Czech University of Life Sciences Prague

Faculty of Economics and Management

Department of Information Technologies

Bachelor Thesis

Users' perceptions of online privacy issues

Iskandar Musaev

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Economics and Management

BACHELOR THESIS ASSIGNMENT

Iskandar Musaev

Informatics

Thesis title

Users’ perceptions of online privacy issues

Objectives of thesis

The main objective of the thesis is to investigate differences between students and working professionals in perceptions of online privacy issues.

Partial objectives:
- To review existing models of privacy and security online.
- To prepare and conduct a survey among users.
- To test the differences between students and working professionals, evaluate results and interpret findings.

Methodology

The methodology of solving the theoretical part of the diploma thesis will be based on the study and analysis of the literature. Based on the knowledge gained in the theoretical part of the work, the practical part will conduct a survey among students and working professionals. Statistical data of survey data and interpretation of the results through a relevant theory will be done. The author will formulate the conclusion by synthesizing findings from the literature review and survey data analysis.

The proposed extent of the thesis

50 pages

Keywords

Privacy, policy, perception, safety, awareness, students, professionals.

Recommended information sources

ALMARZOOQI, Fatima Mohamed; MOONESAR, Immanuel Azaad; ALQUTOB, Raeda. Healthcare professional and user perceptions of eHealth data and record privacy in Dubai. Information, 2020, 11.9: 415.

CHIN, Erika, et al. Measuring user confidence in smartphone security and privacy. In: Proceedings of the eighth symposium on usable privacy and security. 2012. p. 1-16.

TSAI, Janice Y., et al. The effect of online privacy information on purchasing behavior: An experimental study. Information systems research, 2011, 22.2: 254-268.

ULMAN, Milos, et al. IT Ethics Perceptions and Behavior: An International Comparison. Journal of Computer Information Systems, 2021, 61.5: 418-427.

ZEISSIG, Eva-Maria, et al. Online privacy perceptions of older adults. In: International Conference on Human Aspects of IT for the Aged Population. Springer, Cham, 2017. p. 181-200.

Expected date of thesis defence

2022/23 SS – FEM

The Bachelor Thesis Supervisor

Ing. Miloš Ulman, Ph.D.

Supervising department

Department of Information Technologies

Electronic approval: 14. 7. 2022

doc. Ing. Jiří Vaněk, Ph.D.

Head of department

Electronic approval: 27. 10. 2022

doc. Ing. Tomáš Šubrt, Ph.D.

Dean

Prague on 15. 03. 2023

Declaration

I declare that I have worked on my bachelor thesis titled "*Users' perceptions of online privacy issues"* by myself and I have used only the sources mentioned at the end of the thesis. As the author of the bachelor thesis, I declare that the thesis does not break any copyrights.

In Prague on date of submission ______________________________

Acknowledgment

I would like to express my sincere gratitude to Ing. Miloš Ulman, Ph.D., for his invaluable guidance, support, and encouragement throughout the completion of my bachelor’s thesis. His expertise, insightful comments, and constructive feedback have been instrumental in shaping my research and improving the quality of my work.

Users' perceptions of online privacy issues

Abstract

The purpose of this study was to study the factors that play a role in the decision-making process of users on the Internet. The focus was on the privacy and security precautions that users can take while using the sites and online services. There are many definitions of privacy, and the concept of privacy has changed over time and with new technologies constantly evolving. When we talk about privacy, we mean the privacy of information. This includes deciding what personal information may be disclosed to others and understanding how that personal information is obtained by others and how other parties use that information.

Keywords: Privacy, policy, perception, safety, awareness, students, professionals.

Vnímání problémů online soukromí uživatelů

Abstrakt

Cílem této studie bylo prostudovat faktory, které hrají roli v rozhodovacím procesu uživatelů na internetu. Důraz byl kladen na ochranu soukromí a bezpečnostní opatření, která mohou uživatelé přijmout při používání webových stránek a online služeb. Existuje mnoho definic soukromí a pojem soukromí se v průběhu času a s neustále se vyvíjejícími novými technologiemi měnil. Když mluvíme o soukromí na internetu, máme na mysli informační soukromí. Tento pojmem zahrnuje rozhodování o tom, jaké osobní údaje mohou být sděleny ostatním, a pochopení toho, jak tyto osobní údaje získávají ostatní a jak je ostatní strany používají.

Klíčová slova: Soukromí, politika, vnímání, bezpečnost, informovanost, studenti, odborníci.

Table of content

1 Introduction

The study is focused on assessing the topic of “Online privacy concerns and protection” among the working people and students. Today, social networks offer so many flexible services in a more convenient way that involves an online communication online, however, the communication deems to exchange an information that is more or less deals with a personal data. There have previously been observations made on the risks and difficulties associated with people utilizing the internet in their personal affairs. Users of online information technology are presently showing a greater concern for the environment of the internet. The lives of users are being invaded by fundamental issues such as spam, cookies, the clickstream, and real-time location tracking, all of which have a significant probability of compromising the users' online privacy and security. There have been a lot of studies that looked at how people feel about their privacy and security online. One study conducted by (Okazaki, S., Eisend, M., Plangger, K,. Ruyter, Ko de., and Grewal. D., 2020). They provide an explanation of the fundamental differences between privacy and security. This was done so that the researchers could then evaluate whether internet users of different “age” share similar worries.

Even though, this research is intended to discover whether there is a difference in the perception of “Online privacy concerns and protection” among different “Occupation”. The author focuses on “students” and working group population to identify how different those perceptions are.

There are many studies that research (Jordan, G., Leskovar, R., Marič, M., 2018) the “Online privacy and security concerns” in regards on online purchase, where again, the disclosure of information is present, and the authors mainly focused on criteria of user experience in online shopping. But none of these studies have ever mentioned the difference between “Occupational” background, thus, this research has a primary objective to identify how people of different “Occupational” background perceive an “online privacy concerns and protection” and whether there are differences in the perception of such issues.

2 Objectives and Methodology

2.1 Objectives

The main objective of the thesis is to investigate differences between students and working professionals in perceptions of online privacy issues.

Partial objectives:

- To review existing models of privacy and security online.
- To prepare and conduct a survey among users.
- To test the differences between students and working professionals, evaluate results and interpret findings.

2.2 Methodology

The methodology of solving the theoretical part of the diploma thesis will be based on the study and analysis of the literature. Based on the knowledge gained in the theoretical part of the work, the practical part will conduct a survey among users from different regions. Statistical data of survey data and interpretation of the results through a relevant theory will be done. The author will formulate the conclusion by synthesizing findings from the literature review and survey data analysis.

3 Literature Review

This chapter is dedicated to describing the factors that go align with a human perception of online privacy as it is one of the basic human needs to be considered.

3.1 Online Privacy

As it has been mentioned above, the basic human needs within e-commerce and in overall online platforms is a privacy. With a growing technological development and progress of internet, many researchers have conducted its studies on the related issue of "Privacy". (Westin, 2015.) defined "Privacy" as the extent and allowance of personal, institutional, group information, that could be shared with others and with what purposes. Another definition that was provided by Belanger et al. (2002) stated that "Privacy" is the willingness of an individual to share its own information over the Internet, where eventually, the internet stores might use that information in their advantage, process it, utilize, and contribute to the growth in privacy concerns. However, it is obvious that privacy concerns and a process of sharing a personal information should be taken into consideration by e-commerce businesses as people don't want their personal information to be leaked and used against them (Belanger et al, 2002).

Ham (2016) claimed that not many people are aware of how personal information might affect their advertising as, based on their personal information sharing, the algorithm tailors' promotional messages which makes a great risk to consumer privacy.
On the other hand, companies use personal information in their advantage in order to offer as many products as possible, that could eventually fit the consumer's needs. However, it is almost impossible for consumers to track how their personal information is used, collected, and shared (Lee & Cranage, 2011).

Ham (2016) states that privacy is one of the most important ethical, legal, social, and political issues of information. He identified different dimensions that consumers might fear while having an online experience.

When consumers feel like their privacy is being violated, the reaction might be absolutely different. (Kim, 2008) identified it as (IPPR) information privacy–protective response and further describes it as user's behavioral responses to their perception of information privacy threats because of how companies collect information.

3.2 The privacy parodox

It can be observed that on the Internet, people reveal a large amount of personal data when shopping online or provide complete information about their personal lives on social networks. On the other hand, many people are concerned about their personal information and its distribution and say that they value this information very much and will never share it. This dilemma is called the privacy paradox.

The phenomena of privacy paradox illustrate that individuals share and disclose their personal information via internet and consequently contradict their privacy attitudes. For example, SNN users reveal their personal information to the point that their disclosing behavior does match their preferences of concerns and privacy (). Hence, it is right to conclude that a high privacy requirement should not be taken for granted as a strong predictor of low disclosure behaviors (Tufekci, 2008).

From a scientific point of view, if there is indeed a privacy paradox phenomenon, most of the empirical researchers question its existence as, "per se", which indicates a lack of knowledge about the limited link between privacy concerns and disclosure behavior. Which eventually leads to a little protection of own information by users, however, their initial behaviors were seriously concerned about their privacy.

A lack of technical knowledge about the policies and procedures behind mobile apps and websites can greatly contribute to the mechanism of the privacy paradox in people who fear losing their personal data.

3.2.1 Privacy paradox factors

Many researchers demonstrate that privacy is the most concernable factor for most citizens who are involved in the digital world. Tufekci (2008) claimed that individuals are willing to sell their personal information for a small reward. conducted economic research and found out that individuals value their personal information (browsing history) for about 7 EUR, which is the cost of a Big Mac meal.

Tufekci (2008) claims that individuals that engage with social media platforms or use different technologies can be alluring, and it is based on the principle that the more people share their personal data, the more they enjoy the benefits that are offered within the platform social media.

One and probably the main factor of the privacy paradox is the "Clickthrough" policy that seems to be designed to encourage potential customers to share their data. Because of the length and complexity of such policies, it discourages customers to read the whole chapter of plain text, but simply by clicking agree with terms and conditions, they unconsciously agreeing with the terms, in order to quickly connect to a social media website.
Advertisers want to avoid the controversy, when customers disagree with policies and leave its pages, but rather keep customers in the "buying atmosphere". Other factors might influence the process, such as "small icons for privacy policy" or similar effects that will make customers engage in the consent process (Obar and Aeldorf-Hirsch, 2018).

3.3 Online Security

The section illustrates on how to keep a secure E-commerce environment and generally brings an idea of what an online security is. The process of keeping with a technological development requires also new security trends to be secured and protective with the data and database. It is quite a challenging task to protect information nowadays, from cyberthreats which are taken by so many forms. One of the most problematic factors of cybersecurity is the evolving nature of security risks. Due to a technological development and emergence, new options of attacks are developed, hence in order to make sure that your organization keeps – up with the updating, against those attacks is quite a challenging task (Techtarget, 2022).

Those companies that have a lot of data about their customers are more prone to cyberattacks and cybercriminals who would want to steal personally identifiable information (PII) is another concern. As an example, an organization that underwent the ransomware attack of PII, ought to do everything possible to prevent a cloud breach.

Unintentional internal attacks could happen because of employees of an organization, due to the fact that employees use their own devices which might have a virus. A constant training courses as well as regular security awareness will help employees to keep their company safe from cyberthreats.

The author lists different types of cybersecurity attacks which include:

Malware – is a form of malicious software (worms, viruses, Trojans, spyware) is embedded into the file or program that intends to harm or damage a user's computer.

Ransomware – is another form of malware. Intends to lock the user's computer with the typical encryption process and afterwards, demand payment from the victim to decrypt and unlock the computer.

Social engineering – is an attack that relies on individual interactions to trick users into breaking security procedures to obtain sensitive information which is usually protected.

Phishing – is a form of random emailing or text message that are structured in a way to steal sensitive information of monetary matters such as (credit card, login information).

Spear phishing – is a type of phishing attack that is intended to target a user, organization, or business.

"*As risky as the Internet Is, companies still have no choice but to be there. The lure of new markets, new customers base, new revenue sources and new business models are just so great that companies will flock to the Internet regardless of the risks*" – (Schneir, 2005).

Company's competition which are involved in Internet or (e-commerce) and serve customers online, is very high. To survive, companies try to protect and secure its customer via their software, websites and other systems of information. However, some companies do not consider the online security as a thing but do know about the consequences that might come along.

3.3.1 Networking security

Network security includes a list of requirements, recommendations and policies that are used in the network infrastructure to increase its level of protection and resiliency (Hyland and Sandhu, 1998).

The second important function is to analyses the operation of the company's infrastructure and prevent unauthorized access (UAS) to information resources by intruders. Regardless of the scale and type of business (small, medium, or large), the use of network infrastructure implies the integration of hardware and software solutions that ensure the health and safety of the network; however, intruders might use all types of attacks at once, and the system can crack. It is up to the company on how to secure its networking. Better yet, to limit access only for workers and make only one networking publicly, where none of the data is stored Culnan et el, 2000).

Protection of equipment connected to the network infrastructure. As protective measures, they use anti-virus solutions with regular database updates, firewalls with traffic filtering and blocking unwanted subscribers and etc. The equipment must be fault-tolerant and provide for the possibility of rapid recovery. It implies the presence of redundant components in critical nodes. Systematic monitoring of the entire infrastructure of the company to detect vulnerabilities. Also, the system should provide detailed information about any software or hardware component of the equipment.

Continuous monitoring of network channel bandwidth. This guarantees timely blocking of unwanted traffic, and also allows manual load balancing.

Critical nodes of an organization's infrastructure must provide high availability against any threat or attack on the company. This is achieved by creating a second independent site (DPC), which replicates data from the first in synchronous mode.

3.3.2 Security Concerns

Along with the privacy concerns, the security of consumer's information has been identified as a barrier which prevents an expansion of e-commerce Wildt (2016). He discovered six technological reasons that hamper the rise of the e-commerce, those are: slow downloads, interface restrictions, search issues, insufficient evaluations of web-applications and its effectiveness, security of web and the lack of internet standards. However, security concerns are closely related to the privacy concerns, although its differently constructed.

Vijayasarathy (2004) conceptualized the privacy and security concerns as a "control over secondary use of information" and "environmental control". Whereas environmental control referred to as an ability of to influence the conduct of other people in the environment during a market transactions or commercial exchanges. Environmental control in e-commerce refers to the implementation of security measures to ensure the safe exchange of private data across transactions. Due to the technological development such as encryption and authentication, e – commerce has gained more trustworthiness. When data is encrypted with the mathematical method to scramble the message, the recipient of such message supposed to have a code or a message to encrypt it an. Today, every internet browser has got a SSL[1] technology. Online companies, whose business is heavily dependent on online transactions, based their security system on (SET), which is quite similar to the SSL, but the online company will not get access to credit card information.

Authentication is another way of digital certification which became quite common for the past decade. It helps to identify the transaction parties, digital signature. Technically speaking, it is an electronical signature that is used to verify the identity of a sender or a receiver of a certain document.

Even though, this bachelor thesis doesn't deal with the transactions, it deals with the security of personal data and the credit card information is a part of it.

3.4 Data Leakage

Data leakage is the phenomena that evolves the unintentional or intentional disclosing of data, without the permission of the author to whom this data belongs. Usually, the data leakage consists of a sensitive information which was gained by unauthorized party, either by accident or by deliberate acts (Shabtai's et al, 2012). Based on his review, the data leakage is a big problem.

3.4.1 Data leaks in organizations

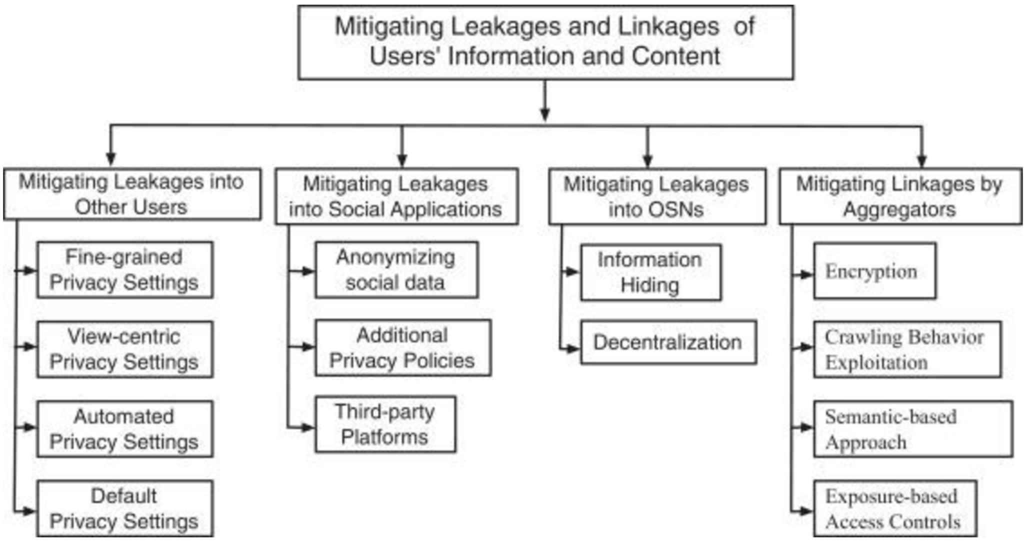
Data leakage in the organizations and its process was researched by et el., Shabtai (2012) who revealed that companies don't scan the outgoing communication for confidential information to prevent data leakage for the third parties. Emails, messages through teams,

[1] SSL - (literally secure sockets layer), abbreviated SSL, is a protocol, or a layer inserted between the transport layer (e.g. TCP/IP) and the application layer (e.g. HTTP) that provides communication security by encryption and authentication of communicating parties.

website forms, file transfers and other types of electronical communication are not tracked neither monitored. One this type of information is gone, the company is in big troubles. Conwill (2009) claims that an insider exists in every enterprise, and the enterprise might have no idea who it could be, which raises a risky concern. A survey done by RSA/IDC discovered that most chief security officers (CSO's) were not worried about the internal threats but rather they were worried about the external ones. The survey also discovered that most of the attacks were originated from internal core of the enterprise, around 5792 incidents, where more than half of those incidents were caused by insiders, who abused the access to control the right and only 18 % of attacks were deliberate.

Users provide their personal information while making a mobile payment or buying online in order to receive services or goods. Shopping information, purchase tracking, and browsing history may all be compromised. Users frequently submit personal information in exchange for practical services. Mobile devices will be attacked by malware to steal personal information.

Figure 1: Mitigating solutions for the users' leakage and linkage of formation and content.


Mitigating Leakages and Linkages of Users' Information and Content
Mitigating Leakages into Other Users
Fine-grained Privacy Settings
View-centric Privacy Settings
Automated Privacy Settings
Default Privacy Settings
Mitigating Leakages into Social Applications
Anonymizing social data
Additional Privacy Policies
Third-party Platforms
Mitigating Leakages into OSNs
Information Hiding
Decentralization
Mitigating Linkages by Aggregators
Encryption
Crawling Behavior Exploitation
Semantic-based Approach
Exposure-based Access Controls

Source: Adopted from (Iamnitchi, 2017)

Given the amount of sensitive information users expose on Online Social Networks and the different types of relationships in their online social circles, the challenge OSNs face is to provide the correct tools for users to protect their own information from others while taking full advantage of the benefits of information sharing (Iamnitchi, 2017).

Users of online social networks face numerous risks when using social applications. First, the application may be malicious; it may collect a large amount of user data for unwanted use.

3.4.2 Data Leakage prevention

Data leakage prevention is a solution that prevents data leakage by monitoring, filtering and recoding outgoing communication to more confidential data (McCormick, 2008). The statement was verified by Blasco and Jorge (2013) who described the DLP as a solution which is based on analyzing, monitoring and controlling the confidential data usage across computing systems to prevent data leakage either intentionally or accidental. According to McCormick, all the necessary procedures are needed to be taken as follows:

- Some organizations create a database where all files are stored, however, before that, organization scans them for keywords which are later tagged with a flag. If any of those tagged files are trying to exit the organizations, it will be automatically blocked.
- By using a deep packet inspection, it could also prevent the data leakage.
- Another solution is to encrypt the data which is about the leave the enterprise so that an unauthorized person will not be able to open the file.
- By applying the firewall of next generation that has other features that work hand in hand with DLP, to prevent the leakage. Such as "Application Control which blocks different applications and webapps based on ID application, instead of TCP/UDP sessions (ECS, 2014.)

3.5 User's perceptions and behavioral intentions

3.5.1 Theory of Reasoned Action

The Theory of Reasoned Action (TRA) aims to clarify the link between attitudes and behaviors within an act. It's mainly accustomed predict how individuals will behave supported by their pre-existing attitudes and behavioral intentions. The TRA assumes that individuals are usually rational and will consider the implications of their actions prior to deciding whether to perform a given behavior (Yousafzai, Foxall and Pallister 2010).

The TRA model was proposed in 1975 by Fishbein and Azjen. It focuses on the construction of a system of observation of two groups of variables, which are:

- attitudes defined as a positive or negative feeling in relation to the achievement of an objective.
- subjective norms, which are the very representations of the individuals’ perception in relation to the ability of reaching those goals with the product (Salgues, 2016).

3.5.2 Theory of Planned Behavior

The Theory of Planned Behavior (TPB) is a theory based on the psychology of an individual that links beliefs to behavior. The theory shows that three main components, namely attitudes, subjective norms, and perceived behavioral control, together form a person's behavioral intentions.

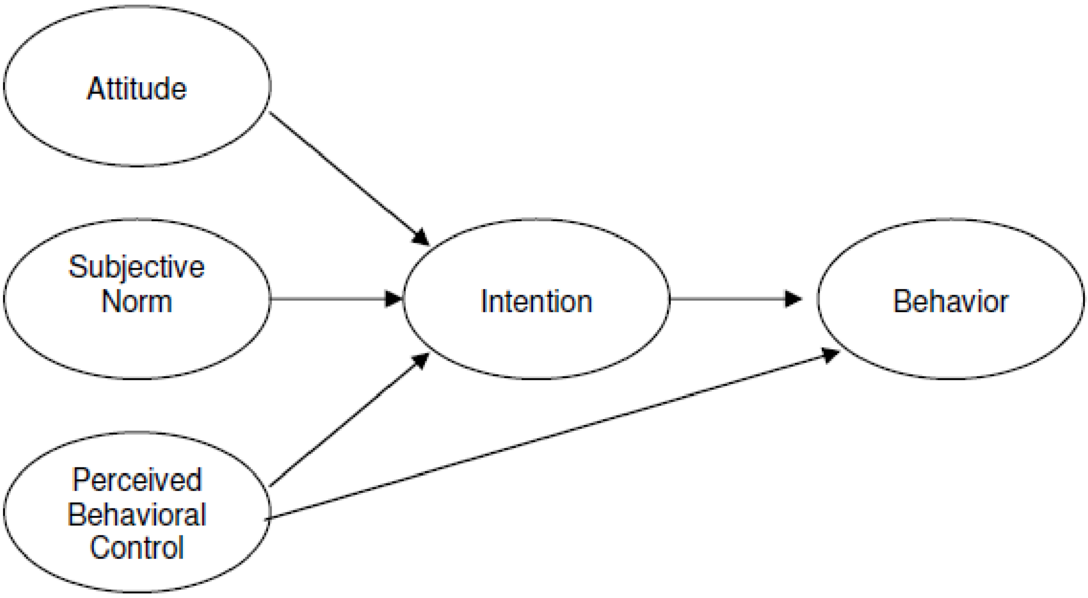
The main basic assumption of the TPB model is that most behavior people engage in is under their own control and is rational. Moreover, the decision factor in a person's actual behavior is the tendency to behave, that is, behavioral intention. In addition, a person's personality, age, occupation, gender, etc. have no direct impact on his/her behavioral intention. In fact, these variables can only affect behavioral intention indirectly through attitude and subjective norms (Chen and Tung 2014).

Any consumer's intents to engage in online commerce are influenced by the views they have developed toward online retailing and sales. Salient ideas influence how we feel. Along with beliefs, behavioral intentions are also influenced by one's control over the process of engaging in e-commerce and their perception of how others will see them. According to the TRA theory, the intentions of users are mainly determined by their behavior and subjective norms, see chapter above.

However, TPB implies the fact, when a user visits an online website of a company, he/she firstly attentive to the trustworthiness of the attributes. Such indicators as online reputation, popularity of a company and feedbacks left in regards of that company are considered as „indirect “attributes. Thus, Ajzen and Fishbein (1980) concluded that such observations format „descriptive beliefs “. However, beliefs are formed not only by direct observations. Interactions with the customer support through the online chat-bot, might increase a belief perception. If a consumer experienced a satisfactory buying process, it might build a positive belief attitude towards online purchasing and etc. Likewise, if a user bought something online in an offline setting, the overall perception will also be positive. Beliefs that are formed by processing previously learned experience, is related to such behavior and called as „Inferential beliefs "(Fin, 1981).

Beliefs could also be acquired by gathering knowledge from various sources, such as the media, books, magazines, online businesses, search engines, acquaintances, coworkers, or family members. These sources can include both descriptive and inferential beliefs.

Figure 2: The Planned Behavior Theory



Source: Ajzen (1985).

In the **Figure – 2**, the TPB illustrates how risk perception and trust beliefs shape the consumer's attitude and what intentions are formed and how eventually they impact the decisions of participants whether go ahead and make purchases, registrations, verifications of payments etc. Malhotra et al. (2004) determined a scale that measures privacy concerns of the internet users, and it is based on a dimensional base. His study included the factors of security concerns of online users, their attitudes, subjective norms and control of own behavior. The main phenomenon of his study was an intention to take part in the transaction processes with online companies, not just giving personal data to the third parties.

Beliefs can also be created by gathering knowledge from various sources, such as the media, books, magazines, online businesses, search engines, acquaintances, coworkers, or family members. These sources can include both descriptive and inferential beliefs.

Individual evaluations of the characteristics of the internet business determine beliefs about things like risk or confidence in a provider. When presented with new things, people evaluate their own characteristics based on how those characteristics relate to existing objects, attributes, or qualities toward which they already have attitudes. For instance, a customer will evaluate an online business' security measures based on what they have seen in other online businesses, their prior online shopping experiences, and their Internet experience (Fishbein and Ajzen, 1975).

The attitude a person has about anything is shaped by numerous of beliefs that have been formed by certain experiences. All beliefs, though, do not influence attitudes in the same ways. A person's attitude at any given time only appears to be influenced by a relatively small number of beliefs. Subjective norms are those that stand out. A belief that is important at one point in time might not be important at another. It is conceivable that subjective norms may shift and be replaced by new ones. User’s might have high trust beliefs and low attitude towards online shopping and vice versa, some individuals might have low trust beliefs, however, treat an online shopping as a daily activity (Ajzen et al, 1980).

3.6 Summary of main findings

The author has described the theoretical background that concerned “Online privacy” and different aspects that either directly or indirectly are involved in the topic. The theory discovered the “Privacy paradox” factors and why people engage in the online shopping, what kind of reward they might expect and the unnecessary risk that might come alone with such rewards. Additionally, the author covered the topics on a prevention of a data leakage, to identify common steps on how people could prevent a data leakage on a first place.

Based on the above-mentioned theories, such as: “Theory of planned behavior” and “Theory of reasoned actions” they are different, and each assumes that people act in a certain way based on its own purpose. In the “Theory of planned behavior” such personal attributes as” gender, occupation and age are not considered and people usually act with a help of additional data, for example, when purchasing online, people would be more attentive to the to such attributes as “Trustworthiness” of a certain information., whereas: According to the Theory of reasoned actions, the intentions of users are mainly determined by their behavior and subjective norms.

4 Practical Part

The researchers will outline their technique for gathering primary data in the section that follows. The study goal will be reiterated at the beginning of the chapter, which will then move on to the research strategy, philosophy, method, and data analysis that were all chosen, before concluding with a discussion of the research method's quality.

4.1 Research purpose

The purpose of this research is to investigate differences between students and working professionals in perceptions of online privacy issues.
Variables are stated as following:

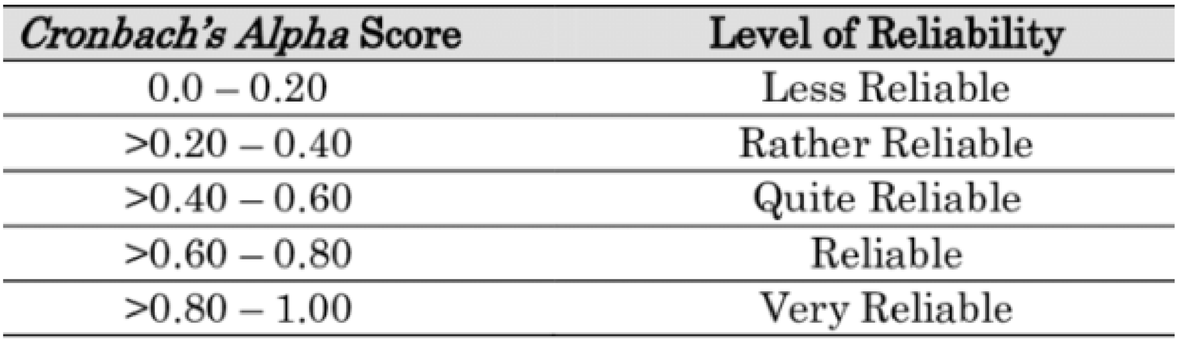
- Privacy Concerns and Protection (hereinafter PCP).
- Disclosure of sensitive information (hereinafter DSI)
- Control over Personal Information (CoPI)
- Identity Theft (IT)

Further, the author wants to test the following hypothesis to see the co – independence between different factors such as:

1) H1: There is no dependency between the students and employed people and the way each occupation perceives the "data protection"
2) H2: There is no dependency between the "Occupation" and "Frequency of being hacked" online.
3) There is no dependency between the "Occupation" and "Adding strangers to the friend's list" on a social media.

The author uses IBM SPSS Statistics 24 to analyze all the necessary data.
The Cronbach's alpha test will be applied in order to make sure that the answers of participants are reliable and consistent. The **Picture – 1**, illustrates the evaluation of the Cronbach's alpha test.

Picture 1: Cronbach's alpha Score



Cronbach's Alpha Score	Level of Reliability
0.0 – 0.20	Less Reliable
>0.20 – 0.40	Rather Reliable
>0.40 – 0.60	Quite Reliable
>0.60 – 0.80	Reliable
>0.80 – 1.00	Very Reliable

Source: Ahbika (2017).

4.2 Breakdown of the participation rate

The following chapter is devoted to illustrating the participation rate from the perspective of: (Gender, Age, Occupation, Country of residency, country of origin, and how many times participants were hacked).

To start – off, the participation rate regarding gender is the following, see **Table - 1**. There is a slight dominance of males by 5.3 %. Overall, there were 206 participants who took part in the survey.

Table 1: Demographical data of participants

Demographic data of respondents	**N – 206**	**Percentage**
Gender		
Male	114	55.3
Female	93	44.7
Age		
15 - 25	112	54.37
26 – 35	52	28.64
36 – 45	24	11.65
46 – 55	12	5.83
55 +	6	2.91
Place of birth		
CIS	86	41.75
EU	59	28.64
Middle East and Africa	61	29.61
Occupation		

Student	99	48.06
Working Full Time	81	39.32
Freelance	12	5.83
Unemployed	9	4.37
Retired	5	2.43

Source: Own calculation, based on the gathered data.

The following question related to the place of birth or (country origin). The results are the following. Majority of participants came from CIS[2] countries. However, three people out of 206, didn't answer the question fully, due to unavailability of the answer. This point, however, will not impact the research and the model in general.

The following question is related to the age of the participants. The majority of the participants were aged between 15 – 25 (54.4 %). Followed by the age (26 – 35) (25.2 %) and 36 – 46 (11.7 %). The next question related to the occupation of participants. Most of the participants were students (48.1 %), followed by (38.3 %) of people working – full-time and, freelance (5.8 %).

4.3 Descriptive analysis of variables

The chapter is devoted to illustrating the data that was gathered by the author, with the help of a questionary. Where planned sampling was about to reach 200 participants, however, the author managed to receive 206.

Descriptive statistics of each dimension will further be presented, whereas:

Table 2: Descriptive statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
PCPAve	206	1.00	5.00	2.4721	1.02148	.566	.169	-.474	.337
DSIAve	206	1.00	5.00	1.4709	.62380	1.784	.169	4.724	.337
COPIAve	206	1.00	5.00	2.3034	.88217	.654	.169	-.039	.337
ITAve	206	1.00	5.00	1.8754	.82318	.723	.169	-.062	.337

[2] Commonwealth of Independent States

Valid N (listwise)	206								

Source: Own processing, SPSS IBM.

If we take a look at the mean, we could see that most people agree with all the statements, within each dimension. Especially, the dimension o Disclosure of sensitive information, with an average of 1.47. However, the author run an analysis of Normality and see how residuals are distributed. In this case, the author should look at the Kolmogorov and Smirnov test, because the sampling size is more than 40.

Table 3: Test of Normality

Tests of Normality

	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
PCPAve	.139	206	.000	.948	206	.000
DSIAve	.251	206	.000	.762	206	.000
COPIAve	.141	206	.000	.949	206	.000
ITAve	.162	206	.000	.892	206	.000

a. Lilliefors Significance Correction

Based on the gathered results, we look at the significance level for all variables. The significance of *p – the value* equals to .000, which is less than the .05 alpha level. Since the variables are non–normally distributed, the author should apply the Ordinary Regression Analysis. This means that the author should transfer the variables into the *Log function.* The table below demonstrates that variables are still not–normally distributed.

Table 4: Log- Normality test

Tests of Normality

	Kolmogorov-Smirnova			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
log_PCPAve	.078	206	.004	.969	206	.000
log_DSIAve	.284	206	.000	.818	206	.000
log_COPI	.121	206	.000	.969	206	.000
log_ITave	.182	206	.000	.900	206	.000

a. Lilliefors Significance Correction

Source: Own processing, based on SPSS IBM.

4.4 Test of model of Fit

In the next chapter, the author uses Ordinal Regression Analysis to run the model to evaluate whether the data fits the model.

Table 5: Test of fit model

Model Fitting Information

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Intercept Only	992.157			
Final	796.810	195.347	3	.000

Link function: Logit.

Source: Own processing, based on SPSS IBM.

The p-value is higher because the significance level is between.001 and.05, which means that the data fit the model well. The model is statistically significant.

The next test which the author will run is a test of parallel lines, to check whether the model doesn't violate the proportional odds. This test helps to identify whether the assumptions hold themselves accountable and assumes that the parameters are the same for all categories is reasonable.

Table 6: Test of parallel lines

Test of Parallel Lines[a]

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	796.810			
General	837.104b	10.887c	45	.367

The null hypothesis states that the location parameters (slope coefficients) are the same across response categories.

a. Link function: Logit.

b. The log-likelihood value cannot be further increased after maximum number of step-halving.

c. The log-likelihood value of the general model is smaller than that of the null model. This is because convergence cannot be attained or ascertained in estimating the general model. Therefore, the test of parallel lines cannot be performed.

Source: Own processing, based on SPSS IBM.

The test demonstrates that the significance is .367, which is higher than p – value, .05. Meaning that the model doesn't violate the assumption of proportional odds.

4.5 Correlation analysis

Correlation analysis is shown in **Table – 7**, their correlation analysis demonstrates the positive relationship among all selected variables. However, the roots seem to have a weak correlation with each other, the highest positive correlation is 0,2 between Control over personal information and Identity Theft.

Table 7: Correlation analysis

Correlations

		PCPAve	DSIAve	COPIAve	ITAve
Pearson Correlation	PCPAve	1.000	.331	.742	.651
	DSIAve	.331	1.000	.351	.472
	COPIAve	.742	.351	1.000	.560
	ITAve	.651	.472	.560	1.000
Sig. (1-tailed)	PCPAve	.	.000	.000	.000
	DSIAve	.000	.	.000	.000
	COPIAve	.000	.000	.	.000
	ITAve	.000	.000	.000	.
N	PCPAve	206	206	206	206
	DSIAve	206	206	206	206
	COPIAve	206	206	206	206
	ITAve	206	206	206	206

Source: Own processing, based on SPSS IBM.

The next step is to run the coefficient analysis and eventually demonstrate the model. The author already knows that the data fits the model perfectly, as it is seen in the Table – 5. Thus, the coefficients are the following, see, **Table – 8**.

Table 8: Coefficients

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	.242	.141		1.723	.086		
	DSIAve	-.051	.080	-.031	-.639	.523	.767	1.304
	COPIAve	.642	.060	.554	10.687	.000	.677	1.478
	ITAve	.441	.068	.355	6.449	.000	.600	1.667

a. Dependent Variable: PCPAve

Source: Own processing, SPSS IBM.

The coefficients of the model are seen in the **Table – 8**. First, the author highlights the significance of the parameters, with .000 and .000. Both values are less than 0,05 alpha, meaning that both are significant and contribute to the dependent variable of Privacy Concerns. However, one variable of Disclosure of personal information seems to be insignificant and doesn't contribute to the dependent variable. However, again, based on the model of fit, the overall contribution of variables seems to have an impact on the dependent variable overall

With one increase in the dimension of Disclosure of sensitive information, the dependent variable would decrease by -0.051, and with one increase unit of Control over personal information, the dependent variable increases by 0.642.

In the last column, the author pays attention to the multicollinearity between dependent variables. Initially, the author sets **5** as high multicollinearity. Thus, multicollinearity is not present in the model.

4.6 Hypothesis testing

The following chapter is devoted to demonstrating whether the stated hypothesis holds itself accountable.

1) H1: There is no dependency between the students and employed people and the way each "Occupation" perceives the "Data protection"
2) H2: There is no dependency between the "Occupation" and "Frequency of being hacked" online.
3) There is no dependency between the "Occupation" and "Adding strangers to the friend's list" on a social media.

The author applies Chi–Square test, to analyze the dependency between two categories, "occupation" and "privacy concern", "frequency of being hacked" and "adding unknown people to a friend's list".

Table 9: Contingency table between occupation and Privacy Concerns

What is your occupation? * Privacy Concerns and Protection [I apply different passwords for different social media because it decreases the risk that my account will be hacked.] Crosstabulation

		Com. Agree	Agree	Neutral	Disagree	Com. Disagree	Total
What is your occupation?	Student	37	34	3	21	4	99
	Working full – time	40	16	3	13	9	81
	Freelance	6	4	1	1	0	12
	Unemployed	4	2	1	1	1	9
	Retired	3	1	0	0	1	5
Total		90	57	8	36	15	206

Source: Own processing, SPSS IBM.

Table 10: Chi-Square test for H1.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	15.666[a]	16	.037
Likelihood Ratio	16.931	16	.020
Linear-by-Linear Association	.300	1	.014
N of Valid Cases	206		

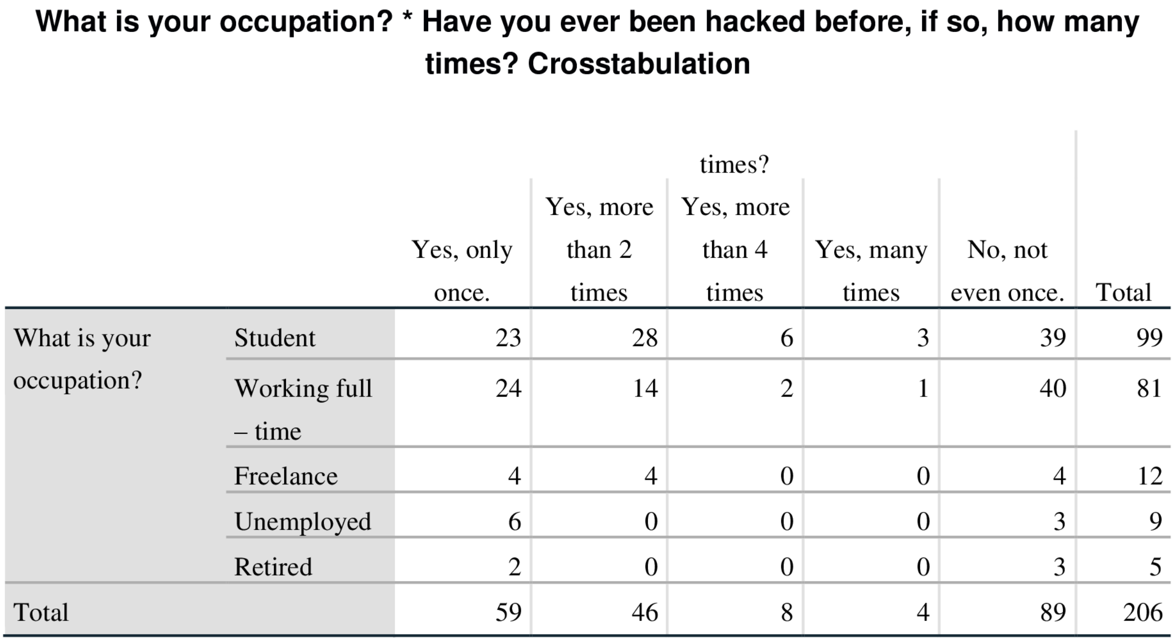
a. 0 cells (0 %) have expected count less than 5. The minimum expected count is 5.32

Source: Own processing, SPSS IBM.

Based on the Asymptotic Significance (2 – sided) value, it indicates that .037 is less than .05, which indicates that there is a dependency between the occupation and the fact that students apply different passwords for social media and thus, avoid the risk of being hacked. ***Thus, the first hypothesis is rejected, indicating that there is a dependency between two factors.***

The following hypothesis assumes that students are less prone to be hacked that working people. The cross-tabulation/ contingency table is shown below with the results.

Table 11: Occupation and frequency of hacking

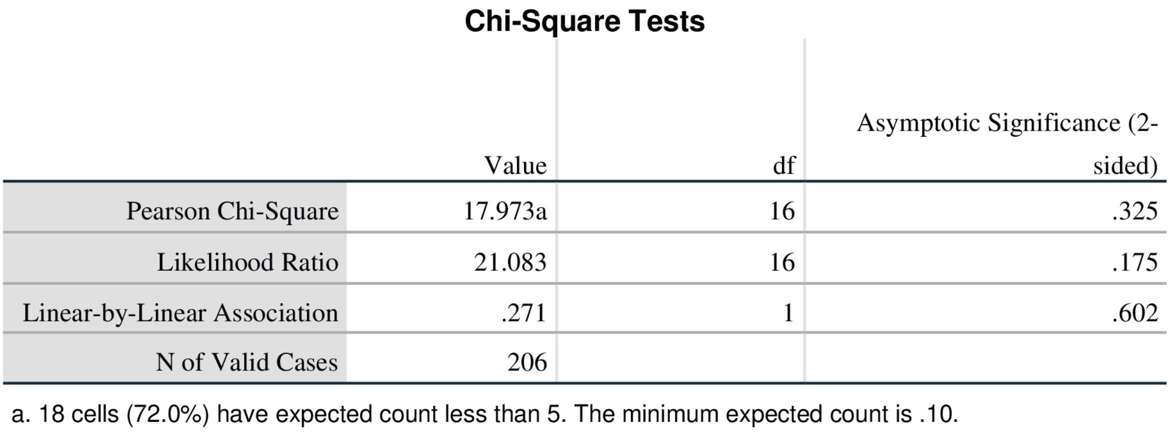


What is your occupation? * Have you ever been hacked before, if so, how many times? Crosstabulation

		times?					
		Yes, only once.	Yes, more than 2 times	Yes, more than 4 times	Yes, many times	No, not even once.	Total
What is your occupation?	Student	23	28	6	3	39	99
	Working full – time	24	14	2	1	40	81
	Freelance	4	4	0	0	4	12
	Unemployed	6	0	0	0	3	9
	Retired	2	0	0	0	3	5
Total		59	46	8	4	89	206

Source: Own processing, SPSS IBM.

Table 12: Chi-Square test for H2.



Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	17.973a	16	.325
Likelihood Ratio	21.083	16	.175
Linear-by-Linear Association	.271	1	.602
N of Valid Cases	206		

a. 18 cells (72.0%) have expected count less than 5. The minimum expected count is .10.

Source: Own processing, SPSS IBM.

Based on the results of the Chi–Square test, the results demonstrate that there is no dependency between the frequency of being hacked and occupation. The Asymptotic Significance (2-sided) is .325, which is higher than the p – value.

Thus, the hypothesis – 2, is accepted, indicating no dependency between two factors.

Table 13: Occupation and Adding unknown people.

What is your occupation? * I do not add unknown contacts to my friend's list or people who I personally do not know.

		Com. Agree	Agree	Neutral	Disagree	Com. Disagree	Total
What is your occupation?	Student	42	27	14	14	2	99
	Working full – time	45	17	5	11	3	81
	Freelance	6	2	2	2	0	12
	Unemployed	5	3	1	0	0	9
	Retired	5	0	0	0	0	5
Total		103	49	22	27	5	206

Source: Own processing, SPSS IBM.

Table 14: Chi-Square for H3.

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	13.632[a]	16	.126
Likelihood Ratio	17.307	16	.136
Linear-by-Linear Association	4.295	1	.138
N of Valid Cases	206		

a. 16 cells (18.0%) have expected count less than 5. The minimum expected count is .12.

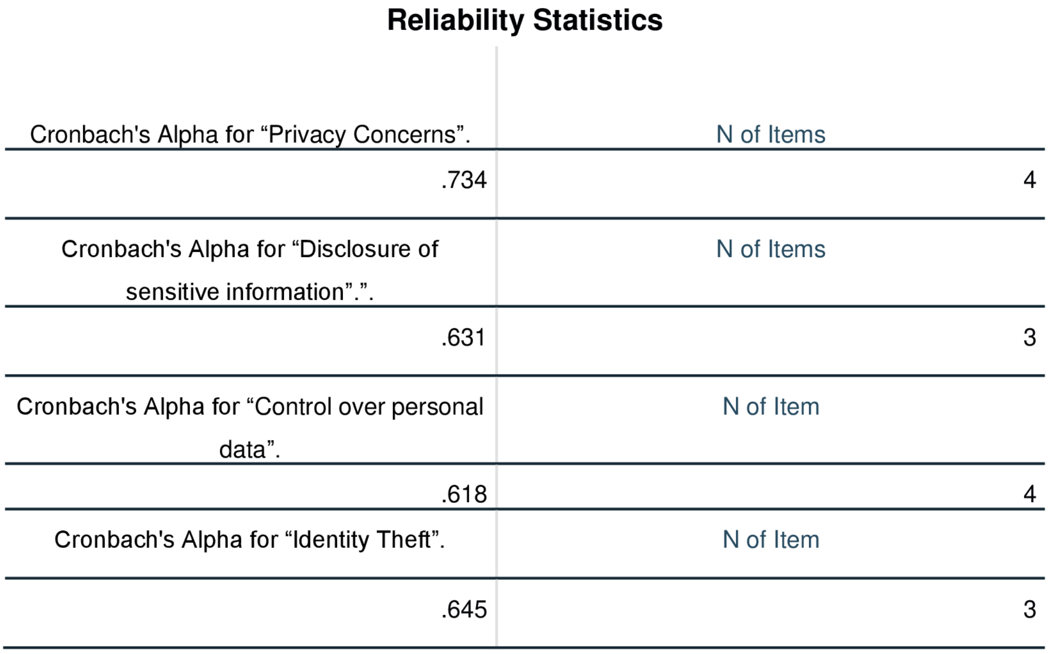
Source: Own processing, SPSS IBM.

Table – 14, demonstrates the results of the Chi–Square test. The Asymptotic Significance (2-sided) demonstrate that the significance of .136 is higher than .05, which indicates no dependency between the that "Students" do not add unknown people to their friend's list. ***The hypothesis – 3, is accepted, indicating no dependency between two factors.***

4.7 Cronbach’s alpha test

The author runs an additional test of Cronbach’s alpha to see the internal consistency within dimensions. It will help the author to analyze whether the questions were asked in a proper way and were structured well enough.

Table 15: Cronbach’s alpha test



Reliability Statistics

Cronbach's Alpha for “Privacy Concerns”.	N of Items
.734	4
Cronbach's Alpha for “Disclosure of sensitive information”.”.	N of Items
.631	3
Cronbach's Alpha for “Control over personal data”.	N of Item
.618	4
Cronbach's Alpha for “Identity Theft”.	N of Item
.645	3

Source: Own processing, SPSS IBM.

Based on Cronbach’s alpha results, the data demonstrates 0.734 reliability, which is “Acceptable”. For “Online Privacy and Concerns”, for “Disclosure of sensitive information” is 0.631 which is good, for “Control over personal information” 0.618 which is good, and for “Identity theft” is 0.645.

5 Results and Discussion

5.1 Hypothesis summary

Based on the analysis of a survey, the author gathered data from 206 respondents. The author also stated several hypotheses to see the dependencies between two factors that are mostly related to the occupation of participants across different regions and their attentiveness to data protection. Which turned out to be rather true, see **Table – 10**. Students are more attentive to data protection than working people and factors that could potentially influence that fact are many. According to Gopnik (2020):

- More advanced skills in terms of data protection among students
- More free time that students have.
- More experienced with technological development.
- People who belong to the "working group" are less worried about data protection as they are more focused on career growth and personal development.

The second hypothesis that considered the dependencies between two factors was, the frequency of being hacked and occupation. Again, based on the H1, which was accepted it demonstrates that "students" pay more attention to "data protection" and hence have fewer chances to be hacked. The question related to the "How many times have you been hacked" helped the author to analyze the dependency, between the frequency of being hacked and "occupation". The second hypothesis was rejected, hence there is no dependency between the "Occupation" and "How many times people have been hacked". It slightly undermines the acceptance of H1; however, the author only bases the results on the statistical data which is limited to 206 samples.

The third hypothesis considered two factors the "Adding strangers to the friend's list" is dependent on the "occupation", which means that people of a young age are less likely to add someone whom they don't know to their friend's list. This hypothesis again is somehow related to the hypothesis H1. Hence, the author assumed that there is a dependency between those two factors. The result demonstrates that there is no dependency between those two factors "Students are less likely to add someone they don't know, whereas, working people do not add strangers to their friend's list either". It is clearly seen in the **Table – 13**. All people answered in the same way.

5.2 Model Summary

The **Table – 8**, demonstrates the coefficients of the model.

Coefficients^a

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	.242	.141		1.723	.086		
	DSIAve	-.051	.080	-.031	-.639	.523	.767	1.304
	COPIAve	.642	.060	.554	10.687	.000	.677	1.478
	ITAve	.441	.068	.355	6.449	.000	.600	1.667

a. Dependent Variable: PCPAve

Source: Own processing, SPSS IBM.

Thus, the model and its coefficients are properly related. It is because of its correlation analysis that demonstrates positive and negative roots of certain explanatory variables on the dependent variable. The “Disclosure of sensitive information” negatively impacts on the “Privacy Concerns and Protection” variable, which is logical. The more disclosure of personal and sensitive information is, the less privacy protection is”. The same applies to the” Control over personal information” and “Identity theft” which positively impact “Privacy concerns and protection”. The more people have control over personal data, the more protected they are. The more people pay attention to the sharing of “Geolocation” and adding unknown people the more protected they are.

5.3 Cronbach’s alpha summary

Based on the analysis of a survey, the author gathered data from 206 respondents. The author also stated several hypotheses to see the dependencies. The results demonstrate the fact that the model was well structured, and all independent variables contributed to a dependent variable.

Eventually, the author run the additional test within dimensions. Meaning that each dimension is presented by several questions, See **Chapter – 9**. Since the Cronbach’s alpha indicates internal reliability. In this research, it helped the author identify, how well the questions were structured, See Table – 1. If the evaluation is less than .60, it indicates either a poor understanding of questions or illogically structured questions. In the author's case, all Cronbach's alpha tests across all dimensions were above .60 %. Indicating that, the questions were structured well – enough to make sure that participants understand the questions and follow the logic behind them, See, Chapter – 4.7.

5.4 Comparison with other studies

For a better understanding, the author pulled the descriptive statistics for a better orientation.

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
PCPAve	206	1.00	5.00	2.4721	1.02148	.566	.169	-.474	.337
DSIAve	206	1.00	5.00	1.4709	.62380	1.784	.169	4.724	.337
COPIAve	206	1.00	5.00	2.3034	.88217	.654	.169	-.039	.337
ITAve	206	1.00	5.00	1.8754	.82318	.723	.169	-.062	.337
Valid N (listwise)	206								

Source: Own processing, SPSS IBM.

The study of an author confirms the results of the (Deliri & Albanese, 2015; Fire et al, 2014) who compared the means of the study, based on the answers of participants. They concluded that, even though the "Online Privacy concerns and protection might be" something that people agree is "Important" or "Very important" the "Disclosure of personal information" seem to have a better "mean" overall. Hence, it is much more important for participants to have full control over their personal data. The results are shown in the table above, within the descriptive statistics "Mean category" and fully correspond to the previous results of the researchers. However, Identity theft in case of the author, gained more "Important" scaling than in the research of Johnson et al. (2012).

5.5 Limitations and implications of the study

There is certain limitation of this study, first, even though the planned number of samplings was received. The author still insists on a wider sampling size, which might identify the other factors of dependencies. The author also limited the research and excluded the dimensions of "Third party applications" and "Data Leakage".

Moreover, the author was limited to with the sources that discovered the co – independency between the "Age" and "Frequency of being hacked", "Occupation" and "Frequency of being hacked" and etc. None of the previous studies mention such a dilemma. Thus, the relevance of such topic is still undiscovered. However, the result possesses a certain level of dependency, again, the wider sampling size could have shown to the author better resulting overall. The author had to go back to the research of Deliri & Albanese (2015), in order to somehow link the findings with the already existing findings, and see, whether they correspond or not.

6 Conclusion

The thesis is dedicated to the topic of “Online Privacy Concerns and Protection” with the objective to identify the how people of different “Occupation” perceive “Online Privacy and concerns and protection” which is very important nowadays due to robust technological development. Online users should pay more attention to “Privacy concerns and protection” as it could undermine their personal life, disclose personal data and even steal it. The research covered the theoretical background of recent researchers such as Ham (2016) (Iamnitchi, 2017) etc. The theory has helped the author to deeply understand how and what “Online Privacy” represents. Additionally, the author also touched on such theories as “Theory of reasoned actions” and why people act online in a certain way and what motivates them to do things online.

In the empirical part, the author mainly focused on structuring proper questions in order to get reliable data, that potentially could help the author to analyze it, and explain how different factors such as “Age, occupation and gender” might somehow contribute to the decisions that are related to “Privacy Concerns and Protection”. Finally, the author planned to receive at least 200 answers from around the globe, and managed to receive more than planned, it was 206 participants. The gathered data helped the author to make a conclusion. After evaluation, the author has built the statistical model, that has shown the correlation matrix and binary relations between dependent variables and independent variables.
It also has helped to either confirm or reject the stated hypothesis at the beginning of the thesis, see Chapter – 4.1.

Finally, after all the assessments, the author run an additional test to make sure that the questions were structured properly. With that test, the author made sure that the data is reliable and could explain how people perceive “Privacy Concerns and Protection”. Additionally, those questions could be applied by different scientific researchers even in the future, with the same methodological tool of Likert – Scale method.

7 References

Ahbika, K. (2017). *mprovement of Quality, Interest, Critical, and Analytical Thinking Ability of Students through the Application of Research Based Learning (RBL) in Introduction to Stochastic Processes Subject. [online]. [Accessed: 14-03-2023]. Available at: 10.29333/iejme.*

Ajzen, I. (1985). *From Intentions to Actions: A Theory of Planned Behavior.* . New-York.: Springer Verlag.

Ajzen, I. (2002). *Perceived Behavioral Control, Self-Efficacy, Locus of Control, and the Theory of Planned Behavior.* Journal of Applied Social Psychology.

Ajzen, I., and Fishbein, M. (1980). *Understanding Attitudes and Predicting Social Behavior.* Englewood Cliffs.: Prentice-Hall.

Belanger, F. & Hiller, J.S. and Smith, W. (2002). Trustworthiness in Electronic Commerce: The Role of Privacy, Security, and Site Attributes. *The Journal of Strategic Information Systems.*, p. 245-270.

Blasco, J., HernandezCastro, C., Tapiador, E. & Ribagorda, A. (2012). *Bypassing information leakage protection with trusted applications. [online]. [Accessed: 6-12-2022]. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0167404812000120.* Computers & Security.

Boshmaf, Y., Muslukhov, I., Beznosov, K., & Ripeanu, M. (2011). *The socialbot network: When bots socialize for fame and money.*

Brandimarte, L., Acquisti, A., & Loewenstein, G. (2013). *Misplaced confidences: Privacy and the control paradox.* Social Psychological and Personality Science, 4(3), 340-347.

Carrascal, J.P. and Riederer. C. (2013.). *Your browsing behavior for a big mac: economics of personal information online.*

Colwill, C. (2009). *Human factors in information security: The insider threat – Who can you trust these days?* Information Security Technical Report.

Dinev, T., Xu, H., Smith, J. H., & Hart, P. (2013). *Information privacy and correlates: An empirical attempt to bridge and distinguish privacy-related concepts.* European Journal of Information Systems.

ECS. (2014.). Retrieved from DLP Key Features. [online]. [Accessed: 14-10-2022]. Available at: http://ecs.arrow.com/suppliers/documents/RSASolutionBrief-enVisionSolutions.pdf: http://ecs.arrow.com/suppliers/documents/RSASolutionBrief-enVisionSolutions.pdf

Gopnik, A. (2020). *Childhood as a solution to explore–exploit tensions. [online]. [Accessed: 14-03-2023]. Available at: 10.1098/rstb.2019.0502.* Department of Psychology, University of California.

Ham, C.D. (n.d.). Exploring how consumers cope with online behavioural advertising. [online]. [Accessed: 12-10-2022]. Available at: https://www-tandfonline-com.proxy.library.ju.se/doi/abs/10.1080/02650487.2016.1239878#aHR0cHM6Ly93d3ctdGF. *International Journal of Advertising*, 35.

Johnson, M., Egelman, S., & Bellovin, S. M. (2012). *Facebook and privacy: It's complicated. In Proceedings of the Eighth Symposium on Usable Privacy and Security, 1-15.*

Jordan, G., Leskovar, R., Marič, M. (2018). *Impact of Fear of Identity Theft and Perceived Risk on Online Purchase Intention. [online]. [Accessed: 26-02-2023]. Available at: https://doi.org/10.2478/orga-2018-0007.* Organizacija 51(2), 146–155.

Kim, S.S. (2008). *Internet Users' Information Privacy-Protective Responses: A Taxonomy and a Nomological Model.*

Koohikamali, M., Peak, D. A., & Prybutok, V. R. (2017). *Beyond self-disclosure: Disclosure of information about others in social network sites.* Computers in Human Behavior, 69, 29-42.

Lee, C., & Cranage, D. (2018). Personalisation–privacy paradox: The effects of personalisation and privacy assurance on customer responses to travel Web sites. Tourism Management.

McCormick, M. (2008). *Data theft: a prototypical insider threat. Insider Attack and Cyber Security.*

Nosko, A., Wood, E., & Molema, S. (2010). *All about me: Disclosure in online social networking profiles: The case of Facebook.* Computers in Human Behavior, 21 (2), 306-348.

Obar and Oeldorf-Hirsch. (2018). The Biggest Lie on the Internet: Ignoring the Privacy Policies and Terms of Service Policies of Social Networking Services. p.33.

Okazaki, S., Eisend, M., Plangger, K,. Ruyter, Ko de., and Grewal. D. (2020). *Understanding the Strategic Consequences of Customer Privacy Concerns: A Meta-Analytic Review. [online]. [Accessed: 27-02-2023]. Available at: https://doi.org/10.1016/j.jretai.2020.05.007.* Science Direct.

Rouse, M. (2005). *What is pop-up ad? [online]. [Accessed: 22-12-2022]. Available at: http://whatis.techtarget.com/definition/pop-up-ad.*

Schneir, B. (2005). *Managed Security Monitoring: Network Security for the 21st Century.* Retrieved from http://www.counterpane.com.

Shabtai, A. and Kanonov, U. (2012.). "Andromaly": A behavioral malware detection framework for android devices. *Journal of Intelligent Information Systems.*, 38.

Shabtai, A., Elovici, Y. & Rokach, L. (2012). A Survey of Data Leakage Detection and Prevention Solutions. In A Survey of Data Leakage Detection and Prevention Solutions. pp. 5-11.

Son, J.Y. and Kim, S.S. (2008). Internet Users' Information Privacy-protective Responses: A Taxonomy and a Nomological Model. [online]. [Accessed: 10-10-2022]. Available at: http://web.a.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=1&sid=2514b829-460d-4661-a7fd-6cafbcee3307%40sessionmgr. *MIS Quarterly, Volume 32, Issue 3.*, 501-521.

Sun, Y., Fang, S., & Hwang, Y. (2019). *nvestigating privacy and information disclosure behavior in social electronic commerce.* Sustainability, 11(12), 3221.

TREPTE, S. L.-D., Aufl., 1., & . . Sabine TREPTE a Leonard REINECKE. Berlin, H. S. (n.d.).

Tufekci, Z. (2008). *Can You See Me Now? Audience and Disclosure Regulation in Online Social Network Sites. [online]. [Accessed: 06-12-2022]. Available at: 10.1177/0270467607311484.*

Vijayasarathy, L. R. (2004). *Predicting consumer intentions to use on-line shopping the case for an augmented technology acceptance model. [online]. [Accessed: 22-12-2022]. Available at: http://dx.doi.org/10.1016/j.im.2003.08.011.* Information & Management, 41(6), 747-762.

Wang, Y., & Nepali, R. K. (2015). *Privacy impact assessment for online social networks.* International Conference on Collaboration Technologies and Systems (CTS), 370-375.

Westin, F.A. (2015.). *Privacy and Freedom. ISBN: 978-19354-399-74.* Ig Publishing.

Wildt, A. (2016). *How to Use Personalization in Email Marketing Campaigns. [online]. [Accessed: 19-01-2022]. Availability: https://www.campaignmonitor.com/blog/email-marketing/how-to-use-personalization-in-email-marketing-campaigns/.*

8 List of pictures, tables, graphs, and abbreviations

List of pictures

List of tables

List of Pictures

9 Questionnaire survey items

1) What is your gender?
a) Male
b) Female

2) Place of birth
a) European Union
b) CIS Countries
c) Middle East and North Africa

3) What is your age?
a) 15 – 25
b) 26 – 35
c) 36 – 45
d) 46 – 55
e) 55 +

4) What is your occupation?
a) Student
b) Working full – time
c) Freelance
d) Unemployed
e) Retired

5) Current region of living
a) European Union
b) CIS Countries
c) Middle East and North Africa

6) Have you ever been hacked before, if so, how many times?
a) Yes, only once
b) Yes, more than 2 times
c) Yes, more than 4 times
d) Yes, many times
e) No, not even once.

Privacy Concerns and Protection.

1 = strongly agree

2 = rather agree

3 = I do not know

4 = rather disagree

5 = strongly disagree

Question	1	2	3	4	5	Source
I apply different passwords for different social media because it decreases the risk that my account will be hacked.						Boshmaf et el (2011) L. Wang et al. (2019)
I do not share my personal data on social media (only date of birth, name and surname) as there is a higher chance to be a subject of phishing letters, hacking, ransomware and many more.						
When I create an account for the first time on a website, I carefully read the policies and rules of that website.						
I only register on websites, which have a security certificate, it makes me feel secure and I can share more personal data such as credit card information, residency, siblings, occupation, and company name.						

Disclosure of sensitive information.

Question	1	2	3	4	5	Source
The social media should have an option to hide a sensitive information about myself If I want to						Dinev et al. (2013 Nosko et al (2010) Koohikamali et al (2017)
In case of a leakage of my personal data, the service provider or it's online–bot or representative should inform me as soon as possible.						
The service provider can only share my personal data at the request of my government or police for official purposes, however, I should be informed about it.						

Control over Personal Information.

Question	1	2	3	4	5	Source
I only use the two-factor – authentication (a password and a confirmation code) to log-in to my personal account.						Brandimarte, L., Acquisti, A., & Loewenstein, G. (2013).
I do fully understand that my personal information could be viewed by another user at any time, on a registered social media such as (Facebook, Instagram, Vkontakte, LinkedIn and etc).						

When I connect to free WiFi hotspots or networks, I do understand that my control over my personal information could be lost.						
I do not use a “remembered password” on my PC/phone/laptop, as it might undermine my control over my personal information.						

Identity Theft.

Question	1	2	3	4	5	Source
I usually pay attention to emails from the service provider about a new login or suspicious activity on my account and follow their instructions.						Y. Wang & Nepali, (2015) Johnson et al., (2012)
I do not add unknown contacts to my friend’s list or people who I personally do not know.						
I don’t share my geolocation through any application or website as it might lead to identity theft.						

Source: mentioned in the tables above.