INFORMATION PRIVACY FROM A RETAIL MANAGEMENT PERSPECTIVE

Wanda Presthus  
Westerdals Oslo ACT – Faculty of Technology, prewan@westerdals.no

Linda Andersen  
Westerdals Oslo ACT – Faculty of Technology, linda.andersen@protonmail.com

Follow this and additional works at: http://aisel.aisnet.org/ecis2017_rp

Recommended Citation
http://aisel.aisnet.org/ecis2017_rp/126
INFORMATION PRIVACY FROM A RETAIL MANAGEMENT PERSPECTIVE

Research paper

Wanda Presthus, Westerdals Oslo ACT – Faculty of Technology, prewan@westerdals.no
Linda Renate Andersen, linda.andersen@protonmail.com

Abstract

Information privacy concerns a person’s right to access and control personal data. Advances in technologies like smart phones, beacons, and video surveillance influence a customer’s privacy in both physical spheres and online shopping. We wanted to investigate: What kind of information do retailers accumulate from in-store versus online shopping, and how does this affect their customers’ information privacy? We interviewed managers at large retailers that have both physical and online stores. As anticipated, we found that case companies accumulate more data from online than in-store shopping, however the retailers are in the process of integrating data from both sources. We make the following contributions: (i) an updated mapping of Norwegian retailers’ exploitation of technology and customer data, and (ii) the following argument: Mason’s PAPA framework from 1986 still addresses the main issues regarding companies’ data collection from both online and physical stores. However, we need to focus on how companies handle the information. Pertaining to this, we offer two guidelines for retail managers. This study should be of interest to researchers within privacy, Business Intelligence, Big Data, and Data Science. It is also relevant for retail managers and customers both inside and outside Norwegian borders.

Key words: information privacy, retail, in-store shopping, online shopping, case study.

1. Introduction to Information Privacy in the retail context

“I don’t want to live in a world where everything I say, everything I do, everyone I talk to, every expression of creativity and love or friendship is recorded. Privacy matters. Privacy is what allows us to determine who we are and who we want to be.” Edward Snowden, 2013 (NSA Whistleblower).

Privacy issues were brought to the forefront in 2013 when former CIA employee, Edward Snowden, revealed details of a classified United States government program from National Security Agency (NSA) on extensive Internet and telephone surveillance. These revelations put information privacy on the agenda and diffused across national borders. In April 2016, the General Data Protection Regulation was adopted and has several implications for a retail company. Failure to protect customer data will have larger consequences and higher fines. Moreover, data must be collected according to a specific business goal and within the scope of the directive (Council of the European Union, 2016).

For the retail manager, it can be challenging to balance information privacy with staying competitive. Retailers need consumer information to deliver products, study consumer profiles, and offer personalized services (Hui, Teo, and Lee 2007; Awad and Krishnan 2006). The ability to collect, analyse and respond to user information is of growing importance (Awad and Krishnan 2006), and for companies to survive, they depend on vast amounts of information to gain insights about existing and potential new customers (Culnan and Armstrong 1999). Current technologies enable both benefits and unintended consequences (Conger, Pratt, and Loch 2013). In terms of benefits, consumers have access to improved and personalized services. Unintended consequences include that the consumer cannot remain
anonymous anymore and the information is subject to be shared with third parties such as advertisers (Conger, Pratt, and Loch 2013; Hui, Teo, and Lee 2007).

While many studies have investigated information privacy as an individual, sectoral and national level phenomenon (Greenaway and Chan 2005; Smith, Dinev, and Xu 2011), there is a gap in our understanding of organisational approaches regarding what kind of information that is gathered from in-store and online shopping, and how this can affect information privacy. Some studies do exist, such as Greenaway and Chan (2005) who investigated the organisational level of information privacy behaviours, and Smith (1993) who studied how corporations were handling sensitive personal information. However, most research on privacy has not addressed broader organisational and managerial issues (Culnan and Williams 2009). There seems to be a mismatch where organisations are concerned with information privacy, but researchers rarely consider this level of analysis in their studies (Bélanger and Crossler 2011). Collection of personal information became a theme in the 1970’s (Smith, Burke, and Milberg 1996), even before the World Wide Web created a new sales channel for online retailing in the late 1990’s. Many companies attempted to take advantage of this method for reaching customers, and it was soon argued that most successful retailers would have both online and physical stores (Browne, Durrett, and Wetherbe 2004).

Based on this background, we present a twofold research question: (i) what kind of information do retailers accumulate from in-store versus online shopping, and (ii) how does this affect their customers’ information privacy? The information needed for this study was difficult to acquire, and demanded confidence from the retailers. Thus, to alleviate these problems, we chose to approach retailers in our own country: Norway.

2. Related literature

In this section, we present prevailing publications on information privacy within the Information Systems context, followed by technologies that retailers use for gathering data from in-store and online shopping. We also briefly point to consequences of loss of information privacy.

2.1 Information Privacy and Frameworks

We searched all Basket of Eight journals, the Business Source Premier Database, and Google Scholar. This first stage identified particularly two relevant articles from Bélanger and Crossler (2011) and Smith, Dinev and Xu (2011). Privacy was defined as “the right to be left alone” in 1890 by Warren and Brandeis. Information privacy is a subset of the overall concept privacy, also referred to as physical privacy (Smith, Dinev, and Xu 2011; Dinev et al. 2013). Bélanger and Crossler (2011) presented results from analysing 500 articles on information privacy research in information systems. Together with Smith, Dinev and Xu (2011), they stated that the focus on individual level of analysis has been at the expense of our understanding at organisational level. This could be explained by how it is easier to collect and analyse data from a large number of individuals with surveys and interviews rather than with for example exhaustive interviews with employees or stakeholders from organisations (Bélanger and Crossler 2011; Smith, Dinev, and Xu 2011). Information privacy concepts are mostly understood in terms of individual-level constructs, like attitudes and concerns. Bélanger and Crossler found it surprising that there is a low number of studies in the context of e-business and e-commerce and the different information privacy attitudes within organisations, and suggest using a case study, action study or ethnography to approach the topic. Bélanger and Crossler recommended that information systems researchers should be more involved in information privacy tools and technologies, as computer scientists have mostly addressed this issue.

A second search in “The Basket of Eight” and Google Scholar was conducted for literature on online store versus in-store. This topic can be described in different ways, and we used the following keywords: Online versus Offline, Bricks versus Clicks, Online versus Store, Online versus Traditional shopping,
Physical versus Virtual store, and In-Store versus Online. For literature on technologies, we mainly consulted Institute of Electrical and Electronics Engineers (IEEE) publications.

Overall, we observed that research articles focusing on privacy issues in online stores versus in-stores have escaped attention in the academic research. We found some research comparing online stores and in-stores, such as consumer reactions towards clicks and bricks (Browne, Durrett, and Wetherbe 2004) and consumer attitudes toward traditional stores and online stores (Kacen, Hess, and Chiang 2013). A recurrent issue is trust regarding the purchase and transaction when shopping online. Uslaner (2004) disclosed consumers’ concern around website security, how the collected information is used, websites installing spyware on their computer, and the risk of getting credit card numbers stolen.

Most people are unaware of what goes on behind the scenes when we visit a website (Grossman 2013). For example, when shopping online, we initiate a trail of actions. The customer is first segmented (gender, age, hobbies), followed by an advertisement exchange that is facilitated by cookies. Next, a calculation determines how much the customer is worth to the advertiser. The highest bid wins. The winner places an advertisement on the web page that the user will be able to see the next time he or she loads the page on the computer. The whole process takes only 200 milliseconds! More details can be obtained from the public document called The Great Data Race (The Norwegian Data Protection Authority 2015). Furthermore, The Great Data Race also reveals that an average of 43 different companies have a presence on Norwegian online newspapers and record what we do. Between 100 and 200 cookies are stored on our web browser when visiting the front page of six Norwegian newspapers. These advances in online technologies have raised new concerns about privacy (Drennan, Sullivan, and Previte 2006). Although online retailers inform users about their privacy policies available on their website, this information often remains invisible to consumers who rarely make the effort to read and understand what “This website uses cookies” actually mean.

In the physical world, our actions are traced when using our credit cards, bonus cards or when signing up for a customer membership. The anecdote from Target reads: A father gets very upset when Target sends coupons in the mail for baby clothes and cribs, addressed to his teenage daughter in high school. He has an outburst toward the manager about how outrageous it is, before he receives the news from his daughter about the pregnancy some days later (Hill 2012). By using data from previous sales, retailers can predict what their customers desire. The majority of the literature agrees with these concerns. However, Mateosian asked: “Does it matter if Target knows you are pregnant before your husband does?” (Mateosian, 2013, p. 60). He did not provide an answer but claimed that such issues usually fall out of business discussions, even at the strategic level.

Conceptual frameworks are numerous, but Mason is seminal within Information Systems. Mason (1986) defined privacy, accuracy, property and accessibility (henceforth PAPA) as the four ethical issues of the information age. This was initially published as an essay, launched for creating a debate around these issues. It has grown into a theoretical social framework concerning the major growth of information technology in different research areas (Masrom, Mahmood, and Zainon 2013; Peslak 2006). Conger, Loch and Helft (1995) primarily confirmed the four ethical issues with small variations, but they chose to add the fifth factor of motivation. Motivation reflects a recognized responsibility for actions that affect others (Conger, Loch, and Helft 1995). Even more extensions were provided by Woodward, Martin and Emboden (2010). They presented the new factors of low risk property misuse, high risk property misuse, personal responsibility and privacy. Parrish (2010) used the PAPA as a theoretical lens in his research to highlight issues regarding information sharing on social networking sites. More recent research conducted by Chatterjee, Sarker and Valacich (2015) drew upon the PAPA framework to define unethical information technology use. Unethical use includes hacking, phishing and spoofing, which have become a major security concern for society, individuals and organisations. In this paper, we chose to use the original framework by Mason, because it is the most adaptive to our organisational perspective. The subsequent studies with extensions have, in our opinion, a more individual aspect.
2.2 Technologies that render tracking and surveillance possible

According to Peslak (2006) we currently live with a degree of tracking and surveillance that would have been unimaginable just a few generations ago. This subsection will describe seven main technologies linked to tracking and surveillance: beacons, various camera solutions, people-counting systems, MAC addresses, cookies, web beacons, and Google Analytics.

For physical retail stores, it is essential to interact with the customer. Pierdicca, Contigiani and Frontoni (2015) refer to pervasive smart environments and embedded location systems like the beacon technology. Beacons are small devices that transfer identification through low energy Bluetooth signals and radio waves. They are used to provide low-cost indoor localization to investigate and improve consumer movements (Sturari et al. 2016). Cameras, also referred to as video surveillance (Porikli et al. 2013) are common in public buildings and many stores, and have become one of the most popular sensors for indoor tracking. Cameras can provide high-resolution localization results and gather huge volumes of data. Porikli et al. (2013) claim that use of video surveillance is expected to increase in retail, transportation and education. It is called camera monitoring when a person has the responsibility of monitoring for suspicious activities (Porikli et al. 2013). Recent technologies and market trends have demanded a significant need for Intelligent Video Systems and Intelligent Video Analytics (Liu, Chen, and Kubota 2013). Intelligent Video Systems is a solution that uses technology to automatically process, manipulate and/or perform actions to either live or stored video images (Liu, Chen, and Kubota 2013). Intelligent Video Analytics can help public and commercial organisations to transform video surveillance into a real-time, reactive process. Another type of tracking technology is the people-counting system. It automatically estimates the number of people in open or closed places, and can be used in combination with video surveillance (Zhao, Delandrâ, and Chen 2009; Connell et al. 2015). A connection between active radio-beacons and an intelligent retail environment with cameras could allow customer indoor tracking and localization (Sturari et al. 2016). Finally, a media access control address (MAC), also called a physical address, is found on our smart phones and computers. This unique identifier can be collected and used to identify the owner of the device (Cunche 2014).

Online stores use a variety of tracking technologies. A cookie is a text file with a tracking number that is downloaded onto a user’s hard drive. Cookies can include login-information, user preferences and/or online shopping cart information (Lozada, Kritz, and Mintu-Wimsatt 2013). Cookies are the main technology applied with different purposes. Transient cookies (also referred to as session cookies) are small files that contains information about a user in terms of visited pages and disappears when the user closes the browser. The persistent cookie is stored on the user’s hard drive, either permanently or within a specific time frame (Case, King, and Gage 2015). It is typically used for collecting overview of which browser the user prefers. Statistical cookies are used for retrieval of numbers of visitors to the online store. These are only some examples of how cookies can be used. First-party cookies are set by the website owner and the domain that the user visits directly. Third-party cookies are set by some other domain and are normally set up to track and send information between your computer and the multiple websites a user visits (Roesner, Kohno, and Wetherall 2012). Customers are starting to be somewhat aware of cookies, but third party cookies can be more difficult to spot. They transmit the collected data to fourth parties, which shares this information with online marketers. A new, more persistent type is called Flash Cookies. They have the ability to carry on the collection of data and tracking even after the user deletes all cookies (Lozada, Kritz, and Mintu-Wimsatt 2013). A web beacon (Lozada, Kritz, and Mintu-Wimsatt 2013) is a tiny graphic, one pixel wide and one pixel long, embedded in a web page or an email. Web beacons can record the user’s Internet protocol address and confirm when an email or a web page is viewed. Measuring visitor statistic is a core activity for any web site provider (Pakkala, Presser, and Christensen 2012). Google Analytics offers a free tool for measuring and analysing visitors’ statistics, such as: number of visitors, search history, and how long they stay on the web page. Culnan and Williams (2009) argue that organisations can successfully secure personal data about customers, but still make bad decisions about the further use of the collected information, which result in information privacy problems. Solove (2006) points to the potential secondary uses, such as profiling,
targeted advertising, spam and other intrusive advertising. This makes security a prerequisite for privacy. Both security and privacy are necessary for having a well-functioning democratic society.

Summing up, we found limited research on information privacy regarding the comparison of in-store and online shopping. Based on the introduction we think that this research gap is unfortunate, and we will respond to the call for research by conducting a case study, which we describe next.

3. Method and presentation of case companies

Our overall approach is a case study (Yin 2014; Eisenhardt and Graebner 2007) as this is scarce in information privacy research (Bélanger and Crossler 2011). We selected six large retailers that offer shopping in both physical as well as online stores. They were chosen due to their large size and relevance, which we based on figures from Statistics Norway (Statistics Norway, 2016). The case companies wanted to be completely anonymous, and consequently, we cannot reveal much about them. Complying with the companies’ request for anonymity, we clustered the six different retail companies into “case company A, B, C”. A represents sports and leisure products including clothes, shoes, bicycles, skies, tents, fishing gear, B is in entertainment such as books and DVDs, and C constitutes telecommunication. Similarly, the nine participants from the six different companies are referred to by numbers. The data collection (Table 1) was mainly obtained through in-depth, semi-structured interviews carried out during spring 2016. The average time of each interview was about 60 minutes. The questions covered each participant’s background, role and personal consideration of privacy in the company. We also inquired about the various types of technologies the company used; what kind of data they collected and how it was stored; and the persons responsible for the data. In addition, we interviewed two experts from the Norwegian Data Protection Authority about beacons, web beacons and video analysis, and one expert from fluxLoop (sic), a company specialising in beacon-technology.

All participants allowed us to record the interviews in order for us to transcribe them, and we erased the recordings upon approval. In addition to these interviews, we attended meetings, reviewed documents, and conducted observations. The Norwegian Data Protection Authority allowed us to join their meeting with the security vendor Robert Bosch AS to learn more about their security solutions within intelligent video analysis. We studied the existing EU directive and reports such as the Great Data Race. Finally, we explored the physical stores by visiting them in person and the online shops by systematically browsing the websites.

<table>
<thead>
<tr>
<th>Action</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review of related research</td>
<td>Themes for the interview guide aimed at experts and retail managers</td>
</tr>
<tr>
<td>Data collection</td>
<td>Semi-structured interviews, documents, observation</td>
</tr>
<tr>
<td>Data analysis</td>
<td>Themes and trends (Miles and Huberman 1994)</td>
</tr>
<tr>
<td>Findings and discussion</td>
<td>Findings packaged in a 2x2 matrix adapted from Parrish (2010), and</td>
</tr>
<tr>
<td></td>
<td>discussed against extant research, mainly Mason (1986)</td>
</tr>
<tr>
<td>Contributions</td>
<td>One documented case study based on three case companies and two guidelines for retail managers</td>
</tr>
</tbody>
</table>

Table 1: Summary of method

4. Findings and analysis

This section is presented according to themes and trends, which we identified using techniques by Miles and Huberman (1994). The four themes are: the importance of privacy, use of technologies, how the companies collect and store data, and the value of gaining insight. We also identified two future trends: integration between in-store and online customer data, and increased use of beacons.

All participants agreed that information privacy is very important, but only one case company knew about the new EU General Data Protection Regulation. This company demonstrated concern about misunderstanding the new EU General Data Protection Regulation due to its early stage of adoption; it
will be implemented in May 2018. Nonetheless, they were in the process of preparing by hiring lawyers, starting internal projects, and assigning responsibilities to dedicated employees.

Table 2 presents an overview of tracking and surveillance technologies used in the case companies:

<table>
<thead>
<tr>
<th>Case company</th>
<th>In-stores</th>
<th>Online stores</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>People-counting systems, Queue systems, Video surveillance</td>
<td>Adobe analytics (persistent and session cookies), AdForm, server (persistent cookies), Google Analytics, Hotjar, Optimizely. Performed test: web beacons – concluded too vulnerable</td>
</tr>
<tr>
<td></td>
<td>Ongoing project: Beacons</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Smiley buttons</td>
<td>Google Analytics, Silverpop, Apptus Esales</td>
</tr>
<tr>
<td>C</td>
<td>People-counting systems, Beacons</td>
<td>Google analytics, Visual website optimizer, AdForm, ClickTale, Hotjar, Webtraffic, Scorecardsearch, TradeDouble, DoubleClick Floodlight, AdTech, Euroads, Turn, Google tag manager for administration</td>
</tr>
</tbody>
</table>

Table 2: Technology use in-stores and online stores

Tracking technology in-stores is dominated by people-counting systems, as used by case companies A and C. This technology has existed for about 10 years and is in constant development (such as 3D cameras), and offered by multiple vendors. Video surveillance is typically installed for security and allows companies the opportunity to provide authorities with evidence in robberies or embezzlements. We acknowledge that video surveillance can potentially be used to track both employees and customers. This technology can identify face features and perform advanced analytics. Our case companies informed that surveillance cameras are indeed installed in the shopping areas, but not in restrooms or changing rooms. There are information signs by the entrance doors in all of the stores. Company B has no tracking technology installed, as one might argue that queue systems and "smiley-buttons" do not qualify as such. Queue systems register objects in the area where they are installed, with the purpose of informing the manager about the number of objects in that specific area. “Smiley buttons” is a physical object placed typically at the entrance/exit of a store, where the customer can choose between different buttons to express their level of satisfaction. The company described their stores as “black holes” when it comes to collecting information. While they are looking into different solutions, their focus remains on their online store. Case company A informed us about an ongoing project regarding beacons, and indicated great expectations to an increased utilisation of technology. Customer flow analysis was not in use within any of the case companies. The managers in case company A regarded this to be a cost-benefit discussion and did not see the need for that yet, but believe it could prove useful in the future.

Data collection within the case companies is mainly on an aggregated level, meaning that it is not connected to a specific individual, as shown in Table 3:

<table>
<thead>
<tr>
<th>Case Company</th>
<th>How data are collected from in-stores</th>
<th>Individual information</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Demographics if customer purchases something</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>Demographics if customer wants product sent to home address</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Demographics if customer purchases something</td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Data collection in-stores

Data collection in-stores is done primarily through people-counting systems and queue systems, where managers can get an overview of visitors, buyers and non-buyers, hit rates, meaning the number of sales divided by the number of customer who enters the store/website. “Hot” and “Cold” spots indicate where customers usually are moving around in the store, which is information that can be used for product placements. This is non-identifiable information where customers are registered as objects, or spots,
based on their body temperature. Demographics are collected only when customers choose to have a product sent to their home address or engage in a telephone subscription.

Regarding online stores, most of the data collection is also on aggregated level as shown in Table 4:

<table>
<thead>
<tr>
<th>Case Company</th>
<th>How data are collected from online stores</th>
<th>Individual Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Navigation, clicks, “looked at”, purchased or not, content response toward customers, visitors, traffic</td>
<td>Demographics if creating an account</td>
</tr>
<tr>
<td>B</td>
<td>Navigation history, purchase history, traffic, who opens their newsletters, clicks, recognize earlier logged on customers, Facebook data</td>
<td>Demographics if creating account, IP address</td>
</tr>
<tr>
<td>C</td>
<td>Device type, language, operating system</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Table 4: Data collection in online stores

Case companies A and B have much of the same data collection on an aggregated level, such as navigation history: where users click on the website, their purchase history, in addition to overview of traffic and visitors on the website. Case company C collects device type (for example mobile phone, PC, or tablet) and operating systems, which integrates with the online store, and the type of language.

We inquired about the storage of the collected information, and how these actions are communicated to customers. We found that data from people-counting systems is stored up to eight years, but the data are completely anonymous. They use the information for trends and to provide an overview. Case company B stated a two-week long storage time for cookies on customer hard drive and three years for the Silverpop e-mail system. Video surveillance followed legal requirements in a strict manner, and one case company had their own security chief. Communication regarding the anonymous data collection, such as with people-counting systems, was not common practice in any of the case companies. However, when we browsed the case companies’ online stores, we observed that they all have information pages that inform customers about the use of technology and collection of data.

All case companies received value from gaining insight into customer data. Technology is necessary for gaining insight, because you would have needed thousands of employees to conduct the same amount of work as case company A stated. They aim to understand customer behaviour and patterns in order to better adapt their services for the customer. Insights are also valuable regarding competitors, efficient marketing and using the operational business processes to its full potential. The use of beacons aids in understanding the customers’ journey through the store, as well as personalising offers.

We asked all interviewees about their thoughts regarding future trends, and we made two major discoveries: (i) integrating data from both in-store and online shopping, and (ii) increased use of beacons. All case companies expressed a continuous focus on collection and analysis of data for understanding their customers. They are on the lookout for new technology and many of them are testing, or planning to test, new devices for both in-store and online shopping. A closer cooperation between in-store and online shopping will make it easier to spot trends, according to case company A: “We will be smarter and use customer data.” Case company B seemed to be equally prepared: “Data gathering will increase. We have to make ourselves interesting for the customer.” The use of beacons is increasing: case company C installed beacons in their stores before Christmas of 2015 after a successful pilot-project, and case company A has an ongoing project regarding beacons. Web beacons, on the other hand, were assessed as too intrusive when it comes to customer privacy.

Increased information about customers is also a topic according expert interviewee 1, who stated that collection of information today is poorly communicated to the customers. There is reason to believe that video analysis will become more sophisticated in the future, according to expert interviewee 2. Expert interviewee 3 claimed that there is a lot of talk about beacons but a lack of experience in the area.
Summing up, we present a taxonomy of the case companies. Figure 1 classifies the case companies according to which extent they collect data on the x-axis and what level of data management along the y-axis in a 2x2 matrix.

**Figure 1**: Overview of the case companies

**Case company A** in-store is placed in Cell 1 due to a high amount of data collection. This is primarily aggregated data, such as queue statistics and customer counting. The in-store also demonstrates a high level of data management, as they continuously stay up to date with laws, regulations and their employees’ technical skills. The online store is in Cell 2, as a result of high amount of data collection but a lower level of data management, because their data collection lacks a clear purpose for further analysis and utilisation.

**Case company B** in-store demonstrated a low amount of data collection, but a high level of data management, which place them in Cell 3. Data are collected when customers pay by card or choose to sign up for “loyalty cards”. We place the online store in Cell 1, as they collect a high amount of data as well as high level of data management.

**Case company C** in-store is placed in Cell 1 because of a high amount of data collection and data management. Like case company A, the data is mostly aggregated and come from queue statistics and customer counting. Information about the individual is only collected when customers sign up for subscriptions. The online store is placed in Cell 1, close to Cell 2. Their very high amount of data collection entails a high need for management, which they are in the process of fulfilling in the form of mandatory e-learning courses for their employees, privacy officers and hiring external help.

Due to our findings related to the case companies’ process of closer data integration between in-store and online shopping environment, we believe that Figure 1 will look different in the years to come. Since our findings revealed increasing integration, we believe that cell 4 will remain empty. We hope...
that companies will move out of cell 2, because it is the one with high amount of data collection but low level of data management, thus making it the “bad cell” when it comes to information privacy.

5. Discussion
The discussion is guided and structured by the twofold research question.

5.1 What kind of information do retailers accumulate from in-store versus online shopping?
Information collection about individuals has increased since the 1970’s (Smith, Milberg and Burke 1996), and both the extant literature and our findings indicate that this is still the case and will continue to increase in the future. Collecting information is essential for the retail industry, as expressed by participant 3: “Everyone can compete on price and the product range, but the customer experience is where we can differentiate ourselves and give us a lasting competitive advantage.”

In-Stores
Two out of nine participants acknowledged the use of camera surveillance. Both stated that this is communicated to the customers and that the purpose is for preventing robbery and shoplifting. While the type and use of surveillance can vary, it can be argued that some surveillance is necessary and obligatory for the industry for security reasons as long as it performed with care. Participant 1 stated that their surveillance is supervised by a security chief with strict guidelines. Video analysis is taking it one step further: “I understand it like an automated understanding of content and meaning in images, that can be used in a vast range” (expert interviewee 2). Video analysis is a multifaceted phenomenon, also referred to as intelligent video analysis and image analysis. Customer flow analysis is a way of using video analysis and can be described as a timeline based on the customer’s path through the store. An analyst can roughly put the timelines on top of each other and get movement patterns. However, the case companies did not use these technologies. Expert interviewee 2 said that the technical opportunities are present, but there is no knowledge of extensive use in Norway. Six out of nine participants informed us about analysis being conducted on customer data. Case company A had begun establishing dedicated departments: “I am leader of the Insight and Analytics-department and responsible for business and digital analytics. Relatively new department. Created as a result of a greater digital focus.” Participant 3 further explained its establishment in 2015, and the purpose is to make the most out of collected data to improve customer services and experiences.

People-counting systems are used in two of the three case companies. The participants described the technology as very basic. It provides anonymous, aggregated data on the number of customers per hour and per day, and when connected with a cash register, it can show the number of purchases made. Additional information could include: how many customers there were in a specific zone (“hot” and “cold” zones), where they went afterwards, and how many preferred to take the stairs instead of the escalator. No images are taken during this tracking. The retailers store this data for years, but no privacy issues have been raised at this stage of the technology development according to our findings. “People-counting systems are the same. Exactly the same as 9 years ago. What has changed, is more focus on queues statistics and how we use this data” (participant 1).

Six out of nine participants stated that they have a connection between the in-store and online store, while two could tell us about this being planned or under development. This could make the anonymous and aggregated data collection from in-stores an increasing privacy issue.

Beacons became a subject during data collection. Case company C implemented beacons in all physical stores in 2015. The beacons are controlled by the main office: “We recently started using beacons, but this is not used continuously. We have small campaigns. Beacons is used in combination with our app and bluetooth needs to be activated. The head office controls everything about beacons” (participant 6).
The campaign in question offers the customers, with their app and bluetooth activated, free charging for their phones if the customer was close to one of their locations. The data collected by beacons were sent to a database accessibly by a web platform. The company integrated this data with the rest of what they call their “data lake”. The challenge is to use beacons in a secure manner: “Our big question, and what we are wondering about regarding the personal privacy, is how we ensure that this is done securely enough” (expert interviewee 1).

Retailers are in the initial stages of implementing beacons, and our findings indicate no privacy threats at this stage, but this is a technology with many possibilities. By connecting beacons with a network, it could be used as an analogue cookie and trace customers’ paths in the store (or museums, fitness centres, malls). Most beacons have randomized identifications for security reasons to prevent for example spoofing, which means cloning and creating another beacon with the same identification, according to expert interviewee 3. However, a more malicious concern is piggybacking which is when someone can collect your beacon profile and implement it in their own application (Ashan and Vaidya, 2015).

**Online Stores**

We found more extensive use of technology in online stores than in-stores. All three case companies use Google Analytics. One may ask if there are any hidden agendas behind this free tool. Is it possible to track individuals with the collected data? Regardless, our findings show that the case companies do not exploit this. Participant 2 stated that IP addresses are collected but processed as personal data. Participant 3 explained that the company has a different approach, as they pay for an expensive version of Adobe Analytics, due to the “...advantage that we own all the data, no third party and nothing shared” (participant 3).

Adobe analytics use both persistent and session cookies, which customers are informed about on their website. According to participant 4, they have an information page online, but “few people read this page, even though we refer to it.” It provides organisations with a challenging balance between providing appropriate privacy protection and at the same time embracing new technologies and innovations (Greenaway, Chan, and Crossler 2015).

The Founder of the White Hat Security argued: “If people knew how much of their personal information they unwittingly make available to each and every website they visit; they would have been disturbed.” (Grossman 2013, p. 63). Nevertheless, we found that most of the customers are willing to share their information without much concern regarding their privacy. Our case companies had not registered any increased awareness among customers. The exception was participant 5, who informed us that they had some customers who wanted the retailers to delete the information. Our findings show that it is possible to minimize tracking online. The customer has the option of using safety modes. The most common are Incognito-mode in Google Chrome and InPrivate window in Microsoft Edge. By installing Ghostery, the user can see who tracks her web browsing, or Privacy Badger that protects her from trackers on the web.

Information collection is more extensive from online shopping than in-stores, but one might wonder if customers are still more anonymous when shopping in-store? This leads us toward the second part of the research question.

5.2 How does this affect customers’ information privacy?

This subsection will present findings through the theoretical lens of PAPA (Mason 1986). Mason argued that two forces threaten our privacy: the advancement of information technology and the value of information to those who can use it to their advantage. As we adapted the PAPA framework to our study the acronym spells out like this: Privacy refers to the information that the customer is willing to give to the retailer. Accuracy refers to the companies’ responsibility of accuracy and fidelity of the information. Property refers to ownership and exchange of information in the organisation. Accessibility refers to what information organisations have a right to obtain and store.
Privacy
We found that customers are willing to give up their demographic information when purchasing a product or service in-stores in all three case companies. The customer is not informed about the people-counting systems in case company A and C, but the case companies assured us that the data are completely anonymous. Online stores collect demographics from customers if an account or purchase is made. Case company A has provided customers with the option of shopping anonymously online but would still need names and addresses in order to deliver products via mail. Information privacy is important for management and they “do not want to test their customers limits in this area” according to participant 2. However, not all case companies shared this perception: “It could be that we make more out of it than we disclose to the customer, to be completely honest. We say electronic communication, but we use the data for what it’s worth to identify the customer and create a profile” (participant 4). One participant stated that everything that can be measured will be measured: “as a retailer, we are interested in collecting data within the written and unwritten rules”. This honest quote aside, we think that the case companies show good ability of assessment. They do not want to be in the lead by applying the newest technology, but await its’ maturity, as was the case with case company A’s skepticism about web beacons.

Accuracy
Information systems that collect and store customer data should provide accuracy, as they are often used for analysis and decision making (Mason 1986). While technological advancements allow organisations to collect data in more ways than before, it also inhibits the formation of authentic norms within the society (Parrish 2010). We applied this concern by looking to responsibilities regarding the collected information in the retail companies. The case companies could tell that they hold the responsibility, either by referring to their head office or particular roles such as CTO, Ecommerce manager or operations managers. No third-party responsibilities were identified.

Property
Mason’s (1986) concern about property is about taking information from individuals and embedding it into intelligent systems and the associated ownership. This might be the one issue from PAPA that has changed the most. Organisations are establishing and expanding their focus on analytics. For example, participant 3 was head of the Insight and analytics department, which was established in January 2015. Ownership of information is assigned to the main office and allocated to appropriate roles. Davenport and Harris (2010) present an example of how our digital “trash” may be sold to a third company. The example concerns a grocery chain making more money selling their customer data to retail syndication firms than they did selling meat, and that they later bought back their data mixed with competitors’ data. None of our case companies shared such stories, but they plan to exploit data better. Participant 4 said that they have more data than they use, and further explained how they want a closer integration between the physical store and online store to increase the data processing between them. Although this integration will most likely challenge information privacy, our research indicated that the data will remain within the organisation and not shared with a third party.

Accessibility
Access concerns the individuals’ access to education for coping with information, access to technological tools that store and process information (Mason 1986; Parrish 2010). Placed in an organisational context, it can relate to case company C’s organisational practices: all employees are obligated to take mandatory privacy and security courses, and they hire external help when necessary. We found this to be the least applicable concern of the four within PAPA. Access was an issue in 1986 when PAPA was introduced and Mason (1986) went as far as to say that the level of access was in a state of retreat. The development of information technology and how it enables access to information and between people, can partly explain the strengthening position of access today. Parrish (2010) argues that Web 2.0 technologies have made it easier to create and consume information.
We end this section by pointing to the metaphor known as the *boiling frog*. Is the increasing collection of information about customers happening so fast that they are unaware and unable to grasp the consequences it brings? Very often, we cannot anticipate the impact of something that happens to us in bits and pieces over time. Instead, we are shocked to open our eyes and find how much a series of small changes and developments have turned our shopping environment into something different (Kaykas-Wolff 2014). As we anticipated, technology makes it easier for the companies to collect more data from online shopping than in-store. Customers seem to have no problem with cookies, profiles and other online tools, according to the case companies. However, these data are aggregated and the case companies demonstrated more interest in identifying trends and customer segmentation than targeting the individual. Nonetheless, for the time being, if a person wants to purchase a book and leave no trail of evidence, she should choose the physical store, cover her head with a wide brimmed hat, pay with cash, and kindly refuse any offer of loyalty cards and the like.

6. Contribution

This study contributes mainly to the information privacy literature by empirically assessing Norwegian retailers’ data collection and management about their customers from online and in-store shopping. We applied Mason’s (1986) PAPA framework and were excited about testing its’ relevance. Privacy and property proved to be most relevant in this study. We agree with Parrish (2010) that the PAPA framework alone is not adequate because it lacks guidelines for action, and that we need more research that includes specific principles for further use. Pertaining to this, we present two practical guidelines aimed for both retail managers as well as future research:

**Guideline 1: Consider information privacy as business opportunity**

Companies should aim for transparency and honesty in how the customer data are gathered, processed and protected within the organisation, and promote these actions as a competitive advantage for existing and potential customers: “Choose our store, because we will protect your information privacy.” Companies can successfully secure the collected personal information from customers, but still make unfortunate decisions about the further use, which can result in information privacy problems (Culnan and Williams 2009). We think that the company’s privacy strategy is only as strong as its’ weakest link – which in this case, is the employee or manager with the least knowledge. Investing in secure and mature technology, training and maintaining of employees’ technical skills, and conducting regular risk assessment, could provide companies with "privacy as good business”.

**Guideline 2: Provide customers with the possibility of getting involved**

Our findings show that online stores deal with more extensive use of technologies than physical stores. Our participants told us that few customers actually read the privacy policies and information pages online, however, case company B predicted that: *Customers will get more aware of what is happening when they shop online in the future*. The online store could expose the business routines and provide various functions to give the customer an overview of his or her data, with the possibilities to correct and delete as much as the purchase transactions will allow. As mentioned earlier, online shopping requires some information like credit card numbers and an address – at least at present time.

7. Limitations and suggested further research

This study comes with at least five limitations, of which the last four can be subject to further research. Firstly, we observed that information privacy was a sensitive subject, and it was difficult to approach companies. Therefore, we only contacted Norwegian companies, and we acknowledge that the companies that granted us insights probably were the “good ones”. Second, the new General Data Protection Regulation requires the Norwegian Data Protection Authority to cooperate across borders (Council of the European Union, 2016) which indicates studies with a multi-national research team. Third, *smiley buttons* mainly track employees. Customers can express their level of satisfaction and all registered data are sent to the manager. The manager can map this with the employee(s) working at that time. This was beyond the scope of our study, but further research on potential privacy issues from an
employee perspective could be interesting. Our findings revealed a trend of closer integration between data collected from in-store and online shopping, and further research should follow up on potential consequences. Finally, more research can be directed towards how organisations can profit from offering privacy and how retailers can align their business processes with information privacy.

8. Conclusion
This case study has investigated the twofold research question: (i) What kind of information do retailers accumulate from in-store versus online shopping, and (ii) how does this affect their customers’ information privacy? As anticipated, our findings show that more customer data are collected from online shopping than in-store. Less expected was our finding about the companies’ plan to integrate data from both sources in the near future. We assessed the PAPA framework and found that it is still valid in the sense that it addresses companies’ data collection from both online and in-store shopping. However, we need more research that provide actionable principles for organisations regarding data management. In reference to this, we offer two practical guidelines aimed for managers in the retail industry. Information privacy is likely to gain even more interest and importance in the years to come (Bélanger and Crossler, 2011), and we claim that it is subject for several research projects within Information Systems.

Acknowledgements
First, we thank the Norwegian Data Protection Authority and the managers at the retail companies for granting us interviews. Second, we thank the associate editor and the three anonymous reviewers at ECIS 2017 for providing valuable feedback, which raised the quality of our study. Finally, we are grateful for the comments provided by Pamela Humphries Butler and Tomas Sandnes.
References


