Journal, 56*(2), 287-344. https://doi.org/10.1111/ablj.12120

Wang, L., Hu, H., Mei, M. Q., & Yan, J. (2019). Privacy calculus or heuristic cues? The dual process of privacy decision making on Chinese social media. *Journal of Enterprise Information Management*, *33*(2), 353-380. https://doi.org/10.1108/JEIM-05-2019-0121

West, R., Mayhorn, C., Hardee, J., & Mendel, J. (2009). The weakest link: A psychological perspective on why users make poor security decisions. In M. Gupta, & R. Sharman (Eds.), *Social and human elements of information security: Emerging trends and countermeasures* (pp. 43-60). IGI Global. https://doi.org/10.4018/978-1-60566-036-3.ch004

Westin, A. (1967). Privacy and freedom. New York: Athenaeum.

White, T. B. (2004). Consumer disclosure and disclosure avoidance: A motivational framework. *Journal of Consumer Psychology, 14*(1&2), 41-51. https://doi.org/10.1207/s15327663jcp1401&2_6

Winter, J. S., & Davidson, E. (2019). Big data governance of personal health information and challenges to contextual integrity. *The Information Society, 35*(1), 36-51. https://doi.org/10.1080/01972243.2018.1542648

Wirtz, J., Lwin, M. O., & Williams, J. D. (2007). Causes and consequences of consumer online privacy concern. *International Journal of Service Industry Management, 18*(4), 326-348. https://doi.org/10.1108/09564230710778128

Wolford, B. (2020). *Does the GDPR apply to companies outside of the EU?* GDPR. https://gdpr.eu/companies-outside-of-europe/

Wu, K., Huang, S. Y., Yen, D. C., & Popova, I. (2012). The effect of online privacy policy on consumer privacy concern and trust. *Computers in Human Behavior, 28*(3), 889-897. https://doi.org/10.1016/j.chb.2011.12.008

Yao, M. Z., & Linz, D. G. (2008). Predicting self-protections of online privacy. *CyberPsychology & Behavior, 11*(5), 615-617. https://doi.org/10.1089/cpb.2007.0208

Yao, M., Z., Rice, R. E., & Wallis, K. (2007). Predicting user concerns about online privacy. *Journal of the American Society for Information Science and Technology, 58*(5), 710-722. https://doi.org/10.1002/asi

Yao, M. Z., Zhang, J. (2008). Predicting user concerns about privacy in Hong Kong. *CyberPsychology & Behavior, 11*(6), 779-781. https://doi.org/10.1089/cpb.2007.0252

YouTube. (n.d.). *Recommended videos: How does YouTube's recommendation system work?* Retrieved June 19, 2022, from https://www.youtube.com/intl/en_us/howyoutubeworks/product-features/recommendations/#signals-used-to-recommend-content

Zakon, A. (2020). Optimized for addiction: extending product liability concepts to defectively designed social media algorithms and overcoming the communications decency act. *Wisconsin Law Review, 2020*(5), 1107-1146.

Zarsky, T. Z. (2002-2003). Mine your own business: Making the case for the implications of the data mining of personal information in the forum of public opinion. *Yale Journal of Law and Technology, 5*, 1-56.

Zarsky, T. Z. (2011). Governmental data mining and its alternatives. *Penn State Law Review, 116*(2), 285-330.

Zarsky, T. Z. (2017). . Incompatible: the GDPR in the age of big data. *Seton Hall Law Review, 47*(4), 995-1020.

Zhang, X., Tian, X., & Han, Y. (2020). Influence of privacy fatigue on social media users on their privacy protection disengagement behaviour: A PSM based analysis. *Journal of Integrated Design and Process Science, 25*(1), 78-92. https://doi.org/10.3233/JID200015

Zhou, C., Li, K., & Zhang, X. (2022). Why do I take deviant disclosure behavior on internet platforms? An explanation based on the neutralization theory. *Information Processing and Management, 59*(1). https://doi.org/10.1016/j.ipm.2021.102785

Ziar, R. A., Omar, R., Ahmad, I., & Niazy, S. (2019). Information privacy paradox and fatigue in IoT. *Kardan Journal of Engineering and Technology, 1*(1), 37-46.

Zimmer, M. (2010). ‘‘But the data is already public’’: on the ethics of research in Facebook. *Ethics and Information Technology, 12*, 313-325. https://doi.org/10.1007/s10676-010-9227-5

Zukowski, T., & Brown, I. (2007). Examining the influence of demographic factors on internet users’ information privacy concerns. In *Proceedings of the 2007 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries*, (pp. 197-204). https://doi.org/10.1145/1292491.1292514

Appendix

Appendix A: Questionnaire used on Dutch respondents

Dit onderzoek wordt uitgevoerd door dr. Jason Pridmore, dr. Daniel Trottier en Anouk Mols vanuit de Erasmus Universiteit Rotterdam, in samenwerking met dr. Jessica Vitak (University of Maryland), dr. Michael Zimmer (University of Wisconsin-Milwaukee). We nodigen u uit om deel te nemen aan ons onderzoek over uw houding ten opzichte van dataverzameling.

De invultijd van deze enquête is 10-15 minuten. De enquête start met een aantal algemene vragen over uw achtergrond, waarna wij uw mening vragen over 32 specifieke scenarios rondom dataverzameling. Het onderzoek sluit af met algemene vragen over data collectie en vertrouwen in organisaties.

Om deel te nemen aan deze studie, dient u 1) tenminste 18 jaar oud te zijn, 2) een inwoner van Nederland, 3) in het bezit zijn van een smartphone, en 4) regelmatig gebruik te maken van ten minste één social media of messaging applicatie (bijvoorbeeld WhatsApp, Facebook of Instagram).

Uw antwoorden worden anoniem verwerkt. De volledige toestemmingsverklaring is hier te vinden (inclusief informatie over mogelijke risico's, vertrouwelijkheid en dataverwerking), deze kunt u indien gewenst printen voor uw eigen administratie.

Deelname aan dit onderzoek is vrijwillig. Door deze enquête in te vullen, gaat u akkoord met het toevoegen van uw antwoorden aan de dataset.

Hierbij bevestig ik dat ik ten minste 18 jaar oud ben, in Nederland woon, actief gebruik maak van ten minste één social media service of messaging app, en in het bezit ben van een smartphone.

- O Ja (doorgaan naar enquete)
- O Nee (U kunt helaas niet deelnemen aan deze enquête, onze excuses voor het ongemak. Bedankt voor uw tijd.)

[Onder toestemmingsverklaring] [verplicht]

Selecteer de social media of messaging applicatie waar u het meest gebruik van maakt:

- Facebook
- Snapchat
- Twitter

- Instagram
- Reddit
- WhatsApp
- LinkedIn
- Youtube
- Anders, namelijk: …
- Ik heb nog nooit gebruik gemaakt van een social media of messaging app

Op de volgende pagina's krijgt u 32 korte scenario's te zien. Deze scenario's worden willekeurig gegenereerd, de basis blijft hetzelfde waarbij de variabelen telkens veranderen. De tekst die verandert, is onderstreept.

Lees elk scenario aandachtig door. Na het lezen van elk scenario, wordt u gevraagd om te antwoorden op twee vragen. U zult scenario's zien die veel op elkaar lijken, het is de bedoeling dat u op alle scenarios reageert.

Soms zijn de combinaties duidelijk, soms zullen ze onwaarschijnlijk lijken. Wij willen graag weten hoe u aankijkt tegen de verschillende combinaties die van invloed zijn op het gebruik van gegevens.

[Actor] verzamelt **[tijdsduur]** jouw **[soort informatie]**. Ze zijn van plan om deze gegevens te gebruiken om **[gevolgtrekking]** met als doel **[doel van de gevolgtrekking].**

Voorbeeld: De poltie verzamelt een week aan jouw tekstberichten. Ze zijn van plan om deze gegevens te gebruiken om je mentale gesteldheid te beoordelen met als doel productiviteit te verhogen.

- Dit gebruik van mijn data is gerechtvaardigd.
  (Slider scale, 1-100, *Volledig mee oneens* eens tot *Volledig mee eens*)
- Dit gebruik van mijn data baart mij zorgen.
  (Slider scale, 1-100, *Volledig mee oneens* eens tot *Volledig mee eens*)

Actor

- de politie
- de afdeling Personeelszaken van jouw bedrijf
- je huisarts
- een sociale media/messaging app (gebaseerd op het antwoord op de eerste vraag)
- een online data handelaar
- je lokale overheid

Tijdsduur

- een week aan
- een jaar aan
- de volledige historie

Informatiesoort

- tekstberichten
- foto- en video berichten
- internet zoekgeschiedenis
- e-mails
- telefoon’s locatiegegevens
- social media berichten
- telefoon’s belgeschiedenis
- fysieke activiteit (via telefoon statistieken)

Gevolgtrekking

- je mentale gezondheid te beoordelen
- te bepalen hoe gezond je bent
- door jou bezochte locaties in kaart te brengen
- erachter te komen wie je vrienden zijn
- je seksuele oriëntatie te achterhalen

- je politieke overtuiging te achterhalen

Doel van gevolgtrekking

- criminaliteit te voorkomen
- terrorisme te bestrijden
- de verspreiding van ziektes terug te dringen
- jou gepersonaliseerde advertenties te tonen
- verkeersstromen in je regio te verbeteren
- drankmisbruik te verminderen
- een nationale database van burgers te creëren
- productiviteit te verhogen

U identificeert zich als: (selecteer één optie)

Man

Vrouw

Anders

Voorkeur om zelf te omschrijven als _____

Voorkeur om vraag niet te beantwoorden

Hoe oud bent u vandaag (in jaren)? ____

Wat is uw hoogst behaalde opleiding

Basisonderwijs

Middelbare school (LBO, MAVO, VMBO, HAVO of VWO)

MBO of HBO studiepunten behaald, maar geen diploma

Middelbaar beroepsonderwijs (MBO)

Associate degree (2-jarige HBO opleiding)

Bachelor's degree (HBO of universiteit)

Master's degree (universiteit)

Gepromoveerd / training na hoger beroepsonderwijs (bijv. MBA)

Voorkeur om vraag niet te beantwoorden

Hoe vaak gebruikt u uw smartphone op een normale dag, voor verschillende doeleinden?

(slider scale: Nooit --- Constant)

Hoe zeker bent u dat u over de kennis beschikt die nodig is om uw online privacy en veiligheid te kunnen beschermen? (bijv. door uw zoekgeschiedenis te verwijderen)
(slider scale: Heel onzeker --- Heel zeker)

Hoe zeker bent u over uw algemene kennis van de verschillende soorten data die uw telefoon deelt met mobiele applicaties?
(slider scale: Heel onzeker --- Heel zeker)

Hoe zeker bent u over uw vermogen om te bepalen wat en hoe uw informatie online wordt gedeeld?
(slider scale: Heel onzeker --- Heel zeker)

In hoeverre bent u het eens of oneens met de volgende stellingen over uw gebruik van mobiele telefoon-applicaties? (five-punts-schaal: 1=Zeer mee oneens, Enigszins mee oneens, neutraal, enigszins mee eens -5=zeer mee eens)

- Ik geloof dat de locatie van mijn mobiele telefoon ten minste een deel van de tijd wordt gemonitord.
- Ik maak me zorgen dat mobiele applicaties teveel informatie over mij verzamelen.
- Ik maak me zorgen dat mobiele applicaties mijn activiteiten op mijn mobiele telefoon kunnen monitoren.
- Ik heb het gevoel dat mijn gebruik van mobiele applicaties er voor zorgt dat anderen meer over mij weten dan waar ik mij goed bij voel.
- Ik geloof dat mijn gebruik van mobiele applicaties ervoor zorgt dat informatie die ik als privé beschouw meer beschikbaar is voor andere partijen dan ik zou willen.
- Ik heb het gevoel dat mijn gebruik van mobiele applicaties ervoor zorgt dat er informatie over mij verzameld is die mijn privacy schaadt als het gebruikt zou worden.

Selecteer als antwoord op deze vraag ‘Neutraal’

- Ik maak mij zorgen dat mobiele applicaties mijn persoonlijke informatie kunnen gebruiken voor andere doeleinden zonder dat ze mij daarover informeren of mijn toestemming vragen.
- Wanneer ik persoonlijke informatie geef om mobiele applicaties te gebruiken, ben ik bang dat applicaties mijn informatie voor andere doeleinden gebruiken.
- Ik maak mij zorgen dat mobiele applicaties mijn persoonlijke informatie met andere partijen delen zonder dat ze daarvoor mijn toestemming vragen.

In hoeverre bent u het eens of oneens met de volgende stellingen over uw Internet-gebruik? (five-punts-schaal: 1=Zeer mee oneens, Enigszins mee oneens, neutraal, enigszins mee eens - 5=zeer mee eens)

- Over het algemeen vind ik privacy belangrijk.
- Er is niets dat ik kan doen om mijn privacy en beveiliging online te beschermen.
- In de online wereld bestaat privacy niet meer.
- Er is niets dat ik kan doen om te voorkomen dat mijn account wordt gehackt.
- Ik heb controle over de informatie die ik online deel.
- Ik heb niets te verbergen.
- Ik ben niet interessant, dus het is onwaarschijnlijk dat ik een doelwit wordt van surveillance.
- Ik zou mijn privacy kunnen inruilen voor gemak.
- Ik zou mijn persoonlijke gegevens kunnen inruilen voor lagere kosten.
- Het is onwaarschijnlijk dat ik het doelwit zal worden van online oplichterij of hackers.
- Ik ben uiteindelijk verantwoordelijk voor mijn privacy en de veiligheid van mijn online informatie.
- Technologie bedrijven moeten verantwoordelijkheid nemen voor mijn online privacy.

In hoeverre bent u het eens of oneens met de volgende stellingen over uw vertrouwen in verschillende organisaties?

(Scale: 1=Zeer mee oneens, Enigszins mee oneens, neutraal, enigszins mee eens -5=zeer mee eens)

- Meestal vertrouw ik erop dat de Europese overheid doet wat juist is.
- Meestal vertrouw ik erop dat mijn nationale overheid doet wat juist is.
- Meestal vertrouw ik erop dat mijn lokale overheid (incl. rechtshandhaving) doet wat juist is.
- Meestal vertrouw ik erop dat Nederlandse bedrijven doen wat het beste is voor de consument.
- Meestal vertrouw ik erop dat Amerikaanse bedrijven doen wat het beste is voor de consument.
- Meestal vertrouw ik erop dat de sociale media of chat app die ik het meest gebruik doet wat het beste is voor de consument.
- Meestal vertrouw ik erop dat nieuwsmedia op de juiste wijze verslag doen.