LUISS

DEPARTMENT: Management

MAJOR: International Management

SUBJECT: Advanced Marketing Management

DATA PRIVACY AND SELF-DISCLOSURE: AN EXPLORATORY ANALYSIS OF ONLINE CONSUMERS' BEHAVIOUR

SUPERVISOR

Prof. Marco Francesco Mazzù

CO-SUPERVISOR

Carmela Donato

CANDIDATE

Susanna Staiano

721491

ACADEMIC YEAR 2020 /2021

INDEX

INTRODUCTION

One of the hottest topics in marketing today is self-disclosure of personal data. Especially after the scandal that involved Facebook and its creator, the topic in question is considered extremely topical as never before.

More and more individuals are concerned about the disclosure and especially the uses that companies make of their data, and of course at the same time business entities are increasingly interested in the collection of this sensitive data.

Sensitive data are the ones that concern the most intimate sphere of the individual and, therefore, need special protection. In the digital age, privacy protection has become one of the most important goals to be achieved. Public filming, photographs, magazine subscriptions, subscriptions to online platforms: everything travels fast on the net, making personal data public domain (or almost).

In order to put a stop to the exaggerated diffusion of information, the Italian law has provided that some of them can be processed only with the express consent of the person concerned or with the prior authorization of the Privacy Guarantor. These are the so-called sensitive data. These are particular personal data that, for their delicacy, require a particular discipline. Specifically, with the expression "Sensitive Data", we see that these are included within the personal data, sensitive data are those that reveal the racial and ethnic origin, religious, philosophical or other beliefs, political opinions, membership of parties, trade unions, associations or organizations of a religious, philosophical, political or trade union, the state of health and sex life. Sensitive data are subject, due to their sensitivity, to a particular legal treatment.

As I said, these data are increasingly coveted by companies, as it is commonly recognized that being able to collect them in large quantities, brings great benefits, as it allows to design and produce products and services tailored to a certain target of consumers.

These benefits are not only for companies, as by receiving offers made to measure, consumers also receive certain benefits. For this very reason, regardless of the concerns and risks involved in releasing personal data, consumers are often inclined to move on, thus deciding to provide third parties with their data.

The objective of this research is therefore to investigate this topic trying to understand how, over time, the concept of privacy and personal data has changed and, identify those factors (also called facilitating factors) that today push individuals to underestimate and disclose their data.

These factors will be investigated through an enormous amount of work, aimed at understanding first of all the context of reference and everything that revolves around it, to arrive at the end result of being able to formulate a research question that we will then go to analyse and verify or disprove, through the analysis of data.

So, after a first chapter that will be necessary to give a general background on the subject, in which will be mainly seen the meaning that today is attributed to privacy, the regulations in force and the subjects are interested in its development, in the second chapter we will address specifically the issues that have been most affected in recent years. We will discuss the dichotomy that exists between the declared attitudes and behaviours actually implemented by consumers, and in particular we will see the cost-benefit analysis that drives individuals to make certain decisions. Finally, in the second chapter we will see what factors are currently recognized by experts as important for most individuals to negotiate their sensitive data. In particular the factors we will see are: trust, personalization and control.

Finally, in the last part of this work an experimental argumentation will be developed on disclosure of personal data online, more specifically in the online shopping world, with the objective of clarifying which safeguards must be guaranteed on personal data in compliance with the General Regulation on the Protection of Personal Data. This will be followed by an analysis that indicates that when a company's website looks visually trustworthy, consumers trust the company more and show greater intentions to provide their personal information. Traditionally, most research on interpersonal communication indicates that when someone discloses more, the other participant in the communication also discloses more. My goal in this paper is to study the impact of online information disclosure on consumer trust and self-revelation, I will try to demonstrate that, the latter are largely influenced by specific endogenous and exogenous drivers that lead them to underestimate the pervasiveness with which personal data are treated.

CHAPTER ONE

Introduction to the topics of "Self - disclosure" and "Privacy concern"

1. PREMISE

The first part of the following research thesis begins with the introduction of the topic we have chosen as our subject and the arguments about why it is topical and relevant. The topic we will discuss is the "self - disclosure" of sensitive and personal data, i.e. data that is particularly valuable to those in possession of it.

It is believed that this topic is of particular importance as it refers to a transversal phenomenon of our society. Nowadays, any individual, whether voluntarily or involuntarily, can be involved in the process of sharing his/her data with third parties. The evolution of society and continuous technological innovations provide the conditions for the importance we give to this issue to make the phenomenon of objective relevance and interest to many. In an economic system in which the infrastructures of power, at various levels, cannot do without a massive volume of information, understanding the modalities and motivations that allow individuals to share their data becomes absolutely relevant.

Our conceptual journey will start with the definition of personal data and the description of its role within society. This first phase, which is fundamental to grasp the value we want to distribute through the paper, will be accompanied by the introduction of reference concepts such as those of personal data processing and the right to privacy, which are particularly important especially if analysed according to the impact they have on the behaviour of individuals. Once the reference concepts have been outlined, proceeding along the path outlined, we will identify the main areas of influence of self-disclosure, also highlighting the related stakeholders, i.e. the recipients of our work. Self-disclosure and adjacent issues potentially impact consumers' concerns about personal data privacy. It will be of interest to study the behavioural response that individuals organize as a consequence of these inputs.

We will then introduce the theme of the "*privacy paradox*" that will open the central discussion of this paper: what are the drivers that determine the dichotomy between potential intentions and actual consumer behaviour?

The common thread of all these elements will be the role of technology that, on the one hand enhances the analytical capabilities of third parties, on the other hand makes users hyper-connected and therefore exposed to the risk of privacy violations.

2. SELF-DISCLOSURE

We live in a society where the role of data and the information derived from its interpretation become central. Globalization and the evolution of the web radically change the way we communicate, exchange information and organize any kind of activity. In this context, social networks and all online services such as sharing platforms and websites encourage the construction of individual or group relationships, they become an active part of an international phenomenon (Hsin-Yi Huang, 2016).

These new media are capable of generating hyper-connections that, as well as conveying specific content, contribute positively to the sharing of users' personal data with the network. As a result of this, the amount of sensitive data available has never been as large as it is today, and it will not stop increasing.

"The amount of data currently present doubles every two years and is expected to reach 40 trillion GB in 2022" (Gantz and Reinsel, 2012). From such a premise one can imagine how wide the range of interest of this paper can be, which will be oriented precisely towards the theme of how users' data are shared on the web and towards identifying the psychology behind this process.

We define "self - disclosure" as "that process through which one transmits data that was previously unknown and thus becomes shared knowledge" (Jourard and Lasakow, 1958) and identifies it as a precondition underlying any social relationship. “Self - disclosure", moreover, consists of two dimensions:

- The quality of information shared;
- The depth and value of them. This causes some information to be shared more easily (gender, age, hobbies), others with more difficulty (personal photos, other specific content, emotions and feelings).

A great deal of research has been conducted on the topic of "self - disclosure. In general, these have adopted a theoretical perspective of social exchange (Ajzen, 1977), suggesting that self-disclosure, like any other interpersonal behaviour, is approached and interpreted in terms of a cost-benefit relationship by individuals (Moon 2000). We know in fact that nowadays there is an enormous demand for disseminating and sharing specific data about individuals for new and exciting uses. As I was just saying, with the latest technological revolution, the vast majority of this data is available electronically. Individuals are always driven to share their social capital due to the myriad of technological innovations that are still taking place. However, the tension that exists between consumers' desire to communicate online and the concerns that arise about how the data they release

will be treated, existed even before the boom in social networks. Just think of the scandal under the eyes of everyone, which hit Facebook, the social networking giant. These tools are perceived by society as providers of new benefits, but they are a great source of data uploaded, stored and shared. (Hallam and Zanella, 2017). The most distinctive types of information and clearly also the ones that often arouse the most interest, and consequently provoke more in individuals to release them within the web and beyond, are for example a) name, b) Social Security Number and c) date of birth. But why today is the subject of the disclosure of personal and specific data considered so important and central by the experts? It is now globally recognized that companies and organizations have an extreme need to collect this type of data in order to achieve a significant competitive advantage over their main competitors and, consequently, to grow their business. The release of this information allows these players to take advantage of the great opportunities that the web offers. To give an example of what I'm saying, but it is not the only one, a very important benefit that the Internet puts in the hands of companies, is precisely that of being able to customize the communications to be undertaken with individual consumers. In order to make that possible we really need to go back to our primary issue: the release of information and data by individuals. In addition to attitudes and concerns about privacy, it's crucial to consider the behaviours that people can take to safeguard their privacy. For instance, have you ever provided false or incomplete personal information when registering on a website, rather than providing your real name and address? Most people would probably answer yes. There is likely to be a complex relationship between attitude and behaviour in this context.

3. PERSONAL DATA, THE TREATMENT AND THE CONCEPT PRIVACY

Before describing what the implications of such a process are, however, it seems necessary to me to describe more specifically a fundamental concept. What do we mean when we refer to "personal data"?

The GDPR legislation that has just come into force, a very relevant legislation that we will discuss below, provides us with an effective description: "Personal data means any information relating to an identified or identifiable natural person ("data subject"); an identifiable person is one who can be identified, directly or indirectly, by reference in particular to an identifier such as a name, an identification number, location data, an online identifier or to one or more features of his or her physical, physiological, genetic, mental, economic, cultural or social identity" (Art. 4(1) GDPR, 2015).

They are, therefore, those data which, if combined with other information present in various servers, are able to specifically identify an individual. This definition opens the door to many topics that we will slowly go to deepen.

The continuous growth of the amount of data available on the web becomes a trend that offers different opportunities for various subjects. Kozinets in 2002, for example, identifies the revelation of consumer information as an "ideal source" for those who need to carry out marketing research, even if it is carried out through the web. The natural consequence of this, at least from a business point of view, is that, together with data, the need to make decisions according to the information derived from them grows.

The role of technology, in this sense, becomes crucial. Indeed, without state-of-the-art applications, the interpretation of all this available data would not be possible. Many of these technological applications are commonly identified as "Big Data" (Dutta and Bose, 2015). The digital enhancement of various organizations and the introduction of these applications significantly increase the "thirst" for personal information, which can be crucial in gaining a competitive business advantage.

The objective of these organizations becomes, therefore, that of implementing the technological levers on which to rely to collect, analyse and interpret the information they need to make decisions. This example immediately gives an idea of how dangerous, or at least intrusive, the treatment of one's own sensitive data by third parties can be.

Again according to the GDPR, processing is defined as "any operation or set of operations which is

performed upon personal data or sets of personal data, whether or not by automatic means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction" (Art 4(2) GDPR, 2015) .

Data processing must be carried out in a manner compatible with the purposes determined for its collection, analysis and distribution. It is necessary, therefore, to establish and explain in a transparent manner the purposes of the treatment in order to allow the person concerned to express his/her informed consent. This is regulated in order to contrast the possible negative consequences that could arise from the violation of the personal sphere of others.

In fact, in order to pursue their own interests, companies could become protagonists of actions at the limit of what is allowed.

In 1994, Bloom proposed two issues that highlight an ethical problem and that a marketer should consider:

- "Should a business be allowed to acquire personal information about the individual without their permission?"
- "Should a business be allowed to disclose personal information about the individual to other parties without the individual's permission"? Although these questions have not been answered in the literature, marketers do not always consider it important to engage in respectful behaviour towards their consumers (Singer, 2012).

Revealing your sensitive information online means exposing yourself to the danger of violation of your personal sphere. Dealing with this issue, therefore, we realize how much it is connected to the concept of the right to privacy, especially when understood as a violation of data revealed independently by the consumer. Quoting one of the greatest representatives of contemporary privacy rights, (Alan Westin, 1968) "privacy is the legitimate claim of the individual to determine the extent to which he wishes to share himself with others and his control over the time, place and circumstances for communicating with others. It is also the individual's right to control information about himself. Privacy is synonymous with the right to be left alone." Although this represents a very accurate definition of the right to privacy, it is possible to say that "there is no uniform definition of the right to privacy" (Francesca Fabris, 2009). The concept is too complex to be delimited, and its meaning has undergone a profound evolution over time. In ancient times, speaking of the right to privacy referred to the right to intimacy, or the right to voluntarily separate from public life.Instead, the first

to speak of privacy as we understand it today, at least from a doctrinal point of view, were Samuel Warren and Louis Brandeis.

In 1890, the two attorneys, delivered to society an essay entitled "The Right to Privacy. The Implicit Made Explicit" as a result of a lawsuit brought against indiscretions about the married life of Warren's wife. The two men found themselves reflecting on what information regarding the personal life of an individual should be in the public domain and what, instead, deserved a different protection.

It is from here that we begin to speak of the right to privacy as "the right to the confidentiality of personal information and private life, that is, an instrument placed to safeguard the private sphere of the individual, to be understood as the faculty to prevent third parties from intruding into one's personal sphere" (Internet & Law, 1995).

3.1 AREAS OF INTEREST

So far we have defined the phenomenon that is the object of the thesis and introduced some fundamental concepts. Through these first lines it should be possible to catch a glimpse of the main areas of influence of the treated topic. In this paragraph we are going to highlight the major implications that the concepts introduced and defined have on different areas in order to emphasize the value that this topic has for different stakeholders.

There are at least two obvious implications of what has been said so far.The first one refers to the impact that "self - disclosure" has on legal fields, the second one considers economic ones. From the legal point of view it would seem natural to think that the context in which we find ourselves needs to foresee figures that guarantee the protection of the user's rights. In fact, the development of Web 2.0 focused on the interaction of users and social networks, must increase the attention of jurisprudence on the relationship between the responsibility of the actors operating on the network and the protection of individual rights. "The UN" recently stated that "rights must enjoy the same protection online as they are accorded offline and that digital identity is no less personal than real identity" (UN 2016). A clear example of this is "web reputation", i.e. the right of every person to have a truthful, up-to-date and correct representation of his or her identity. Traditional media guarantee this right through the right of rectification, editorial responsibility and journalistic ethics: the same must happen on the web. But are there tools in web portals that guarantee such protection? For a long time, in an environment like this one characterized by a high presence of "irresponsible" subjects operating in it, the instruments of protection have not kept up with technological innovations.

The presence of an Authority is necessary, which has both the task of investigating the multiple aspects of the treatment of data by third parties, and that of monitoring and controlling the actions of these parties. Whoever is interested in the processing of other people's personal data, in fact, cannot operate on the web without any constraints. In this historical period, we are spectators of various sentences that open a new season of "data regulation". The "EU Court of Justice" with reference to the right to be forgotten (Google Spain case) establishes that "the search engine operator, being responsible for the processing of personal data, is obliged to suppress not only inaccurate data, but also those "inadequate, irrelevant or excessive" or even the "data stored for a period of time longer than necessary". The "European Parliament" in 2014 passed a resolution deleting any relationship between the operation of search engines and the offering of numerous "web-based services". These are just some of the examples that demonstrate the new jurisprudential approach towards the processing of personal data. Generalizing, it is possible to reiterate that the issue of "self - disclosure" is of strong interest to the entire category of policy makers.

It is precisely from this viewpoint that the "General Data Protection Regulatory" is born, the legislation which has just come into force in Europe, which places itself as guarantor of the protection of its consumers, whether they are in relations with organizations that are part of the EU or whether they are intercontinental. According to the Digital Economy Centre of Rome, this legislation represents a real Copernican revolution. "If until a few weeks ago at the centre of the universe there were the platforms that manage data, since GDPR came into force, instead, at the centre of everything there are users" (Digital Economy Centre, 2016) .

We are reporting that in recent years users have transformed themselves from consumers into producers of data, which are increasingly assuming economic and social value. It is in this context that the EU intervenes directly in the management and control of the data infrastructure built by and around its citizens.

An emblematic case that highlights the great value of data and demonstrates the irresponsibility of some market players is that of Cambridge Analytica, a consulting and marketing firm that profiled about 50 million Facebook users by irregularly taking from the same app a gigantic amount of their personal data.

According to Christopher Wylie, the whistle-blower who recounted the details of the scandal, the consulting firm would have exploited such personal data with the aim of targeting the reach of Trump's propaganda during the US election campaign. The data regarding geographical locations, pages followed and interests were extracted by exploiting the app "this is your digital life" and the

possibility that users had of accessing it via the Facebook Social Login function. This gave them access to the information not only of anyone who had used the app, but also of all their friends on the social network.

The violation of Facebook's terms of use occurred when the same data was being shared with Cambridge Analytica: The social network prohibits app owners from sharing data and information collected from users with third-party companies. The ability of these actions to affect the results of the election campaign allowed Trump to win the challenge and allows us to understand even more about the door of the value that personal data has in today's society.

But the full awareness that personal data has such an important economic value is not only in the minds of the authorities. As we were saying, also for the world of economy it is possible to affirm the centrality of the theme of "self - disclosure" and also here the facets are many.

Let's start, for example, by analysing the need for companies to base their decisions on data. The increase in the amount of personal information available on the web can only implement this need. These data are needed to be able to establish a personalised and trustworthy relationship with consumers throughout their journey towards brand loyalty. The main challenge for businesses in today's markets, in fact, is to bring value to the end customer (R. Howell et al, 2017) and, to do this, marketing divisions respond by becoming increasingly interactive. We define Interactive Marketing as a business approach that transcends direct marketing or web marketing, "Interactive marketing permeates virtually every aspect of the marketing organization" (Moe and Ratchford,2018). The advancement of technology in leaps and bounds and interactive marketing allow businesses to collect information that was previously not possible to analyse as it is now. It is possible to observe how people are searching for a Brand, what they say to each other among peers, how they interact with the same Brand before, during and after the purchase.

Thanks to these observations, it is possible to improve many processes, obtaining a unique position with respect to competitors: prices are dynamic, communication contents are absolutely personalized, relationships are traceable. What is challenging for this type of functions is not the difficulty of data analysis, but the ability to anticipate the continuous technological innovation that leads to the search for new ways to do it.

Here we can mention a new kind of trend that will also be important for our research: the analysis of quantitative data related to the monitoring of the actions of the daily life of online users. In a broader sense, however, it is certain that any kind of sensitive data can be useful for a company. We talk about

data any kind of data that has been collected through conversations with consumers, stakeholders or employees. Everything that determines the acquisition of knowledge is a source of inestimable value.

All of these concepts extend to any type of business activity that takes place. Mobile & E-Commerce, Digital Consumer, Retail, New Energy, Commodity trading are all Industries where the role of data sharing is highly focused. Any subject related to these worlds should automatically have as reference the theme of "self - disclosure". In this paragraph we have tried to outline the perimeter of influence of the theme that we are dealing with, trying to bring out its potential relevance and topicality. During this short path, however, was also conducted another type of work that we are now going to highlight. Through the definition of the areas close to the theme of "self - disclosure", in fact, it was indirectly possible to identify those who are the main stakeholders of our research work, whereby stakeholders we mean subjects or entities for which the topics covered are of objective interest. We summarize them quickly:

- **Policy Maker**: we have said that the dangerous context in which individuals are immersed leads to the need for useful tools to ensure consumer protection;
- **Authority**: the role of institutions of power must be central to the defence of citizens' rights;
- **New-technology**: referring to all the structures and applications used by new forms of communication;
- **Companies belonging to different industries;**
- **Consumers**: directly involved in the process of data - disclosure and central figures in our scientific research. We will elaborate on them in the following paragraphs.

The relevance of a topic is naturally influenced by the number and extent of stakeholders involved. We don't feel like judging the level of ours but, based on the evidence that has emerged so far and that will emerge again, we believe we can say that the topic we are dealing with is truly topical and central to the world scene, both present and future.

3.2 PRIVACY CONCERN AND PRIVACY PARADOX

What has been said so far about everything that orbits around the phenomenon of "self - disclosure" leads to an obvious fact. The individual, immersed in a context in which the institutional world does not renew itself as quickly as the market does, sees the dangers around him increase disproportionately. His personal information on the Net may not be safe.

We continue the path traced by this chapter by going on to describe the phenomenon that originates from this evidence. What results from a decrease in user security is a general increase in the level of consumer concern over the security of their sensitive data. Some research supports this view.

In a recent survey conducted by "F-Secure" in 2015, a company that produces antivirus software to ensure the security of digital devices, 9 thousand people were interviewed on the subject of privacy concerns. Some of the most important data tells us that:

- 66% of them say they are concerned about the third-party exposure their data may suffer online;
- 55% of respondents said they had already changed their online behaviour.
- 59% said they would be willing to switch to other search providers to avoid profiling based on online searches. There is an increase in concern about protecting their privacy.

We then introduce the concept of "privacy concern", i.e. the concerns that arise from the fear of seeing detailed information about one's person (e.g. demographic or experiential) violated, which one shares with third parties in order to obtain specific benefits (Krishnamurthy ,2001). The general fear of the violation of the privacy of one's data reflects a condition of stable tension that is not linked to a specific subject, web portal or social network. The feeling is transversal to any content.

Despite this, some scholars argue that the intensity of these concerns increases in relation to the value individuals assign to the type of information shared (Milne and Gordon 1993; Son and Kim 2008).When a consumer has a high level of general concern about privacy, "he or she develops a particular negative attitude toward any form of intrusive, data-driven communication" (e.g., Martin et al., 2017); the sense of danger toward misuse of his or her data increases and, as a result, trust in those who handle it decreases.

With these words Martin, Bora and Palmantier in 2019 assume a direct and negative relationship of privacy concerns on the decision to share information. This statement would seem to be out of line

with what has been said about the sensitive and continuous increase of data available on the network; there are many scholars who have tried to investigate the reasons for such a controversial phenomenon. In fact, it would be really interesting to understand why in a context of general concern for privacy the number of sensitive data available increases instead of decreasing.

From a technical point of view one could support this by highlighting the difficulty of consumers in defending their personal information sphere from the new technological tools that exist, on the other hand one realizes that there are not a few situations in which such information is delivered voluntarily. Hence, although a certain degree of risk should result in an increased focus on planning strategies to protect one's data, this rarely translates into actual proactive behaviour by the consumer himself (Joinson et al., 2010; Pötzsch, 2009; Tsai et al., 2006). "It is a well-documented fact that users tend to have an all-encompassing online privacy behaviour that ultimately results in a dichotomy between privacy attitudes and behaviour" (Acquisti, 2004; Barnes, 2006).

This discrepancy between the concerns expressed and the actual behaviour of the user is a phenomenon represented by the name of "Privacy Paradox", a topic widely debated among the scientific community and which represents yet another relevant implication of the phenomenon of "self - disclosure".

Let us summarize what can be said about the Privacy paradox phenomenon, also through the words of the researcher Monika Taddicken (2013):

- Privacy concerns are positively correlated with the social relevance of social networks, which in turn depends on the amount of information that is shared with it;
- Privacy concerns are negatively correlated with the number of social apps used;
- The propensity for self-disclosure is negatively correlated with online privacy concerns.

In general, therefore, an ideal degree of privacy is perceived when the individual level of self-revelation meets the desired level of privacy. Generally, a high willingness to self-reveal should therefore be correlated with a lower need for privacy and, therefore, a decrease in related concerns, such as a fear that too much personal information will be accessible online by third parties (Altman, 1975).

But when does the coincidence between these two levels of need occur? This can be explained by investigating the factors that underlie each individual's attitude. What is the vulnerable source that can be leveraged to make consumers inclined to share their personal information on the network, thus decreasing their concerns?

The objective of this quantitative survey will be to identify credible reasons that can explain this phenomenon, trying to fill the gaps in the scientific literature.

1. CONSUMER'S PSYCOLOGY

As we have understood from the last paragraph, the focus of our research will be on the psychology behind the data sharing choices made by consumers. The focus of our time, in fact, is to understand how individuals react to the stimuli in the context just described. In this section we will elaborate on this orientation with the aim of completing the presented framework.

"The key question is no longer whether consumers are willing to share their private information but how they react now that this same information is accessible and freely available to numerous parties" (Kelly D. Martin - Patrick E. Murphy, 2016).

How does the economic infrastructure that cannot disregard data, the information derived from it and its control impact on people's attitudes and behaviour? We have said that the advent of the internet contributes significantly to increasing the phenomenon of "self-disclosure"; in fact, the evolution of the web and the introduction of social networks have substantially revolutionized our way of communicating, exchanging information and organizing any type of activity. Online users have the opportunity to take advantage, in real time, of the contents that interest them most and to share them with other network users. In this way, communication becomes participatory, because anyone can contribute to the diffusion of content on the Internet, which becomes accessible to all. The new media allow to obtain a high degree of interconnection and are able to convey new specific contents, overcoming both the temporal and spatial limits of services based on traditional channels. The evolution of the users' peculiarities motivates the increase of the phenomenon in question but it is not yet clear what are the specific reasons behind it.

What we can state is that privacy behaviour is a highly contextual phenomenon (Morando et al., 2014). We could never expect individuals to demonstrate the same behaviour in different contexts.

Consider, for instance, the studies by Norberg et al. (2007) and Tsai et al. (2011). In the former, a

sample of students was asked to discuss their propensity to share specific personal information; 12 weeks later a "confederate" visited the same campus under the guise of conducting pilot research for a new banking project. The students provided much more information about themselves in the second case where the classroom context was much more influential on their propensity for self-disclosure.

The second study on the other hand involves an online shopping experiment where the shopping search engine displayed privacy policy information about the online store. In this case individuals were even shown to pay a plus for privacy. These results allow us to conclude, generalizing, that results obtained in different contexts are contradictory.

Another evidence is that one of the most supported visions in this field is the one that sees at the basis of the propensity to share personal data the ability to manage a trade-off between costs and benefits (Milne and Gordon, 1993). The application of this trade-off has been described in several theories and research streams as in the case of Homans (1961) and his "Social exchange theory: the individual is engaged in an exchange situation if he expects to receive more value than he loses. This theory has often been used in the context of informational exchanges.

This means that individuals are willing to share their information in exchange for rewards, sometimes very small ones. An anecdote, told in the study by Carrascal et al. (2013), tells us that even individuals are willing to browse a particular site in exchange for €7, the equivalent of a Big mac meal.

Different types of costs and benefits have been considered in the scientific literature.

Among the benefits, some are (Ansari and Mela, 2003):

- The need of individuals to retrieve personally interesting information can be considered as one of the most important factors that encourages a user to approach a web page;
- The search for a viable source of entertainment and quality content that meets the value expected by the user;
- Monetary incentives encourage the disclosure of individuals who are looking for new finances, this allows them to easily overcome their concerns. Among the costs (Aaker and Bruzzone, 1985):
- Registration costs, understood both as economic and as the efforts in terms of time and effort that a user has to put in to access a specific platform;
- The high level of intrusiveness that drives away individuals who thus perceive the interlocutor as irritating or boring;
- Privacy concerns we've talked about abundantly before.

There are many studies that address the issue of trade - off, such as that of Smith, Dinev and Xu (2011) which focuses on the so-called privacy's Calculus, a detailed analysis that consumers make according to the costs and benefits of the actions they are asked to perform. These are some evidences related to the world of psychology of the individual but, in fact, there could be many other factors that influence self - disclosure.

Indeed, many empirical studies have shown that the propensity to share personal information is significantly higher in the context of computer-mediated communication than in traditional face-to-face and that there might be several drivers that play a crucial role in this process (e.g., Bargh et al., 2002; Suler, 2004; Taddei at al., 2010; Weisband & Kiesler, 1996).

There are those who support the value of the driver of anonymity, although this thesis has been disproved by Sabina Misoch in 2015 that in her essay "Stranger on the internet: Online self-disclosure and the role of visual anonymity" demonstrates through two qualitative studies the groundlessness of the relationship between self-disclosure and the factor considered (Visual anonymity).

Other research investigates and supports the idea of the presence of gender differences in online self-disclosure behaviour (Special & Li-Barber, 2012; Wu & Lu, 2013).Others do not find evidence for this (Barak & Gluck-Ofri, 2007).Further studies have revealed relevant age-related differences in behaviour by going to compare the behaviour of children, adolescents, and adults offline and online (e.g., Valkenburg, Sumter, & Peter, 2011). In addition, it is interesting to note that factors such as company reputation, consumer trust and level of control over data can also create confidence and mitigate the negative impact of privacy concerns (Xie et al., 2016).

We are reiterating in this way the possibility of referring to different types of drivers in the process that justifies the privacy paradox and that there are many scholars who have focused on this issue and that, finally, we have not yet come to a secret recipe for identifying the right customer drivers to leverage. Through this research, the goal is to be able to fill this gap, through an analysis that as we will see now will be well outlined in a specific context of reference.

2. ONLINE SHOPPING

The ultimate goal of this master's thesis is beginning to take shape: to respond to the gap in the scientific literature concerning the identification of determinants in the process of "self - disclosure". As we have seen, there are many studies carried out in this regard and, each of these, we have noticed that they are characterized by a specific field of investigation.

In order to differentiate our proposal, therefore, it was decided to apply the same approach focusing on a particular field, relevant and trendy in our days.

The research context will be online shopping; it is now well known that buying online is no longer an alternative but a habit for many consumers. We research and buy everything online, from shopping in specialty channels to medical appointments, from leisure events to simple consumer goods.

We live in an "always on" everyday life where exchanging information and searching for products and services is just a click away, at any time of the day. For these reasons, consumers, driven by various factors, very often underestimate websites and the various website platforms and, above all, the type of information and data that are placed on them.

In particular, most research on interpersonal communication indicates that when someone discloses more, the other participant in the communication also discloses more (Denise D. Schoenbachler Geoffrey L. Gordon, 2002). Although much research investigates the impact of online information disclosure on consumer trust and self-disclosure, there are few studies that address the potential reasons for consumers to disclose information about their identity, specifically, on their first visit on online shopping platforms. In fact, this study examines the influence of self-disclosure during the very start of the buying journey of online shoppers and highlights the important role of “visual trust” during the first visit to a website.

One immediately understands the relationship that this context can develop with the phenomenon of "self - disclosure" and all the concepts we have discussed above. With the definition of the context in which we will apply the theories at the basis of this research we come almost to the end of the opening chapter that we are about to summarize. Before we do so, let's devote one last space to a concept that might come in handy for our topics.

3. PERMISSION MARKETING

A little space before closing the chapter is dedicated to a deepening of a theme that can play an important role in the proposed scenario: the permission marketing. The value assigned to the personalization of the offer and communication can be fundamental for the success of your business. In fact, there are many potential consumers that can be reached with customized messages. This way of contacting the prospect or customer, however, can cause the same individuals to perceive companies as intrusive to their privacy. This creates a challenge for companies who, in order to meet the favor of the people they address, rely on a current concept in modern marketing: the "Permission Marketing".

This concept refers to direct marketing activities that require the consent of the consumer to be contacted by specific companies. Permission marketing can play a key role in the fight against privacy dangers, that is, as a legal requirement. In this way it can become an adaptive solution for all consumers. Consumers will be able to choose for themselves which companies/apps/platforms deserve to share their data, obviously excluding others. In this way reaching a consistent number of permissions can become an absolute competitive advantage in the market.

This is another reason why companies have the primary objective of identifying the factors that influence the decision of their customers. It is therefore necessary to enter into the perspective of costs and benefits to understand what influences positively the process and what impacts negatively.

4. CONCLUSIONS

In this chapter we have defined the object of our research. We have described the centrality we will give to the phenomenon of the privacy paradox and the general need to investigate the psychological causes underlying it. We have identified the main stakeholders of our paper and specified the area in which we will deepen our research.

Finally, the main conceptual tools necessary for the understanding of the treated topic were provided. All this allowed us to identify the main thematic areas of our paper and to define as the main focus that of the dichotomy that occurs between the concern that consumers have for the privacy of their personal data and their actual online behaviour. Our research will endeavour to explain this pattern within a specific context, online shopping platforms.

We will now move on to the second chapter, in which we will go into even greater depth on the topics discussed through an analysis of the reference literature. This will allow us to understand which are the scientific gaps to be filled and therefore to specifically address our work.

CHAPTER 2

Scientific literature and reference work

1. PREMISE

In the first part of this research, we outlined the areas of interest and highlighted the implications that the topics discussed have on them. The topics we refer to are the "self - disclosure", particularly relevant in the introduced social context, and the "privacy paradox", an important phenomenon generated by the distance between the concerns for "privacy" and the actual "behaviours" implemented by individuals.

These concepts we have seen to be particularly interesting both for those who want to exploit their knowledge of the relational dynamics of consumers in order to obtain an economic advantage in the market, and for those who have to guarantee the respect of people's rights.

In this second part, we will deepen some of these themes through the analysis of the existing scientific literature on the subject. The objective will be, in fact, that to acquire greater knowledge about the treated themes and therefore to identify those theoretical postulates that later will be indispensable to build a well-founded scientific research.

Specifically, we summarize the elements we need to arrive prepared for the next research phase. It will be deepened the centrality of the role of data, a concept of primary importance in our society and that can prove to be transversal to the interests of all stakeholders. In particular, it will be important to understand the use that companies and individuals can make of this important tool and the positive and/or negative implications that this can entail at different levels, including that of privacy.

Next, we will discuss the theories behind the concept of privacy: what is the value that society associates with this topic?

We will also describe the different types of privacy with particular attention to that of information privacy, which is very akin to the theme of self-disclosure.

Going forward, we will investigate the models so far produced by the community of scholars to explain the privacy paradox phenomenon: we will argue the existing dichotomy, better define the psychological approach to the cost-benefit trade-off, and identify the main factors that scholars describe as crucial in influencing an individual's propensity to share their personal data.

As we will see, such an approach to the scientific literature has been put into practice because it is important for the purposes of my research to perimeter what has already been studied and to identify the paths to take in order to generate value. In fact, we will say later that one of the main intentions of this scientific research is to allow a solution to be found to a gap in the scientific literature.

2. THE ROLE OF DATA IN TODAY'S SOCIETY

The society in which we live is characterised by major changes. People's needs are changing, the ways in which they relate to each other are changing, technology is evolving in great strides, the market is becoming global and the socio-economic context of individuals and businesses is being transformed. All these events allow us to speak of an era of innovation in which the combination of technology and knowledge management becomes a distinctive element of anyone who wants to succeed (Daud, 2012).

We are in the 4.0 era where data, now considered as the new oil of the digital era, are the cause of a great impact on many levels due to the revolution of networks and platforms whose use becomes viral. The main impact of the evolution is on human relationships. All individuals active on the web, being hyper-connected, develop a new way of thinking and deeply modify their social interaction practices.

Digital life integrates with the real one, covering any kind of activity: from the university student who uses social networks to keep in touch with his classmates, to the manager who is always attentive to new trends, to companies that want to improve their relationships with their consumers. Why is it safe to say that data plays an important role in this era?

Such an assumption can be affirmed if we consider the presence of particular characteristics within society. The social networks that change the ways of relating, the great transparency of new media, the traceability of every online operation and other elements are all conditions that provide the freedom for data to proliferate and become accessible to third parties.

In addition, the global economy, a superstructure of society, sees firms constantly striving to find new ways to gain competitive advantage. "The ability of firms to access information and create valuable knowledge provides a unique competitive advantage over competitors in this era of evolution" (Sarvan et al., 2011).

What this means is that business interest in data is increasing significantly. This, together with the increase in the number of data and their accessibility makes the role and therefore the value of the data itself more and more central. That's why the collection of data to obtain new knowledge and, consequently, the role of data itself see their specific weight increase.

Therefore, we can say that yet another relevant change of this era concerns another type of relationship: we are talking about the relationship between economic organizations and the market, which from one-way becomes bi-directional. The emergence of social and new economic configurations requires organizations to constantly adapt their strategies and behaviours according to the information provided by the business environment (Caputo et al., 2015), referring with this last term to all partners and subjects that have effect on the success of a company: Customers, suppliers, shareholders, employees, sub-contractors, potential local and global regulators and others (Perko et al., 2015).

In the past, in order to acquire information, it was necessary to have a direct meeting with the subject of reference; nowadays, thanks to the new technological opportunities, it is possible to generate a much faster and regular flow of knowledge.

Through this paragraph we have touched on what the value of data can be in our society. Let's now focus on the role that data can have for two specific stakeholders: individuals and businesses.

2.1 THE ROLE OF DATA FOR BUSINESS PURPOSES

In recent decades, the business information system has evolved in form from a transaction storage system to a business decision support system at different levels (Jeble et al., 2018). Traditional systems depended on internal information sources within organizations and the databases thus produced were used for decisions related to e.g. inventory management, pricing, identifying losses. Internal data was integrated through complex systems with information derived from partners and customers.

The advent of the Internet makes the process of data integration easier. Information systems find new life through new technological solutions, also as a consequence of the massive volume of data now available. The technologies used in this sense are commonly referred to as "Big Data". These tools allow to organize and process the new volume of data efficiently and effectively through statistical applications.

By using Big Data, managers are able to better predict the right market trends in a smart way, having the possibility to make more success-oriented decisions. But what exactly is ***Big Data***?

There are several avenues and several authors who try to precisely define the concept of Big Data.

Boyd and Crawford in 2012 define big data as a "cultural, technological phenomenon" while Fan et al. in 2014 define it as "an ocean of information". Kitchin on the other hand refers to it as "large volume of structured and unstructured data". A final example, that of Waller et al. in 2013 defines Big Data as datasets that are too large to be analysed through traditional computer systems.

Ultimately, we can say that Big data is all those new data that support business decisions and that, in our time, can be made up of different types of content: voice data, text, log files, images and videos. There are many methods and tools to analyse big data. Companies like Amazon and Netflix have developed algorithms to identify a correlation between consumer searches and past purchases to predict future ones.

From what has been said up to now, one realizes that the term Big data is relatively new, while the act of gathering information and, eventually, analysing it, already originated in the first years of 2000.In those years the analyst Doug Laney articulated the discourse on this quantity of data to be analysed stating that the big data was characterized by the so-called "3 Vs": Volume, speed and variety.

After subsequent definitions, current Big Data integrates this postulate by transforming it into the "5

V's":

- **Volume**: large amount of primary data expressed in terabytes or petabytes;
- **Velocity**: expresses the speed and degree of data accumulation that characterizes any cutting-edge organization;
- **Variety**: refers to a large number of data sources such as enterprise systems, social media, text, video, audio, etc;
- **Veracity**: refers to the quality of the data which is essential for the decisions made to be effective;
- **Value**: economic and social results acquire a higher value.

Once we understand what Big Data are and what their characteristics are, let's try to understand why they are so useful in the business decision-making process.

The information requirements of business leadership have changed in recent years. There are large datasets from structured, unstructured, or semi-structured sources that provide multiple avenues for companies to draw value and make better strategic, tactical, and operational decisions. Data on business transactions, when mined cumulatively, provide key insights for decision makers to sell products and items.

In fact, the treatment of Big Data enables to reconstruct and predict economic crises, epidemics, trends and fashions. In particular, every phenomenon that can be described by data becomes analysable at a microscopic level. Every aspect of human activity can be studied and can be found, bringing to light information and knowledge that cannot be identified in advance. From an economic point of view, by exploiting this opportunity today's companies have the chance to put themselves in a position of advantage over those who are not capable to keep up.

The potential of monetizing these large data flows has been intuited but underestimated by many and economically exploited by few. As it happened for *Google*, which in the beginning was a simple search engine able to collect information from searches made to improve its service and return data as useful and accurate as possible, today it has become a real source of data of all kinds concerning our tastes, preferences, political opinions. The idea is a winning one, it has spread and forms the basis for an entire digital economy: the offer of a free service on one hand and the granting of use of data on the other.

The following table provides clear examples of how in the world of e-commerce, a good amount of data can produce insights and therefore information that help to predict good decisions.

Big Data source	**Big Data Driven Insights**	**Actionable Decision**	**References**
Amazon	- Intention of a consumer to purchase a particular product; - Identify the products sold for each type of clientele.	- Remind the consumer at the next access that on that same page are now present discounts; - Product recommendation.	Amazon.com website
Zannier Group	- Real-time intentions and desires of consumers by relying on two ERPs	- The business can distil retail activity into significant customer purchasing patterns to influence real-time sales and inventory decisions.	ZannierGroup.com
Nordstrom	- Identifies trends and intentions of consumers to buy through Pinterest pins.	- Reminds customers of what to buy (product recommendation); - Creates a trend based catalogue.	BigDatamadesimple

Another way for companies to exploit the potential of data can be to analyse the customer journey of the consumer in order to be able to respond to every need at the exact time and place.

It is clear from these insights how the role of data is of great importance to business operations.

2.2 THE VALUE OF DATA FOR CONSUMERS

We have come to understand that the role of data for businesses is critical on many levels. The same is true for individuals. Technological innovations impact all of society, and individuals also have the opportunity to gain from new conditions.

Around social networks a lot of conversations are generated between users involving the exchange of any kind of information, allowing people to have direct contact with the sources of the information available. Consumers, too, can take advantage of the opportunity to gather information and improve their purchasing decisions.

This is just one example of why the evolving role of data can also be important for individuals. For the purposes of our research, however, what is important to explore is the role or value that individuals place on personal information in the era of big data.

This attributed value is a consequence of the context we are analysing. An environment in which both the public and private sectors have shown interest in Big Data to take value from it (Lim et al., 2018), the implications this has on the consumer may not be so beneficial. Indeed, on the part of consumers, there is increased apprehension about their right to self-determination of information, which may now be violated more frequently than before.

There are few studies that address the issue of the effects of information management on business performance from a consumer-centric perspective. A study by Martin et al., contributes in this sense to the existing literature by going to explicate the feelings of consumers resulting from their vulnerability online, the latter concept that embraces a multitude of salient aspects such as those of privacy concerns or violation of identity or financial damage.

The more companies focus their efforts on collecting and using consumer data, the greater their concern for privacy and potential harm. Vulnerability is something negative that companies do not take into consideration during the process of collecting and interpreting consumer data.

"The most negative effect of third-party processing of one's data is anxiety about potential financial or reputational damage, and the actual disadvantage generated is not" (Fisher, 2013). Martin's study outlines the vulnerability of user data along an axis at the beginning of which is the "vulnerability of data access". This is the first form of consumer weakness that is filed by the company within a detailed digital dossier that implies a "dispersal of information among a wide variety of entities" (Solove, 2003).

Customers limit how and to whom they share information precisely to avoid this weakness, they use disclosure management as a weapon in response to emerging concerns (Acquisti et al., 2012). Although this is the case, companies already possess and continue to search for new volumes of information and this only increases risk sensitivity because individuals are aware that their identity has already been breached a first time ("Data breach vulnerability").

The second form of vulnerability described by the study is the "Spillover vulnerability" which arises when you have the impression that a company similar to another that has already broken through your data wall is about to perform the same operation. This form of vulnerability is less dangerous than the data breach.

Finally, there is the "data minfest vulnerability" which occurs when the consumer's privacy is actually abused. This type of disclosure and fraudulent activities represent the most dangerous form of vulnerability because it transforms the danger of harm to an actual state of harm (Anderson et al., 2013).

In contrast to this research, some empirical studies on *customer data management* have argued that individuals become loyal to a company and therefore share information as a result of their data management process (Moon and White, 20012).

In each case, the consumer response to perceived vulnerability generates negative outcomes for businesses that can take the form of falsification of personal data, the spread of negative Word of Mouth, and the ongoing mutability of their behaviours. Martin's research shows that transparency in the use of data and its control can be "suppressors" of these effects.

Consumer psychology and behavioural response to feelings of vulnerability can be investigated by "Gossip Theory". Gossip is an evolving communication about a third party and researchers recount that about 2/3 of all global public communications are geared towards this type of information (Dunbar, 2004). When people talk about someone else they feel safe and minimize their feelings of vulnerability; when they are the target of Gossip they can react very negatively through a dangerous series of behaviours such as deterioration of trust in the company (Turner, 2003).

Applying Gossip theory to a business context reveals that the vulnerability of customer data can lead to feelings of emotional violation and reduced cognitive evaluations of trust.

In summary, the evidence from this research are 3:

- Consumers perceive the potential harm from processing their data and this leads to their negative behavioural response;
- Gossip theory can be a lens through which to identify these negative behaviours and justify their existence. This is possible because individuals are clear in their minds about how they are perceived and evaluated by others (Richman and Leary, 2009), even if others are businesses. Therefore, when gossip becomes salient it produces a range of negative emotions and responses from the target towards the source (Baumeister and Leary, 1995).
- The elements that, according to Gossip Theory, allow these negative responses to be dominated are the transparency of the data and the granting of control over it.

As we mentioned, one of the possible negative behaviours that online users can engage in is falsifying their personal information. An investigation by Girish Punj seeks to understand whether such an attitude should be considered unethical or justifiable.

As individuals believe that personal information is subject to misuse by third parties, they are inclined to identify unethical attitudes in the activities of businesses. This can lead to a levelling of behaviour between individuals and the market.

Consumers, too, are looking for ways to counteract the phenomena they are forced to endure through behaviour that does not necessarily meet their attitudinal standards.

In general, The disclosure of personal information online is the core of many business interactions between consumers and companies (Maury & Kleiner, 2002). The main routine of companies becomes the analysis of users' online activities to understand the main patterns of their actions, so much so that many e- commerce businesses are built around the monetization of data provided by consumers (Ashworth and Free, 2006).

For these and other reasons, when a consumer shares personal information they are concerned about the traceability of that information (Palmer 2005). When consumers feel they may be violated they try to fight back, and the avenues for doing so may not always prove to be linear and valuable.

This is what happens in the individual in relation to data. We are talking about dangerous consequences because the individual sees their own behaviour and moral standards change because of the role that data plays in society. They are not always willing to concede their personal information without reacting. The value that the individual assigns to the data, therefore, is a value that is sensitive

to the way they are treated. If people trust others they are willing to expose themselves, if they feel in danger they can create problems for business performance.

2.3 DATA AND THE CONCEPT OF PRIVACY

We can summarize what has been said so far by saying that the introduction and evolution of the digital world (instant messaging, video streaming, social networks) has profoundly revolutionized the market, especially the telecommunications and media markets. In the past these markets were characterized by a vertical integration of the different operators, now the relationships are horizontal.

Widespread digitization, accompanied by an unbridled rush to collect and integrate different data sources induces a new quality of data analysis that we have called Big Data.

Especially in the business of sharing economy companies force consumers in the marketplace to expose part of their privacy and to share a lot of information with the public. Information about users' location, details about digital profiles, personal information are given and the companies' management of these data is not always responsible.

Understanding personal data becomes an asset of absolute value in the new economy that allows, for example, companies to personalize the offer to consumers or to target advertisements (Monopol kommission, 2015). These are just some of the utilities derived from data, consequently, it is possible to say with certainty that data assume an economic value that is essential to have excellent performance in the markets.

This value of the data is very difficult to determine. Just think of the acquisition of WhatsApp by Facebook in which it was purchased a service basically already offered for approximately 19 billion USD. This was possible because in addition to the service Facebook bought personal information about 600 million users. It can be said in fact that the world's largest social network has paid about 31.7 USD per user or, rather, per data.

However, when dealing with these issues, one realizes that the most important and negative implication of data is on the level of privacy of the individual, who sees the confidentiality of his or her own sensitive data increasingly questioned.

From a regulatory point of view, the extent to which data represents the basis for power in the marketplace and thus how much access to data should be regulated was analysed.

When data is used to gain a competitive advantage and thus to offer personalized products and services that are better suited to meet the needs of individuals, data collection and processing activities cannot be considered abuse (Acquisti and Varian, 2005), at least from an economic point of view. On the other hand, when data collection and processing is aimed at monetization through sharing with other parties or when companies apply intrusive and unlawful methodologies to obtain them, the abuse of companies becomes evident.

The normative frame work must, therefore, consider not only the financial dimensions of the data, but also aspects that impact on the personal rights of the individual. Precisely for this reason, despite the fact that some territorial areas of the world (e.g. USA) do not provide for any kind of restriction against those who violate the privacy of consumers, except in particular cases, the general worldwide orientation is that which is expressed through the concepts of "right to be forgotten" and "right to data portability", despite the fact that it is really complicated to make such principles effective.

3.0 DATA PRIVACY

We have argued that the effects of widespread access to consumers' personal information are diverse: they include vulnerability to fraud, invasion of privacy, receipt of unwanted marketing communications, high levels of targeting, and intrusive marketing communications that disrupt the rhythm of daily activities.

These trends have enabled the development of a focus by researchers, social critics, and regulators on privacy issues.

In general, the question has shifted from whether consumers are willing to share their private information to how they react now that the same information is widely accessible and available to various third parties.

One respondent in a recent privacy survey says, "I share data every time I leave the house, whether I want to or not. The data is no longer the problem; it's the people who collect and process it who create the danger. Besides, it's too late to put the genie back in the bottle" (Rainie and Duggan, 2016). Although there is no widely accepted definition of privacy, as also stated in the first chapter, the perspective that sees privacy understood as an "individual right" can be shared, and discussions of the "right to privacy" are recently common. The theories that are developed to explain the concepts are also many. Let us analyse some of them.

3.1 PRIVACY AND RATIONALITY OF INDIVIDUALS IN MAKING DECISIONS

Traditional theories suggest that consumers should be able to manage their privacy. Empirical and theoretical research, on the other hand, shows that consumers often lack sufficient information to make appropriate privacy decisions and are wont to negotiate long-term privacy for short-term benefits. (Grossklags, 2005).

From the earliest days to recent incarnations, economic studies of privacy have viewed the individual as an economically rational individual who, when deciding whether or not to disclose personal information, is best able to do so.

In accordance with this assertion, individuals are utility maximisers and are fully informed or otherwise base their decisions on probabilistic outputs randomly distributed in the environment in which they are embedded (Posner et al., 1978). This approach generates a policy debate, in which some believe that individuals and organizations should be allowed to manage privacy trade-offs without intervention by external regulators, allowing anyone to operate according to their own interests.

Although numerous studies have reported a growing concern for privacy among the world's population (Ackerman, 1999), recent studies, evidence, and experiments have shown that actual consumer behaviour is not affected by this trend.

When there are real benefits individuals are willing to negotiate their data in order to get it. The study by Purchasing and Grossklocks attempts to ground some theoretical models to show that the consumer in this process is not rational.

An individual's decision-making process with respect to privacy is influenced by a multitude of factors. Among them, incomplete information, limited rationality, and systematic deviations in psychology suggest that the assumption of perfect rationality does not adequately capture the actual privacy behaviour of consumers.

First, incomplete information influences privacy decision-making due to external factors (when third parties share an individual's information, the individual may be pulled out of the transaction without even realizing it, Varian, 1997), information asymmetries (information relevant to the decision-making process, such as how the data can be used to achieve great success, is not always shared), risks, and uncertainties. The costs and benefits associated with intrusion and privacy protection are

complex, multi-layered and context-specific. They are often constrained by subscription to other products and/or services and are generally only recognizable after privacy has been breached. Second, even if the information reaching individuals were complete, they would be unable to process it optimally.

Especially in the presence of complex and articulated consequences associated with the protection or release of personal information, our innate limited rationality would negatively influence our ability to acquire, store and process all relevant information, leading us to apply only simple theoretical models, approximate strategies and heuristic procedures (Simon, 1998). These elements replace quantitative theoretical approaches with qualitative assessments and "aspirational" solutions that prevent perfect optimization. This impacts the outcome of decisions. Third, even if individuals had perfect rational capacities and adequate abilities to optimize privacy strategies, they would probably still be diverted from a rational strategy. Indeed, numerous economic and psychological investigations have revealed various forms of systematic deviations from rationality that plague individuals and thus their decisions (Kahneman, 2000). For example, in addition to cognitive and computational obstacles, individuals are affected by motivational limitations and misrepresentation of personal utility. Experiments have shown different valuations with respect to losses and gains (generally, losses are felt more than gains when they have the same absolute value), and documented decreased sensitivity for high values of deviation from the status quo. Psychological research has also described how individuals err in predicting their future preferences or draw inaccurate conclusions from past choices according to such research, thus, any of these factors can influence decisions about the domain of privacy, even if not all of them are present. Empirical evidence of this influence on privacy decisions does not automatically imply that individuals make choices against their own interests. It simply implies the possibility of making erroneous decisions that do not necessarily lead to the desired outcome.

This study on the rationality of individuals is very important to understand that the focus of privacy and consumers' willingness to share personal information is precisely on the drivers that lead people to assume a particular behaviour as appropriate, or at least on the psychological dynamics that determine users' choices.

3.2 PRIVACY WITHIN ORGANIZATIONS

We have started to delve into the concept of privacy which, as mentioned, is closely related to the process of data collection and use, of reference in our time. In the previous paragraph, moreover, the theme of privacy and how this concept is integrated into the individual panorama of the consumers was addressed. Through a conceptual passage, now, we go to analyse how the concept of privacy is managed by the companies. Some research, in fact, examines the economics of privacy or, rather, how businesses manage consumer privacy. Understanding the avenues through which companies handle consumer privacy is of paramount importance.

In a survey bridging consumer and marketer privacy preferences, Milne and Bahl (2010) identify both synergies and disconnects between the two views. Unsurprisingly, consumers are more likely to prefer hard-to-surface boundaries between themselves and firms than marketers. This survey compares groups of consumers who are reluctant to new data analytics technologies and groups who are more receptive. Interestingly, the receptive groups are even more likely to accept these new technologies than the marketers themselves, who report a strong desire to remove the barriers mentioned earlier. These results suggest that organizations' privacy practices are likely to deviate, at least to some extent, from customers' desires, again underscoring the need for further research on corporate privacy policy. Other research by Rus et al. (2002) uses theoretical models to map the consumer economics of online privacy. The authors hypothesize a free market system where privacy is unregulated, and identify that in such a setting the degree of consumer privacy erodes until a privacy market is born. This is envisioned as a market where individuals can pay to obtain a certain level of privacy but, as the level of privacy continues to erode, the quality of firm value provides deteriorating exchanges.

Similarly, Conitzer et al. (2012), generate a model in which it is possible to see how firms can use information derived from consumers to discriminate the price of upcoming purchases. The propositions derived from this model indicate that exchanges in which buyers can maintain anonymity are more profitable and that consumers benefit more when anonymity is something that is costly.

Third-party privacy gatekeepers also influence the price to consumers. The last two authors cited find that these gatekeepers worked to the detriment of consumers by negotiating with companies to encourage free anonymity. Together these studies show that organizations depend on basic data privacy protection to regulate the functionality of the overall market system.

Previously, we highlighted the importance of the centrality of data for today's society, but also and above all for businesses. We then went on to describe the role of data for individuals, realizing that this concept is strongly correlated to that of privacy. Through this last analysis, we realized how even in the world of business organizations it is essential to pay attention to the impact of this issue.

3.3 THE ECONOMIC VALUE OF PRIVACY

As we have seen, the concept of privacy in its entirety has repercussions on all the areas taken as reference in this work; let us now analyse the economic value that can be attributed to it.

Also in this case there are many studies in the literature that deal with this subject. But is it really possible to attribute a real and monetary value to the concept of privacy?

Many of the experts who have tried to answer this question, attempting to carry out studies to determine the real value of personal information, have failed to justify the high privacy concerns expressed by individuals, confirming the difficulty in assessing such a phenomenon. A series of experimental auctions were conducted to identify the value that people place on their private data, specifically weight and age information. In these auctions, participants quoted a price for their data and the person who asked for the least received the second lowest asking price (Huberman, 2005). The average asking price for age was $57.56 versus $74.06 for weight. The experiment revealed a tendency for weight information to be valued more highly when it is perceived as embarrassing and, as expected, younger people were more willing to reveal their age than older subjects.

Another interesting study carried out on the real value given by individuals to their personal information is that of Carrascal (2013) who aimed to determine the monetary value of different types of personal information. In this example, the scholars prompted users to value their personal information at the time and place it was generated. In the first phase of the experiment, the plug-in used to record these valuations was intended to collect data on the browsing behaviour of each subject in question. This data was later used to calibrate the plug-in's behaviour in the second phase, in which the plug-in displayed popups as participants browsed the Internet. The popups contained two types of questions: questions about personal information assessment and questions about participants' perceptions and knowledge of privacy. The information assessment questions were framed as auctions. One question, for example, was "What is the minimum amount of money you would accept for selling 10 of the photos you uploaded to this website to a private company?" The experiment ended with a post-study questionnaire, the results of which show significantly low ratings of personal

information, where users value their browsing history at an average of €7. On the other hand, for offline personal information, such as age, address and economic status, the average valuation is around 25 euros. Users provided higher ratings of data related to interactions in social networks (12 euros) and financial websites (15.5 euros), compared to activities such as searching (2 euros) and shopping (5 euros). These studies provide evidence that supports the hypothesis of a paradoxical dichotomy between privacy attitudes and privacy behaviour. People disclose personal information when they benefit from it, but at the same time they are significantly affected by the way this information is handled. In fact, they are seriously concerned about the secondary use of personal information and these concerns lead to prudent behaviour.

3.4 THE PRIVACY CONCERN

As we have seen, in the first part of this paper we introduced the concept of privacy concern which, as we can see from what has been said so far, is transversal to the whole paper. We have defined the online world as a latent context in which most consumers are apprehensive about their data and personal information, considering the latter at risk.

There are so many words said so far that assume an exponential increase in consumer attention to this issue; the increase in individual's sensitivity to the possible breach of their unique data is now an established piece of information.

Having said this, it seems right to go into this topic in more detail, identifying what the scientific literature has proposed on the subject.

The issue of privacy of personal data is recognized as fundamental in both online and offline contexts but the literature is only beginning to grasp the true role of privacy concerns since this variable has been introduced into online models.

Fortes and Rita in 2007 deepen this assumption contextualizing it in the world of E-commerce. Their study focuses on the direct impact that privacy concerns have on online purchase intention. The researchers investigated this issue because they believed that in the scientific literature there was a gap due to the lack of a framework robust enough to explain the reasons that lead privacy concerns to change consumer behaviour. Their goal was then to provide one. The two started from the concept of information privacy, defined as the ability of the individual to control the conditions that regulate the collection by third parties of their personal data (Westin, 1967). The first instrument that the

literature provides for the measurement of privacy concerns related to personal information is the scale called "Concern for information Privacy" (Smith et al., 1996). In this occasion Smith et al. constructed a framework that justified privacy concerns through 5 dimensions:

- Collection, which refers to the concern that data may be collected on a large scale;
- Unauthorized secondary internal use, which concerns situations in which companies use the data collected for reasons other than those explicitly stated;
- Unauthorized secondary external use, which covers situations where data already shared is used by other parties for unspoken purposes;
- Improper accesses, which refers to the misappropriation of data by someone who does not have access to it;
- Errors, which points to the concern of those who do not believe that protection against improper acts is adequate.

With the exception of a few pioneers, until the early 2000s, no study had provided specific results to explain the phenomenon of privacy concerns in the context of the Internet. The exception comes with the study of Malhotra et al. (2004) that applies with modifications the framework of 5 dimensions to the new context. The result is a framework composed of 3 fundamental dimensions:

- ***Collection***, an individual's level of concern about the amount of personal data held by others, relative to the benefits received;
- ***Control***, which reflects the ability of consumers to be active in the process of third parties using and accessing their data;
- ***Awareness***, which reflects the knowledge that the individual has about how the organization handles their data.

Subsequent literature has shown great interest in issues related to online privacy, having seen this construct used in numerous consumer behaviour studies, many of which are anchored in the 'theory of planned behaviour' (TPB) and the 'technology acceptance model' (TAM). These studies show that concern for privacy has a positive influence on perceived risk (Van Slyke et al., 2006) and a negative influence on trust, future purchase *intention,* and online purchase behaviour.

Regarding the study of Fortes and Rita, we will see shortly that the two determined the degree of positive and negative influences of privacy concerns in the world of E-commerce; with regard to the following elaboration it could be instead useful to deepen the two models just mentioned of TPB and TAM through a digression that allows to describe the way in which the individual decides for his own behaviour.

Theory of planned behaviour (TPB)

The theory of planned behaviour is an extension of the theory of reasoned action (TRA), Fishbein and Ajzen (1975). ARF predicts that behaviour is preceded by the intention to put it into practice; this intention is simultaneously determined by attitude toward the behaviour and subjective norms. To overcome the limitations of the theory of reasoned action (TRA), Icek Ajzen introduces a new element, thus formulating the theory of planned behaviour. The added variable consists of perceived behavioural control, that is, the perception that an individual has of being able to enact the desired behaviour. This control goes to affect the intention to perform a given behaviour and the actual behaviour itself. Perceived behavioural control should be differentiated from actual control, which is the actual control the person has over the behaviour. Perceived behavioural control is an indirect measure of it that relates only to subjective perception, not to the individual's actual control over the behaviour. Moreover, perceived behavioural control is a uniquely situational variable, thus related to the context of the individual behaviour considered, which differs from self-efficacy, which is indeed considered a part of it. In any case, self-efficacy is one of the two components of perceived behavioural control. Perceived behavioural control also has a direct influence on the behaviour, but the intention to perform the behaviour also has a major influence on the performance of the behaviour. In this context Ajzen introduces perceived behavioural control into the scheme of relationship and influence of attitude on behaviour.

Technology acceptance model (TAM)

Based on ARF, the technology acceptance model (TAM) was developed with the generic purpose of explaining and predicting the use of information systems by end users. According to Davis, Bagozzi, and Warshaw (1989), TAM is assumed to be a theoretically robust model that can be applied to explain and predict the adoption by multiple user populations of a wide range of computer-based technologies. The model allows tracking the subsequent impact of external variables on an individual's beliefs, attitudes, intentions, and behaviours. According to Davis (1989) and Davis et al.

(1989), the TAM supports two specific beliefs in information technology adoption: perceived usefulness, defined as the individual's belief that the use of certain technologies will improve his or her performance, and perceived ease of use, which refers to the individual's belief that the use of a particular technology will be effortless. Due to its simplicity and ease of application, TAM has become one of the most widely used methods of technology-related behaviour found in the literature, and has been widely adopted in the context of EC (e.g., Crespo et al., 2009; Ha & Stoel, 2009; Palvia, 2009). In this area of research, TAM has been used in its original form and combined with other models of consumer behaviour, including the TPB. Through an integration of the two models described, it might be possible to increase the power needed to explicate and predict users' online behaviour.

After this brief digression we return to analyse what are the implications of Fortes and Rita's research about the concepts of privacy concerns.

Based on the literature review, the two conclude that the concepts of privacy concerns, namely trust and perceived risk are related to TPB and TAM models, which have been widely used, alone or together, in the context of online purchase behaviour. This conceptual framework is articulated to answer the question: what is the impact of privacy concerns on online purchase behaviour?

The confirmed hypotheses allow us to conclude the research by stating that privacy concerns on the internet have a negative impact on various beliefs about the use of E-Commerce. This confirms what we are saying about how the economic superstructure in which we live generates privacy dangers. These dangers are real, they are perceived by consumers, and they can cause a change in the behavior that is put into practice both online and offline.

4.0 BETWEEN EXPRESSED CONCERN AND INDIFERRENCE: CONSUMERS' ACTUAL ONLINE BEHAVIOR

The rise of the Semantic Web has brought with it numerous opportunities, including almost unlimited access to information, connectivity through 24-hour social networking, and large-scale data aggregation. This phenomenon is so crucial that it plays a role in the daily lives of billions of people around the world. At the same time, the advent of big data and digital technologies has also raised significant privacy and security concerns. Considering mobile applications in particular, recent literature argues that consumers' decision to use mobile technologies is primarily driven by considerations of the popularity, usability, and price of a given technology (Kelley et al., 2013; Kim et al., 2008). At the same time, however, research indicates that consumers are concerned about their privacy (Smith et al., 2011). This discrepancy between expressed concern and actual user behaviour is a phenomena known as the privacy paradox: users claim to be very concerned about their privacy, but do very poorly to protect their personal data. There are currently a number of theories explaining the privacy paradox. Some have explained this paradoxical behaviour from a rational perspective, arguing that users weigh the cost-benefit ratio of disclosing information online both consciously and rationally (Simon, 1955). Others have interrogated this rational view by arguing that individuals are limited in their rational decision making by various cognitive biases, resulting in a predeterminable cost-benefit calculation (Simon, 1982). Interestingly, both perspectives result in a risk-benefit calculation that ultimately chooses benefits over risks. Furthermore, an unbalanced decision-making process serves as the basis for a third perspective, in which the decision-making process is based on the prevailing benefits and, consequently, on no or negligible risk assessment.

Most research on the privacy paradox considers general activities performed on the Internet, with a particular focus on e-commerce and social networking activities.

It is a well-documented fact that users have a propensity toward online privacy-compromising behaviours that ultimately result in a dichotomy between privacy attitudes and actual behaviour (Acquisti, 2004; Barnes, 2006). Some degree of risk perception involves increased knowledge of useful strategies for privacy protection, but this is not always realized (Oomen and Leenes, 2008). Thus, although there are many users who show a theoretical interest in their privacy and maintain a positive attitude toward privacy-protective behaviour, this rarely translates into actual protective behaviour (Joinson et al., 2010; Pötzsch, 2009; Tsai et al., 2006). Furthermore, while the intention to limit data disclosure exists, actual disclosure often significantly exceeds the intention (Norberg et al., 2007).

Research on online service providers has shown that concrete privacy decisions and abstract risk awareness are not interchangeable. Privacy decisions do not change in line with changing preferences, which could explain the disparity between stated privacy preferences and actual behaviour (Flender and Müller, 2012). Although users are aware of privacy risks on the Internet, they tend to share private information in exchange for commercial value and personalized services (Acquisti and Grossklags, 2005; Sundar et al., 2013). A similar pattern of users is observed in the context of social networking activities.

The use of various privacy protection strategies, such as restricting access to Wall posts, limiting tags, and sending photos via private messages instead of posting open content, are contrary to the purpose of controlling the flow of information between friends and colleagues.

The implementation of such strategies, however, shows little concern for third-party data collection in the background (Young and Quan-Haase, 2013). Privacy concerns should logically lead to restricting the provision of information in social networks; however, as many users provide personal information seemingly without hesitation, the opposite effect can be observed (Hughes-Roberts, 2012).

When looking at the disclosure of personal information in an app purchase process, Buck et al. (2014) highlight that information collected through one's social group and the is more appropriate than information provided by third parties. While users are able to articulate their privacy needs, the actual decision to use apps does not align with their claims.

Oetzel and Gonja (2011) go further, stating that "privacy is not yet built into the social presentation of a smartphone and thus on these devices one may not have the perception of control over one's privacy."

The debate developed so far has shown a variety of facets of the privacy paradox. In each case, the cited literature discusses the discrepancy between privacy concerns and actual information disclosure from a practical point derived from observations of the contexts of general internet activities, e-commerce, social networking sites, and mobile applications.

This study discusses an issue that is very important to us, especially for the type of research we intend to conduct. As introduced in the previous chapter, the reference context of the phenomenon that we are going to analyse, in fact, is that of online shopping in the fashion industry, a trend that year after year gathers more and more adhesions; according to data from the B2C e-Commerce Observatory, in 2019 the online fashion industry touched 3.3 billion euros in Italy, with a growth of +16% compared

to the previous year. For this reason, therefore, we are going to delve into the issues of the privacy paradox within this context; with the aim of analysing in more depth the perception of online shoppers about their privacy and personal data.

4.1 FACTORS FAVOURING SELF-DISCLOSURE

One of the few conceptual certitudes in the literature about why the dichotomy between potential disclosure intentions and actual online behaviour exists is that consumers are influenced by a cost-benefit trade-off approach to decision making. We have argued extensively above this issue such that our research will be influenced by the theories underlying this approach. Let's go even deeper into the indicated trade-off with the aim of understanding which are the main cost and benefit drivers that the literature has so far investigated and identified. We have already mentioned these in the opening chapter, waiting for this moment to argue further.

Customization

Although personalization provides an advantage to consumers through new ways of communication, recommendations on products and services and more, at the same time it can produce negative effects. Obtaining benefits from this type of driver involves a more extensive release of personal information, reducing the level of privacy.

Personalization has taken on new significance in the digital age and marketing environments. this role has been explored in detail by Chung et al. (2016). The privacy literature originally focused on personalised email communication (White et al., 2008) to examine how website advertising promotes click-through and purchase intent (Bleier and Eisenbeiss 2015; Goldfarb and Tucker 2011). A key question in this regard relates to whether the information is provided consciously or not. This is because marketing now has a great ability to obtain data even covertly through a variety of tools (Aguirre et al., 2015).

Personalization can lead to greater engagement through click-through, but information collection must be disclosed otherwise consumers develop vulnerabilities. The effectiveness of personalization in promoting click-through when consumers are concerned about privacy is absolutely limited (Bleier and Eisenbeiss 2015).

Control

In an experimentation Tucker (2014) demonstrates that people respond more positively to personalized messages when they have more control over their personal privacy settings.

Xu et al. (2012) argues that control over consumers' self-perceived information is the focal mechanism through which various approaches such as self-protection, industry self-regulation, and government mandates lower privacy concerns.

Martin et al. (2016) claim that control can suppress individuals' feelings of vulnerability, promoting trust and reducing negative feelings.

Trust

In contexts where privacy is prominent, trust can provide positive marketing outcomes through its ability to convince consumers to consensually disclose information. In addition to being referred to as an antecedent for privacy issues, trust has been examined as a driver that enables consumers to actively participate in the online and offline activities of organizations (Aiken and Boush 2006). Thus, trust also plays a dominant role in the scientific literature. Wirtz and Lwin's (2009) research for example suggests precisely that trust simultaneously promotes information disclosure and the evolution of increasingly close firm-consumer relationships.

Recently it has been stated that trust plays a positive role in alleviating privacy concerns, and, for this reason, the driver "trust" will be the subject of our research, in particular, we will analyse what triggers consumer trust for a given website/e-commerce.

Incentive and lottery

Consumers can be willing to divulge their sensitive data in exchange for monetary inventions such as can be coupons, sweepstakes and more. These items are sometimes considered sufficient to balance privacy concerns. Typically, companies use these tools to obtain information such as phone numbers or dates of birth.

Milne and Gordon's (1993) study gives proof that monetary incentives are even more significant to shoppers than message relevance. Such inducements can dramatically increase consumers'

willingness to receive cell phone advertising (Tsang, Ho, and Liang 2004) and to provide personal information online (Hui, Teo, and Lee 2007).

In the area of relationship marketing, several studies support the value of incentives in guiding consumers to obtain and maintain relationships with companies (De Wulf et al., 2001)

Nevertheless, not all empirical results fully support the positive role of incentives (Xie, Teo, and Wan 2006). This may be because consumers experience incentives individually and are induced to choose only those incentives (e.g., Pick et al., 2016).

The distinction that may exist in the effects of monetary and lottery incentives is also analysed. While monetary incentives provide an immediate financial benefit, in a lottery there is only a small chance of winning. Therefore, we expect a perceived difference between direct monetary incentives and lotteries.

5. CONCLUSIONS

This part of the research plays a fundamental role for the purposes of our research that has as one of the main objectives to contribute with value to expand the horizons of the scientific community. We wanted to outline the advantages and limitations of the literature produced regarding the topics discussed so as to identify a gap that can be covered.

Furthermore, since our research has an initial exploratory nature as repeatedly stated, it was essential, before empirically analysing the research and analysis process, to delve into the topics in order to refine their knowledge and be more comfortable in handling such topics.

In this sense, what emerges most from this process of analysis of the scientific review are some basic evidences.

The literature regarding self-disclosure is extensive and well-established. The privacy paradox has been analysed in many of its conditions and effects but some limitations emerge that we will try to exploit.

The theory of discrepancy between concerns and actual online behaviours will be extended by investigating how individuals manage their approach to online shopping and throughout their buying journey. This investigation will also try to deepen the knowledge and the effects of some behavioural drivers that have been only moderately studied so far.

CHAPTER 3

Scientific research: "Drivers that encourage self-disclosure of sensitive data in the world of online shopping".

The continuous exposure of Web 4.0 users to new risks of privacy violation puts online users in a condition of uncertainty about the sharing of their personal data on the web. If on one hand the benefits associated to the online activity could justify the loss of control of one's own unique information, on the other hand the high probability that third party business subjects will access to the collection and treatment of such data makes the user's choice about the final behaviour to undertake complex.

We have defined "self - disclosure" as "the process of transmitting data that, first were unknown, then become shared knowledge" (Jourard and Lasakow, 1958).

The high level of interest from businesses and third parties to have such knowledge available means that for data disclosure to occur, a number of factors must be present to motivate consumers to release. The business implications of these issues are many.

Precisely because of its relevance, it was chosen to investigate within this thesis the phenomenon of the behavioural dichotomy between self-disclosure and privacy concerns.

In the first part of the research, the topic was framed in such a way as to enhance its relevance for the different addressees: we told about the centrality of data in the current social and economic context, and how the amount of available information is growing exponentially. We described the meaning of data privacy and how it has evolved and spread to become a very sensitive issue for the individual. Continuing, the phenomenon of the privacy paradox was described, generated by the distance that exists between the concerns about this issue and the real propensity of the individual to share data online. Finally, it was highlighted the importance of investigating the psychology of the consumer and to identify the factors behind the phenomenon.

In the second part, on the other hand, a precise analysis of the scientific literature on the subject was carried out: the aim was to identify theoretical models that would be a reference for research. In fact, it was necessary to have conceptual foundations that could subsequently be associated with empirical evidence. What emerged as most important is the approach that individuals take to their behaviour: users are willing to negotiate their personal data in exchange for specific benefits. In order to give meaning to the privacy paradox dichotomy, therefore, it was necessary to research and eventually

measure the drivers that allow the individual to lean on one side or the other of the balance of costs and benefits.

An element already validated by the literature is the possibility that the context in which information is shared can be a determining factor in the implementation of the individual's behaviour.

For this reason, it has been decided that the survey that we will present in this last part of the research will be limited to a specific context that is particularly current and relevant to the topic. This context will be that of online shopping platforms, particularly in the fashion sector. This will give the paper a unique topicality and a high level of differentiation compared to related studies already carried out.

Finally, in this chapter, after specifying the objective of the investigation, we will get into the heart of the scientific research.

1. OBJECTIVE OF THE RESEARCH

The research process of the experimental thesis began with our need to implement the knowledge of the analysed context and to touch on what users think and say about the problems of privacy concern and self-disclosure in relation to personal dimensions of the individual.

The intention is to describe how the final relative behaviour of users of online shopping platforms/ s is conditioned by a set of factors, positive and negative, that interact with each other, and that push consumers themselves to disclose personal and delicate set of information.

Quantitative analysis, in practice, has been the main method of a series of studies aimed at giving a logical solution to the phenomenon of privacy paradox. This exploratory process has allowed us to further improve familiarity with the issues, while at the same time giving us the possibility to create a direct link between the theoretical abstractions that describe the topic and the real feelings of those who experience this phenomenon. It is in this way that specific research questions will be defined.

This research required such an initial approach because, although the topic has been extensively addressed by the scholarly community, the phenomenon of the "privacy paradox" still lacks an official formulation and a formally established theory underlying it. But how was this quantitative investigation carried out?

2. RESEARCH QUESTIONS

As stated several times, the general goal of this thesis is to investigate the phenomenon we have talked about so much, namely that of "self - disclosure", within the context of online shopping platforms.

The intention is to describe how the final relative behaviour of the user of online shopping platforms is conditioned by a set of factors, positive and negative, that interact with each other.

In this paragraph we will deepen this objective, making it more specific, and we will define the research questions of reference.

The most important thing to emerge from the literature is the need to approach the phenomenon by measuring various factors and considering them as "weights on a scale of effects".

In fact, even considering the cost-benefit perspective discussed above, one realizes that the final behaviour of the user depends on a balance of different drivers. The way and intensity in which these relate can determine behaviours that are at odds with the growing concerns for privacy typical of our age.

To conduct our research and to differentiate it, we identified the following variables as the focus of our research:

- *Independent variable*: Brands' website appearance & trustworthiness
- *Dependent variable:* Willingness to disclose information
- *First moderator*: Consumers social media usage
- *Second moderator:* Cookies Disclosure

Based on these three drivers I formulated the three fundamental hypothesis of these research:

- *Online shoppers are positively influenced by the aesthetic of websites, hence, are more likely to disclose themselves when the brand and its related website look visually trustworthy. (Main effect)*
- *Online shoppers who regularly use social media are positively influenced by them, hence, the frequent use of social media moderates the main effect presented above. (First moderator)*
- *When Cookies policies are showcased on the website, online shoppers are more likely to disclose themselves. (Second moderator)*

In this way we have defined in detail the objective of our thesis, in the next paragraphs we will describe in more detail the steps that allowed us to arrive at a result.

3. METHODOLOGY

The quantitative analysis was carried out through a collection of primary data obtained thanks to the administration of two online questionnaires divided into a pre-test and a final test, and thanks to a subsequent statistical analysis of their meaning carried out with the support of the SPSS software.

Both questionnaires were generated according to the previously observed dimensions and the items defined through the analysis of the scientific literature. Following this criterion, the questionnaires are composed of different sections, each of which focuses on the reference items. These items are described by Likert scales (from 1 to 7), below we will see in detail the dimensions, items, scales and sources of the questionnaires:

The objective of the surveys was to collect primary data for each of the identified factors. The object of the surveys was to collect primary data for each of the identified factors. The questionnaire was intended for a heterogeneous user audience targeting different age groups, with the aim of better understanding whether this factor could possibly affect the final result.

The questionnaires were created through the Qualtrics platform and shared mainly through social networks (Whatsapp, Linkedin, Facebook and Instagram). The methodologies used allowed us to administer the questionnaire to approximately 260 individuals.

The questionnaires were generated both in Italian and English language as we did not want to limit the target sample with geographical criteria. The only objective was to collect responses from those familiar with the world of online shopping whatever the demographic characteristics of the respondent.

3.1 SURVEY: pre-test

As already mentioned, to ensure the effectiveness of the final test, it was necessary to administer a pre-test to a smaller audience of about 60 users. The pre-test was based on the administration of two images and our main objective was to verify that these images were perceived as different, so that we could then continue with the final test.

Below are the dimensions, items and images used for the following pre-test:

Construct	Sample Items
➢ Trustworthiness https://www.sciencedirect.com/science/article/pii/S109499680270159X)	- This website looks like one I can believe in; - I do NOT think there is much risk involved with purchasing items on this website; - I can trust this website to keep my best interest in mind; - I find it necessary to be cautious with this website (reverse scored).
➢ Socio demographics GENDER - What is your gender? AGE - What is your age? EDUCATION - What is the highest degree or level of school you have completed? If currently enrolled, highest degree received. EMPLOYMENT STATUS - Are you currently?	- Male; - Female; - Not binary, third gender; - Prefer not to tell. - 18-25 - 25-30 - 30-35 - >35 - High school graduate, diploma or equivalent; - Bachelor's degree; - Master's degree; - Professional degree; - Doctorate degree. - Employed for wages - Self-employed - Out of work and looking for work - Out of work but not currently looking for work - A student

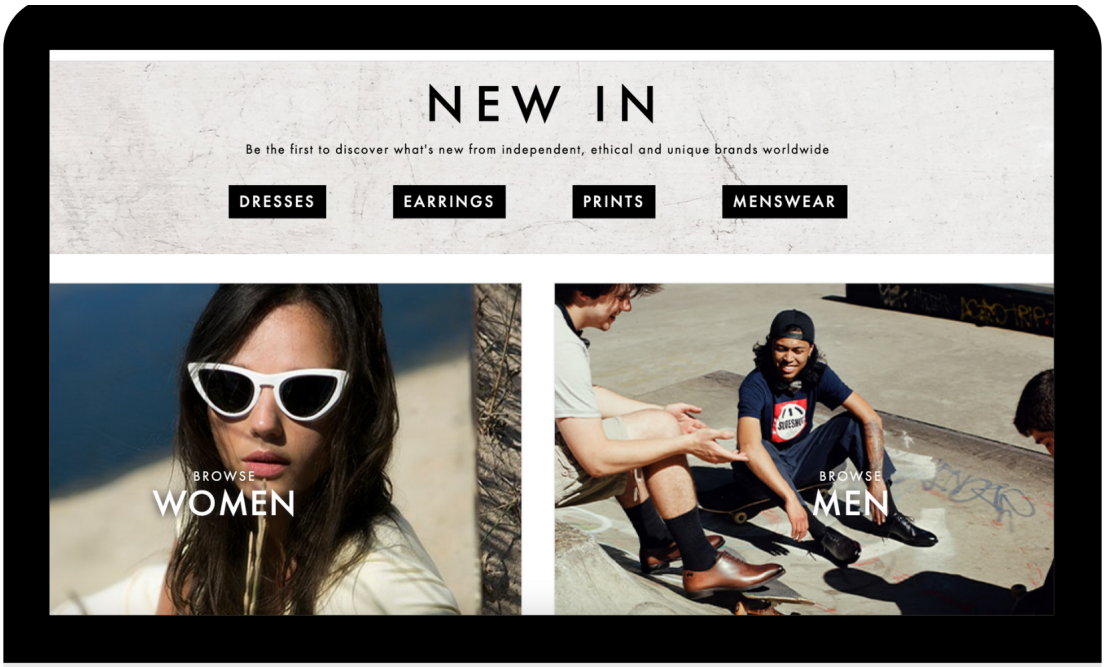


Figure 01. trust website

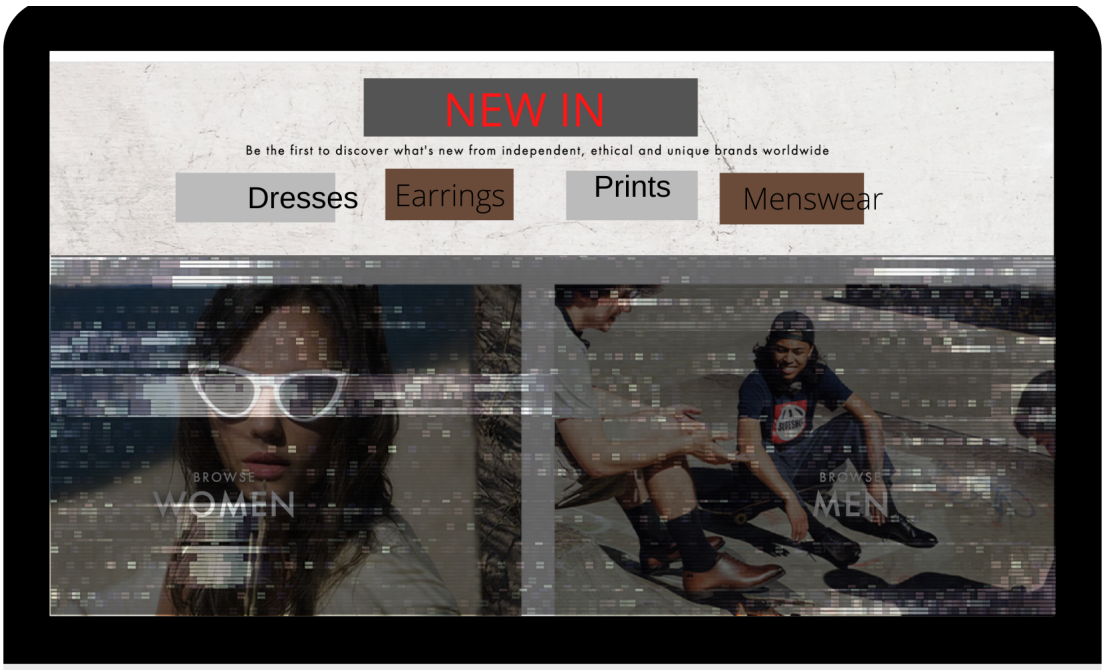


Figure 02. no trust website

The first image represents a well-maintained website which aims to be trustworthy and therefore inspire confidence in consumers. The second image, on the other hand, represents a poorly maintained website which aims to be unreliable and therefore not inspire consumer confidence.

By building the questionnaire through the Qualtrics platform, we were able to administer the two images in a randomised manner in order to increase the effectiveness of the test. Once the target audience had been reached, the data were collected in numerical form, using an excel report, and then reported to the SPSS software to perform an independent samples t-test.

Below are the results obtained:

The number of individuals participating in the questionnaire was found to be 50% between the ages of 18-25, 30% between 25-30 and the remaining 20% between 30-35. In addition, the majority of users were male and predominantly students.

The scale reliability has been tested through the Cronbach analysis. The results are excellent, with a value of =.09

Moving to the results of the pre-test, as it is possible to observe from the images below, the t-test with independent samples was statistically significant, that is, observing the second table we will see that the two-tailed Sign. is <.05, (figure 3) therefore, the two images, also defined as stimuli, were perceived by the users as different. In particular, the well maintained website is perceived trustworthy with a mean of 5.94, while the other with a mean of 2.14, thus confirming that the two stimuli are differently perceived.

Test t

Statistiche gruppo

	STIMOLO	N	Media	Deviazione std.	Media errore standard
TRUSTMEAN	1	31	2,1398	1,46761	,26359
	2	30	5,9444	,92261	,16845

(Figure 3)

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie						
									Intervallo di confidenza della differenza di 95%	
		F	Sign.	t	gl	Sign. (a due code)	Differenza della media	Differenza errore standard	Inferiore	Superiore
TRUSTMEAN	Varianze uguali presunte	3,695	,059	-12,075	59	<,001	-3,80466	,31508	-4,43514	-3,17418
	Varianze uguali non presunte			-12,163	50,750	<,001	-3,80466	,31282	-4,43274	-3,17658

(Figure 4)

Statistiche elemento-totale

	Media scala se viene eliminato l'elemento	Varianza scala se viene eliminato l'elemento	Correlazione elemento-totale corretta	Alpha di Cronbach se viene eliminato l'elemento
Trust Scale_1	8,00	20,900	,979	,995
Trust Scale_2	8,05	20,614	,992	,986
Trust Scale_3	8,02	20,683	,987	,990

(Figura 5)

Thanks to these results it was possible to proceed with the administration of the final test, which we will discuss in the next section.

3.2 SURVEY: main test

As we have seen in the previous paragraph, before moving on to the main test, it was necessary to validate and ensure that the images used were reliable and meaningful. Once this had been validated, it was possible to move on to the construction of the main test, or final test, which will allow us to subsequently validate the hypotheses formulated previously.

For the construction of the questionnaire, the Qualtrics software was used once again and we once again followed the dimensions and the items previously reported with the addition of two other items that we will report. Following this criterion, the questionnaire consists of three different sections, each of which focuses on the reference items. These items are described by Likert scales (from 1 to 7; of which 1 represents strongly disagree and 7 represent strongly agree).

The questionnaire was shared through social networks, forums and blogs used; thanks to these methodologies we were able to reach and administer the questionnaire to an audience of about 210 individuals of 12 different nationalities, the majority of which turns out to be of an age between 18-25, a factor that will later prove useful to give validity to our hypotheses.

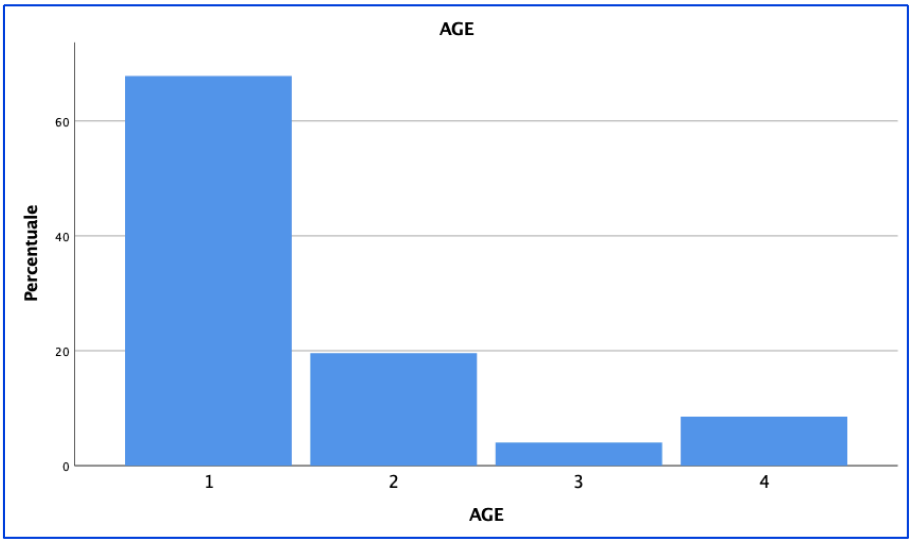


Figura 06

The questionnaire was generated again in both Italian and English, as we did not want to limit the target sample in any way by geographical criteria.

Below are details of the dimensions, items, scales and sources of the questionnaire:

Construct	Sample Items
➢ Social media usage Davis(1989) Cheng et al.(2006)	- Using social platform makes it easier for me to find what I am looking for - I find social media useful to find targeted products - Overall, I find using social media to be useful
➢ Intention to disclose information Salisbury et al.(2001) Cheng et al.(2006)	- I am willing to provide this company with information about me. - I am willing to provide this company with information about my product needs. - I feel secure putting my information on this website; - Online platforms are a safe place for me to shop; - I think the information given to the website will not be misused.

➢ Socio demographics GENDER - What is your gender? AGE - What is your age ? EDUCATION - What is the highest degree or level of school you have completed? If currently enrolled, highest degree received. EMPLOYMENT STATUS - Are you currently?	- Male; - Female; - Not binary, third gender; - Prefer not to tell. - 18-25 - 25-30 - 30-35 - >35 - High school graduate, diploma or equivalent; - Bachelor's degree; - Master's degree; - Professional degree; - Doctorate degree. - Employed for wages - Self-employed - Out of work and looking for work - Out of work but not currently looking for work - A student

In summary, the experiment is therefore composed of 3 parts

- Section 1, which investigates the first moderator, i.e. "Social Media Usage";
- Section 2, which simultaneously investigates both the main effect "Website appearance" and the second moderator "Cookies Disclosure"; on the intention to disclose information.
- Finally, the last section was devoted to the collection of socio-demographic data.

4. DISCUSSION: analysis of the results

After exporting the questionnaire responses in numerical form from the Qualtrics software, the data set was subsequently cleaned of outputs that contained potential missing values, values that would make our research less credible, and made easier to interpret by assigning the right labels. After this, it was possible to use the SPSS software to obtain our final results. Let us analyse them construct by construct.

The first macro-dimension we are going to analyse is social media usage. As mentioned, we begin by determining the reliability of the item scales and analysing the results of the averages.

The Cronbach's alpha (see figure 7) of the scale in question is .850 - very good. Furthermore, no items should be deleted as there is no value that by eliminating it would guarantee a better result.

Statistiche di affidabilità

Alpha di Cronbach	N. di elementi
,850	3

(Figure 7)

Once the reliability analysis has been carried out on the social media usage scale, we now check the reliability of the intention to disclose information scale (see figure 8). In this case, the Cronbach's alpha is 900, but if we omit the fourth item, we obtain a better result, i.e. 913 (figure 9); this means that to calculate the average value we will use all the items except the fourth one.

Statistiche di affidabilità

Alpha di Cronbach	N. di elementi
,900	5

(Figure 08)

Statistiche elemento-totale

	Media scala se viene eliminato l'elemento	Varianza scala se viene eliminato l'elemento	Correlazione elemento-totale corretta	Alpha di Cronbach se viene eliminato l'elemento
website protection_1	15,76	33,881	,826	,861
website protection scale_2	14,98	35,131	,750	,879
website protection scale_3	15,74	34,323	,861	,853
website protection scale_4	14,95	42,609	,569	,913
website protection scale_5	15,41	37,323	,766	,875

(Figure 9)

Thanks to these results, we were able to proceed with the analysis of the effect of the independent variable that is the mere visual reliability of the website, and the first moderator, i.e. the individuals' use of social media on the consumers' willingness to disclose information. In particular, we first focused on what happens when the group of users is subjected in a randomized manner to the two images seen above in the pre-test (Figures 1 & 2) and how their regular use of social media was going to influence their response towards their willingness to disclose personal information.

To do this, a two-way anova analysis was performed, i.e., we relied on the use of the general univariate linear model; however, before arriving at the model construction, it was necessary to reproduce the values of the social media usage scale in binary codes and, therefore, calculate the median in order to establish a low or high level of social media usage. The median was calculated to be 6 (Figure 10), meaning that the majority of respondents regularly use social media.

Statistiche

SCuse_mean

N	Valido	93
	Mancante	0
Mediana		6,0000

Figura 10

Once this was established we recoded the variable social media usage in 1= low usage and 2= high usage, and we moved on to the 2-way anova analysis, which showed us that 45 respondents were subjected to the reliable image (figure 1) while 48 respondents were subjected to the unreliable image (figure 2).

		N
CONDIZIONE	2	45
	4	48
SCuse_REC	1	68
	2	25

(Figure 11)

CONDIZIONE	SCuse_REC	Media	Deviazione std.	N
2	1	4,8750	1,09985	32
	2	4,5577	1,11409	13
	Totale	4,7833	1,10088	45
4	1	2,7639	1,33222	36
	2	3,4375	1,57799	12
	Totale	2,9323	1,41114	48
Totale	1	3,7574	1,61671	68
	2	4,0200	1,44503	25
	Totale	3,8280	1,56905	93

(Figure 12)

From Figure 12 we can observe that respondents subjected to the trustworthy image would be willing to disclose their information with a mean of 4.78; on the contrary, subjects subjected to the unreliable image would be willing to give their information with a mean of 2.93. With a significance of p=,000 and therefore less than p=0.5 we can widely affirm the validity of our first hypothesis: Online shoppers are positively influenced by the aesthetic of websites, hence, are more likely to disclose themselves when the brand and its related website look visually trustworthy.

But what changes when we also consider the first moderator, social media usage?

Let's take a look at the results in detail:

- People subjected to the trustworthy image and who use social media less said they would be willing to disclose their data with an average of 4.87; while those who use social media more with an average of 4.5. (Figure 12)
- On the other hand, those subjected to the unreliable image with less use of social media reported an average of 2.76; while those who use social media the most reported an average of 3.43. (Figure 12)

Checking for significance, we have a non-significant result, indeed the p-value is 0,97 (Figure 13). This result indicates that the use of social media does not moderate the disclosure of information by

the participants in the questionnaire, therefore, our second hypothesis cannot be considered validated. However, it is possible that this finding is the result of an analysis of a sample that is not very heterogeneous either in terms of age, since the majority is between 18 and 25, or in terms of the use of social media which, as we saw earlier, is very high given the median of 6, thus also those identified with a low level of social media usage, in reality are used to social media.

Test di effetti tra soggetti

Variabile dipendente: IntDisclosure_mean

Origine	Somma dei quadrati di tipo III	gl	Media quadratica	F	Sign.	Eta quadrato parziale
Modello corretto	84,594[a]	3	28,198	17,686	<,001	,373
Intercetta	1114,647	1	1114,647	699,095	<,001	,887
CONDIZIONE	47,615	1	47,615	29,864	<,001	,251
SCuse_REC	,579	1	,579	,363	,548	,004
CONDIZIONE * SCuse_REC	4,478	1	4,478	2,808	,097	,031
Errore	141,903	89	1,594			
Totale	1589,250	93				
Totale corretto	226,497	92				

a. R-quadrato = ,373 (R-quadrato adattato = ,352)

(Figure 13)

We continue our analysis by adding our second and final moderator, Cookies. We are going to examine what happens when we add this additional moderator and thus have four different conditions. Here are the additional images used (figures 14 & 15)

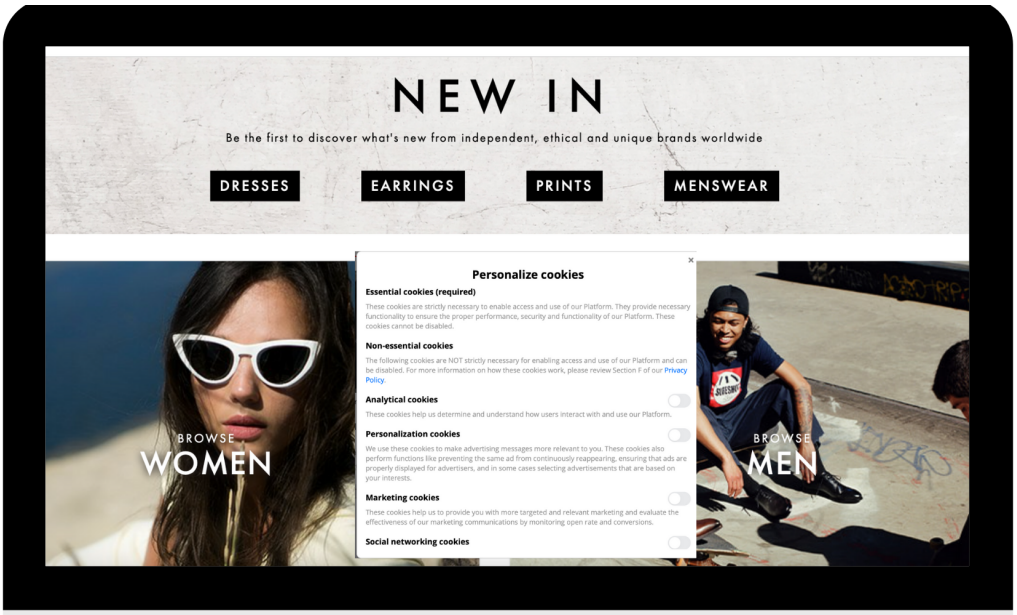


Figure 14 (trust /cookies)

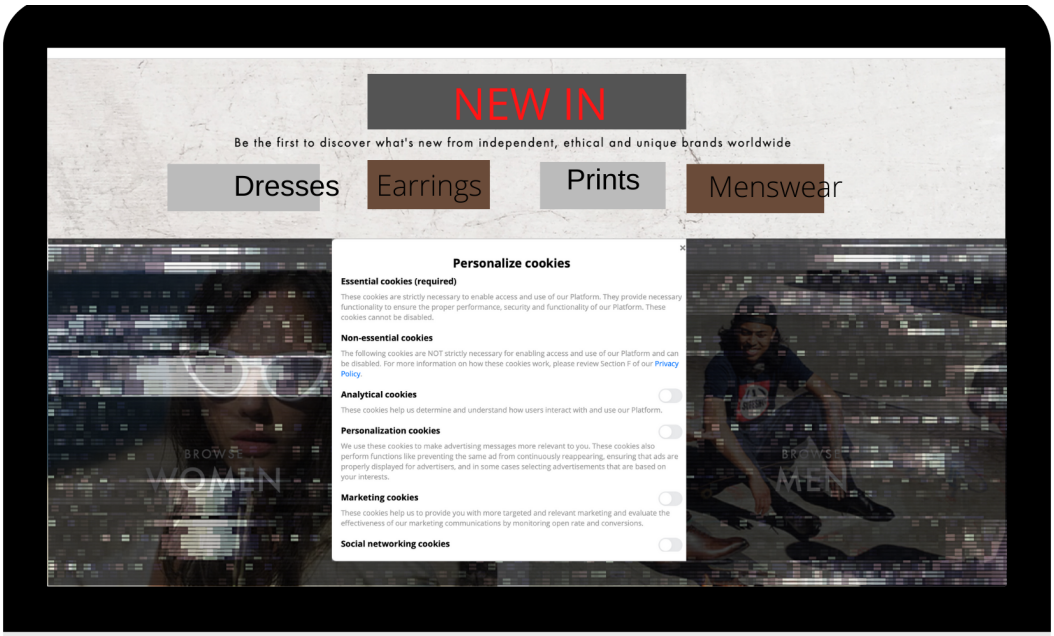


Figure 15 (no trust/cookies)

Statistiche descrittive

Variabile dipendente: IntDisclosure_mean

TRUSTvsNOTRUST	COOKIESYESNO	Media	Deviazione std.	N
1	1	2,9323	1,41114	48
	2	2,9545	1,46731	55
	Totale	2,9442	1,43440	103
2	1	4,7833	1,10088	45
	2	4,4216	1,58862	51
	Totale	4,5911	1,38655	96
Totale	1	3,8280	1,56905	93
	2	3,6604	1,68861	106
	Totale	3,7387	1,63189	199

(Figure 16)

In order to analyse data, a two-way anova has been conducted. The presence of cookies has been recoded with 2, while their absence with 1. From figure 16 we can deduce the following results:

- Respondents subjected to the unreliable image without cookies would be likely to give their information with an average of 2.93;
- Respondents subjected to the unreliable image but with cookies would be inclined to give

their information with a mean of 2.95;

- Respondents to the trustworthy image without cookies would be willing to disclose their information with a mean of 4.78;
- Respondents subjected to the trustworthy image with cookies would be willing to disclose their information with a mean of 4.42.

Again, the main effect of the Independent variable is significant, with the trustworthy website provoking a greater willingness to disclose personal information among respondents. Regarding, the presence of cookies, their main effect is not significant, indeed the p value=.399 . Furthermore, looking at the moderation between the type of website and the presence or not of the cookies, the result is non-significant with a p-value of p=.340 (figure17) - which indicates that the interaction between the moderator Cookies and the independent variable is not significant and therefore, the variable cookies does not appear to influence the responses of respondents, thus going to exclude our third hypothesis. However, using this variable led us to an interesting finding, namely, if we dwell on the average of respondents subjected to the reliable image with cookies we will see that the latter is lower (4.42) than the average of those subjected to the reliable image but without cookies (4.48). This could mean that the vision of the Cookies variable has negatively influenced the responses of some respondents, perhaps arousing in them a concern about their personal data that they did not have before.

Test di effetti tra soggetti

Variabile dipendente: IntDisclosure_mean

Origine	Somma dei quadrati di tipo III	gl	Media quadratica	F	Sign.	Eta quadrato parziale
Modello corretto	137,922[a]	3	45,974	23,024	<,001	,262
Intercetta	2817,243	1	2817,243	1410,918	<,001	,879
TRUSTvsNOTRUST	136,181	1	136,181	68,201	<,001	,259
COOKIESYESNO	1,426	1	1,426	,714	,399	,004
TRUSTvsNOTRUST * COOKIESYESNO	1,824	1	1,824	,914	,340	,005
Errore	389,365	195	1,997			
Totale	3308,875	199				
Totale corretto	527,287	198				

a. R-quadrato = ,262 (R-quadrato adattato = ,250)

(Figure 17)

To conclude the analysis of the questionnaire, and to dispel some doubts, we decided to carry out a final calculation by setting the first moderator, i.e. the use of social media as a covariate of the model comprehensive of cookies and no cookies. Below are the results obtained:

Test di effetti tra soggetti

Variabile dipendente: IntDisclosure_mean

Origine	Somma dei quadrati di tipo III	gl	Media quadratica	F	Sign.	Eta quadrato parziale
Modello corretto	155,305[a]	4	38,826	20,249	<,001	,295
Intercetta	185,004	1	185,004	96,486	<,001	,332
SCuse_REC	17,383	1	17,383	9,066	,003	,045
TRUSTvsNOTRUST	138,043	1	138,043	71,994	<,001	,271
COOKIESYESNO	1,457	1	1,457	,760	,384	,004
TRUSTvsNOTRUST * COOKIESYESNO	1,176	1	1,176	,613	,435	,003
Errore	371,982	194	1,917			
Totale	3308,875	199				
Totale corretto	527,287	198				

a. R-quadrato = ,295 (R-quadrato adattato = ,280)

(Figure 18)

As we can see from Figure 18, using the social media variable as a covariate of the dataset, the significance is p=.003 - a good result that leads us to conclude that, despite the previous results, we can say that the use of social media, influences users' responses.

5. LIMITS AND FUTURE RESEARCHES

As mentioned above, we can be satisfied with the results obtained but, at the same time, we are aware that these could be supplemented and improved through further investigations. This thesis, in fact, has some limitations.

Despite the clarity and validity of the tools used to conduct this research, there were some errors and limitations that can be corrected and overcome for future research on Data Privacy and Self-Disclosure in the online world.

Among the main limitations found during the research we find the number of individuals and the heterogeneity of the sample that submitted to the questionnaire. In fact, as we have seen during the discussion of the main test, the number of people we managed to reach was 210 and the majority of them were between 18 and 25 years old; a factor that could negatively have affected the results, especially regarding the moderator variable "usage of social media". Furthermore, with reference to the limitations found, when analysing the influence of the presence of the Cookies variable, the simple screenshot image recreated was not optimal for the type of search. Indeed, a real website with the possibility for the user to accept or not the cookies would be better in order to recreate an environment close to the reality

In any case, the topic is an evolving one that is beginning to gain attention within the scholarly community. Given the limiting elements and considering the possible evolution of such a context and phenomenon, we realise that for those who wish to continue this research it would be easy to see areas for implementation and/or integration.

Personally, I would suggest conducting research that:

- Overcomes the age limitations and have a more heterogeneous sample of respondents;
- Uses more channels with the aim of reaching a wider audience;
- Relies on actual online shopping visual simulation, in order to better analyse the relationship between the Cookies variable and consumers' behaviour and gain more detailed results.

CONCLUSIONS

After a thorough process of investigation, we arrive at the conclusions of our research by trying to interpret the results as objectively as possible.

The paper started from a long way back in order to achieve the defined objective of answering the research questions posed about balancing the effects that influence online self-disclosure.

A very complex context has been described in which the number of personal data available for processing inexorably increases along with the concerns that third party activities generate.

Many scholars have explored the relationship between self-disclosure and privacy concerns as a function of the negotiation between positive and negative drivers, and we too have identified a number of factors that influence consumer behaviour on a daily basis.

Our research tries to overcome some limitations of the scientific literature: The drivers usually identified as determining the occurrence of self-disclosure are control, personalisation, trust and monetary incentives. Our intention was to set aside what we have already achieved and focus on new dimensions (website appearance, social media usage and cookies) trying to realize a new and valid conceptual framework in a research never done before. With the same aim, we set ourselves the objective of contextualising the survey in today's society by exploiting the online shopping trend.

The results partially proved us right.

The statistical analyses carried out allow us to answer the three proposed research questions in a positive way. In fact, as we have seen in the presentation of the results of the main test, our first research question had a positive response in the results, which clearly indicate that the level of self-disclosure of users in the online shopping process is highly influenced by the mere aesthetic appearance of a website and how reliable it is at first glance. With regard to the second research question, we were not able to obtain the desired feedback due to the non-heterogeneity of the sample in question but, thanks to further verifications, we were able to partially demonstrate that the variable of the use of social media in some way influences the level of self-disclosure of online shoppers. Finally, with regard to the last research question, again we found limitations that prevented us from demonstrating the relationship between the presence of Cookies and the level of self-disclosure.

The study conducted can be considered as an extension of the privacy paradox study which states that the level of disclosure of one's own data by the consumer is influenced by a number of positive or negative drivers.

Focusing on the context in which this was demonstrated, it is important to understand other implications of this research. Online is a typical context of the new way of doing business. Digitisation is by now native in every field and knowing that in this context, where the risk of data violation is very high, there are levers that allow users to remain loyal and operational is a piece of information of primary importance.

MANAGERIAL IMPLICATIONS

Consequently, one can imagine the implications that our findings can provide in the managerial field. Business organisations are always looking for drivers to manipulate in order to gain a competitive advantage in the market and now, in the online shopping industry.

The focus of the research was to demonstrate that website appearance and daily social media usage factors can easily influence online consumer behaviour and can be positive factors for companies that have access to user data and personal information.

In this perspective, going back to the phenomenon investigated, that of the pirate paradox, we can assume that companies have an extra element to understand the psychology of the consumer that lies at its base. This is a contradictory phenomenon but one that is based on solid psychological concepts characteristic of people's subjective personalities.

REFERENCES

1. Acquisti A. John L. (2013) What Is Privacy Worth?. Journal Of Legal Studies: 34 - 41
2. Acquisti A. John L. Loewenstein G. (2012) The Impact of Relative Standards on the Propensity to Disclose. Journal of Marketing Research 49: 160 - 174
3. Andrade, E., Kaltcheva, V., Weitz, B., (2002) Self-disclosure on the web: The impact of privacy policy, reward, and company reputation. Advances in Consumer Research 29: 350 – 353
4. Barth S., De Jong M. (2017) The privacy paradox - Investigating discrepancies between expressed privacy concerns and actual online behavior - A systematic literature review. Telematics and Informatics 34: 1038 – 1058
5. Benamati, J., Zafer, O., Smith, J., (2017) Antecedents and outcomes of information privacy concerns in a peer context: An exploratory study. Journal of information science 43: 583 - 600
6. Blank, G., Lutz C., (2018) Benefits and harms from Internet use: A differentiated analysis of Great Britain. New Media & Society 20: 618 – 640
7. Chen D. Fraiberger S. Moakler R. Provost F. (2017) Enhancing Transparency and Control When Drawing Data-Driven Inferences About Individuals. Big data 5: 56 – 87
8. Christofides, E., Muise A. Desmarais S., (2009) Information Disclosure and Control on Facebook: Are They Two Sides of the Same Coin or Two Different Processes? CYBERPSYCHOLOGY & BEHAVIOR 12: 441 - 445
9. Demmers, J., Van Dolen, M., Weltevreden, J., (2018) Handling Consumer Messages on Social Networking Sites: Customer Service or Privacy Infringement? International Journal of ElectronicCommerce 22: 8-35
10. Ginosar A. Ariel Y. (2017) An analytical framework for online privacy research: What is missing? Information & Management 54: 948 – 957
11. Gross, G., Acquisti, A., (2005) Information Revelation and Privacy in Online Social Networks (The Facebook case).Workshop on Privacy in the Electronic Society
12. Grossklads, J., Acquisti, A., (2005) Privacy and Rationality in Individual Decision Making. Security & Privacy: 24 – 30
13. Guragai B. Hunt N. Neri M. Taylor E. (2017) "Accounting Information Systems and Ethics Research: Review, Synthesis, and the Future. Journal Of Information Systems 31: 65 – 81

14. Hallam, C., Zanella, G., (2017) Online self-disclosure: The privacy paradox explained as a temporally discounted balance between concerns and rewards. Computer in Human Behavior 68: 217 - 227
15. Hsin-Yi, H., (2016) Examining the beneficial effects of individual's self-disclosure on the social network site. Computer in Human Behavior 57: 122 – 132
16. Jordaan, Y., Van Heerden, G., (2017) Online privacy-related predictors of Facebook usage intensity. Computer in Human Behavior 70: 90 – 96
17. Junglas, I., Johnson N., Spitzmu¨ller, C., (2008) Personality traits and concern for privacy: an empirical study in the context of location-based services. European Journal of Information Systems 17: 387
18. .Kokolakis,S.,(2017)Privacyattitudesandprivacybehaviour:A review of current research on the privacy paradox phenomenon. Computers & Security 64: 122 – 134
19. Krafft, M., Arden, C., Verhoef, P., (2017) Permission Marketing and Privacy Concerns - Why Do Customers (Not) Grant Permissions? Journal of Interactive Marketing 39: 39 - 54
20. Krishnan M. (2006) The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. MIS Quarterly 30: 13 - 28
21. Krishen, A., Raschke, R., Close, A., Kachroo, P., (2017) A power-responsibility equilibrium framework for fairness: Understanding consumers' implicit privacy concerns for location-based services. Journal of business research 73: 20 – 29
22. Li, Y., (2012) Theories in online information privacy research: A critical review and an integrated framework. Decision Support Systems 54: 471 – 481
23. Li, H., Luo, X., Zhang, J., (2017) Resolving the privacy paradox: Toward a cognitive appraisal and emotion approach to online privacy behaviours. Information & Management 54: 1012 - 1022
24. Markos, E., Milne, G., Peltier, J., (2017) Information Sensitivity and Willingness to Provide Continua: A Comparative Privacy Study of the United States and Brazil. Journal of public policy & marketing 36: 79
25. Martin, K., Borah, A., Palmatier, R., (2017) Data Privacy: Effects on Customer and Firm Performance. Journal of Marketing 81: 36 – 58
26. Martin, K., Murphy, P., (2017) The role of data privacy in marketing, Journal of the Academy of Marketing Science 45: 135 – 155

27. Mosteller, J., Poddar, A., (2017) To Share and Protect: Using Regulatory Focus Theory to Examine the Privacy Paradox of Consumers' Social Media Engagement and Online Privacy Protection Behaviours. Journal of Interactive Marketing 39: 27 – 38
28. Nam T., "Does ideology matter for surveillance concerns? (2017) Telematics and Informatics 23: 134 -156
29. Nam T. (2018) Untangling the relationship between surveillance concerns and acceptability. Journal of Information Management 38: 262 - 269
30. Norberg, P., Horne, D., (2007) The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. The Journal of Consumer Affairs 41: 100 – 126
31. Nottingha, Q., Collignon, S., Warkentin, M., Ziegelmayer, J.,(2015) The interpersonal privacy identity (IPI): development of a privacy as control model. Information Technology and Managementn17: 341 – 360
32. Pagani, M., Malacarne, G., (2017) Experiential Engagement and Active vs. Passive Behavior in Mobile Location-based Social Networks: The Moderating Role of Privacy. Journal of Interactive Marketing 37: 133 – 148
33. Potoglou D. Dunkerley F. Patil S. Robinson N. (2017) (COMPUTERS IN HUMAN BEHAVIOR, 2017) Public preferences for internet surveillance, data retention and privacy enhancing services: Evidence from a pan-European study. Computers In Human Behavior 75: 811 – 825
34. Prince, C., (2018) Do consumers want to control their personal data? Empirical evidence. International Journal of Human-Computer Studies 110: 21 – 32
35. Schuster S. Van Den Berg M. Larrucea X. Slewe T. Ide-Kostic P. (2017) Mass surveillance and technological policy options: Improving security of private communications. Computer Standards & Interfaces 50: 76 - 82
36. Spottswood, E., Hancock, J., (2017) Should I Share That? Prompting Social Norms That Influence Privacy Behaviors on a Social Networking Site. Journal of Computer-Mediated Communication, 22. 55– 70
37. Wu H. Zhang H. Cui L. Wang X. (2017) A Heuristic Model for Supporting Users' Decision-Making in Privacy Disclosure for Recommendation. Security and Communication Networks: 1 – 13
38. Zhao,L.,(2014)DisclosureIntentionofLocation-RelatedInformationinLocation BasedSocialNetwork Services. International Journal of electronic commerce 16: 53 – 59
39. https://www.sciencedirect.com/science/article/pii/S109499680270159X)
40. Salisbury et al.(2001) Cheng et al.(2006)

SUMMARY

LUISS

DEPARTMENT: Management

MAJOR: International Management

SUBJECT: Advanced Marketing Management

DATA PRIVACY AND SELF-DISCLOSURE: AN EXPLORATORY ANALYSIS OF ONLINE CONSUMERS' BEHAVIOUR

SUPERVISOR

Prof. Marco Francesco Mazzù

CO-SUPERVISOR

Carmela Donato

CANDIDATE

Susanna Staiano

721491

ACADEMIC YEAR 2020 /2021

ABSTRACT

The following master's thesis will deal with the topic of 'self-disclosure' of sensitive and personal data, i.e. data that are particularly valuable for those in possession of them. This topic is considered to be of particular importance as it relates to a transversal phenomenon in our society. Nowadays, any individual, whether voluntarily or involuntarily, can be involved in the process of sharing their data with third parties.

The evolution of society and continuous technological innovations provide the conditions for the importance we attach to this issue to make the phenomenon of objective relevance and interest to many. In an economic system in which the infrastructures of power, at various levels, cannot do without a massive volume of information, understanding what are the modalities and motivations that allow the individual to share his data becomes absolutely relevant.

We define "self-disclosure" as "that process through which data that was previously unknown is transmitted and thus becomes shared knowledge" (Jourard and Lasakow, 1958) and identify it as a precondition for any social relationship. The concept just expressed is closely related to the issue of privacy in the ultimate sense of this term: "privacy is the legitimate claim of the individual to determine the extent to which he wishes to share himself with others and his control over the time, place, and circumstances for communicating with others. It is also the individual's right to control information about himself. Privacy is synonymous with the right to be left alone' (Alan Westein).

Self-disclosure and adjacent issues potentially impact on consumers' concerns about personal data privacy. In fact, the individual sees a disproportionate increase of dangers around him. His or her personal information on the network may not be safe. The growth of 'privacy concerns' does not always imply a proactive response from the consumer who, often, performs dichotomous behaviours with respect to the above concerns.

We define the phenomenon that generates discrepancy between potential intentions and actual user behaviour as the "Privacy Paradox". Through first a quantitative research based on two questionnaires, one pre-test and a final main test with an analysis of 260 answers, this paper will investigate some drivers that can give a meaning to this phenomenon.

Scientific research: "Drivers that encourage self-disclosure of sensitive data in the world of online shopping".

The continuous exposure of Web 4.0 users to new risks of privacy violation puts online users in a condition of uncertainty about the sharing of their personal data on the web. If on one hand the benefits associated to the online activity could justify the loss of control of one's own unique information, on the other hand the high probability that third party business subjects will access to the collection and treatment of such data makes the user's choice about the final behaviour to undertake complex.

We have defined "self - disclosure" as "the process of transmitting data that, first were unknown, then become shared knowledge" (Jourard and Lasakow, 1958).

The high level of interest from businesses and third parties to have such knowledge available means that for data disclosure to occur, a number of factors must be present to motivate consumers to release. The business implications of these issues are many.

Precisely because of its relevance, it was chosen to investigate within this thesis the phenomenon of the behavioural dichotomy between self-disclosure and privacy concerns.

In the first part of the research, the topic was framed in such a way as to enhance its relevance for the different addressees: we told about the centrality of data in the current social and economic context, and how the amount of available information is growing exponentially. We described the meaning of data privacy and how it has evolved and spread to become a very sensitive issue for the individual. Continuing, the phenomenon of the privacy paradox was described, generated by the distance that exists between the concerns about this issue and the real propensity of the individual to share data online. Finally, it was highlighted the importance of investigating the psychology of the consumer and to identify the factors behind the phenomenon.

In the second part, on the other hand, a precise analysis of the scientific literature on the subject was carried out: the aim was to identify theoretical models that would be a reference for research. In fact, it was necessary to have conceptual foundations that could subsequently be associated with empirical evidence. What emerged as most important is the approach that individuals take to their behaviour: users are willing to negotiate their personal data in exchange for specific benefits. In order to give meaning to the privacy paradox dichotomy, therefore, it was necessary to research and eventually measure the drivers that allow the individual to lean on one side or the other of the balance of costs and benefits.

An element already validated by the literature is the possibility that the context in which information is shared can be a determining factor in the implementation of the individual's behaviour.

For this reason, it has been decided that the survey that we will present in this last part of the research will be limited to a specific context that is particularly current and relevant to the topic. This context will be that of online shopping platforms. This will give the paper a unique topicality and a high level of differentiation compared to related studies already carried out.

OBJECTIVE OF THE RESEARCH

The research process of the experimental thesis began with our need to implement the knowledge of the analysed context and to touch on what users think and say about the problems of privacy concern and self-disclosure in relation to personal dimensions of the individual.

The intention is to describe how the final relative behaviour of users of online shopping platforms/ s is conditioned by a set of factors, positive and negative, that interact with each other, and that push consumers themselves to disclose personal and delicate set of information.

Quantitative analysis, in practice, has been the main method of a series of studies aimed at giving a logical solution to the phenomenon of privacy paradox. This exploratory process has allowed us to further improve familiarity with the issues, while at the same time giving us the possibility to create a direct link between the theoretical abstractions that describe the topic and the real feelings of those who experience this phenomenon. It is in this way that specific research questions will be defined.

This research required such an initial approach because, although the topic has been extensively addressed by the scholarly community, the phenomenon of the "privacy paradox" still lacks an official formulation and a formally established theory underlying it. But how was this quantitative investigation carried out?

RESEARCH QUESTIONS

As stated several times, the general goal of this thesis is to investigate the phenomenon we have talked about so much, namely that of "self - disclosure", within the context of online shopping platforms.

The intention is to describe how the final relative behaviour of the user of online shopping platforms is conditioned by a set of factors, positive and negative, that interact with each other.

In this paragraph we will deepen this objective, making it more specific, and we will define the research questions of reference.

The most important thing to emerge from the literature is the need to approach the phenomenon by measuring various factors and considering them as "weights on a scale of effects".

In fact, even considering the cost-benefit perspective discussed above, one realizes that the final behaviour of the user depends on a balance of different drivers. The way and intensity in which these relate can determine behaviours that are at odds with the growing concerns for privacy typical of our age.

To conduct our research and to differentiate it, we identified the following variables as the focus of our research:

- *Independent variable:* Brands' website appearance & trustworthiness
- *Dependent variable:* Willingness to disclose personal information
- *First moderator*: Consumers social media usage
- *Second moderator:* Cookies Disclosure

Based on these variables I formulated the three fundamental hypothesis of this research:

- *Online shoppers are positively influenced by the aesthetic of websites, hence, are more likely to disclose themselves when the brand and its related website look visually trustworthy. (Main effect)*
- *Online shoppers who regularly use social media are positively influenced by them, hence, the frequent use of social media moderates the main effect presented above. (First moderator)*
- *When Cookies policies are showcased on the website, online shoppers are more likely to disclose themselves. (Second moderator)*

In this way we have defined in detail the objective of our thesis, in the next paragraphs we will describe in more detail the steps that allowed us to arrive at a result.

METHODOLOGY

The quantitative analysis was carried out through a collection of primary data obtained thanks to the administration of two online questionnaires divided into a pre-test and a final test, and thanks to a subsequent statistical analysis of their meaning carried out with the support of the SPSS software.

Both questionnaires were generated according to the previously observed dimensions and the items defined through the analysis of the scientific literature. Following this criterion, the questionnaires are composed of different sections, each of which focuses on the reference items. These items are described by Likert scales (from 1 to 7), below we will see in detail the dimensions, items, scales and sources of the questionnaires:

The objective of the surveys was to collect primary data for each of the identified factors. The object of the surveys was to collect primary data for each of the identified factors. The questionnaire was intended for a heterogeneous user audience targeting different age groups, with the aim of better understanding whether this factor could possibly affect the final result.

The questionnaires were created through the Qualtrics platform and shared mainly through social networks (Whatsapp, Linkedin, Facebook and Instagram). The methodologies used allowed us to administer the questionnaire to approximately 260 individuals.

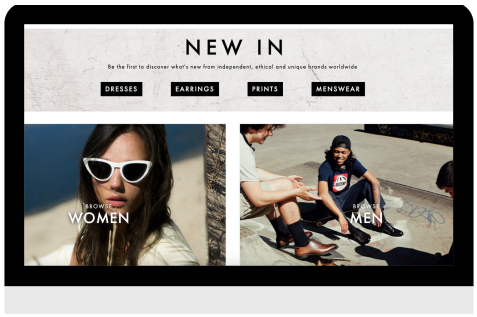
The questionnaires were generated both in Italian and English language as we did not want to limit the target sample with geographical criteria. The only objective was to collect responses from those familiar with the world of online shopping whatever the demographic characteristics of the respondent.

SURVEY: pre-test

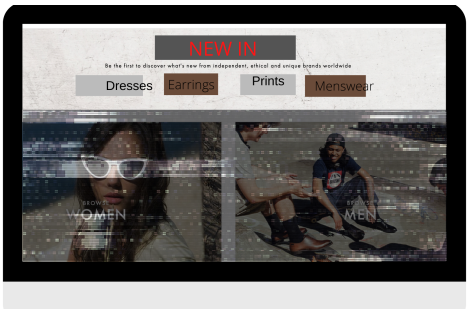
As already mentioned, to ensure the effectiveness of the final test, it was necessary to administer a pre-test to a smaller audience of about 60 users. The pre-test was based on the administration of two images and our main objective was to verify that these images were perceived as different, so that we could then continue with the final test.

Below are the dimensions, items and images used for the following pre-test:

Construct	Sample Items
➢ Trustworthiness https://www.sciencedirect.com/science/article/pii/S109499680270159X)	- This website looks like one I can believe in; - I do NOT think there is much risk involved with purchasing items on this website; - I can trust this website to keep my best interest in mind; - I find it necessary to be cautious with this website (reverse scored).
➢ Socio demographics GENDER - What is your gender? AGE - What is your age? EDUCATION - What is the highest degree or level of school you have completed? If currently enrolled, highest degree received. EMPLOYMENT STATUS - Are you currently?	- Male; - Female; - Not binary, third gender; - Prefer not to tell. - 18-25 - 25-30 - 30-35 - >35 - High school graduate, diploma or equivalent; - Bachelor's degree; - Master's degree; - Professional degree; - Doctorate degree. - Employed for wages - Self-employed - Out of work and looking for work - Out of work but not currently looking for work - A student



(Figure 1)



(Figure 2)

The first image represents a well-maintained website which aims to be trustworthy and therefore inspire confidence in consumers. The second image, on the other hand, represents a poorly maintained website which aims to be unreliable and therefore not inspire consumer confidence

Below are the results obtained:

The scale reliability has been tested through the Cronbach analysis. The results are excellent, with a value of =.09

Moving to the results of the pre-test, as it is possible to observe from the images below, the t-test with independent samples was statistically significant, that is, observing the second table we will see that the two-tailed Sign. is <.05, (figure 3) therefore, the two images, also defined as stimuli, were perceived by the users as different. In particular, the well maintained website is perceived trustworthy with a mean of 5.94, while the other with a mean of 2.14, thus confirming that the two stimuli are differently perceived.

Test t

Statistiche gruppo

	STIMOLO	N	Media	Deviazione std.	Media errore standard
TRUSTMEAN	1	31	2,1398	1,46761	,26359
	2	30	5,9444	,92261	,16845

(Figure 3)

Test campioni indipendenti

		Test di Levene per l'eguaglianza delle varianze		Test t per l'eguaglianza delle medie						
		F	Sign.	t	gl	Sign. (a due code)	Differenza della media	Differenza errore standard	Intervallo di confidenza della differenza di 95% Inferiore	Superiore
TRUSTMEAN	Varianze uguali presunte	3,695	,059	-12,075	59	<,001	-3,80466	,31508	-4,43514	-3,17418
	Varianze uguali non presunte			-12,163	50,750	<,001	-3,80466	,31282	-4,43274	-3,17658

(Figure 4)

Statistiche elemento-totale

	Media scala se viene eliminato l'elemento	Varianza scala se viene eliminato l'elemento	Correlazione elemento-totale corretta	Alpha di Cronbach se viene eliminato l'elemento
Trust Scale_1	8,00	20,900	,979	,995
Trust Scale_2	8,05	20,614	,992	,986
Trust Scale_3	8,02	20,683	,987	,990

(Figure 5)

SURVEY: main test

Before moving on to the main test, it was necessary to validate and ensure that the images used were reliable and meaningful. Once this had been validated, it was possible to move on to the construction of the main test, which would allow us to subsequently validate the hypotheses formulated previously.

For the construction of the questionnaire, the Qualtrics software was used once again and we once again followed the dimensions and the items previously reported with the addition of two other items that we will report. Following this criterion, the questionnaire consists of three different sections, each of which focuses on the reference items. These items are described by Likert scales (from 1 to 7; of which 1 represents strongly disagree and 7 represent strongly agree).

The questionnaire was shared through social networks, forums and blogs used; and we were able to reach and administer the questionnaire to an audience of about 210 individuals of 12 different nationalities, the majority of which turns out to be of an age between 18-25, a factor that will later prove useful to give validity to our hypotheses.

Construct	Sample Items
➢ Social media usage Davis(1989) Cheng et al.(2006)	- Using social platform makes it easier for me to find what I am looking for - I find social media useful to find targeted products - Overall, I find using social media to be useful
➢ Intention to disclose information Salisbury et al.(2001) Cheng et al.(2006)	- I am willing to provide this company with information about me. - I am willing to provide this company with information about my product needs. - I feel secure putting my information on this website; - Online platforms are a safe place for me to shop; - I think the information given to the website will not be misused.

➢ Socio demographics GENDER - What is your gender? AGE - What is your age ? EDUCATION - What is the highest degree or level of school you have completed? If currently enrolled, highest degree received. EMPLOYMENT STATUS - Are you currently?	- Male; - Female; - Not binary, third gender; - Prefer not to tell. - 18-25 - 25-30 - 30-35 - >35 - High school graduate, diploma or equivalent; - Bachelor's degree; - Master's degree; - Professional degree; - Doctorate degree. - Employed for wages - Self-employed - Out of work and looking for work - Out of work but not currently looking for work - A student

DISCUSSION: analysis of the results

After exporting the questionnaire responses in numerical form from the Qualtrics software, the data set was subsequently cleaned of outputs that contained potential missing values, values that would make our research less credible, and made easier to interpret by assigning the right labels. After this, it was possible to use the SPSS software to obtain our final results. Let us analyse them construct by construct.

The first macro-dimension we are going to analyse is social media usage. As mentioned, we begin by determining the reliability of the item scales and analysing the results of the averages.

The Cronbach's alpha (see figure 7) of the scale in question is .850 - very good. Furthermore, no items should be deleted as there is no value that by eliminating it would guarantee a better result.

Statistiche di affidabilità

Alpha di Cronbach	N. di elementi
,850	3

(Figure 7)

Once the reliability analysis has been carried out on the social media usage scale, we now check the reliability of the intention to disclose information scale (see figure 8). In this case, the Cronbach's alpha is 900, but if we omit the fourth item, we obtain a better result, i.e. .913 (figure 9); this means that to calculate the average value we will use all the items except the fourth one.

Statistiche di affidabilità

Alpha di Cronbach	N. di elementi
,900	5

(Figure 08)

Statistiche elemento-totale

	Media scala se viene eliminato l'elemento	Varianza scala se viene eliminato l'elemento	Correlazione elemento-totale corretta	Alpha di Cronbach se viene eliminato l'elemento
website protection_1	15,76	33,881	,826	,861
website protection scale_2	14,98	35,131	,750	,879
website protection scale_3	15,74	34,323	,861	,853
website protection scale_4	14,95	42,609	,569	,913
website protection scale_5	15,41	37,323	,766	,875

(Figure 9)

Thanks to these results, we were able to proceed with the analysis of the effect of the independent variable that is the mere visual reliability of the website, and the first moderator, i.e. the individuals' use of social media on the consumers' willingness to disclose information. In particular, we first focused on what happens when the group of users is subjected in a randomized manner to the two images seen above in the pre-test (Figures 1 & 2) and how their regular use of social media was going to influence their response towards their willingness to disclose information.

To do this, a two-way anova analysis was performed, i.e., we relied on the use of the general univariate linear model; however, before arriving at the model construction, it was necessary to reproduce the values of the social media usage scale in binary codes and, therefore, calculate the median in order to establish a low or high level of social media usage. The median was calculated to be 6 (Figure 10), meaning that the majority of respondents regularly use social media.

Statistiche

SCuse_mean

N	Valido	93
	Mancante	0
Mediana		6,0000

(Figure 10)

Once this was established, we recoded the variable social media usage in 1=low usage and 2= high usage, then we moved on to the 2-way anova analysis, which showed us that 45 respondents were subjected to the reliable image (figure 1) while 48 respondents were subjected to the unreliable image (figure 2).

		N
CONDIZIONE	2	45
	4	48
SCuse_REC	1	68
	2	25

(Figure 11)

CONDIZIONE	SCuse_REC	Media	Deviazione std.	N
2	1	4,8750	1,09985	32
	2	4,5577	1,11409	13
	Totale	4,7833	1,10088	45
4	1	2,7639	1,33222	36
	2	3,4375	1,57799	12
	Totale	2,9323	1,41114	48
Totale	1	3,7574	1,61671	68
	2	4,0200	1,44503	25
	Totale	3,8280	1,56905	93

(Figure 12)

From Figure 12 we can observe that respondents subjected to the trustworthy image would be willing to disclose their information with a mean of 4.78; on the contrary, subjects subjected to the unreliable image would be willing to give their information with a mean of 2.93. With a significance of p=,000 and therefore less than p=0.5 we can widely affirm the validity of our first hypothesis: Online shoppers are positively influenced by the aesthetic of websites, hence, are more likely to disclose themselves when the brand and its related website look visually trustworthy.

But what changes when we also consider the first moderator, social media usage?

Let's take a look at the results in detail:

- People subjected to the trustworthy image and who use social media less said they would be willing to disclose their data with an average of 4.87; while those who use social media more with an average of 4.5. (Figure 12)
- On the other hand, those subjected to the unreliable image with less use of social media reported an average of 2.76; while those who use social media the most reported an average of 3.43. (Figure 12)

Checking for significance, we have a non-significant result, indeed the p-value is .097 (Figure 13). This result indicates that the use of social media does not moderate the disclosure of information by the participants in the questionnaire, therefore, our second hypothesis cannot be considered validated. However, it is possible that this finding is the result of an analysis of a sample that is not very heterogeneous either in terms of age, since the majority is between 18 and 25, or in terms of the use of social media which, as we saw earlier, is very high given the median of 6, thus also those identified with low level of social media usage, in reality are used to social media.

Test di effetti tra soggetti

Variabile dipendente: IntDisclosure_mean

Origine	Somma dei quadrati di tipo III	gl	Media quadratica	F	Sign.	Eta quadrato parziale
Modello corretto	84,594[a]	3	28,198	17,686	<,001	,373
Intercetta	1114,647	1	1114,647	699,095	<,001	,887
CONDIZIONE	47,615	1	47,615	29,864	<,001	,251
SCuse_REC	,579	1	,579	,363	,548	,004
CONDIZIONE * SCuse_REC	4,478	1	4,478	2,808	,097	,031
Errore	141,903	89	1,594			
Totale	1589,250	93				
Totale corretto	226,497	92				

a. R-quadrato = ,373 (R-quadrato adattato = ,352)

(Figure 13)

We continue our analysis by adding our second and final moderator, Cookies. We are going to examine what happens when we add this additional moderator and thus have four different conditions. Here are the additional images used (figures 14 & 15)

Figure 14 (trust /cookies)

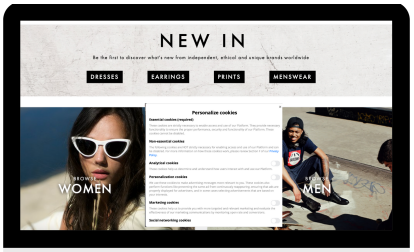
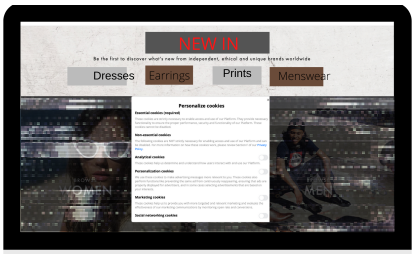


Figure 15 (no trust/cookies)



Statistiche descrittive

Variabile dipendente: IntDisclosure_mean

TRUSTvsNOTRUST	COOKIESYESNO	Media	Deviazione std.	N
1	1	2,9323	1,41114	48
	2	2,9545	1,46731	55
	Totale	2,9442	1,43440	103
2	1	4,7833	1,10088	45
	2	4,4216	1,58862	51
	Totale	4,5911	1,38655	96
Totale	1	3,8280	1,56905	93
	2	3,6604	1,68861	106
	Totale	3,7387	1,63189	199

(Figure 16)

Test di effetti tra soggetti

Variabile dipendente: IntDisclosure_mean

Origine	Somma dei quadrati di tipo III	gl	Media quadratica	F	Sign.	Eta quadrato parziale
Modello corretto	137,922[a]	3	45,974	23,024	<,001	,262
Intercetta	2817,243	1	2817,243	1410,918	<,001	,879
TRUSTvsNOTRUST	136,181	1	136,181	68,201	<,001	,259
COOKIESYESNO	1,426	1	1,426	,714	,399	,004
TRUSTvsNOTRUST * COOKIESYESNO	1,824	1	1,824	,914	,340	,005
Errore	389,365	195	1,997			
Totale	3308,875	199				
Totale corretto	527,287	198				

a. R-quadrato = ,262 (R-quadrato adattato = ,250)

(Figure 17)

In order to analyse data, a two-way anova has been conducted. The presence of cookies has been recoded with 2, while their absence with 1. From figure 16 we can deduce the following results:

- Respondents subjected to the unreliable image without cookies would be likely to give their information with an average of 2.93;
- Respondents subjected to the unreliable image but with cookies would be inclined to give their information with a mean of 2.95;
- Respondents to the trustworthy image without cookies would be willing to disclose their information with a mean of 4.78;
- Respondents subjected to the trustworthy image with cookies would be willing to disclose their information with a mean of 4.42.

Again, the main effect of the Independent variable is significant, with the trustworthy website provoking a greater willingness to disclose personal information among respondents. Regarding, the presence of cookies, their main effect is not significant, indeed the p value=.399. Furthermore, looking at the moderation between the type of website and the presence or not of the cookies, the result is non-significant with a p-value of p=.340 (figure17) - which indicates that the interaction between the moderator Cookies and the independent variable is not significant and therefore, the variable cookies does not appear to influence the responses of respondents, thus going to exclude our third hypothesis. However, using this variable led us to an interesting finding, namely, if we dwell on the average of respondents subjected to the reliable image with cookies we will see that the latter is lower (4.42) than the average of those subjected to the reliable image but without cookies (4.48). This could mean that the vision of the Cookies variable has negatively influenced the responses of some respondents, perhaps arousing in them a concern about their personal data that they did not have before.

To conclude the analysis of the questionnaire, and to dispel some doubts, we decided to carry out a final calculation by setting the first moderator, i.e. the use of social media as a covariate of the model comprehensive of cookies and no cookies. Below are the results obtained:

Test di effetti tra soggetti

Variabile dipendente: IntDisclosure_mean

Origine	Somma dei quadrati di tipo III	gl	Media quadratica	F	Sign.	Eta quadrato parziale
Modello corretto	155,305[a]	4	38,826	20,249	<,001	,295
Intercetta	185,004	1	185,004	96,486	<,001	,332
SCuse_REC	17,383	1	17,383	9,066	,003	,045
TRUSTvsNOTRUST	138,043	1	138,043	71,994	<,001	,271
COOKIESYESNO	1,457	1	1,457	,760	,384	,004
TRUSTvsNOTRUST * COOKIESYESNO	1,176	1	1,176	,613	,435	,003
Errore	371,982	194	1,917			
Totale	3308,875	199				
Totale corretto	527,287	198				

a. R-quadrato = ,295 (R-quadrato adattato = ,280)

(Figure 18)

As we can see from Figure 18, using the social media variable as a covariate of the dataset, the significance is p=.003 - a good result that leads us to conclude that, despite the previous results, we can say that the use of social media, influences users 'responses

These results allow us to come to the final paragraphs in which we will share our views on what has been scientifically demonstrated.

LIMITS AND FUTURE RESEARCHES

Despite the clarity and validity of the tools used to conduct this research, there were some errors and limitations that can be corrected and overcome for future research on Data Privacy and Self-Disclosure in the online world.

Among the main limitations found during the research we find the number of individuals and the heterogeneity of the sample that submitted to the questionnaire. In fact, as we have seen during the discussion of the main test, the number of people we managed to reach was 210 and the majority of them were between 18 and 25 years old; a factor that could have negatively affected the results, especially regarding the moderator variable " social media usage". Furthermore, with reference to the limitations found, when analysing the influence of the presence of the Cookies variable, the simple screenshot image recreated was not optimal for the type of search. Indeed, a real website with the possibility for the user to accept or not the cookies would be better in order to recreate an environment close to the reality

Personally, I would suggest conducting a research that:

- Overcomes the age limitations and have a more heterogeneous sample of respondents;
- Uses more channels with the aim of reaching a wider audience;
- Relies on actual online shopping visual simulation, in order to better analyse the relationship between the Cookies variable and consumers' behaviour and gain more detailed results.

CONCLUSIONS

After a thorough process of investigation, we arrive at the conclusions of our research by trying to interpret the results as objectively as possible.

The paper started from a long way back in order to achieve the defined objective of answering the research questions posed about balancing the effects that influence online self-disclosure.

A very complex context has been described in which the number of personal data available for processing inexorably increases along with the concerns that third party activities generate.

Many scholars have explored the relationship between self-disclosure and privacy concerns as a function of the negotiation between positive and negative drivers, and we too have identified a number of factors that influence consumer behaviour on a daily basis.

Our research tries to overcome some limitations of the scientific literature: The drivers usually identified as determining the occurrence of self-disclosure are control, personalisation, trust and monetary incentives. Our intention was to set aside what we have already achieved and focus on new dimensions (website appearance, social media usage and cookies) trying to realize a new and valid conceptual framework in a research never done before. With the same aim, we set ourselves the objective of contextualising the survey in today's society by exploiting the online shopping trend.

The results partially proved us right.

The statistical analyses carried out allow us to answer the three proposed research questions in a positive way. In fact, as we have seen in the presentation of the results of the main test, our first research question had a positive response in the results, which clearly indicate that the level of self-disclosure of users in the online shopping process is highly influenced by the mere aesthetic appearance of a website and how reliable it is at first glance. With regard to the second research question, we were not able to obtain the desired feedback due to the non-heterogeneity of the sample

in question but, thanks to further verifications, we were able to partially demonstrate that the variable of the use of social media in some way influences the level of self-disclosure of online shoppers. Finally, with regard to the last research question, again we found limitations that prevented us from demonstrating the relationship between the presence of Cookies and the level of self-disclosure.

The study conducted can be considered as an extension of the privacy paradox study which states that the level of disclosure of one's own data by the consumer is influenced by a number of positive or negative drivers.

Focusing on the context in which this was demonstrated, it is important to understand other implications of this research. Online is a typical context of the new way of doing business. Digitisation is by now native in every field and knowing that in this context, where the risk of data violation is very high, there are levers that allow users to remain loyal and operational is a piece of information of primary importance.

REFERENCES

1. Acquisti A. John L. (2013) What Is Privacy Worth?. Journal Of Legal Studies: 34 - 41
2. Acquisti A. John L. Loewenstein G. (2012) The Impact of Relative Standards on the Propensity to Disclose. Journal of Marketing Research 49: 160 - 174
3. Andrade, E., Kaltcheva, V., Weitz, B., (2002) Self-disclosure on the web: The impact of privacy policy, reward, and company reputation. Advances in Consumer Research 29: 350 – 353
4. Barth S., De Jong M. (2017) The privacy paradox - Investigating discrepancies between expressed privacy concerns and actual online behavior - A systematic literature review. Telematics and Informatics 34: 1038 – 1058
5. Benamati, J., Zafer, O., Smith, J., (2017) Antecedents and outcomes of information privacy concerns in a peer context: An exploratory study. Journal of information science 43: 583 - 600
6. Blank, G., Lutz C., (2018) Benefits and harms from Internet use: A differentiated analysis of Great Britain. New Media & Society 20: 618 – 640
7. Chen D. Fraiberger S. Moakler R. Provost F. (2017) Enhancing Transparency and Control When Drawing Data-Driven Inferences About Individuals. Big data 5: 56 – 87
8. Christofides, E., Muise A. Desmarais S., (2009) Information Disclosure and Control on Facebook: Are They Two Sides of the Same Coin or Two Different Processes? CYBERPSYCHOLOGY & BEHAVIOR 12: 441 - 445
9. Demmers, J., Van Dolen, M., Weltevreden, J., (2018) Handling Consumer Messages on Social Networking Sites: Customer Service or Privacy Infringement? International Journal of ElectronicCommerce 22: 8-35
10. Ginosar A. Ariel Y. (2017) An analytical framework for online privacy research: What is missing? Information & Management 54: 948 – 957
11. Gross, G., Acquisti, A., (2005) Information Revelation and Privacy in Online Social Networks (The Facebook case).Workshop on Privacy in the Electronic Society
12. Grossklads, J., Acquisti, A., (2005) Privacy and Rationality in Individual Decision Making. Security & Privacy: 24 – 30
13. Guragai B. Hunt N. Neri M. Taylor E. (2017) "Accounting Information Systems and Ethics Research: Review, Synthesis, and the Future. Journal Of Information Systems 31: 65 – 81

14. Hallam, C., Zanella, G., (2017) Online self-disclosure: The privacy paradox explained as a temporally discounted balance between concerns and rewards. Computer in Human Behavior 68: 217 - 227
15. Hsin-Yi, H., (2016) Examining the beneficial effects of individual's self-disclosure on the social network site. Computer in Human Behavior 57: 122 – 132
16. Jordaan, Y., Van Heerden, G., (2017) Online privacy-related predictors of Facebook usage intensity. Computer in Human Behavior 70: 90 – 96
17. Junglas, I., Johnson N., Spitzmu¨ller, C., (2008) Personality traits and concern for privacy: an empirical study in the context of location-based services. European Journal of Information Systems 17: 387
18. .Kokolakis,S.,(2017)Privacyattitudesandprivacybehaviour:A review of current research on the privacy paradox phenomenon. Computers & Security 64: 122 – 134
19. Krafft, M., Arden, C., Verhoef, P., (2017) Permission Marketing and Privacy Concerns - Why Do Customers (Not) Grant Permissions? Journal of Interactive Marketing 39: 39 - 54
20. Krishnan M. (2006) The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. MIS Quarterly 30: 13 - 28
21. Krishen, A., Raschke, R., Close, A., Kachroo, P., (2017) A power-responsibility equilibrium framework for fairness: Understanding consumers' implicit privacy concerns for location-based services. Journal of business research 73: 20 – 29
22. Li, Y., (2012) Theories in online information privacy research: A critical review and an integrated framework. Decision Support Systems 54: 471 – 481
23. Li, H., Luo, X., Zhang, J., (2017) Resolving the privacy paradox: Toward a cognitive appraisal and emotion approach to online privacy behaviours. Information & Management 54: 1012 - 1022
24. Markos, E., Milne, G., Peltier, J., (2017) Information Sensitivity and Willingness to Provide Continua: A Comparative Privacy Study of the United States and Brazil. Journal of public policy & marketing 36: 79
25. Martin, K., Borah, A., Palmatier, R., (2017) Data Privacy: Effects on Customer and Firm Performance. Journal of Marketing 81: 36 – 58
26. Martin, K., Murphy, P., (2017) The role of data privacy in marketing, Journal of the Academy of Marketing Science 45: 135 – 155

27. Mosteller, J., Poddar, A., (2017) To Share and Protect: Using Regulatory Focus Theory to Examine the Privacy Paradox of Consumers' Social Media Engagement and Online Privacy Protection Behaviours. Journal of Interactive Marketing 39: 27 – 38
28. Nam T., "Does ideology matter for surveillance concerns? (2017) Telematics and Informatics 23: 134 -156
29. Nam T. (2018) Untangling the relationship between surveillance concerns and acceptability. Journal of Information Management 38: 262 - 269
30. Norberg, P., Horne, D., (2007) The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors. The Journal of Consumer Affairs 41: 100 – 126
31. Nottingha, Q., Collignon, S., Warkentin, M., Ziegelmayer, J.,(2015) The interpersonal privacy identity (IPI): development of a privacy as control model. Information Technology and Managementn17: 341 – 360
32. Pagani, M., Malacarne, G., (2017) Experiential Engagement and Active vs. Passive Behavior in Mobile Location-based Social Networks: The Moderating Role of Privacy. Journal of Interactive Marketing 37: 133 – 148
33. Potoglou D. Dunkerley F. Patil S. Robinson N. (2017) (COMPUTERS IN HUMAN BEHAVIOR, 2017) Public preferences for internet surveillance, data retention and privacy enhancing services: Evidence from a pan-European study. Computers In Human Behavior 75: 811 – 825
34. Prince, C., (2018) Do consumers want to control their personal data? Empirical evidence. International Journal of Human-Computer Studies 110: 21 – 32
35. Schuster S. Van Den Berg M. Larrucea X. Slewe T. Ide-Kostic P. (2017) Mass surveillance and technological policy options: Improving security of private communications. Computer Standards & Interfaces 50: 76 - 82
36. Spottswood, E., Hancock, J., (2017) Should I Share That? Prompting Social Norms That Influence Privacy Behaviors on a Social Networking Site. Journal of Computer-Mediated Communication, 22. 55– 70
37. Wu H. Zhang H. Cui L. Wang X. (2017) A Heuristic Model for Supporting Users' Decision-Making in Privacy Disclosure for Recommendation. Security and Communication Networks: 1 – 13
38. Zhao,L.,(2014)DisclosureIntentionofLocation-RelatedInformationinLocation BasedSocialNetwork Services. International Journal of electronic commerce 16: 53 – 59
39. https://www.sciencedirect.com/science/article/pii/S109499680270159X)
40. Salisbury et al.(2001) Cheng et al.(2006)