iscte

INSTITUTO
UNIVERSITÁRIO
DE LISBOA

GDPR implications on social networks: perceptions of the IT specialists and non-specialists

Simão Afonso Filipe Branco Antunes Dias

Mestrado em Gestão de Sistemas de Informação

Orientador:
PhD Joaquim Reis, Assistant Professor,
ISCTE-IUL

Junho, 2020

iscte

TECNOLOGIAS
E ARQUITETURA

Departamento de Ciências e Tecnologias de Informação (ISTA)

GDPR implications on social networks: perceptions of the IT specialists and non-specialists

Simão Afonso Filipe Branco Antunes Dias

Mestrado em Gestão de Sistemas de Informação

Orientador:
PhD Joaquim Reis, Assistant Professor,
ISCTE-IUL

Junho, 2020

Acknowledgments

Não considero possível a realização de um projecto tão importante como uma dissertação um projecto individual, mas sim, um projecto e esforço colectivo feito por múltiplas pessoas, com o propósito de o sucesso de um único indivíduo. Sendo eu esse indivíduo, gostaria de deixar os meus sinceros agradecimentos.

O meu primeiro agradecimento é ao Professor Joaquim Reis, por sempre me apoiar e, com boa disposição, orientar-me e guiar-me a fim de atingir o meu objectivo. Obrigado Professor pela sua disponibilidade, atenção e por me orientar num dos projectos mais importantes da minha vida.

Um obrigado a todos os docentes que, de uma forma ou de outra, contribuíram para o sucesso da minha jornada como aluno do ISCTE.

Um obrigado também à minha namorada, Joana, por me aturar e apoiar nos dias em que parecia que tudo poderia correr mal, demonstrando sempre amor, carinho e compreensão pelo meu estado de espírito, sendo o sol dos meus dias cinzentos.

Aos meus irmãos por de uma forma ou de outra me apoiarem e animarem, principalmente nos dias em que isso parecia impossível. Obrigado por trazerem sempre o melhor de mim "ao de cima".

Um obrigado à AXA GO, empresa onde trabalho, e aos meus colegas, por sempre compreenderem a minha situação, disponibilizando-se sempre a ajudar-me. Agradeço toda a compreensão pelo esforço requerido por mim para escrever a dissertação e todo o apoio dado no sentido de me permitir ser bem-sucedido.

Por mim, um grande obrigado aos meus pais, pela fé inabalável no meu sucesso e por todos os contributos para o fim desta dissertação. São e sempre serão os alicerces para todo o meu sucesso e tudo o que tenho é graças a vocês. Desde a disponibilidade para aturar os meus dias mais difíceis, à capacidade e intenção de os melhorar, os meus pais são o fator comum em todas as minhas "aventuras" bem-sucedidas e, além disso, são o meu abrigo para as não tão bem-sucedidas. Obrigado por me motivarem quando já não me restava mais forças para lutar, pelas tardes que passaram ao meu lado a trocar ideias, pelos fins-de-semana de descontração antes de dias duros de trabalho, pelas palavras meigas, outras mais duras, mas sempre com o intuito de me fazer uma pessoa melhor, em todos os sentidos.

Sem todas as pessoas mencionadas em cima, nada disto seria possível. A todos o meu mais sincero agradecimento.

Resumo

A Sociedade do Conhecimento em que vivemos é caracterizada pelo crescimento exponencial de dados e pela capacidade tecnológica de os recolher, tratar e usar, com fins nem sempre observando princípios, éticos, deontológicos ou legalidade.

O Regulamento Geral de Proteção de Dados (RGPD) constitui-se como um instrumento de proteção da preservação de dados pessoais, num contexto de uma crescente adesão às redes sociais, em que os seus utilizadores, nem sempre detendo as necessárias competências em matéria de literacia digital, poderão expor os seus dados pessoais, desconhecendo e/ou não fazendo uso de estratégias de proteção de dados.

A presente investigação visa saber se o nível de literacia digital – que conduziu à divisão dos participantes entre especialistas e não especialistas em TI – tem influência no que respeita à informação revelada, às estratégias de proteção de dados usadas e à alteração de comportamentos, em função do nível de conhecimento sobre o RGPD, por parte dos utilizadores das redes sociais.

A adoção de uma metodologia quantitativa, com uma pesquisa descritiva e também exploratória, a aplicação de um questionário, usando o *Google Forms*, a partir de uma estratégia de recolha de dados do tipo *snowball*, permitiu a resposta a 608 participantes.

Partindo do pressuposto que o nível de literacia digital determinaria uma menor exposição dos dados pessoais, um maior uso de estratégias de proteção de dados e um melhor conhecimento do RGPD, por parte dos participantes com maior nível de literacia digital (Especialistas em TI), os resultados obtidos permitem concluir que as diferenças não são significativas.

Palavras-Chave: Redes Sociais; Privacidade na Internet; Estratégias de proteção da privacidade; Literacia digital; Regulamento Geral de Proteção de Dados (RGPD)

Abstract

The Knowledge Society in which we live is characterized by the exponential growth of data and the technological capacity to collect, treat and use, for purposes not always observing principles, ethics, deontology or legality.

The General Data Protection Regulation (GDPR) can be used as an instrument for the protection of personal data, in a context of increasing adherence to social networks, in which its users, not always having the requirements in digital literacy topics, can export their personal data, unaware and / or not using data protection strategies.

This investigation aims to find out if the level of digital literacy - which led to the division of participants between IT specialists and IT non-specialists - has an influence regarding the information revelation, the data protection strategies used and the behavior changes, depending on the level of knowledge on the GDPR, by users of social networks.

Adopting a quantitative methodology, with descriptive and also exploratory research, a questionnaire application, using Google Forms, using a snowball data collection strategy, allowed to obtain 608 participants' answers.

Based on the assumption that the level of digital literacy determines a lower exposure of personal data, a greater use of data protection strategies and a better knowledge of the GDPR, by participants with a higher level of digital literacy (IT specialists), the obtained results allowed to conclude that differences are not significant.

Keywords: Social network sites (SNSs); Internet privacy; privacy protection strategies; Digital literacy; General Data Protection Regulation (GDPR).

Index

Table of Index

Picture's Index

List of Abbreviations and Acronyms

DAP – Data Protection Act

EEA – European Economic Area

ETL – Extract, Transform and Load

EU – European Union

GDPR – General Data Protection Regulation

ICT – Information and Communication Technology

IT – Information Technology

RGPD – Regulamento Geral de Proteção de Dados

SNSs – Social Networks Sites

SPSS – Statistical Package for the Social Sciences

Chapter 1 – Introduction

The Introduction contemplates the scope, the motivation, the research questions and the objectives, the methodological approach and, finally, the structure and organization of the dissertation.

1.1. Scope

The thematic addressed by this master's dissertation is the perception of the implications of the GDPR (General Data Protection Regulation) on Social Networks Sites (SNSs).

The dissertation aims to understand the differences in information disclosure behaviors between IT specialists and IT non-specialists, how these behaviors changed with the entry into force of the GDPR, as well as the strategies used to protect the privacy of personal data. For the purpose of segmenting the sample under study, IT specialists are those who have professional training or experience in the IT area and IT non-specialists are all the other elements of the sample. Finally, and in view of the need that the SNSs management entities had to adjust their mode of operation, in compliance with the GDPR, this dissertation seeks to identify the way in which IT specialists and IT non-specialists perceive this process of change.

The scope of the dissertation is considered relevant in a context of a knowledge society, where the exponential growth of data and information produced, as well as the technological capacity of collection, processing, storing and availability of this information, often with illicit access and use, makes the possibility of protecting personal data asymmetric.

1.2.Motivation

This master's dissertation will focus on the most recent personal data protection law of all citizens of EU countries and the European Economic Area (EEA), GDPR, seeking to understand the implications of it on SNSs. The relevance of the research around this thematic gain emphasis due the big amount of data collected by those social networks, whose volume of personal information collected from its users is substantial (substantial enough for the use of the term Big Data), that it needs data privacy laws, such as GDPR, to control it and guarantee that there is not any unfair use of the information provided by those mentioned users.

According to Fan & Zhao (2015), Big Data can be defined as the amount of data too large to be stored, managed and processed efficiently through conventional software. To handle large volumes of information such as this, there are Big Data tools, processes, methods, and techniques that are extremely important, such as predictive analytics and user behavior Analytics techniques. The techniques of predictive analytics are closely linked with data mining, predictive modelling tools, among others, and those tools allow analysing current or historical data, in order to predict certain events (Hazen, Boone, Ezell, & Jones-Farmer, 2014).

The application of the GDPR entails implications for the use of these tools, processes, methods and techniques, since they process information covered by the GDPR, i.e. information on all EU citizens (Tankard, 2016).

Therefore, it is important to know if the users of the SNSs are aware, not only of the information they provide every day when using these sites, but also the most recent legal provision (GDPR) that the SNSs management entities must comply with, regarding collection, storage and use of that data.

In this sense, this dissertation has the objective of making a diagnosis about this problem, looking for differences between IT specialists and IT non-specialists, that may lead to reflection and discussion about whether the simple entry into force of the GDPR constitutes effective means to protect SNS users in safeguarding the privacy of their personal data against access and unlawful use of that data.

Being a recent issue, there is not a significant amount of scientific literature to support the investigative strategies to be followed, which substantiate the relevance of the delimitation of the theme, to which, given the broad adherence to the SNSs, as well as the breadth of application of the GDPR, make the realization of the present study urgent and pertinent.

In addition to the scarcity of scientific production, personal affinity with the theme, for reasons of the master's professional activity, social relevance is a driving force that induced the choice. The social relevance is linked, as mentioned above, to the broad adherence of the SNSs and to the sensitivity regarding the privacy of personal data.

1.3. Research question and objectives

In order to answer the research question, the investigative strategy aims to fulfill the following general objectives:

- To know how the level of digital literacy (IT specialists versus IT non-specialists) influences the behavior, concerns and protection strategies adopted in terms of privacy on social networks;
- Find out about possible changes in behavior, regarding privacy on social networks, due to the entry into force of the GDPR;
- Discuss, on a reflexive basis, whether the differences between the group of belonging, with regard to digital literacy, induce a greater or lesser concern to the way that the management entities of social networks proceed with compliance with the GDPR.

In chapter 3 the general objectives now presented will be broken down into specific objectives.

Finally, all the research carried out will have an ultimate purpose, which is to answer the research question:

> **"In what way does the level of digital literacy, on the domain of GDPR, affects the SNSs users' behavior regarding the privacy and granting of access to their personal data?".**

1.4.Methodological approach

The methodological options resulted in the research design presented in chapter 3.

The present investigation assumes a double orientation: first, it is a descriptive research, seeking to establish relationships between variables regarding the personal data availability in the SNSs; secondly, regarding the implications of the entry into force of the GDPR and the possible change in behavior of the personal data disclosed, the research has an exploratory nature, having in mind that there are no known studies regarding this theme.

In the first part of the empirical component, with the existence of previous studies, the investigations by Govani and Pashley (2005), Tufekci (2008) and Young and Quan-Haase (2009, 2013) will be the references. For the second part, due to the lack of studies focusing the implications of the entry into force of the GDPR, we resorted to some scientific and technical production on the GDPR.

Following a quantitative methodology, a questionnaire was built adjusting the questions formulated in previous studies, and already mentioned, with the technical assessments on privacy of personal data introduced by the GDPR and presented by authors and specialized organizations.

All the treatment, analysis and discussion of the results was processed by the SPSS Statistics software in both descriptive statistics and the statistical tests, and its interpretation arising from the scientific production from Marôco (2011).

1.5.Structure and organization of the dissertation

The present dissertation is organized in five chapters that aim to reflect the different phases until its conclusion.

The first chapter, Introduction, presents the scope of the investigation, the motivation, the research question and objectives, the methodological approach, as well as the structure and organization of the dissertation.

The second chapter reflects the theoretical framework, called literature review, where the concepts and studies previously carried out in this context are presented.

The third chapter is dedicated to the methodology used in the data collection and treatment process, as well as the analysis methods used. The research design, the objectives (general and specific) and the research question, and hypotheses that guide the research will be presented.

The fourth chapter where the results will be analyzed and discussed. The use of SPSS Statistics will allow an analysis of the results obtained and its comparison with those of other previous studies. Finally, with the completion of the appropriate tests, hypotheses will be validated or not.

In the fifth and last chapter, the conclusions of this research are presented, as well as the limitations and future work.

Chapter 2 – Literature Review

Literature review is an essential step in the investigative strategy, with no sense of allowing scientific literature, the main authors and studies, any gaps in the level of research in the study area, as well as tips on the methodological options to be followed in the investigation (Quivy and Campenhoudt, 2018).

In order to fulfill the objectives and answer the research question, the present literature review concepts such as 'Digital Literacy', 'Big Data', 'Social Network Sites' and 'Data Privacy', having as a reference the implications of the entry into force of a new regulatory framework - the General Data Protection Regulation (GDPR).

2.1. Digital Literacy

The notion of Digital Literacy has been changing over the last twenty years, the arguments for how it should be defined date back to 1980s (Buckingham, 2010). The concept of Digital Literacy, as it is now generally used, was introduced by Paul Gilster (Gilster, 1997), and it is described as the ability to understand and to use information from a variety of digital sources.

This definition is aligned with the definition of literacy itself (the ability to read, write and otherwise deal with information using the technologies and formats of the time), applied to the concept of the digital strand (Lankshear & Knobel, 2008).

This being said, this definition should be considered a broader explanation of the concept of Digital Literacy, therefore, the concept must be more than the ability to use digital sources effectively. It is also about the mindset and ability to master the use of those technologies and take benefits from it, applying them in your life (Gilster, 1997).

The Digital Literacy, according to Shapiro and Hughes (Shapiro & Hughes, 1996), can be divided into seven dimensions:

- *Tool literacy* – understanding and use of practical and conceptual tools of IT, including software, hardware and multimedia;
- *Resource literacy* – understanding the forms and access methods of information resource, especially networked resource;
- *Social-structural literacy* – knowing that and how information is socially situated and produced;

- *Research literacy* – ability to use IT tools for research and scholarship, namely either for work or education purposes;
- *Publishing Literacy* – ability to format and publish researched ideas, either using websites or other means of communication that are based on digital resources;
- *Emerging technology literacy* – following the most recent innovations on IT and having the ability to understand and adapt to them;
- *Critical Literacy* – ability to evaluate the pros and cons of information technologies (benefits and costs).

2.1.1. Different Generations

The Generation Gap in digital literacy is evident, and there are two different generations whose languages are different. Those two generations are designated as Digital Natives and Digital immigrants (Fieldhouse & David, 2008).

Digital Natives or Net Generation is the term to designate those who were born between 1980-1995, also designated as millennials (Helsper & Eynon, 2010). To them the concept of digital and technology is part of their life, since ever, not having other alternative (Fieldhouse & David, 2008). For that generation, new technologies have been such a defining feature, that those technologies define the way of communication, socializing, creating and learning (Helsper & Eynon, 2010).

Digital immigrants are those who were born before that period, having on their experience a reality of an era pre-technology, using the term "digital" as differentiator between electronical and digital versions of the same activity (Fieldhouse & David, 2008).

There are two key distinctions between both of them:

- *Adaptation* – most of the digital immigrants will always retain certain habits and ideas from the past, even though they might be trying to adapt to the new digital reality in some tasks (Prensky, 2001). Some of those tasks, such as socializing and researching, are some of the examples of tasks that are a struggle for digital immigrants.
- *Language/communication* – Digital Natives have different ways of processing and using information that, for example, does not fit well on the current educational

practices (Prensky, 2001). Therefore the methods of education become inefficient and need to be changed, having the educators to adapt their mean of communications to the needs of the Digital Natives (Helsper & Eynon, 2010).

Authors like (Oblinger, Oblinger, & Lippincott, 2005) affirm that age is not key to define the difference between Digital Natives and Digital Immigrants, but the exposure to technology is. Digital Immigrants that have the need to use technology on their day to day, either for the means of studying or working, socializing, or for any other reason, tend to be as used to technology as Digital Natives are.

2.1.2. Communication and Social Networks Sites

Digital Natives tend to exhibit more tendencies of communication and team working with either teams or peers, than Digital Immigrants (Oblinger et al., 2005).

Digital Natives give high importance to social networks sites and digital means of communication, they use them extensively, either for networking or socializing with their friends list, virtual communities or even to share their personal lives.

Current technology allows them to use the Internet as a way of expressing their feelings and thoughts, meet new people and it even allows them to meet new cultures.

Signing at a digital network is the most basic social networking skill of digital literacy (Knobel & Lankshear, 2008). All the process that is around the creation of a profile, which components the user allows the others (either friends or unknown people to the user) and sharing of texts or images/videos, are considered a way of expressing knowledge on the technology in use (on the dimensions of Social-structural and Publishing Literacy (Shapiro & Hughes, 1996).

However, Digital Immigrants might also interest on using social networks sites. Geographical distance, time-consuming obligations, among other factors, might be some impediments that lead to the need of use of social networks sites (Leist, 2013). Social Networks Sites may overcome this by allowing social engagement/contact regardless of geographical location or time.

When Social networks are mentioned some names have mandatory mention (for example Facebook), however, contacting people who are on our friends list is not the only purpose of some Social Networks Sites, for example Linkedin, whose purpose is to, for example, networking with people for the purpose of recruiting or being recruited.

Digital Networks are a perfect example of a technology that can be used by Digital immigrants, due to the simplistic interface and way of accessing it. Some of the only prerequisites in social media are the ability to use a computer or a web-enabled device and ICT (Information Communication Technology)-related knowledge (Leist, 2013), meaning basic skills as browse the Internet, sending emails, among others.

Although we can assess, from the previous paragraphs, that both Digital Natives and Digital Immigrants have the basic abilities to use the Social Networks Sites, they might be unaware that, when they register and provide personal information, they are providing information that has potential commercial value (Peras, Mekovec, & Picek, 2018).

One of the terms that highlights, when mentioning collecting user's data, is Big Data.

2.2.Big Data

With the advance of technology, the amount of generated of both traditional, structured, transactional data as well as more contemporary, unstructured, behavioral data has increased a lot (Erevelles, Fukawa, & Swayne, 2014)

This led to businesses generating more data than they are able to use or take profit from (Fayyad & Piatetsky-Shapiro, 1996)

That amount of generated data, that is too large or complex for traditional data-processing applications software to process, is given the term of Big Data.

Big Data can be defined with 5 dimensions (3 main dimensions and 2 associated dimensions), that are called the Vs (Erevelles et al., 2014), as shown at Figure 1.

- *Volume (Main dimension)* – volume is one of the most distinct ways to characterize Big Data, however, having a big volume of data can sometimes lead to the lack of Velocity and does not mean that data has Value;
- *Velocity (Main dimension)* – considered to be the second main dimension of Big Data (Lycett, 2013) it is one of the most relevant dimensions, especially for Marketing where useful knowledge today is outdated information tomorrow;

- *Variety (Main dimension)* – considering that Big Data is, by definition, a big amount of data, it does not mean necessarily that the data is diverse or useful;
- *Veracity (Associated Dimension)* – although it is not considered one of the main dimensions of Big Data, veracity is a very important characteristic of Big Data. It underscores the need to be aware of data quality (Paper, 2018).In a time where the volume and variety of information is increasing, keeping the veracity of the data is a major issue (Dijcks, 2012);
- *Value (Associated Dimension)* – the ever-increasing amount of data might lead to question its value. It is very important to eliminate not relevant data. The rest of the data, in order to be valuable, must be analyzed by someone who has insight and domain-specific interpretation (Lycett, 2013).

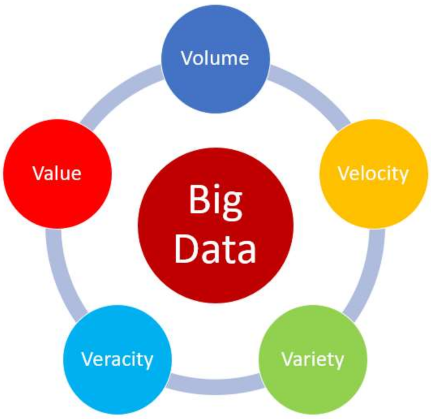


Figure 1- 5Vs of Big Data (Dillon, n.d.)

2.2.1. Value Creation through Big Data

Big Data can lead to value creation, there are several examples of value creation through diverse fields such as the pricing, product, etc. Value creation is whenever there is an increase in the worth of goods, services or businesses. Nowadays, many businesses focus on value creation for customers purchasing its products or services ("What is value creation? definition and meaning - BusinessDictionary.com," n.d.).

2.2.1.1. Pricing

For example, within the pricing section it is possible to understand the importance of dynamic prices on stores such as Amazon or eBay, where the demand is higher, and competition intensifies. It is possible to such companies to dynamically changed their price. This way this is possible is by integrating various sources of information and variables within the data, such as, consumer demand, number of page views, etc. The ability to perform this creates value since it increases the ability of the companies to adapt to consumer behavior (Erevelles et al., 2014)

2.2.1.2. Products

Instead of the classic way to obtain information using surveys, most of the big companies use costumer's behavior data in order to take decisions regarding the quality or improvements of their products. These Big Data techniques are designated as costumer analytics. For example, the company Ford used those techniques on its own revolution in product innovation and design, capturing consumer data from around four million of its vehicles on the road through sensors and remote app-management software (King, 2012).

2.2.2. Big Data processes

Having Big Data just by its own is worthless, the ability to achieve its potential value emerges when leveraged to drive decision making (Gandomi & Haider, 2015).

In order to make sure that every decision taken goes according to the possessed data, it is important for organizations to efficiently process big volumes of data into meaningful insights. This process can be broken down in five stages (Labrinidis & Jagadish, 2012) and those five stages can be grouped in two subprocesses: Data Management and Analytics, represented on the Figure 2.

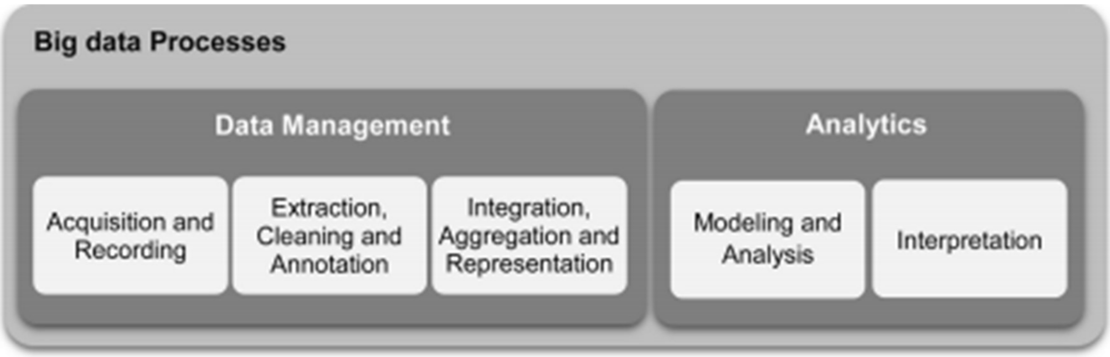

Big data Processes
Data Management
Acquisition and Recording
Extraction, Cleaning and Annotation
Integration, Aggregation and Representation
Analytics
Modeling and Analysis
Interpretation

Figure 2- Big Data Processes (Gandomi, Amir & Haider, Murtaza, 2015)

2.2.3 Data Management

Data management is every process where data, with the support of technologies, is acquired, stored, prepared and retrieved to analysis (Gandomi & Haider, 2015).

This subprocess is divided in three stages:

- *Acquisition and Recording* – current data sets are growing because they gather information from all types of technology that the companies provide to the users. Big Data has changed the way data is stored, from the data storage device and data storage architecture to data access mechanism (Chen & Zhang, 2014);
- *Extraction, Cleaning and Annotation* – before all the data can be analyzed, it must pass through the process of Extract, Transform and Load (ETL) – Figure 3. Only the first two steps (Extract and Transform) are used on this stage. The Extract tasks are responsible to access various sources in order to extract the selected data to analysis purposes. In order to increase the homogeneity of data, we use Transform Tasks, where the data is then standardized using diverse techniques of transformation (cleansing, filtering, merging, etc.);
- *Integration, Aggregation and Representation* – this is the last step of the ETL process where there is the Loading. The Loading tasks are done in order to make sure that the prepared data is charged on the warehouse.

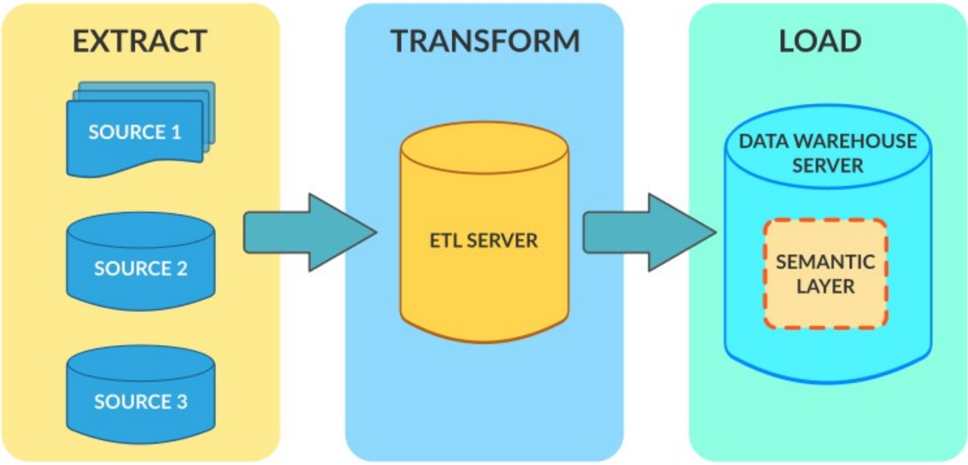


Figure 3- ETL Process (Shimko, 2020)

2.2.4 Analytics

Data Analytics can be defined as technologies (such as database and mining tools) and techniques (such as analytical methods) that organizations can use to analyze large scale, complex data for various applications.

This subprocess is divided in two stages:

- *Modelling and Analysis* – this stage refers to all the techniques used analyze the data that was loaded. Some examples of those techniques are: Text analytics, Audio analytics, Video analytics, Social media analytics, Predictive analytics, etc. (Gandomi & Haider, 2015);
- *Interpretation* – this stage refers to the last part of all the Big Data process and, without any disregard to the other stages, it is one of the most crucial stage of the all process. This stage is where the analysts do the interpretation of the data and results gotten, from the previous stage, and pursuit the decision making that will result on the added value to the company (Gandomi & Haider, 2015).

2.2.4.1. Text analytics

Text analytics/mining refers to all techniques that extract data from textual data. Text analytics involve statistical analytics computational linguistics, and machine learning (Figure 4).

Due to this techniques it is possible for businesses to convert large volumes of data to summaries, that can lead to better decision making (Gandomi & Haider, 2015).

As an example, this technique can be used to me extract information from financial news and, after analyzing, it can be used to predict stock market (Chung, 2014).

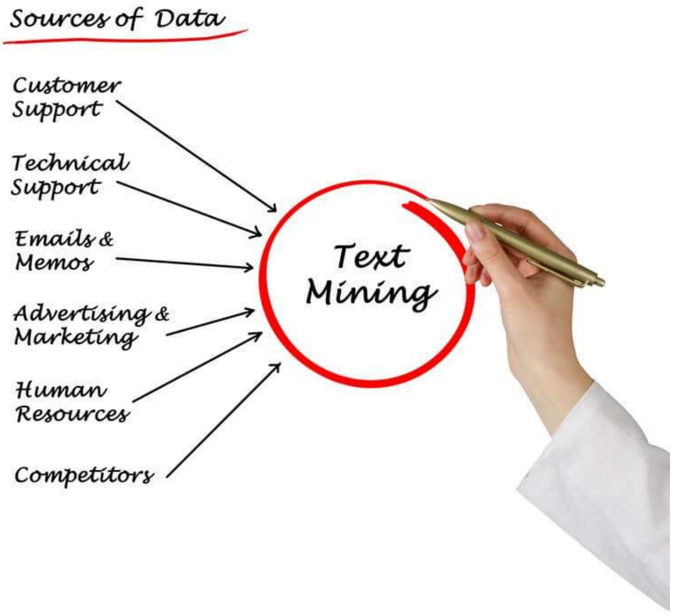


Figure 4- Text Mining (Vadakkanmarveettil, 2014)

2.2.4.2. Audio analytics

Audio analytics analyze and extract information from unstructured audio data.

By analyzing video-calls, customer call centers, etc., it is possible to improve customer experience, evaluate agent's performance, enhance sales turnover rates, monitor compliance with different policies (e.g. privacy and security policies), gain insight into customer behavior, and identify product or service issues, among many other tasks (Gandomi & Haider, 2015).

2.2.4.3. Video analytics

Video analytics, compared to the other methods of mining, is very recent (Panigrahi, Abraham, & Das, 2010). Marketing and operations management are the primary application areas for this method. For example, in retail, video analytics can help in the study of buying behavior of groups, by measuring the time each group of people stays in each segment of a store and, correlating this information with costumer demographics can lead to value creation on the decision making, as result, it can be, for example, decided where to place each product, define prices, layout of store, among others (Gandomi & Haider, 2015)

Another example for this is YouTube. The possibility to have an overview on the most viewed content, separating the videos by tags, and get a statistic of the most viewed tags, can lead to value creation on placing Advertisement of certain products on certain videos tagged, correlating the type of product with the video tag, by this mean, enhancing the contact with the targeted consumer.

2.2.4.4. Social media analytics

Social media is currently one of the most powerful tools to address marketing to costumers.

It provides the ability to collect data from users' web browsers and, by that mean, display advertisement that is fully directed to them. This is possible due to what is called Social Identification theory. This theory clarifies how individuals improve self-esteem and self-affirmation through categorization, identity and comparison (Tajel & Turner, 2004).

There are two types of source that provide social media data to the companies:

- User generated data (any publication, photograph, video, etc. posted by any user);
- Relationships and interactions between the network entities.

The social media analytics, whose purpose is to analyze these two types of sources of information can be divided in two groups (one for each type of source) (Gandomi & Haider, 2015):

- *Content-based analytics* – these analytics focuses are the data posted by users on the social media platforms (e.g. customer feedback, product reviews, images, and videos). Since this is a big volume of unstructured information, most of the techniques describe above (text, video and audio analytics) can be used to process this data.
- *Structure-based analytics* – these analytics focuses are on the relationships among entities, synthesizing the structural attributes of a social network site and extracting intelligence from those relationships.

2.3.Social Networks Sites and Data Privacy

When mentioning Social Networks, one of the most common topics that comes to mind is Privacy. While using them, most of the users are invited to provide private information in order to create a profile. Despite all the concerns about the possible consequences of disclosing private information, users fill those profiles (Young & Quan-Haase, 2009). Mentioned profiles contain information such as: cell phone number, location, profile picture, sexual and political orientation, etc. (Govani & Pashley, 2005).

The privacy of personal data on the Internet has acquired particular relevance with the exponential increase in the ability to collect, aggregate, tagging and cross-indexability, allowing search and accessibility to personal information (Tufekci, 2008; Young & Quan-Haase, 2013), breaking the boundaries of the personal sphere and distorting its meaning when decontextualized (in relation to space and time), generating new types of threats. Referring to the studies by Palen and Dourish (2003), Tufekci (2008, 22) enunciates the different dimensions in which these threats (or challenges) can be seen, based on the capabilities conferred by information technologies in accessing personal information: "threats to spatial boundaries, threats to temporal boundaries because of persistence of data, and intersections between multiple spaces".

Social Networks Sites are increasing on the number of users, and what continues to attract more users is the possibility of chat with their friends, share digital data and connect to more people (Young & Quan-Haase, 2009). Govani and Pashley (2005) also conclude that more information is being displayed on the websites and Young and Quan-Haase add that the more time that users use them, more information are they likely to reveal. Having this in mind, it can be concluded that there is little to no relationship between online privacy concerns and information disclosure.

So, how can the users prevent this? There are some known strategies applied by the users to prevent this from happening. These strategies are mainly the excluding contact information, limit the access to your profile, untaging or removing photographs from your profile, decline connection requests, among others (Young & Quan-Haase, 2009).

However, applying these strategies might lead to not having the most benefits out of the social networks since, for example, not providing your location might stop social networks from recommending you a good restaurant nearby. Thus, the justifying the disconnection between online privacy concerns and information disclosure, creating the concept of privacy paradox (Norberg, Horne, & Horne, 2007).

Privacy paradox shows that even though users have more and more information about the privacy concerns and scandals that surround social networks, they tend to provide more personal information about themselves, as a way of taking full benefits of the social networks and, as exchange, social network companies are able to hold more personal information of the user.

Given the purpose of the present study, at this point we will address the issue of data privacy, the way this question arises when contextualized in the field of social networks sites, using the studies previously carried out, in particular the identification of the relevant variables for this investigation.

2.3.1. Data Privacy

Privacy is considered by Diamantopoulou, Androutsopoulou, Gritzalis e Charalabidis, (2020) as the right of individuals to determine what information is accessible, to whom and when. This matches with some of the ways more experienced users use the privacy settings of their networks, in order to have more control of their information (Boyd & Hargittai, 2010).

A very important concept to have in mind is the Latent-data privacy. Latent-data privacy is the fact that one can suppose some personal information, by having access to other personal information (He, Cai, & Yu, 2018). Meaning that even if social networks' users do not disclose some interest or information, that information can be inferred by the access to other personal information.

Latent-data Privacy concept is intertwined with the all process of data management (mentioned on the previous chapter), since it is using data-mining techniques that it is possible to withdraw personal information that the user didn't want to share.

Data privacy, or information privacy, is an area of data security that has to do with how data is handled properly – consent, notice and regulatory obligations – and is concerned with issues such as how the data is collected, stored and made available to third parties, in compliance with the current regulatory provisions (Petters, 2020). Whenever the issue of data privacy appears associated with the Internet, this concept is identified with another name – Digital Privacy.

One of the sensitive aspects of Digital Privacy has to do with the possibility that each individual has control over the exposure and availability of data related to themselves (Belyh, 2015), in the context of a digital age where everyone is related to everyone, where access to data is facilitated and makes everyone's privacy condition vulnerable.

In view of the complexity of the theme, and of the current regulatory standards (Petters, 2020), digital privacy has received increasing attention, both in terms of scientific production and the appearance of multiple companies specialized in providing consultancy services. The different standards, both European and others, tend to identify the type of data and the processes associated with the collection, storage and availability of data, in a confusing and ambiguous way, which seems to be the result of the lack of a conceptual definition rigorous and objective (Belyh, 2015).

As stated Petters (2020), if the information constitutes one of the major business assets, the use of information of third parties for these companies tends to be particularly relevant, especially when it involves the individual's right to privacy, even quoting an Information and Privacy expert: “Privacy forms the basis of our freedom. You have to have moments of reserve, reflection, intimacy, and solitude”. The malicious and fraudulent use of information by third parties is one of the main risks.

In general, data privacy includes, and as Behyl (2015) states:

- *Online Privacy*: This includes all personal data that is given out during online interactions. Most sites have a privacy policy regarding the use of the data shared by users or collected from users;
- *Financial Privacy:* Any financial information shared online or offline is sensitive as it can be utilized to commit fraud;
- *Medical Privacy:* Any details of medical treatment and history is privileged information and cannot be disclosed to a third party. There are very stringent laws regarding sharing of medical records;
- *Residential and geographic records:* sharing of address online can be a potential risk and needs protection from unauthorized access;
- *Political Privacy:* this has become a growing concern that political preferences should be privileged information.

2.3.2. Data Privacy in Social Networks Sites

Although many social networks sites offer options on the privacy regime for different types of personal information, many of the users, either due to an apparent confidence and feeling of security, or due to lack of care, continue to behave negligently about the availability of such data (Acquisti & Gross, 2006; Govani & Pashley; 2005; Gross & Acquisti, 2005). The consequences on this risky online behavior are identity theft, stalking, harassment, spamming and fraud (Belyh, 2015; Govani & Pashley; 2005; Young & Quan-Haase, 2013).

Next, we will address the central issues regarding data privacy, both in terms of the information revelation and in terms of the strategies used to guarantee non-access by unwanted audiences, with reference to some of the studies previously carried out.

2.3.2.1. Information revelation

Users can share with other users a considerable diversity of extremely accurate data (Acquisti & Gross, 2006), in particular personal data such as age, gender, sexual orientation, tastes and preferences, photos, contact information, relationship status and partner, political affiliation, professional or school career (Govani & Pashley, 2005) and who does so, in many cases, intentionally, disqualifies the risks that the availability of that data brings (Young & Quan-Haase, 2009). In studies carried out, these authors report that the availability of this personal information occurs in more than 60% of cases.

In social networks sites without a professional dimension, such as LinkedIn, the choice of the type of user who can access the profile can be configured, as well as the type of information they can know, giving rise to selective access, making it relevant if it is relevant discuss the issue of privacy of personal data, given the possibility for each user to control who accesses that information (Govani & Pashley, 2005).

In a study by Govani and Pashley (2005), on the configuration of the privacy level by Facebook users, it was concluded that more than 80% (approximately the same value as those who had not read the privacy policy) knew the settings privacy, but more than half of them had not done so. An underestimation of the risk or lack of information about the risk of disclosure of personal information, associated with high exposure on social networks sites, is only considered in view of the real possibility of control, stalking, harassment, spamming and fraud by other users (employers, colleagues, friends or parent), which will have led to less availability of information regarding contact details (telephone number and e-mail address) or home addresses.

In the same line of investigation, we find the studies by Tufekci (2008), focusing on issues such as the general concern with online privacy, unwanted audiences and likelihood of future audiences (employer, romantic partner, government agency), because of searchability and persistence of online records, in continuation of studies by Acquisti and Gross (2006). In this study, carried out with higher education students, Tufekci (2008, 33) refers to another perspective on the conceptualization of privacy regarding personal information:

> «starting from a conceptualization of privacy as a boundary negotiation process and "selective access to the self," we tried to move beyond the dichotomy between "students say they are worried but they don't care" and "students say they are worried but they don't know" and offer a another possibility: Students do try to manage the boundary between publicity and privacy, but they do not do this by total withdrawal because they would then forfeit a chance for publicity. Students attempt to optimize their privacy and restrict who can find them by using monikers that they can share with only those they want to be found by or by restricting the visibility of their profiles to only "friends.».

Almost all SNSs allows for various levels of privacy control, with 'visibility' being one of the most important. As 'name' is one of the most used items in terms of searchability, Tufekci (2008) is surprised by more than 90% using his real name, noting, in line with previous studies, that SNSs users are aware of the level of visibility and searchability of their profiles, just adjusting that level of visibility before unwanted audiences.

The level of visibility of some personal information is extremely high (Tufekci, 2008): more than 70% (in some cases, close to 100%) indicate birth date, e-mail address, affiliation and political views, religion, music, book, and movie preferences, school name, relationship status, sexual orientation and the current city or town in which they live and post an image of themselves and photos of their friends; almost half indicate cell phone number; and only less than 30%, physical address and cell phone number.

Given the persistence of information in a digital setting and the possibility of research by future audiences, seeking to assess the perceived degree of threat, Tufekci (2008) concludes that the level of threat was different, fearing respondents more about its use by future romantic partner than that by future employer or government.

The conclusions of the study by Tufekci (2008), in a technologically mediated society, point to a compromise between the need to be in social networks sites and to be visible and due precaution in privacy management: The mean response to the question about general online privacy concerns was 2.76 (4=high concern, 1=no concern), indicating some, but not extreme, concern." (Tufekci, 2008: 26).

Recognizing the importance that the issue of quantity and type of information disclosed on social networks, Young and Quan-Haase (2009) try to identify the factors and analyze the reasons why users make this information available. Not being the objective of the present investigation to identify the reasons, we focus on the factors that can influence the quantity and type of availability, as well as the visibility of the profiles.

Young and Quan-Haase (2009), in line with previously mentioned studies, identify a positive correlation between the frequency of SNSs and the amount and type of information revealed. This conclusion is confirmed in a subsequent study, noting that more than 80% of users access SNSs several times a day (Young & Quan-Haase, 2013).

In addition to the frequency of access, and using several studies, Young and Quan-Haase (2009) identify three factors associated with information revelation: (1) network size; (2) concern about Internet privacy; and (3) concern about unwanted audiences.

Regarding the network size, the first factor, they found that the larger users' SNSs size, the more information is revealed in their profiles. Information revelation, according to the authors, will be associated with the need for greater interaction and social participation and the formation and maintenance of relationships. According to Young e Quan-Haase (2009: 268),

> "Nearly two-thirds of respondents indicated their sexual orientation, relationship status, and interests (such as favorite books, movies and activities). Large percentages of respondents noted their school name (97.4 per cent), e-mail address (83.1 per cent), birth date (92.2 per cent), the current city or town in which they live (80.5 per cent), and almost all respondents reported posting an image of themselves (98.7 per cent) and photos of their friends (96.1 per cent).".

In turn, the second factor, research establishes a negative association between concern for Internet privacy (over 80% of users show concern about Internet privacy) and information revelation. That is, users with a high level of concern for Internet privacy tended to disclose less personal information on Facebook. As for the concern with unwanted audience, if research has demonstrated that general concern for Internet privacy influences the information revelation behaviors of Internet users, the study by Young and Quan-Haase (2009) has opposite conclusions. This paradox, according to the authors, can be explained by the fact that the closure in relation to the unwanted audience allows a greater exposure of personal information to the users who belong to the network.

Finally, the third factor: profile visibility. Profile visibility refers to the extent to which users' profiles are accessible by other SNSs users. According to Young and Quan-Haase (2009: 272), "the less an individual closed their profile to others, the more information they revealed. This suggests that individuals, who are generally concerned about their privacy and hence close their profile to only friends, will reveal less information than those who do not manage their profiles".

In line with previous studies, Young and Quan-Haase (2009) claim that the less closed the user profile is, more information it makes available to other users, indicating, by contrast, that a greater concern with personal data privacy leads SNSs users to manage the configuration of the privacy settings (Young & Quan-Haase, 2013 Thus, Young and Quan-Haase (2013: 487) refer that "many respondents altered the visibility of their profile data: 79 percent regulated access to tagged photos, 77 percent restricted access to their wall, and 71 percent limited access to their news feed".

According Young e Quan-Haase (2009: 269), regarding concerns about unwanted audiences, "The concern is highest for the following groups gaining access to private data: political parties, sexual predators, employers and university administrators.".

2.3.2.2. Privacy Protection Strategies on Social Networks Sites

The objective of establishing a balance between disclosure, advertising and privacy, without the loss of control over the accessibility of profiles by unwanted audiences, is one of the biggest concerns of SNS users.

The studies carried out by different authors do not allow us to withdraw a trend in terms of specific disclosure and audience management techniques. Different authors recognize the existence of a multiplicity of privacy protection strategies and techniques regarding unwanted audiences, from use of fake or inaccurate information, restricted access to their profile and withheld information that could be used to link them to a physical location, exclusion personal information from their profiles, the use of private email messages to communicate, changing the visibility of their profiles by changing the default privacy settings, using nicknames or monikors in place of real names, by either untagging or removing photographs or by making use of the limited profile option, to restrict certain contacts or groups of contacts from viewing specific types of personal information to block unwanted audiences.

Gross and Acquisti (2005) and Govani and Pashley (2005) found that despite awareness and concern for Internet privacy, users seldom provide false information and very rarely alter their privacy settings. Tufekci (2008) found that SNSs users' concern for unwanted audiences accessing their profiles influenced them to use protective measures, such as altering the visibility of their profiles and using nicknames or monikors in place of real names. Finally, the privacy protection strategies against threats that were found by Young and Quan-Haase (2009, 2013) indicates that SNSs users were more likely to exclude personal information to restrict unknown others from accessing information, to use private email messages to restrict access to content, and to alter the default privacy settings than they were to use fake or inaccurate information or to block contacts.

2.4.General Data Protection Regulation (GDPR)

Information technology is considered a major threat to user's privacy, since it enables not only the ability to gather a large amount of information but also save it in large databases. Thus, the need for a way to ensure that all the information provided by the users must be protected from unauthorized access (Peras et al., 2018).

One of the most recent and important regulation that comes to ensure this is the General Data Protection Regulation (GDPR).

The General Data Protection Regulation is a rebuilt of the Data Protection Act 1998 (DPA) that focus the protection of private information of the consumer, providing more control and permissions, aiming to allow the consumer to decide what to do with that information (Beckett, 2017).

This regulation not only brings new advantages and disadvantages to the Data Management possibilities, but also tries to follow up with the most recent progresses made in technology, changing the way some data is categorized. Taking for example, an IP address will probably now be considered private information, having the companies to follow the new regulation and act by it when working with IP addresses (Beckett, 2017).

2.4.1. GDPR – Subject-matter, objectives and contents

The General Data Protection Regulation - GDPR - was introduced into the Community legal system through Regulation (EU) 2016/679, of the European Parliament and of the Council, of 27 April 2016, and entered into force on 25 May 2018, establishing the regime for the protection of natural persons with regard to the processing of personal data and the free movement of such data and on repealing Directive 95/46 / EC (General Data Protection Regulation) (Balinha, Marques, Lourenço, Fonseca, Martins, & Dinis, 2018; European Commission, 2020; MyDataPrivacy, 2020c). Law No. 58/2019, of August 8, 2019, ensures its enforcement under the Portuguese legal system.

The European Commission (2020) considers that personal data is "any information that relates to an identified or identifiable living individual. Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data.". Additionally, the European Commission (2020) states that the GDPR "protects personal data regardless of the technology used to process that data - it is technologically neutral and applies to both automated and manual processing, provided that the data is organized according to pre-defined criteria (for example, in alphabetical order). It is also irrelevant how data is stored - in a computer system, through video surveillance, or on paper; in all these cases, personal data are subject to the protection requirements set out in the GDPR.".

The GDPR does not apply to data relating to legal persons or to deceased persons, with the exception of sensitive data and neither it is not just about protecting sensitive information against hackers and leakers (GDPR.EU, 2020, MyDataPrivacy, 2020c).

The GDPR also affects business practices. Petters (2020) states that "there are many aspects that companies have to undertake to achieve and maintain compliance with the GDPR. These include but are not limited to: Explicit opt-in consent from users; The right to request data from companies; The right to have your data deleted".

The GDPR is organized into 11 chapters: Chapter 1 (Art 1 – 4): General provisions; Chapter 2 (Art 5 – 11): Principles; Chapter 3 (Art 12 – 23): Rights of the data subject; Chapter 4 (Art 24 – 43): Controller and processor; Chapter 5 (Art 44 – 50): Transfers of personal data to third countries or international organizations; Chapter 6 (Art 51 – 59): Independent supervisory authorities; Chapter 7 (Art 60 – 76): Cooperation and consistency; Chapter 8 (Art 77 – 84): Remedies, liability and penalties; Chapter 9 (Art 85 – 91): Provisions relating to specific processing situations; Chapter 10 (Art 92 – 93): Delegated acts and implementing acts; and Chapter 11 (Art 94 – 99): Final provisions.

2.4.2. GDPR and Personal Data Privacy

One of the most relevant articles for this study is article 4, which defines what personal data is: "'personal data' means any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person" (European Commission, 2020).

The European Commission (2020) presents some examples of personal data: the name and surname; the address of a residence; an e-mail address like name.surname@firm.com; the number of an identification card; location data (for example, the location data function on a mobile phone); an IP address (internet protocol); connection testimonies (cookies); your phone's advertising identifier; the data held by a hospital or doctor, which makes it possible to identify a person unambiguously.

Some Portuguese consultancy companies have sought to clarify what personal data is. MyDataPrivacy (2020a) organizes personal data into categories:

- *Internal*: Knowledge and Beliefs; Authentication; Preferences;
- *External:* Identification; Ethnicity; Sexual; Behavior; Demography; Medical and Health; Physical characteristics;
- *Historical:* History of life;
- *Financial:* Account; Property; Transactions; Credit;
- *Social:* Professional; Public Life; Family; Social networks; Communication;
- *Tracking:* Computer; Contact; Location.

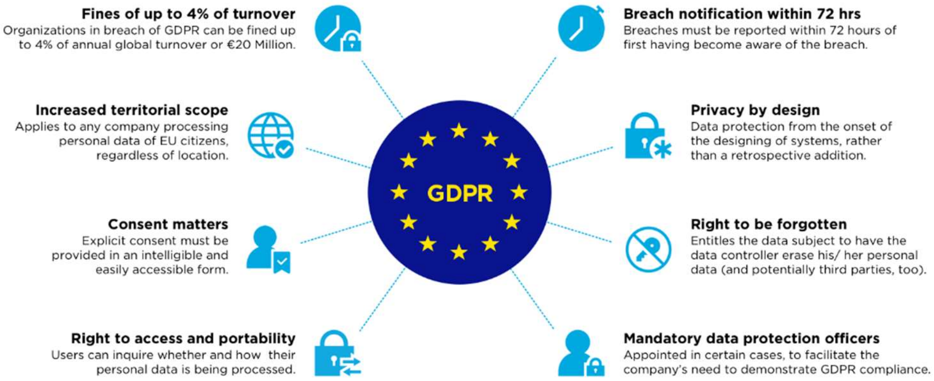
2.4.3 GDPR's Measures

Regarding the way the control of the private information is given to the user, the GDPR implements the following measures (Parliament & Council, 2016) (Figure 5):

- *Increased Territorial Scope*: As mentioned before, on the sub-chapter 1.2, this regulation will extend to all companies, inside the EEA, that process the private data of subjects, regardless of the company's location (chapter 1, article 3);
- *Penalties*: Organizations in breach of GDPR can be fined up to 4% of annual global turnover or €20 Million (whichever is greater) (article 83);
- *Consent*: There was a change as well regarding the terms and conditions of consent, that previously were long and illegible, now they must be clear and easily accessible, using clear and plain language. Also, regarding the empowerment of the consumer, the mentioned consent is also easier to withdraw.

Regarding specific rules of empowerment to the consumer, there was also given to them the following rights:

- *Breach Notification:* Every company that suffers from a breach of data that might contain private information of the users, must notify the mentioned users within 72 hours of becoming aware of that breach;
- *Right to Access:* In order to reinforce the idea of transparency and empowerment of data subjects, they have the right to ask for any confirmation from the data controller as to whether private data concerning them is being processed, and for which purpose. Further, they also the right to be provided with a free of charge digital copy of the mentioned information;
- *Data Erasure:* It must be provided to the data subjects the right to ask for the removal of all private data. The conditions for erasure determined (article 17).



Fines of up to 4% of turnover
Organizations in breach of GDPR can be fined up to 4% of annual global turnover or €20 Million.

Increased territorial scope
Applies to any company processing personal data of EU citizens, regardless of location.

Consent matters
Explicit consent must be provided in an intelligible and easily accessible form.

Right to access and portability
Users can inquire whether and how their personal data is being processed.

Breach notification within 72 hrs
Breaches must be reported within 72 hours of first having become aware of the breach.

Privacy by design
Data protection from the onset of the designing of systems, rather than a retrospective addition.

Right to be forgotten
Entitles the data subject to have the data controller erase his/ her personal data (and potentially third parties, too).

Mandatory data protection officers
Appointed in certain cases, to facilitate the company's need to demonstrate GDPR compliance.

Figure 5- GDPR (Emotiv, n.d.)

Chapter 3 – Methodology

The methodological options to consider in a research project aim to answer questions like ‘What?’ (theme to be investigated), 'What for?' (definition of objectives), 'Why?' (justification), 'How?' (methods, procedures and techniques to be used) and 'Where?' (population or sample), equating as well as how the 'Problem' should be formulated and what are 'Hypotheses' to be tested, and which serve as the basis for the definition of an action plan for the investigation (Pardal & Correia, 1995).

In this sense, the method to be pursued, understood as the sequential set of steps and procedures with having in mind the production of new scientific knowledge or the systematization of the knowledge one has on a given theme or problem, is a necessary requirement for the scientific validity of a research project (Quivy and Campenhoudt, 2018).

Firstly, the need to be able to contribute to the increase of scientific knowledge about how the entry into force of the GDPR would have contributed to the change in behavior regarding privacy of personal data by users of social networks suggested that the type of research to be carried out would be basic or fundamental.

However, and for reasons inherent to the professional activity of the master's student (Informatica PowerCenter Consultant), a reorientation of the research was carried out, more of an applied nature, trying to understand if there would be a difference in behavior between IT specialists - those who had training and / or perform a professional activity in this field - and non-specialists, that is, other citizens, meaning, those who have not had any specialized training in IT or who do not perform a professional activity in that area.

The conclusions to be obtained could, through their disclosure, reinforce the sensitivity of those who, for professional reasons, have access to and use the personal data of third parties.

Having in mind the defined objectives, which will be later presented, having the initial purpose of conducting an exploratory research, aiming to create a picture about the behavior, in terms of personal data privacy, of social network’s users, creating a questionnaire, and based on the option of an applied research, we opted for a descriptive, transversal and deductive research, since it was oriented based on previously defined hypotheses, based on previously performed studies (partial replication), given that it is this type of research that allows us, through systematic observations about the population or a sample, to describe the characteristics of a phenomenon or problem, by establishing relationships between variables, based on a sequential process of collecting, tabulating and analyzing data, obtained through the application, namely, of questionnaires (Pardal & Lopes, 2011).

This reorientation had implications for the design of the research project, particularly in methodological options, with the adoption of a quantitative methodology - in line with the positivist paradigm (Denscombe, 2010; Punch, 2000, 2013; Shah & Corley, 2006) which, looking to establish relationships between variables, allows to acquire knowledge through the observation and measurement of the properties of the objects of interest -, with the nature of the research, the choice of the data collection instrument and with the treatment and analysis of that data.

Following this, the presentation of the research design, research objectives and hypotheses framework that serve as the basis for this research project, will be presented.

3.1. Research Design

The research design includes the path to be followed with regard to the definition of the population, or possible sample, the construction of the data collection instrument and its subsequent application, the measures to be considered, as well as the treatment, the techniques to be applied in the analysis and interpretation of data, in a sequential process of steps, observing a previously established scheduling.

The literature review, carried out in chapter 2, will allow to, in a subsequent time, outline the objectives of the research question, as well as hypotheses table that will guide the way of interpreting the data and finding the conclusions.

3.1.1. Population and Sample

Considering the problematic displayed in the Introduction, that leads to the questioning of how the behaviors regarding data privacy on social networks have changed, or not, due to the entry into force of the GDPR, as well as to understand if there are significant differences in this change in behaviors between specialists and non-specialists in IT, it was considered as population those who, using social networks, are protected by

Portuguese law. Nonetheless, given the multiplicity of social networks, as well as the method chosen for the application of the questionnaire - a topic that to be discussed afterwards - it was decided to restrict the population to those who use the social networks Facebook and LinkedIn.

Given the impossibility of surveying the entire population that uses these social networks, a non-probabilistic sample was considered, for convenience (convenience sampling), recognizing that, because of this fact, it is not possible to establish generalizable conclusions for this entire population.

3.1.2 Data collection instrument

The data selection instrument chosen was the questionnaire that, due to its characteristics, allows greater freedom of response, because it guarantees the anonymity of respondents and the confidentiality of the responses, and the non-interference by the researcher (Quivy and Campenhoudt, 2018).

The questionnaire (Annex A) is organized in 4 parts:

- *The first part,* with 6 characterization questions (Q1. Age; Q2. Gender; Q3 Academic qualifications; Q4. Information on training and / or professional activities in the IT area; Q5. Frequency of access to social networks; Q6. Dimension of connections on social networks);
- *The second part*, that regards privacy issues on social networks, that includes 6 questions (Q7. Degree of concern with privacy on social networks; Q8. Degree of concern regarding access to the profile of unwanted users; Q9. Degree of concern regarding access to the profile by future users; Q10. To whom is granted permission to view the profile; Q11. The use strategies of data protection; Q12. The information available on social networks);
- *The third part*, with only one question (Q13), regarding the degree of information on the GDPR;
- *The fourth and final part*, regarding the behaviors in terms of data privacy terms, owing to the implementation of the GDPR, having 2 questions (Q14. Change in behavior while using social networks, after the implementation of the GDPR, mainly on the data availability; Q15. Perception of changes in the way social

networks operate, by the management entities, having in mind the compliance with the established on the GDPR).

The second part [except the question Q9., that was based on the study by Tufekci (2008), and the questions Q5 and Q6 of the first part derive from, or can be considered a partial replication of the studies carried out by Young and Quan - Haase (2009, 2013); the question Q4 is introduced to segment the respondents, having in mind the objectives and the present research question investigated; the third and fourth parts were also elaborated according to the research purpose.

The typology of the questions chosen was closed questions, namely (Pardal & Correia, 1995):

- *Dichotomous*: from 'Gender' (Male / Female), Q2.; and 'Yes' / 'No', Q4 and Q12;
- *Closed answer:* where or respondent select one of the several options available: Q1., Q1., Q3., Q5., Q6., E Q10.;
- *Evaluation or application:* which allows to produce a judgment with varying degrees of intensity: Q7., Q8., Q9., Q11., Q13., Q14. and Q15.

As partly a partial replication of the studies by Young & Quan-Haase (2009, 2013), the questionnaire was not validated. For an innovative part of the empirical study, that has to do with the changing of behaviors regarding the privacy terms when using social networks sites, and understanding that a non-representative sample does not allow generalization of studies for the population, this same validation is considered unnecessary.

Before applying the questionnaire, a pre-test was carried out with 4 social network users aged between 23 and 57 years old, with different academic and professional profiles, meaning heterogeneous profiles, where problems were not revealed, in terms of intelligibility, in terms of answer.

3.1.3. Measures included in the data collection instrument

The measures contemplated in the following questionnaire, with the necessary adaptations, are the measures used in the studies of Young and Quan-Haase of (2009, 2013), which are shown in the Table 1 as well as in the study of Tufekci (2008).

Table 1- Measures of Information Revelation and Internet Privacy Concerns on SNSs

Young & Quan-Haase (2013)	Young & Quan-Haase (2009)
Frequency of Facebook use	Information Revelation
Privacy settings	Frequency of Facebook Use
Information visibility	Personal Network Size
Privacy protection strategies	Concern for Internet Privacy
Friending practices	Concern about Unwanted Audiences
	Profile Visibility

Considering the research objectives that guided this empirical research, in the construction of the questionnaire the following measures were considered, adapting the measures considered in the studies of Young & Quan-Haase (2009, 2013):

- *Frequency of Social Network Use (Q5.):* Respondents reported their frequency of use on an seven-point scale (1='several times a day'; 2='once a day'; 3='several times a week'; 4='once a week'; 5='several times a month'; 6='once a month'; 7='a couple of times a year');
- *Personal Network Size (Q6.):* Respondents had different levels of response options - The Personal Network Size were coded as 1='Less than 250', 2='Between 250 and 500', 3='Between 500 and 1000', 4='Between 1000 and 3000'and 5='More than 3000';
- *Concern for Social Network Privacy (Q7.):* Respondents were asked to indicate their level of concern - 1=’never thought about it’; 2= ‘not concerned at all’; 3= ‘not too concerned’; 4= ‘somewhat concerned’;’5=‘very concerned’;
- *Concern about unwanted audiences items (Q8.):* Respondents were asked to indicate their level of concern about access by unwanted audiences on various items - 1=’never thought about it’; 2=‘not concerned at all’; 3=‘not too concerned’; 4=‘somewhat concerned’; 5=‘very concerned’;
- *Concern about future audiences items (Q9.):* Respondents were asked to indicate their level of concern about access by future audiences on various items - 1=’never thought about it’; 2=‘not concerned at all’; 3=‘not too concerned’; 4=‘somewhat concerned’; 5=‘very concerned’;
- *Profile Visibility (Q10.):* The profile visibility levels were coded as 1='visible to only my friends', 2 ='visible to some of my networks and all of my friends', 3='visible to all of my networks and all of my friends' and 4='visible to anyone';

- *Privacy protection strategies (Q11.):* The answers to these questions were reported on a 5-point Likert scale (1='never'; 2 'rarely'; 3='sometimes'; 4='often'; 5='always');
- *Information Revelation (Q12.):* The answers to these questions were coded as 0='Yes' and 1='No'.

Question 9 was conceived in line with the study by Tufekci (2008). Respondents were asked to indicate their level of concern about access by future audiences on various items - 1=’never thought about it’; 2= ‘not concerned at all’; 3= ‘not too concerned’; 4= ‘somewhat concerned’; 5= ‘very concerned’.

Additionally, and considering the purpose of the present investigation, the following measures were also considered:

- *GDPR information level (Q13.):* Respondents were asked what level of information about the GDPR - 1=’Never heard of it’; 2= ‘Not informed at all’; 3='Somewhat informed'; 4= ‘Moderately Informed’; 5='Very informed';
- *Changing behaviors on social networks (Q14.):* Respondents were asked how they changed their behavior, when using social networks, after the implementation of the General Data Protection Regulation, regarding the disclosure of personal data - 1=’None’; 2='Very mild'; 3=‘Moderate’; 4='Very'; 5='Totally';
- *Change in the functioning of social networks (Q15.):* Respondents were asked, in their understanding, to what extent the management entities of social networks changed the way they function, in order to comply with the provisions of the GDPR - 1='None’; 2='Very mild'; 3=‘Moderate’; 4='Very'; 5=‘Totally’.

Finally, the segmentation questions:

- *Age (Q1.):* Respondents had different levels of response options - 1=‘<18 years old’; 2=‘18 to 29 years old’;3 =‘30 to 44 years old’; 4=‘45 to 54 years old’; 5=‘55 to 65 years old’; and 6=’> 65 years old’;
- *Gender (Q2.):* ‘Gender’ was coded as 0=’Male’ and 1=’Female’;
- *Highest completed education level (Q3.):* Respondents had different levels of response options - 1= ‘Primary Education (Elementary+ Middle School)’; 2= ‘Secondary Education’; 3= ‘Bachelor Degree’; 4= ‘Master Degree’; and 5='PhD';
- *Has academic training or develops / developed some professional activity in the area of Information Technologies (Q4.):* Was coded as 0='Yes' and 1='No'.

3.1.4. Application technique of the data collection instrument

The questionnaire was created on the Google Forms online platform and published on the social networks Facebook and LinkedIn, on April 8, 2020, for a period of one month, using the snowball technique, that is, giving opportunity to all who wanted to answer and, at the same time, allow disclosure of the questionnaire.

Although the data collection technique used, the snowball technique - also called chain referral sampling - does not allow the generalization of conclusions for the population, it is particularly adjusted when it becomes difficult to access to a large number of participants and geographically dispersed (Biernacki & Waldorf, 1981; Johnson, 2014), or when dealing with sensitive issues (Browne, 2005), and has been used in different areas of research (Johnson, 2014) .

This technique allows obtaining a large number of responses, in a short time and with no significant costs, even with the limitation of obtaining data from a non-probabilistic sample, but adequate when it is not possible to obtain responses from the population of a network, or from a probabilistic sample (Browne, 2005; Johnson, 2014).

3.1.5. Data Treatment, Analysis and Interpretation

As an exploratory / descriptive research, the present empirical study is of a deductive and correlational type, insofar as its objective is to test a theory, based on the stated hypotheses, which will allow it to validate or refute hypotheses (Creswell, 2013), while seeking to deepen knowledge on this topic.

The entire procedure to be followed in terms of treatment, analysis and interpretation will be guided by the nature of the variables and the necessary statistical processes to be applied according to hypotheses formulated (Robson, 2002), relating the state of the art to the problem under study (Cardoso, Alarcão, & Celorico, 2010), in order to obtain an answer to the research question, based on the data tabulation and the analysis of the relationships between the variables.

After collecting the data, which, as mentioned above, was carried out through the *Google Forms* online platform, the database was converted into EXCEL for the SPSS software and the subsequent coding of the variables, in order for the statistical data to be further processed, which will be referred in detail when presenting and discussing the results.

3.2. Research Objectives and Question

The research design was built with the goal of achieving the respective research objectives, the implementation of which will allow an answer to the research question.

For this purpose, the following are considered as General Objectives to be achieved with the present study:

- *General Objective 1 (GO1):* To know how the level of digital literacy (IT specialists versus IT non-specialists) influences the behavior, concerns and protection strategies adopted in terms of privacy on social networks;
- *General Objective 2 (GO2):* Find out about possible changes in behavior, with regard to privacy on social networks, due to the entry into force of the GDPR;
- *General Objective (GO3):* Discuss, on a reflexive basis, whether the differences between the group of belonging, with regard to digital literacy, induce a greater or lesser concern with the way that the management entities of social networks proceed with the compliance with the GDPR.

Following closely the fundamental assumption that the present study considers the existence, with regard to the level of digital literacy, of two major categories - IT specialists and IT non-specialists -, with implication in the research design, it is assumed that, of the General Objectives, result the following Specific Objectives:

- *Specific Objective 1 (SO1):* To inquire how variables such as the size of the network of connections, the frequency of access, the concern with privacy and the level of visibility of the profile influence the information available on social networks;
- *Specific Objective 2 (SO2):* Understand how the degree of concern with access to the profile of social networks influences the information provided;
- *Specific Objective 3 (SO3):* To know the level of concern regarding access by unwanted users, as well as the protection strategies adopted;
- *Specific Objective 4 (SO4):* Understand how the level of information about the GDPR influences the behavior of users of social networks in terms of information made available;

- *Specific Objective 5 (SO5):* To know, in a perfective of the users' perception, the changes introduced by the management entities in the functioning of social networks, to achieve compliance with the GDPR.

In order to achieve the defined objectives - General and Specific - the following tasks were outlined:

- *Characterize the General Data Protection Regulation (GDPR),* namely, its scope and the nature of the data it aims to protect, a task already carried out in chapter 2;
- *Proceed with the analysis and interpretation of the results of the applied questionnaire,* seeking to distinguish behaviors, strategies and perceptions about privacy on social networks, between specialists and IT non-specialists, in the context of the validity of the GDPR, to be carried out in chapter 4.

The purpose of implementing the investigative strategy is to answer the following research question:

In what way does the level of digital literacy, on the domain of GDPR, affects the social network users ‘behavior regarding the privacy and granting of access to their personal data?

3.3. Research Hypotheses

The formulation of hypotheses constitutes the guiding axis of the empirical component of the investigation.

In this sense, and as previously mentioned, the present study considers two sets of hypotheses, in which the first set aims to replicate the studies carried out by Young & Quan-Haase (2009, 2013) and by Tufekci (2008). This first set integrates hypotheses 1 to 5.

Several studies (Govani & Pashley, 2005; Gross & Acquisti, 2005; Young & Quan-Haase, 2009, 2013; Tufekci, 2008) on the issue of privacy on social networks allow us to conclude that the availability of personal data, such as age, e-mail or cell phone contacts, among others, is directly and positively associated with the frequency of access to social networks. Thus, the following hypothesis is proposed:

- **Hypothesis 1 (H1): Frequency of Social Networks Sites use will be positively associated with information revelation on Social Networks Sites.**

Authors such as Jones and Soltren (2005) and Young Quan-Haase (2009) refer that the information provided by users is in greater volume as the number of connections that are part of social networks increases, namely with regard to tastes and preferences among others referred to in chapter 2, so it seems relevant to put the following hypothesis:

- **Hypothesis 2 (H2): Personal network size will be positively associated with information revelation on Social Networks Sites.**

It is consensual among the authors of the studies accessed (Lenhart, Madden, Macgill, & Smith, A, 2007; Tufekci, 2008; Young & Quan-Haase, 2009) that there is concern about protecting privacy on social networks. These studies refer to the very frequent refusal to provide personal information when required on the internet, and also on social networks, fearing the use that it may have, namely for commercial purposes. Because the issue of privacy is associated with the amount and diversity of information made available, the following hypothesis is stated:

- **Hypothesis 3 (H3): Concern for Internet privacy will be negatively associated with information revelation on Social Networks Sites.**

Studies by Acquisti and Gross (2006, apud Young & Quan-Haase, 2009), Tufekci, (2008) and Young and Quan-Haase (2009, 2013) concluded that there is an effective concern with the personal information accessed by unwanted users , either for the protection of family and friends, or for the fraudulent or criminal use that this access may allow. Places of residence, sexual orientation, political affiliations or romantic partners are data that studies reveal to be a major concern in terms of privacy protection. Accordingly, the following hypothesis is presented:

- **Hypothesis 4 (H4): Concern for unwanted audiences will be negatively associated with information revelation on Social Networks Sites.**

The concept of 'profile visibility' is related to the way in which each user of social networks seeks to avoid access to their profile by unwanted users of that same social network and, in this way, to have access to personal information (Tufekci, 2008; Young & Quan-Haase, 2009). The use of restrictive measures regarding the contents of the profile is a way to guarantee privacy, so the following hypothesis is considered:

- **Hypothesis 5 (H5): Profile visibility will be negatively associated with information revelation on Social Networks Sites.**

Presented hypotheses 1 to 5, which constitute a partial replication of the studies carried out by Tufekci (2008) and Young and Quan-Haase (2009), we will present a second set of hypotheses due to the research objectives presented and the formulated research question, defining them as a complement to previous studies, but whose relevance stems from the entry into force of the GDPR, discussing both the change of behaviors and the possible differentiation of these behaviors between specialists and non-specialists in IT.

The studies by Young and Quan-Haase (2009, 2013) establish, among others, relationships between profile visibility and revealed information and between the dimension of the network and revealed information; the study by Tufekci (2008) the relationship between the use of data protection strategies and the information revealed. There are no known studies between the use of data protection strategies and the size of the network, so it is relevant to find new lines of investigation, presenting the following hypothesis:

- **Hypothesis 6 (H6): The use of privacy protection strategies is positively associated with the dimension of the network of social networks sites.**

The GDPR defines rights for individuals in terms of data privacy and establishes duties on companies regarding the use of such personal data (Balinha et al., 2018; Bridges, 2020; GDPR.EU, 2020; MyDataPrivacy, 2020b; Petters, 2020). Due to the functional content of their professional activity and / or training, information technology specialists have special knowledge and duties, in the observance of the compliance with the rules on the privacy of personal data, particularly with the GDPR, as well as the possibility of finding new lines of research, so we present the following hypotheses:

- **Hypothesis 7 (H7): Knowledge of the General Data Protection Regulation is positively associated with the information made available on social networks sites.**
- **Hypothesis 8 (H8): The change in behavior, in terms of information revealed on social networks sites, is positively associated with knowledge of the General Data Protection Regulation.**

- **Hypothesis 9 (H9): The change in behavior, in terms of information revealed on social networks sites, is positively associated with the entry into force of the General Data Protection Regulation.**
- **Hypothesis 10 (H10): Information Technology specialists have, in relation to Information Technology non-specialists, a greater perception of the changes in the practices of social network managers, in the observance of the compliance of the General Data Protection Regulation.**

Chapter 4 – Analysis and discussion of results

In Chapter 4, we will analyze and discuss the results.

Both in the analysis and in the discussion of the results, a comparison will be made with the results obtained with those of previous studies, which will serve as a reference to the present investigation, as well as the differentiation of results between IT specialists and IT non-specialists.

The characterization of the participants in terms of age, gender and educational qualifications will be the first thing.

Afterwards will be the analysis of results considering the 3 dimensions under study: (1) Information Revelation; (2) Privacy Protection Strategies; and (3) General Data Protection Regulation and behaviors.

In the end, a test of hypotheses presented in Chapter 3 will be made, following the same process.

4.1. Data collection and participants

As mentioned in Chapter 3, the questionnaire was created on the Google Forms online platform and published on the social networks Facebook and LinkedIn, on April 8, 2020, for a period of one month, using the snowball technique.

608 responses were collected, with only one respondent under the age of 18 and 18 responses over the age of 65. The remaining respondents are distributed among the remaining age groups, with 58.2% aged between 30 and 54 years (Table 2).

Table 2- Age distribution of participants

		Freq	Perc	ValPerc	CumP
Valid	< 18 years old	1	0,2	0,2	0,2
	18 to 29 years old	94	15,5	15,5	15,6
	30 to 44 years old	227	37,3	37,3	53
	45 to 54 years old	127	20,9	20,9	73,8
	55 to 65 years old	141	23,2	23,2	97
	> 65 years old	18	3	3	100
	Total	608	100	100	

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

As for gender, respondents are mostly female with 60.4% (Table 3). With regard to educational qualifications (Table 4), 72.8% have a Bachelor's degree or higher degree. In turn, when asked if they have training and / or professional experience in the area of Information Technologies, 29.9% answered yes (Table 5).

Table 3- Distribution of participants by gender

		Freq	Perc	ValPerc	CumP
Valid	Female	367	60,4	60,4	60,4
	Male	241	39,6	39,6	100
	Total	608	100	100	

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

Table 4- Distribution of participants by level of education

		Freq	Perc	ValPerc	CumP
Valid	Primary Education (Elementary+ Middle School)	10	1,6	1,6	1,6
	Secondary Education	95	15,6	15,6	17,3
	Bachelor’s degree	327	53,8	53,8	71,1
	Master’s degree	140	23	23	94,1
	PhD	36	5,9	5,9	100
	Total	608	100	100	

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

Table 5- Distribution of participants between IT Specialists and IT Non-Specialists

		Freq	Perc	ValPerc	CumP
Valid	No	426	70,1	70,1	70,1
	Yes	182	29,9	29,9	100
	Total	608	100	100	

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

Continuing with the characterization of the sample, an effort was made to find out if, for the variables previously presented, there are differences between ‘IT specialists’ and ‘IT non-specialists’. Regarding the age distribution of the sample, since the figures are not coincident, there are no significant differences (Table 6).

Table 6- Age distribution of participants between IT Specialists and IT Non-Specialists

			IT Specialists		Total
			No	Yes	
Age	< 18 years old	Count	0	1	1
		% within IT Specialists	0,00%	0,50%	0,20%
	18 to 29 years old	Count	64	30	94
		% within IT Specialists	15,00%	16,50%	15,50%
	30 to 44 years old	Count	157	70	227
		% within IT Specialists	36,90%	38,50%	37,30%
	45 to 54 years old	Count	91	36	127
		% within IT Specialists	21,40%	19,80%	20,90%
	55 to 65 years old	Count	101	40	141
		% within IT Specialists	23,70%	22,00%	23,20%
	> 65 years old	Count	13	5	18
		% within IT Specialists	3,10%	2,70%	3,00%
Total		Count	426	182	608
		% within IT Specialists	100,00%	100,00%	100,00%

Regarding the distribution by gender, there are significant differences, with the female gender prevailing in the 'IT non-specialists', 67.1%, and the male gender in the 'IT specialists', 55.5% (Table 7).

Table 7- Gender distribution of participants between IT Specialists and IT Non-Specialists

			IT Specialists		Total
			No	Yes	
Gender	Female	Count	286	81	367
		% within IT Specialists	67,10%	44,50%	60,40%
	Male	Count	140	101	241
		% within IT Specialists	32,90%	55,50%	39,60%
Total		Count	426	182	608
		% within IT Specialists	100,00%	100,00%	100,00%

As for educational qualifications, the IT specialists present a positive, but not significant difference (Table 8).

Table 8- Distribution by educational qualifications of participants between IT Specialists and Non-Specialists

			IT Specialists		Total
			No	Yes	
Educational qualifications	Primary Education (Elementary+ Middle School)	Count	4	6	10
		% within IT Specialists	0,90%	3,30%	1,60%
	Secondary Education	Count	75	20	95
		% within IT Specialists	17,60%	11,00%	15,60%
	Bachelor's degree	Count	229	98	327
		% within IT Specialists	53,80%	53,80%	53,80%
	Master's degree	Count	90	50	140
		% within IT Specialists	21,10%	27,50%	23,00%
	PhD	Count	28	8	36
		% within IT Specialists	6,60%	4,40%	5,90%
Total		Count	426	182	608
		% within IT Specialists	100,00%	100,00%	100,00%

4.2.Results

At this point, the results will be presented and analyzed. Once again, the 3 dimensions mentioned at the beginning of the chapter will be considered: (1) Information Revelation; (2) Privacy Protection Strategies; and (3) General Data Protection Regulation and behaviors.

4.2.1. Information Revelation

Following the studies of Young and Quan-Haase (2009, 2013), the following points to consider are: (1) Items of Information Revelation; (2) Frequency of Social Networks Sites Use; (3) Personal network size; (4) Concern for Internet privacy; (5) Concern for unwanted audiences; and (6) Profile visibility.

4.2.1.1. Items of Information Revelation

The following table, where the personal data provided by the respondents are presented, follows the classification typology proposed by MyDataPrivacy (2020c), which shows some proximity to the typologies followed by Young and Quan-Haase (2009), which follows the studies carried out by Govani and Pashley (2005) and Tufekci

(2008). It is recalled that the total non-replication of the referred studies is due to the specificities arising from the GDPR.

Table 9- Items of Information Revelation (IT Specialists and IT Non-specialists)

	Total (N = 608)		Non-specialists (N = 426)		Specialists (N = 182)	
	Yes	No	Yes	No	Yes	No
Religious, Philosophical and Ideological Beliefs	140 (23%)	468 (77%)	91 (21,4%)	335 (78,6%)	**49 (26,9%)**	133 (73,1%)
Access passwords, PIN, biometric data	12 (2%)	596 (98%)	**9 (2,1%)**	417 (97,9%)	3 (1,9%)	179 (98,4%)
Tastes, interests and preferences	461 (75,8%)	147 (24,2%)	319 (74,9%)	107 (25,1%)	**142 (78%)**	40 (22%)
Identification (name, photos, unique identifier, address, date of birth, etc.)	246 (40,5%)	362 (59,5%)	170 (39,9%)	256 (60,1%)	**76 (41,8%)**	106 (58,2%)
Ethnicity (race, origin, languages spoken)	199 (32,7%)	409 (63,3%)	138 (32,4%)	288 (67,6%)	**61 (33,5%)**	121 (66,5%)
Sexual (orientation and preferences)	117 (19,2%)	491 (80,8%)	76 (17,8%)	350 (82,2%)	**41 (22,5%)**	141 (77,5%)
Hobbies	347 (57,1%)	261 (42,9%)	234 (54,9%)	192 (45,1%)	**113 (62,1%)**	69 (37,9%)
Medical and health information	31 (5,1%)	577 (94,9%)	21 (4,9%)	405 (95,1%)	**10 (5,5%)**	172 (94,5%)
Physical Characteristics (height, weight, age, hair color, skin, tattoos and gender)	83 (13,7%)	525 (86,3%)	57 (13,4%)	369 (86,6%)	**26 (14,3%)**	156 (85,7%)
Life story	107 (17,6%)	501 (82,4%)	72 (16,9%)	354 (83,1%)	**35 (19,2%)**	147 (80,3%)
Banking, financial and equity information	7 (1,2%)	601 (98,8%)	4 (0,9%)	422 (99,1%)	**3 (1,6%)**	179 (98,4%)
Professional (profession, company, professional experience)	412 (67,8%)	196 (32,2%)	281 (66%)	145 (34%)	**131 (72%)**	51 (28%)
Academic (school, course, training)	480 (78,9%)	128 (21,1%)	328 (77%)	98 (23%)	**152 (83,5%)**	30 (16,5%)
Criminal (criminal activities, convictions and charges)	18 (3%)	590 (97%)	**15 (3,5%)**	411 (96,5%)	3 (1,6%)	179 (98,4%)
Public Life (reputation, religion, political and trade union affiliations)	66 (10,9%)	542 (89,1%)	46 (10,8%)	380 (89,2%)	**20 (11%)**	162 (89%)
Family (family structure, marriages and divorces)	205 (33,7%)	403 (66,3%)	**151 (35,4%)**	275 (64,6%)	54 (29,7%)	128 (70,3%)
Relationship with friends, acquaintances and associations or groups	364 (59,9%)	244 (40,1%)	255 (59,9%)	171 (40,1%)	109 (59,9%)	73 (40,1%)
Communication (e-mail and/or voice messages, blog)	135 (22,2%)	473 (77,8%)	**97 (22,8%)**	329 (77,2%)	38 (20,9%)	144 (79,1%)
IP address and MAC address	19 (3,1%)	589 (96,9%)	12 (2,8%)	414 (97,2%)	**7 (3,8%)**	175 (96,2%)

Contact (information that allows contact via email, telephone number)	121 (19,9%)	487 (80,1%)	83 (19,5%)	343 (80,5%)	**38** **(20,9%)**	144 (79,1%)
Location (GPS and country position information)	119 (19,6%)	489 (80,4%)	82 (19,2%)	344 (80,8%)	**37** **(20,3%)**	145 (79,7%)

To analyze the results on the items that constitute the revealed information, the following methodology was chosen:

- *Firstly*, the responses of all participants and the comparison, in the items in which there is correspondence, are considered, the results obtained and their comparison with the results of previous studies. A division will be made between the responses into steps according to the percentage level of 'YES' responses;
- *Second*, the differences between the answers given by IT specialists and IT non-specialists are analyzed.

In a general appraisal, and in relation to the items of information disclosed, the breakdown into four quartiles was considered:

- *With values above 75% of 'YES' answers:* 'Academic (school, course, training) '(78.9%) and 'Likes, interests and preferences' (75.8%). There seems to be some proximity in the results in these items, because in the studies of Young and Quan-Haase (2009) these also present very high values;
- *With values between 50 and 75% of 'YES' answers:* 'Professional (profession, company, professional experience)' (67.8%), 'Relationship with friends, acquaintances and associations or groups' (59.9%) and 'Hobbies' (57.1%);
- *With values between 25% and 50% of 'YES' answers:* 'Identification (name, photos, unique identifier, address, date of birth, etc.)' (40.5%), 'Family (family structure, marriages and divorces)' (33.7%) and 'Ethnicity (race, origin, languages spoken)' (32.7%). In the studies by Young and Quan-Haase (2009) the information on relationship status has a value above two thirds, higher than 40.5% regarding the answers obtained in our study; regarding information on the item 'Family'. In turn, the item 'Identification' considers a set of data treated differently in the studies by Young and Quan-Haase (2009), where the information related to the name (99.35%), the physical address (7.9%) and the birth date (92.2%) have values above 40.5% of the respondents in our study who provide this type of information;

- *With values below 25% of 'YES' answers:* ‘Religious, Philosophical and Ideological Beliefs'’ (23%), 'Communication (email and / or voice messages, blog)' (22.2%)’, ‘Contact (information that allows contact via email, phone number)' (19.9%), ‘Location (information on GPS position and country)' (19.6%); 'Sexual (orientation and preferences)' (19.2%), 'Life story' (17.6%), 'Physical characteristics (height, weight, age, hair color, skin, tattoos and gender)' (13.7%), 'Public Life (reputation, religion, political and union affiliations)' (10.9%), 'Medical and health information' (5.1%), 'IP address and MAC address' (3.1%), 'Criminal (criminal activities, convictions and charges)' (3%), 'Access passwords, PIN, biometric data' (2%) and 'Bank, financial and equity information' (1.2%). The availability of e-mail address presents values much lower than 83.1% or information about sexual orientation with more than two thirds referred to in studies by Young and Quan-Haase (2009), but the cell phone number (10.5%) presents low values, which can be justified for being questioned in isolation.

When comparing responses between experts and non-experts, except for the items 'Access passwords, PIN, biometric data', 'Communication (email and / or voice messages, blog)’, ‘Criminal (criminal activities, convictions and charges)' and 'Family (family structure, marriages and divorces) ', in all other items, IT specialists have higher values of 'YES' answers than IT non-specialists, which allows us to conclude that information revelation is more significant. Not trying to find, in the present study, the explanatory reasons for these results, it is understood, even by the nature of their training and / or professional experience, that the values of the first 2 items point towards greater privacy.

To test whether the information provided is independent or not of the fact that the participant is an IT Specialist or an IT non-specialist, the chi-square test was used, after verifying the respective assumptions, concluding that there are no statistically differences significant, as can be seen in Annex B, for any of the items.

In terms of synthesis, we can point out the following conclusions:

- Information revelation is lower in our study, which can be explained by the greater heterogeneity of our sample, given that previous studies are limited to a young audience;

- Although, for reasons arising from their training and / or professional experience, it was expected that IT specialists would present lower levels of information revelation than IT non-specialists but, for most items, the opposite happens.

4.2.1.2. Frequency of Social Networks Sites Use

When asked about the frequency on which they access social networks, 95.7% answered that they access them at least once a day, and that 89.5% accesses it 'Several times a day', in the case of very frequent users of social networks (Table 10).

Table 10- Frequency of Social Networks Sites Use

		Freq	Perc	ValPerc	CumP
Valid	Several times a day	544	89,5	89,5	89,5
	Once a day	38	6,3	6,3	95,7
	Several times a week	19	3,1	3,1	98,8
	Several times a month	4	0,7	0,7	99,5
	Sometimes during the year	3	0,5	0,5	100
	Total	608	100	100	

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

As for the frequency of access to social networks, there are no significant differences between IT specialists and IT non-specialists, according to the Table 11.

Table 11- Frequency of Social Networks Sites Use (IT Specialists and IT Non-specialists)

		IT Training		Total
		No	Yes	
Less than 250	Count	156	39	195
	% within IT Specialists	36,60%	21,40%	32,10%
Between 250 and 500	Count	116	52	168
	% within IT Specialists	27,20%	28,60%	27,60%
Between 500 and 1000	Count	89	49	138
	% within IT Specialists	20,90%	26,90%	22,70%
Between 1000 and 3000	Count	48	28	76
	% within IT Specialists	11,30%	15,40%	12,50%
More than 3000	Count	17	14	31
	% within IT Specialists	4,00%	7,70%	5,10%
Total	Count	426	182	608
	% within IT Specialists	100,00%	100,00%	100,00%

4.2.1.3. Personal network size

Is the dimension of the respondents' connections shown in the Table 12. From the interpretation of the data it is concluded that 40.3% have at least 500 connections and 17.6% above 1000.

Table 12- Personal network size

		Freq	Perc	ValPerc	CumP
Valid	Less than 250	195	32,1	32,1	32,1
	Between 250 and 500	168	27,6	27,6	59,7
	Between 500 and 1000	138	22,7	22,7	82,4
	Between 1000 and 3000	76	12,5	12,5	94,9
	More than 3000	31	5,1	5,1	100
	Total	608	100	100	

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

Finally, the dimension of the network, where there are significant differences, in which the 'IT specialists' present a higher number of connections than the 'IT non-specialists' (Table 13).

Table 13- Personal network size (IT Specialists and IT Non-specialists)

			IT Training		Total
			No	Yes	
Personal network size	Less than 250	Count	156	39	195
		% within IT Specialists	36,60%	21,40%	32,10%
	Between 250 and 500	Count	116	52	168
		% within IT Specialists	27,20%	28,60%	27,60%
	Between 500 and 1000	Count	89	49	138
		% within IT Specialists	20,90%	26,90%	22,70%
	Between 1000 and 3000	Count	48	28	76
		% within IT Specialists	11,30%	15,40%	12,50%
	More than 3000	Count	17	14	31
		% within IT Specialists	4,00%	7,70%	5,10%
Total		Count	426	182	608
		% within IT Specialists	100,00%	100,00%	100,00%

4.2.1.4. Concern for Internet privacy

The analysis of responses related to privacy concerns allows us to conclude that the level of concern is quite high (64.6%, 'Moderately concerned'; 22.5%, 'Very concerned'), results that are in line with the studies by Tufekci (2008) and Young and Quan-Haase (2009, 2013), as shown in the Table 14.

Table 14- Concern for Internet privacy (IT Specialists and IT Non-specialists) – Descriptive Statistics

		Freq	Perc	ValPerc	CumP
Valid	Never thought about it	4	0,7	0,7	0,7
	Not concerned at all	10	1,6	1,6	2,3
	Not too concerned	64	10,5	10,5	12,8
	Somewhat concerned	393	64,6	64,6	77,5
	Very concerned	137	22,5	22,5	100
	Total	608	100	100	

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

To test whether the degree of concern with privacy on the Internet is independent or not of the fact that the participant is an IT Specialist or an IT non-specialist, the chi-square test was used, after verifying the respective assumptions, having been concluded that there are no statistically significant differences, as can be seen in Annex C.

4.2.1.5. Concern for unwanted audiences

At this point, we will analyze the results on concern for unwanted audiences. Additionally, the concern of the participants in this study, about future users, will be analyzed.

In order to understand the position of the answers given by the participants of this study with regard to their degree of concern regarding access to their profile by unwanted users, in relation to the following, the descriptive statistics of the various questions related to this was carried out, having this in mind, as shown in the Table 15.

Table 15- Concern for unwanted audiences (IT Specialists and IT Non-specialists) – Items

	Total			IT Non-specialists			IT Specialists		
	N	Means	SD	N	Means	SD	N	Means	SD
8ª	608	3,46	1,2	426	3,42	1,1	182	**3,55**	1,1
8b		2,92	1,3		2,87	1,2		**3,04**	1,2
8c		3,18	1,3		3,16	1,2		**3,23**	1,2
8d		3,55	1,3		3,52	1,2		**3,61**	1,2
8e		4,51	1		4,48	0,9		**4,59**	0,9
8f		3,61	1,4		3,57	1,2		**3,7**	1,2

Considering the results of descriptive statistics, the items 'd - Employers and / or Educational Institutions are using social networks to monitor the extra-curricular activities of their employees or students', 'e - Sexual predators use social network sites to track, monitor and locate potential victims' and 'f - Political parties have begun using social networks to target young professionals and students through the use of advertisements and data mining' are considered the unwanted users with the highest values. Both the highest and the lowest results, but still showing concern, coincide with the study by Young and Qua-Haase (2009).

Still considering the descriptive statistics, the highest levels of concern for unwanted users are evidenced by the ‘IT Specialists’.

To test whether the degree of concern is independent or not of whether the participant is an IT Specialist or IT non-specialist, the chi-square test was used, after checking the respective assumptions.

Then, only the results considered statistically significant will be presented (Table 16). Do the remaining tables with the data related to Question 8 count from the Annex D.

Table 16- Concern for unwanted audiences (IT Specialists and IT Non-specialists) – Significant Item

		The current and / or future employers and / or admission professionals of educational institutions use the personal information made available on social networks to assess whether it is suitable for their companies and / or educational institutions					
		Never thought about it	Not concerned at all	Not too concerned	Somewhat concerned	Very concerned	Total
IT Specialists	No	40	73	67	158	88	426
		9,4%	17,1%	15,7%	37,1%	20,7%	100,0%
	Yes	5	32	40	67	38	182
		2,7%	17,6%	22,0%	36,8%	20,9%	100,0%
Total		45	105	107	225	126	608
		7,4%	17,3%	17,6%	37,0%	20,7%	100,0%

It was found that the concern with the fact that current and / or future employers and / or admission professionals of educational institutions use the personal information made available on social networks to assess whether it is suitable for their companies and / or educational institutions is not independent of whether or not the participants are IT Specialists or IT non-specialists (χ2= 0.45; *p*=.033), as shown in the Table 17.

Table 17- Concern for unwanted audiences (IT Specialists and IT Non-specialists) – Significant Item

		Police officers are using social networks to track underage drinking and other illegal activities					
		Never thought about it	Not concerned at all	Not too concerned	Somewhat concerned	Very concerned	Total
IT Specialists	No	64	138	72	93	59	426
		15,0%	32,4%	16,9%	21,8%	13,8%	100,0%
	Yes	15	62	27	56	22	182
		8,2%	34,1%	14,8%	30,8%	12,1%	100,0%
Total		79	200	99	149	81	608
		13,0%	32,9%	16,3%	24,5%	13,3%	100,0%

It was found that the concern that law enforcement officers use social networks to track minors consuming alcohol and other illegal activities is not independent of whether or not the participants are IT Specialists or IT non-specialists (χ2=9.41; *p*=.051).

In order to understand the position of the answers given by the participants of this study with regard to their degree of concern regarding access to their profile by future users, in relation to the following, a descriptive statistic of the various questions related to this concern was carried out (Table 18).

Table 18- Concern for future audiences (IT Specialists and IT Non-specialists) – Items

	Total			IT Non-specialists			IT Specialists		
	N	Means	SD	N	Means	SD	N	Means	SD
9a	608	3,13	1,11	426	**3,15**	1,13	182	3,10	1,07
9b		3,13	1,26		**3,14**	1,27		3,10	1,22
9c		3,25	1,21		3,23	1,23		**3,32**	1,16

The results obtained reveal a concern with access by future users, according to the results obtained by Tufekci (2008) and Young and Quan-Haase (2009). Other complementary statistical information can be viewed in Annex D.

To test whether the degree of concern for future users is independent or not of whether the participant is an IT Specialist or IT non-specialist, the chi-square test was used, after checking the respective assumptions, and it was concluded that there are no statistically significant differences, as can be seen in Annex E, for any of the items.

4.2.1.6. Profile visibility

The analysis of the responses regarding the profile visibility allows us to conclude that 73% of the profiles are 'Visible only to my friends', with a higher value for IT non-specialists than for IT specialists (Table 19), which is consistent in the face of unwanted, current or future audiences. This value is higher than that obtained in the study by Young and Quan-Haase (2013), which is 64%. the level of concern is quite high (64.6%, 'moderately concerned'; 22.5%, 'very concerned'), results in line with the studies by Tufekci (2008) and Young and Quan-Haase (2009, 2013), as shown you see in the Table 19. Other complementary statistical information can be viewed in Annex F.

Table 19- Profile visibility (IT Specialists and IT Non-specialists)

	Total (N = 608)	IT Non-specialists (N = 426)	IT Specialists (N = 182)
a - Visible to only my friends	444 (73%)	**317** **(74,4%)**	127 (69,8%)
b - Visible to some of my networks and all of my friends	75 (12,3%)	47 (11%)	**28** **(15,4%)**
c - Visible to all of my networks and all of my friends	34 (5,6%)	**26** **(6,1%)**	8 (4,4%)
d - Visible to anyone	55 (9%)	36 (8,5%)	**19** **(10,4%)**

To test whether the profile visibility is independent or not of the fact that the participant is an IT Specialist or IT non-specialist, the chi-square test was used, after verifying the respective assumptions, concluding that there are no differences statistically significant, as can be seen in Annex F items.

4.2.2. Privacy Protection Strategies

Given the concern with privacy on the Internet, and in addition to the restrictions on profile visibility, it was questioned the strategies used to protect privacy. The strategies 'b - Exclude personal information on social networks to restrict people I don't know from gaining information about myself ' (M=3.57), 'c - Send private email messages within social networks instead of posting messages to a friend's wall to restrict others from reading them message' (M=3.7) and 'f - Change my default privacy settings activated by social networks' (M=3.62) correspond to the response options with higher average values (Table 20), following the results obtained in the study by Young and Quan-Haase (2013), with values of 4.08, 4.72 and 4.33, also the highest mean values and which do not differ significantly from the 2009 study. As for the lowest mean value in our study – 'a - Provide false or inaccurate information on social networks to restrict people I don't know from gaining information about me' (M=2.02) – it was also the lowest value of the studies by Young-Quan-Haase (2009, 2013) (M=1.66).

Still considering the descriptive statistics, the highest levels of use of privacy protection strategies are evidenced by the 'IT Specialists' (Table 20). Other complementary statistical information can be viewed in Annex G.

Table 20- Privacy Protection Strategies (IT Non-specialists and Specialists)

	Total			IT Non-specialists			IT Specialists		
	N	Means	SD	N	Means	SD	N	Means	SD
a- Provide false or inaccurate information on social networks to restrict people I don't know from gaining information about me	608	2,02	1,3	426	1,96	1,3	182	**2,16**	1,4
b - Exclude personal information on social networks to restrict people I don't know from gaining information about myself		3,57	1,3		3,55	1,3		**3,62**	1,3
c - Send private email messages within social networks instead of posting messages to a friend's wall to restrict others from reading them message		3,7	1,2		3,67	1,2		**3,77**	1,2
d - Block former contacts from contacting me and accessing my social network profile		3,22	1,3		3,2	1,3		**3,27**	1,3
e - Certain contacts on my social network site only have access to my limited profile		3,22	1,3		3,2	1,4		**3,25**	1,3
f - Change my default privacy settings activated by social networks		3,62	1,3		3,53	1,3		**3,83**	1,2
g - Delete messages posted to my social network wall to restrict others from viewing/reading the message		2,95	1,3		2,9	1,3		**3,06**	1,2
h - Untag myself from images and/or videos posted by my contacts		3,16	1,2		3,11	1,3		**3,27**	1,2

Considering the results of descriptive statistics, the items 'd - Employers and / or Educational Institutions are using social networks to monitor the extra-curricular activities of their employees or students', 'e - Sexual predators use social network sites to track, monitor and locate potential victims' and 'f - Political parties have begun using social networks to target young professionals and students through the use of advertisements and data mining' are considered to be unwanted users with the highest values. Both the highest and the lowest results, but still showing concern, coincide with the study by Young and Quan-Haase (2009).

In order to test whether the data protection strategies used are independent or not from the fact that the participant is an IT Specialists or IT non-specialist (Table 21), the chi-square test was used (Table 22), after verifying the respective assumptions.

Then, only the results considered statistically significant will be presented. The remaining tables with the data related to Question 11 are from Annex G.

Table 21- Privacy Protection Strategies (IT Specialists and IT Non-specialists) – Significant Item (11f "Change my default privacy settings…")

			IT Specialists		
			No	Yes	Total
11f	Never	Count	28	11	39
		% within IT Specialists	6,6%	6,0%	6,4%
	Rarely	Count	74	19	93
		% within IT Specialists	17,4%	10,4%	15,3%
	Sometimes	Count	96	34	130
		% within IT Specialists	22,5%	18,7%	21,4%
	Often	Count	100	44	144
		% within IT Specialists	23,5%	24,2%	23,7%
	Always	Count	128	74	202
		% within IT Specialists	30,0%	40,7%	33,2%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Table 22- Privacy Protection Strategies (IT Specialists and IT Non-specialists) – Significant item (11f "Change my default privacy settings…") – Chi-Quare Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	9,296[a]	4	,054	,054		
Likelihood Ratio	9,464	4	,050	,052		
Fisher's Exact Test	9,242			,055		
Linear-by-Linear Association	7,162[b]	1	,007	,008	,004	,001
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 11,67.

b. The standardized statistic is 2,676.

After performing the chi-square test, the results indicate that the fact that the participant changes privacy settings on social networks is not independent of whether or not he is IT Specialist ($\chi^2(4)=9.296$; $p=0.054$), and the fact of being a IT specialist determines a greater propensity to use this data protection strategy.

4.2.3. General Data Protection Regulation (GDPR) and Behaviors

The second part of this study aims to make an exploratory approach on how behaviors have changed, or not, with respect to data privacy, due to the entry into force of the General Data Protection Regulation (GDPR).

The analysis and interpretation of the results follows a guideline that begins by asking respondents about their level of information about the GDPR. Next, respondents are asked how they have changed, or not, their behavior in terms of privacy of personal data when using SNSs, after the entry into force of the GDPR. Finally, it seeks to assess their perception, as users of SNSs, of the change in the way social network managers have changed the way they operate, in order to comply with the GDPR rules.

For the three questions (13, 14 and 15), an analysis will be made of the descriptive statistics (available in the annex H) and the evaluation of any differences in the answers between IT specialists and IT non-specialists, using the appropriate statistical tests.

Starting with the level of information about the GDPR, and from a global perspective, the responses obtained allow us to conclude that the respondents consider themselves sufficiently informed (78.2% consider themselves 'Moderately informed' or 'Very informed'), according to Table 23.

Table 23- Level of information about GDPR (IT Specialists and IT Non-specialists)

	Total (N = 608)	IT Non-Specialists (N = 426)	IT Specialists (N = 182)
Never heard of it	7 (1,2%)	**6** **(1,4%)**	1 (0,5%)
Not informed at all	13 (2,1%)	**11** **(2,6%)**	2 (1,1%)
Somewhat informed	113 (18,6%)	**89** **(20,9%)**	24 (13,2%)
Moderately Informed	347 (57,1%)	**247** **(58%)**	100 (54,9%)
Very informed	128 (21,1%)	73 (17,1%)	**55** **(30,2%)**
Total	608 (100%)	426 (100%)	182 (100%)

Table 24- Level of information about GDPR (IT Specialists and IT Non-specialists) – Chi-Quare Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	16,778[a]	4	,002	,002		
Likelihood Ratio	16,697	4	,002	,003		
Fisher's Exact Test	15,852			,002		
Linear-by-Linear Association	15,148[b]	1	,000	,000	,000	,000
N of Valid Cases	608					

a. 3 cells (30,0%) have expected count less than 5. The minimum expected count is 2,10.

b. The standardized statistic is 3,892.

After performing the chi-square test (Table 24), the results indicate that the fact that the participant considers himself sufficiently informed about the General Data Protection Regulation (GDPR) is not independent of whether or not he is an IT specialist (χ^2(4)=16.778; *p*=0.002), with IT specialists showing a greater degree of information than IT non-specialists.

Then, the results of question 13 are analyzed, which sought to find out about possible changes in behavior in terms of the use of social networks in view of the entry into force of the GDPR, with the implementation of a set of privacy protection rules for personal data.

From a global analysis of the results, it can be concluded that almost half of the respondents 'Very mild' or 'None' changed their behavior (Table 25), even considering the high level of information about the GDPR, as previously mentioned. Only 17% say they have changed ‘Very’ or ‘Totally’ their behavior.

Table 25- Change in behavior in SNSs (IT Specialists and IT Non-specialists)

	Total (N = 608)	IT Non-specialists (N = 426)	IT Specialists (N = 182)
None	107 (17,8%)	**82** **(19,5%)**	25 (13,8%)
Very mild	176 (29,3%)	123 (29,3%)	53 (29,3%)
Moderate	216 (35,9%)	**157** **(37,4%)**	59 (32,6%)
Very	84 (14%)	49 (11,7%)	**35** **(19,3%)**
Totally	18 (3%)	9 (2,1%)	**9** **(5%)**
Total	601 (100%)	420 (100%)	181 (100%)

Table 26- Change in behavior in SNSs (IT Specialists and IT Non-specialists) - Chi-Quare Test

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	11,829[a]	4	,019	,018		
Likelihood Ratio	11,372	4	,023	,024		
Fisher's Exact Test	11,515			,020		
Linear-by-Linear Association	7,283[b]	1	,007	,007	,004	,001
N of Valid Cases	601					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,42.

b. The standardized statistic is 2,699.

After performing the chi-square test (Table 26), the results indicate that the fact that the participant changes their behavior when using social networks, after the implementation of the General Data Protection Regulation (GDPR), regarding the availability of data personal information is not independent of whether or not we regard an IT Specialist or IT non-specialist ($\chi 2(4)=11.829$; $p=0.019$), with IT specialists being the ones who most changed their behavior on social networks after the entry into force of the GDPR.

Finally, it sought to learn about the respondents' perception of the change in the way the SNSs operate by the management entities, in order to comply with what is defined by the GDPR.

The analysis of the Table 27, considering the totality of responses, allows us to conclude that the perception is that there was no change in the way in which the SNSs operate, on the part of the management entities, in view of compliance with the GDPR, given that only 18.5% report that the change was 'Very' or 'Totally'. The remaining respondents are divided between ‘None’, ‘Very mild ’or ‘Moderate’.

Table 27- Change in the way in which SNSs operate by management entities (IT Specialists and IT Non-specialists)

	Total (N = 608)	IT Non-specialists (N = 426)	IT Specialists (N = 182)
None	34 (5,7%)	**28** **(6,7%)**	6 (3,3%)
Very Mild	213 (35,4%)	**156** **(37,1%)**	57 (31,3%)
Moderate	243 (40,4%)	158 (37,6%)	**85** **(47%)**
Very	95 (15,8%)	66 (15,7%)	**29** **(16%)**
Totally	16 (2,7%)	**12** **(2,9%)**	4 (2,2%)

In order to test whether the perception of the change in the way social networks work by management entities, with the objective of complying with the GDPR, it is independent or not whether the participant is an IT Specialist or IT non-specialist, the chi-square test was used, after verifying the respective assumptions, having concluded that there are no statistically significant differences, as can be seen in Annex H.

4.3.Discussion

At this point, the results will be discussed and hypotheses tested. Once again, the 3 dimensions mentioned at the beginning of the chapter, will be considered: (1) Information Revelation; (2) Privacy Protection Strategies; and (3) General Data Protection Regulation and behaviors.

Lastly, a synthesis regarding the validation of hypotheses that guide the present investigation will be presented and some comments will be elaborated, comparing the results of this process with previous studies.

4.3.1. Information Revelation

Following the study of Young and Quan-Haase (2009) closely, we will pass the test of hypotheses 1 to 5.

In order to test hypotheses 1 ("Frequency of Social Networks Sites use will be positively associated with information revelation on Social Networks Sites"), 2 ("Personal network size will be positively associated with information revelation on Social Networks Sites") and 3 ("Concern for Internet privacy will be negatively associated with information revelation on Social Networks Sites"), a hierarchical multiple linear regression was performed (Table 28, Annex I), after verifying the respective assumptions. In order to test this hypothesis using a linear regression model, it was necessary to calculate the average score of the variables under study. In the first step, the frequency of access to social networks was introduced as a predictor variable, in the second step the dimension of the connections and in the third step the concern with privacy.

Table 28- Result of hierarchical multiple linear regression (Test of hypotheses 1, 2 and 3)

	Total			**IT Non-specialists**			**IT Specialists**		
Independent variable	**Step 1**	**Step 2**	**Step 3**	**Step 1**	**Step 2**	**Step 3**	**Step 1**	**Step 2**	**Step 3**
Frequency of access to social networks	-.027	-.027	-.036	.002	.000	-.028	-.086	-.084	-.076
Dimension of connections		.160***	.149***		.198***	.194***		.060	.044
Concern about privacy			-.190***			-.230***			-.126
F	.448	8.215***	13.361***	.001	8.633***	14.105***	1.341	.996	1.623
R^2_a	.001	.026	.062	.000	.039	.091	.007	.011	.027
R^2*Change*		.026***	.036***		.039***	.052***		.004	.016

The results of step 3 indicate that the frequency of access to social networks is not significantly associated with the information provided when we have all the participants in this study (β=-.036; *p*>.05), or when analyzed separately the IT specialists ((β=-.076; *p*>.05) and IT non-specialists (β=-.028; *p*>.05).

As for the dimension of the connections, this is significantly and positively associated with the information made available, both when we have all study participants (β=.149; *p*<.001), and when only IT non-specialists are analyzed (β=.194); *p*<.001). For these participants, the larger their social network the more information they make available. With regard to IT specialists, the same does not happen, since the size of their social networks is not significantly associated with the information provided (β=.044; *p*>.05).

With regard to the concern with privacy, it is significantly and negatively associated with the information provided, both when having all study participants (β=-.190; *p*<.001), and when only IT non-specialists are analyzed (β=-.230; *p*<.001). For these participants, the greater their concern with privacy, the less information they make available. With regard to IT specialists, the same does not happen, since their concern with privacy is not significantly associated with the information provided (β=-.126; *p*>.05).

It is concluded that hypothesis 1 is not verified and only hypotheses 2 and 3 are partially verified.

In order to test hypothesis 4 ("Concern for unwanted audiences will be negatively associated with information revelation on Social Networks Sites"), three simple linear regressions were performed after checking the respective assumptions. In the first simple linear regression this hypothesis was tested with all participants, in the second only with IT specialists and in the third with participants IT non-specialists in information technologies (Table 29, Annex J).

Table 29- Results of simple linear regressions (Test of hypothesis 4)

	Predictor variable	Criterion Variable	F	R^2	β	*p*
Total	Concern for unwanted audiences	Information Revelation	6.603**	.009	-.104**	.010
IT Non-specialists			7.517**	.015	-.132**	.006
IT Specialists			.366	.002	-.045	.546

When the association between the concern with unwanted users and the information made available to all participants was tested, it was found that there is a significant and negative association (β=-.104; *p*=.010) and that the predictor variable is responsible for .9% of the variability of the criterion variable. The greater the concern with unwanted users on the part of the participants, the less information is made available.

With regard only to IT specialists, the concern with unwanted users does not have a significant association with the information provided (β =-.045; *p*>.050).

Concerning IT non-specialist participants, the concern with unwanted users has a significant and negative association with the information provided (β=-.132; *p*=.006) and the predictor variable is responsible for 1.5% of the variability of the criterion variable.

The greater the concern with unwanted users on the part of the participants, the less information is made available.

Hypothesis 4 is validated.

In order to test hypothesis 5 ("Profile visibility will be negatively associated with information revelation on Social Networks Sites"), three simple linear regressions were performed after verifying the respective assumptions. In the first simple linear regression this hypothesis was tested with all participants, in the second only with IT specialists and in the third with participants IT non-specialists (Table 30, Annex K).

Table 30- Results of simple linear regressions (Test of hypothesis 5)

	Predictor variable	Criterion Variable	F	R^2	β	*p*
Total	Profile Visibility	Information Revelation	33.724***	.053	.230***	< .001
IT Non-specialists			15.289***	.035	.187**	< .001
IT Specialists			19.987***	.100	.316	< .001

When we tested the association between profile visibility permission and the information made available to all participants, it was found that there is a significant and positive association (β=.230; *p*<.001) and that the predictor variable is responsible for 5.3% the variability of the criterion variable. The greater the profile visibility permission, by the participants, the more information is available.

With regard to IT specialists, only the profile visibility permission has a significant and positive association with the information provided (β=.316; *p*<.001) and the predictor variable is responsible for 10% of the variability of the criterion variable. The greater the profile visibility permission, by the participants, the more information is available.

Regarding participants who are IT non-specialists, the profile visibility permission has a significant and positive association with the information provided (β=.187; *p*<.001) and the predictor variable is responsible for 3.5% of the variable variability criterion. The higher the profile visibility permission, by the participants, the more information is available.

Hypothesis 5 is validated.

4.3.2. Privacy Protection Strategies

In order to test hypothesis 6 ("The use of privacy protection strategies is positively associated with the dimension of the network of social networks sites"), three simple linear regressions were performed after verifying the respective assumptions. In the first simple linear regression this hypothesis was tested with all participants, in the second only with IT specialists and in the third with participants IT non-specialists (Table 31).

Table 31- Results of simple linear regressions (Test of hypothesis 6)

	Predictor variable	Criterion Variable	F	R^2	β	*p*
Total	Personal Network Size	Privacy Protection Strategies	.027	.000	-.007	.868
IT Non-specialists			.193	.000	-.021	.661
IT Specialists			.015	.000	-.009	.903

When we tested the association between the personal network size and the data protection strategy (Annex L), it was found that there is no significant association even when we have all the participants (β=-.007; *p*>.050), nor for IT specialists (β =-.009; *p*>.050), nor for IT non-specialists (β=-.021; *p*>.050). Thus, Hypothesis 6 is rejected.

4.3.3. General Data Protection Regulation (GDPR) and Behaviors

As no studies were found on how the entry into force of the GDPR influenced the behavior of users of social networks in terms of privacy, which would allow us to support our investigation, this part of our study will be strictly exploratory.

We will then proceed to the test of hypotheses 7 to 10.

In order to test hypothesis 7 ("Knowledge of the General Data Protection Regulation is positively associated with the information revelation on social networks sites"), three simple linear regressions were performed after verifying the respective assumptions. In the first simple linear regression this hypothesis was tested with all participants, in the second only with IT specialists and in the third with participants IT non-specialists (Table 32, Annex M).

Table 32- Results of simple linear regressions (Test of hypothesis 7)

	Predictor variable	Criterion Variable	F	R^2	β	*p*
Total	Knowledge of the General Data Protection Regulation	Information Revelation	.198	.000	.018	.657
IT Non-specialists			.003	.000	-.003	.953
IT Specialists			.323	.002	.042	.571

When we tested the association between knowledge of the general data protection regulation and the information made available, it was found that there is no significant association even when we have all the participants (β=.018; *p*>.050), nor for IT specialists (β=.042; *p*>.050), nor for IT non-specialists (β=-.003; *p*>.050). Hypothesis 7 is rejected.

In order to test hypothesis 8 ("The change in behavior, in terms of information revelation on social networks sites, is positively associated with knowledge of the General Data Protection Regulation"), three simple linear regressions were performed after checking the respective assumptions. In the first simple linear regression this hypothesis was tested with all participants, in the second only with IT specialists and in the third with participants IT non-specialists (Table 33, Annex N).

Table 33- Results of simple linear regressions (Test of hypothesis 8)

	Predictor variable	Criterion Variable	F	R^2	β	*p*
Total	Knowledge of the General Data Protection Regulation	Behavior Change	35.570***	.059	.243***	< .001
IT Non-specialists			27,956***	.063	.250***	< .001
IT Specialists			6.358*	.034	.185*	.013

When tested the association between knowledge of the general data protection regulation and the change in behavior with all participants, it was found that there is a significant and positive association (β=.243; *p*<.001) and that the Predictor variable accounts for 5.9% of the Criterion Variable's variability. The participants who most changed their behavior were those who had a greater knowledge of the general data protection regulation.

With regard to IT Specialists only, knowledge of the general data protection regulation has a significant and positive association with changing behaviors (β=.185; *p*=.013) and the Predictor variable is responsible for 3.4% of the variability from Criterion Variable. The participants who most changed their behavior were those who had a greater knowledge of the general data protection regulation.

Regarding IT non-specialists, knowledge of the general data protection regulation has a significant and positive association with behavioral changes (β=.250; *p*<.001) and the Predictor variable is responsible for 3.5% of the variability of the Variable Criterion. The participants who most changed their behavior were those who had a greater knowledge of the general data protection regulation.

Hypothesis 8 is accepted.

In order to test hypothesis 9 ("The change in behavior, in terms of information revelation on social networks sites, is positively associated with the entry into force of the General Data Protection Regulation"), three simple linear regressions were performed after the respective assumptions are verified. In the first simple linear regression, this hypothesis was tested with all participants, in the second only with IT specialists and in the third with participants IT non-specialists (Table 34, Annex O).

Table 34- Results of simple linear regressions (Test of hypothesis 9)

	Predictor variable	Criterion Variable	F	R^2	β	*p*
Total	Entry into force General Data Protection Regulation	Behavior Change	49.272***	.076	.276***	< .001
IT Non-specialists			34.965***	.077	.278***	< .001
IT Specialists			12.850*	.067	.259***	< .001

When we tested the association between the entry into force of the general data protection regulation and the change in behavior with all participants, it was found that there is a significant and positive association (β=.276; *p*<.001) and that the predictor variable is responsible for 7.6% of the variability of the criterion variable. Participants changed their behaviors after the entry into force of the general data protection regulation.

With regard to IT specialists only, the entry into force of the knowledge of the general data protection regulation has a significant and positive association with behavior change (β=.259; *p*<.001) and the predictor variable it accounts for 3.4% of the variability of the criterion variable. Participants changed their behavior after the entry into force of the general data protection regulation.

With respect to participants IT non-specialist, the entry into force of the knowledge of the general data protection regulation has a significant and positive association with the change in behavior (β=.278; *p*<.001) and the predictor variable is responsible for 3.5% of the variability of the criterion variable. Participants changed their behavior after the entry into force of the general data protection regulation.

Hypothesis 9 is accepted.

To test hypothesis 10 ("Information Technology specialists have, in relation to Information Technology non-specialists, a greater perception of the changes in the practices of social network managers, in the observance of the compliance of the General Data Protection Regulation"), the *t-student* test was used for independent samples, after verifying the assumptions of normality and homogeneity of variances. As the assumption of homogeneity of variances was not verified, t-student test was used for independent samples with Welch correction was used (Table 35, Annex P).

Table 35- Results of t-student test (Test of hypothesis 10)

	t	gl	p	Means IT Non-specialists	Means IT Specialists
Changing the practices of social network managers	-1.511	377.056		2.71	2.82

The results indicate that there are no statistically significant differences (*t*(377,056)=-1,511; *p*>.050) between IT specialists (M=2.82; SD=.909) and IT non-specialists (M=2.71; SD=.818), in their perception of the change in the practices of social network managers, with a view to complying with the General Data Protection Regulation.

Hypothesis 10 is rejected.

4.3.4. Synthesis and final comments

In summary, the Table 35 presents the result of the validation process of the hypotheses that guided this investigation.

Table 36- Table of Hypotheses

Hypotheses		Results
H1	Frequency of Social Networks Sites use will be positively associated with information revelation on Social Networks Sites	Rejected
H2	Personal network size will be positively associated with information revelation on Social Networks Sites	Partially accepted
H3	Concern for Internet privacy will be negatively associated with information revelation on Social Networks Sites	Partially accepted
H4	Concern for unwanted audiences will be negatively associated with information revelation on Social Networks Sites	Accepted
H5	Profile visibility will be negatively associated with information revelation on Social Networks Sites	Accepted
H6	The use of privacy protection strategies is positively associated with the dimension of the network of social networks sites	Rejected
H7	Knowledge of the General Data Protection Regulation is positively associated with the information revelation on social networks sites	Rejected
H8	The change in behavior, in terms of information revelation on social networks sites, is positively associated with knowledge of the General Data Protection Regulation	Accepted
H9	The change in behavior, in terms of information revelation on social networks sites, is positively associated with the entry into force of the General Data Protection Regulation	Accepted
H10	IT Specialists have, in relation to IT non-specialists, a greater perception of the changes in the practices of social network managers, in the observance of the compliance of the General Data Protection Regulation	Rejected

Chapter 5 – Conclusions and recommendations

Based on the contextualization of the problem and on the performed tasks, in this chapter we will try to demonstrate to what extent the produced study allowed the realization of general and specific objectives of the investigation, make some considerations arising from the process of validating the formulated hypotheses and give the answer to the research question.

Then, recognizing the effective modest contribution of the present study, considering the limitations that conditioned its realization, an effort will be made to enunciate a set of proposals for future research.

In terms of context, we can say that the Knowledge Society in which we live is characterized, not only by the enormous capacity to produce data, with the affirmation of the concept of 'Big data' (Fan & Zao, 2015), but above all by the ability to collect, treat and use them, without often understanding the implications of this phenomenon (Erevelles et al., 2014; Gandomi & Haider, 2015; Shapiro & Hughes, 1996; Tankard, 2016), nor the limits the lawfulness of its use (Gilster, 1997; Lankshear & Knobel, 2008; Leist, 2013).

The concern for data privacy on the Internet (digital privacy) has justified studies and reflections on how it is possible to preserve the individual privacy space (Belyh, 2015; Boyd & Hargittai, 2010; Diamantopoulou, Androutsopoulou, Gritzalis, & Charalabidis, 2020; He et al., 2018), particularly at a time when the great variety of SNSs and the broad adherence, particularly among the younger population (Knobel & Lankshear, 2008), where the level of information revelation is very high (Acquisti & Gross, 2006; Govani & Pashley, 2005; Tufekci, 2008; Young & Quan-Haase, 2009, 2013).

The issue of data privacy on social networks sites (Acquisti & Gross, 2006; Govani & Pashley; 2005; Gross & Acquisti, 2005) aroused interest in the privacy protection strategies adopted (Govani & Pashley, 2005; Gross & Acquisti, 2005; Tufekci, 2008; Young & Quan-Haase, 2009, 2013), especially when it comes to the definition and use of personal data (European Commission, 2020; GDPR.EU, 2020; MyDataPrivacy, 2020b), introducing the issue of regulation (Beckett, 2017; Belyh, 2015; Peras et al., 2018; Petters, 2020).

Alongside these scientific studies, institutions have sought to create a legal framework for the protection of personal data (Balinha et al., 2018; European Commission, 2020;

MyDataPrivacy, 2020c; Parliament & Council, 2016; Peras et al., 2018), in which the GDPR assumes itself as the community regulatory norm.

The entry into force of the GDPR and the level of information by the different players (SNSs users, SNSs management entities and potential users of personal data, namely by companies and for commercial and / or other purposes) (Fayyad & Piatetsky-Shapiro , 1996; Peras et al., 2018; Petters, 2020), as well as how the GDPR may imply changes in terms of behavior, with regard to the provision of personal data by SNSs users, and in the way that SNSs function, mainly on the part of the managing entities, constitutes a field of interest in the field of research.

A significant part of these issues was addressed in the investigation and in the main conclusions that will be presented below.

5.1. Main Conclusions

The issue of personal data privacy in the SNSs, according to previous studies, focuses, firstly, on the type of information disposable, secondly, on the way in which some variables (frequency of SNSs use, personal network size and profile visibility) influence information revelation and, finally, the way in which the concern with privacy in the SNSs, namely with unwanted audiences (current or future), in addition to conditioning information revelation, determines the adoption of different options in terms of privacy protection strategies.

The first conclusion to be drawned is that, when considering the answers given on the quantity and quality of disposable information, the results show, globally, a lesser amount of information made available, particularly on the personal data considered more sensitive, when compared with the results from previous studies (Govani & Pashley 2005; Gross & Acquisti, 2005; Tufekci, 2008; Young & Quan-Haase, 2009, 2013). Although, for reasons arising from their formation and / or professional experience, it is expected that specialists would present lower levels of information revelation than non-specialists but, for most items, the opposite happens.

In the achievement of Specific Objective 1 (“To inquire how variables such as the size of the network of connections, the frequency of access and the level of visibility of the profile influence the information available on social networks”) the hypotheses H1 (“Frequency of Social Networks Sites use will be positively associated with information revelation on Social Networks Sites"), H2 ("Personal network size will be positively associated with information revelation on Social Networks Sites") and H5 ("Profile visibility will be negatively associated with information revelation on Social Networks Sites"), were formulated.

The analysis and discussion of the results led to the rejection of H1, the partial acceptance of H2 and the acceptance of H5.

In this sense, and similarly to the conclusions of the study by Young and Quan-Haase (2009), the frequency of access has no influence on information revelation (H1), with no significant differences between IT specialists and IT non-specialists.

In turn, and with regard to personal network size and information revelation (H2), it is confirmed, as in the study by Young and Quan-Haase (2009) that the larger the personal network size, the more information revelation, in this test hypotheses, it is concluded that IT specialists do not follow the behavior of IT non-specialists, not increasing information revelation with the increase of the personal network size, justifying the partial acceptance of the hypothesis.

Finally, it was sought to find out whether profile visibility influences information revelation. The H5 test allowed us to conclude that the greater the profile visibility permission, by the participants, the more information is available, similar to the conclusions of the study by Young and Quan-Haase (2009), not following the conclusions in the opposite direction by Tufekci (2008), where the reduction of profile visibility, as a way to avoid unwanted audiences, led to greater information revelation. In the relationship between these 2 variables, there are no differences between IT specialists and IT non-specialists.

In order to achieve the Specific Objective 2 ("Understand how the degree of concern with access to the profile of social networks influences the information provided and the privacy protection strategies adopted") three hypotheses were formulated: H3 "Concern for Internet privacy will be negatively associated with information revelation on Social Networks Sites"; H4 "Concern for unwanted audiences will be negatively associated with information revelation on Social Networks Sites"; and H6 "The use of privacy protection strategies is positively associated with the dimension of the network of social networks sites". Regarding the issue of concern with unwanted audiences, two questions were created, distinguishing access by current users from access by future users.

The analysis and discussion of the results led to partial acceptance of H3, acceptance of H4 and rejection of H6.

Contrary to the conclusions of the study by Govani and Pashley (2005), but following the conclusions of Young and Quan-Haase (2009), a greater concern for Internet privacy determines less information revelation. However, if this conclusion applies when we consider the totality of responses and the responses of IT non-specialists, the same does not happen when we consider the responses given by IT specialists, in which the greatest concern is Internet privacy does not lead to less information revelation , which justifies the partial acceptance of hypothesis H3. Anyway, globally, there is a high concern for Internet privacy.

In turn, and with regard to concern for unwanted audiences (H4), although the highest levels of concern are evidenced by IT specialists – for the items considered, the difference in responses is only statistically significant in relation to "The current and / or future employers and / or admission professionals of educational institutions… " –, this greater concern does not result in less information revelation, following the conclusions of the studies by Govani and Pashley (2005), Tufekci (2008) and Young and Quan-Haase (2009). The responses of IT specialists contradict the total results and the results of IT non-specialists in which it is possible to conclude that the greater the concern for unwanted audiences, the lower the information revelation. Regarding the concern with future users, the concern is high, as seen in previous studies, except for the conclusions drawn in the study by Tufekci (2008), where the concern with romantic partners is relevant, but there are no differences between IT specialists and IT non -specialists.

As for H6, the rejection of this hypothesis does not allow to conclude that there is an association between the use of privacy protection strategies and the dimension of the network of SNSs. Regarding the strategies adopted, the options are totally coincident with those obtained in the study by Young and Quan-Haase (2013) ('Exclude personal information on social networks to restrict people I don't know from gaining information about myself', 'Send private email messages within social networks instead of posting messages to a friend's wall to restrict others from reading them message 'and' Change my default privacy settings activated by social networks'). The only item with significant differences is 'f - Change my default privacy settings activated by social networks', in which IT specialists are more likely to use this strategy.

The research carried out to achieve Specific Objective 1 and Specific Objective 2 already allows conclusions to be drawn regarding General Objective 1 ("To know how the level of digital literacy (IT specialists versus IT non-specialists) influences the behavior, concerns and protection strategies adopted in terms of privacy on social networks"). For the purposes of digital literacy, the participants' segmentation was considered: a first group, with greater digital literacy, considering the assumption that they are IT specialists, either because they had specific training in the IT area, or because of their professional experience in this area; a second, with less digital literacy, including the remaining participants.

Seeking to have a general perspective on the differences between the responses of the participants, given their level of digital literacy (IT specialists and IT non-specialists), in terms of information revelation and the way the variables 'frequency of access', 'personal network size ', 'concern for Internet privacy', 'concern for unwanted audiences' and 'profile visibility' influence the quantity and quality of information revelation, it would be expected that the differences would be more significant, given that your greater knowledge about the IT area would allow them to be more aware of possible dangers of greater exposure of personal data in the SNSs.

In fact, and contrary to what would be expected, a greater personal network size or a greater concern for Internet privacy, on the part of IT specialists, do not result in less information revelation (for a significant number of items, the information provided is superior) , as opposed to that seen in IT non-specialists. The opposite situation occurs when we try to find out how the 'personal network size' influences information revelation: in this case, in contrast to what is seen in IT non-specialists, a larger network does not determine an increase in the information made available, an important issue given that IT specialists are , in our study, participants who have a larger personal network size. As for the 'frequency of access' and 'profile visibility', there are no significant differences between IT specialists and IT non-specialists, as far as information revelation is concerned.

As for the concern for Internet privacy, namely that related to unwanted audiences (current and future), and the way it influences information revelation, the results are even more surprising. Although IT specialists reveal a greater concern for Internet privacy than IT non-specialist, as well as unwanted audience – although it is only statistically significant in relation to "The current and / or future employers and / or admission professionals of educational institutions… ” –, this is not reflected in less information revelation, unlike what happens with IT non-specialists.

IT specialists will eventually be able to mitigate the risk of this greater exposure through a greater use of privacy protection strategies, as it has been concluded that IT specialists are more likely to use the strategy 'f - Change my default privacy settings activated by social networks sites'. However, it was not possible to conclude that there is an association between the use of privacy protection strategies and the dimension of the network of SNSs.

After presenting the conclusions of the part of the study that accompany previous studies, it is time to proceed to the elaboration of the conclusions of the second part of the study, which, as we stated earlier, is strictly exploratory, that will be refer to the way in which the GDPR came into force influenced the behavior of SNSs users in terms of privacy, referring to the differences found between IT specialists and IT non-specialists, fulfilling the General Objective 2 (“Find out about possible changes in behavior, with regard to privacy on social networks, due to the entry into force of the GDPR”) and General Objective 3 (“Discuss, on a reflexive basis, whether the differences between the group of belonging, with regard to digital literacy, induce a greater or lesser concern with the way that the management entities of social networks proceed with the compliance with the GDPR”).

Regarding General Objective 2, firstly it is important to analyze how the level of information on the GDPR was associated with information revelation, operated through H7 (Knowledge of the General Data Protection Regulation is positively associated with the information made available on social networks sites) – ' Rejected '–, proceeding with the gauging of any behavioral changes, in terms of information revelation, testing H8 (The change in behavior, in terms of information revealed on social networks sites, is positively associated with knowledge of the General Data Protection Regulation) – 'Accepted'.

All participants have a high level of information about the GDPR, although IT specialists are more informed, without the differences being statistically significant. This significant level of information did not result in a significant change in behavior in terms of information revelation.

When trying to associate the level of information with the change in behavior in terms of information revelation, it was found that, in fact, the higher the level of information, the greater the change in behavior, with an important part of the participants answering 'none' or 'very mild' in terms of changing their behavior. However, IT specialists were the group where this change was most significant, although it is not statistically significant.

Seeking to find out how user behavior changed with the entry into force of the GDPR, H9 (The change in behavior, in terms of information revelation on social networks sites, is positively associated with the entry into force of the General Data Protection Regulation) was accepted, which allows to conclude that, although not very markedly, there was a change in behavior, more significant in IT specialists.

In turn, with the purpose of achieving the General Objective 3, an attempt was made to assess the perception that the participants had about possible changes in the way the SNSs function, on the part of the managing entities. For this, the H10 (Information Technology specialists, in relation to IT non-specialists, a greater perception of changes in the practices of social network managers, in compliance with the General Data Protection Regulation) was tested and rejected. In a very expressive percentage, the participants' perception is "none", "very mild" or only "moderate" changes were made, with no difference in responses between IT specialists and IT non-specialists.

Proceeding to a more aggregate analysis of the conclusions obtained in the implementation of General Objective 2 and General Objective 3, it appears that, in general, the level of information on the GDPR is higher in IT specialists, without this having determined significant changes in behavior in terms of information revelation, although the entry into force of the GDPR has led to some changes, particularly in IT specialists. If there has been a change in the way the SNSs operate, these changes are not differently perceived by the 2 groups (IT specialists and IT non-specialists).

Having achieved the general and specific objectives, it is our purpose to answer the research question: “In what way does the level of digital literacy, on the domain of GDPR, affects the social network users' behavior regarding the privacy and granting of access to their personal data?”.

The assumption that the group of participants who have training and / or professional experience allows us to integrate them into a group with greater digital literacy, called IT specialists, thus having knowledge and greater sensitivity to the risks of greater exposure of their personal data, seeking to ensure greater protection of the privacy of such data is not sufficiently verified.

5.2.Main Contributions

Without concern for generalization (non-representative sample), the non-significant differences between IT specialists and IT non-specialists, considering the answer given to the research question, justifies some concern, given that the training and / or professional experience in the IT area would justify other results regarding the implications of the entry into force of the GDPR.

The referred absence of differences opens room to question, on a reflection basis, whether the knowledge about the GDPR is effective and extensive and whether the training acquired and / or professional experience is a sufficient guarantee for raising awareness about personal data privacy issues, which may raise ethical and deontological questions.

In this sense, the present study can act both as an alert and as a starting point for future work.

5.3.Limitations on the study

At last, but not least, because a dissertation is, as a final product, the beginning of a journey, it is important to mention that the present study has some limitations.

The first limitation stems from the fact that the second part of the study has a purely exploratory nature, despite the attempt to describe the differences between the 2 groups of the sample (IT specialists and IT non-specialists), since it was not possible to find previous studies.

The second limitation has to do with the use of the snowball methodology, which, although allowing the collection of a high number of answers, does not allow the generalization of the conclusions, given the non-probabilistic character of the sample.

The third limitation results from the non-consideration of the effect of moderating variables, such as age, gender and education level, on the causal relationship between independent variables (e.g. level of information about the GDPR) and dependent variables (e.g. changes in behavior in terms of information revelation), leading to underutilization in the exploitation of the data obtained.

Finally, it should be noted that the deepening of the dimensions of the GDPR and its degree of information to SNS users was insufficient, as well as the explanatory causes for any change in the quantity and / or quality of the information revealed by these users, with the objective of achieving greater protection of the privacy of personal data.

5.4.Future Work

The limitations to the study mentioned above constitute a first proposal for future work.

Additionally, other proposals can be mentioned:

- Expand the study to allow a generalization effect, or not, of some conclusions obtained by the present investigation;
- Know the explanatory causes for the behavior, in terms of privacy of personal data, after the entry into force of the GDPR;
- Know the perspective of social network managers on the issues of privacy of personal data in the SNSs;
- Replicate the study, considering a segmented analysis in terms of the different SNSs and the different types of users, considering variables such as age, gender, professional activity and the level of qualifications, with the purpose of defining the use profiles of these SNSs and the availability of personal data.

One last reflection, as a way of closing: the considerable capacities for the collection, processing and use of personal data, not always within the limits of legality or lawfulness, based on the information made available in SNSs, would justify a growing concern on the part of all users of these SNSs, in addition to the legal mechanisms for protecting the privacy of personal data, as is the case of the GDPR, with particular attention to those who have a greater obligation, due to the knowledge they have, the IT specialists.

Bibliography

Acquisti, A., & Gross, R. (2006, June). Imagined communities: Awareness, information sharing, and privacy on the Facebook. In *International workshop on privacy enhancing technologies* (pp. 36-58). Springer, Berlin, Heidelberg.

Balinha, H., Marques, L., Lourenço, A., Fonseca, A., Martins, J. C., & Dinis, J. (2018). O RGPD: a articulação entre a gestão de informação e a gestão de segurança da informação. In *Actas do Congresso Nacional de Bibliotecários, Arquivistas e Documentalistas* (No. 13).

Beckett, P. (2017). GDPR compliance: your tech department's next big opportunityAuthor links open overlay panelPhilBeckett. *Computer Fraud & Security*, 2017(5), 9–13.

Belyh, A. (2015), *What is Data Privacy?***.** Accessed in: April 5, 2020, in: https://www.cleverism.com/lexicon/data-privacy/.

Biernacki, P., & Waldorf, D. (1981). Snowball sampling: Problems and techniques of chain referral sampling. *Sociological methods & research*, *10*(2), 141-163.

Boyd, D., & Hargittai, E. (2010). Facebook privacy settings: Who cares? *First Monday*, *15*(8). https://doi.org/10.5210/fm.v15i8.3086

Bridges, J. (2020), *Data Privacy Guide: Definitions, Explanations and Legislation.* Accessed in: April 5, 2020, in: https://www.varonis.com/blog/data-privacy/.

Browne, K. (2005). Snowball sampling: using social networks to research non-heterosexual women. *International journal of social research methodology*, 8(1), 47-60.

Buckingham, D. (2010). Defining Digital Literacy. In *Medienbildung in neuen Kulturräumen* (pp. 59–71). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92133-4_4

Cardoso, T., Alarcão, I., & Celorico, J. A. (2010). *Revisão da literatura e sistematização do conhecimento*. Porto Editora.

Chen, C. L. P., & Zhang, C.-Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, *275*, 314–347. https://doi.org/10.1016/j.ins.2014.01.015

Chung, W. (2014). BizPro: Extracting and categorizing business intelligence factors from textual news articles. *International Journal of Information Management*, *34*, 272–284. https://doi.org/10.1016/j.ijinfomgt.2014.01.001

Creswell, J. W. (2013). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Denscombe, M. (2010). *Guidelines for Good Practice Ground Rules for Social Research*. Maidenhead, Berkshire, England: McGraw-Hill Education.

Diamantopoulou, V., Androutsopoulou, A., Gritzalis, S., & Charalabidis, Y. (2020). Preserving Digital Privacy in e-Participation Environments: Towards GDPR Compliance. *Information*, 11(2), 117. https://doi.org/10.3390/info11020117

Dijcks, J. (2012). Oracle: Big data for the enterprise(White paper). *Oracle White Paper*, (June), 1–14. Retrieved from http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Oracle+:+Big+

Data+for+the+Enterprise#0

Dillon, K. (n.d.). Big Data & the 5Vs. Accessed in: April 5, 2020, in: https://coservit.com/servicenav/en/big-data-the-5vs/?fbclid=IwAR0CripV2QN1wADWCtsRFIHOn0rNpfcFByvp__1NJWUonxbiX0rpfqqxU5w

Emotiv (n.d.). GDPR. Accessed in: April 5, 2020, in: https://www.emotiv.com/glossary/gdpr/?fbclid=IwAR1wEpkR0Ho_cna746C3i7wU8kve10VKgSj11bBXhm2da17No8SkOYvNpcw

Erevelles, S., Fukawa, N., & Swayne, L. (2014). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*.

European Commission (2020), *What is personal data?*. Accessed in: April 5, 2020, in: https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_pt#referncias.

Fayyad, U., & Piatetsky-Shapiro, G. (1996). *Advances in Knowledge Discovery and Data Mining. From Data Mining to Knowledge Discovery in Databases* (Vol. 17). AI Magazine.

Fieldhouse, M., & David, N. (2008). Digital Literacy as Information Savvy. In *Introduction: Digital Literacies—Concepts, Policies and Practices* (pp. 48–72). Peter Lang Publishing Inc.

Fox, S., Rainie, L., Horrigan, J., Lenhart, A., Spooner, T., & Carter, C. (2000). Trust and privacy online: Why Americans want to rewrite the rules. *The Pew Internet & American Life Project*, 1-29.

Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics, *35*, 137–144. Accessed in: April 5, 2020, https://ac.els-cdn.com/S0268401214001066/1-s2.0-S0268401214001066-main.pdf?_tid=75ba3407-59c6-4f62-8c93-fa198b3af0ec&acdnat=1546794853_d2ad4243e03e9a3cb630471aea21fc6

GDPR.EU (2020), *A guide to GDPR data privacy requirements*. Accessed in: April 5, 2020, in: https://gdpr.eu/data-privacy/?cn-reloaded=1.

Gilster, P. (1997). *Digital Literacy* (1st ed.). Wiley Computer Pub.

Govani, T., & Pashley, H. (2005). *Student awareness of the privacy implications when using Facebook*. Unpublished paper presented at the "Privacy poster fair" at the Carnegie Mellon university school of library and information science, 9, 1-17.

Gross, R., & Acquisti, A. (2005, November). Information revelation and privacy in online social networks. In *Proceedings of the 2005 ACM workshop on Privacy in the electronic society* (pp. 71-80).

Hazen, B. T., Boone, C. A., Ezell, J. D., & Jones-Farmer, L. A. (2014). Data quality for data science, predictive analytics, and big data in supply chain management: An introduction to the problem and suggestions for research and applications. *International Journal of Production Economics*, *154*, 72-80.

He, Z., Cai, Z., & Yu, J. (2018). Latent-Data Privacy Preserving With Customized Data Utility for Social Network Data. *IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY*, *67*(1). https://doi.org/10.1109/TVT.2017.2738018

Helsper, E. J., & Eynon, R. (2010). Digital natives: Where is the evidence? *British*

Educational Research Journal, *36*(3), 503–520. Accessed in: April 5, 2020, https://doi.org/10.1080/01411920902989227

Johnson, T. P. (2014). Snowball sampling: introduction. *Wiley StatsRef: Statistics Reference Online*.

Jones, H., & Soltren, J. H. (2005). Facebook: Threats to privacy. *Project MAC: MIT project on mathematics and computing*, *1*(01).

King, R. (2012). Ford Gets Smarter About Marketing and Design - CIO Journal. - WSJ. Retrieved January 6, 2019, from https://blogs.wsj.com/cio/2012/06/20/ford-gets-smarter-about-marketing-and-design/

Knobel, M., & Lankshear, C. (2008). Digital literacies and participation in online social networking spaces. In *Digital literacies: Concepts, policies and practices* (pp. 249–278). Peter Lang Publishing Inc.

Labrinidis, A., & Jagadish, H. V. (2012). *Challenges and opportunities with big data. Proceedings of the VLDB Endowment* (Vol. 5). https://doi.org/10.14778/2367502.2367572

Lankshear, C., & Knobel, M. (2008). *Digital Literacies: Concepts, Policies and Practices*. Peter Lang Publishing Inc.

Leist, A. K. (2013). Social media use of older adults: A mini-review. *Gerontology*, *59*(4), 378–384. https://doi.org/10.1159/000346818

Lenhart, A., Madden, M., Macgill, A. R., & Smith, A. (2007). *The use of social media gains a greater foothold in teen life as they embrace the conversational nature of interactive online media*.

Lycett, M. (2013). 'Datafication': making sense of (big) data in a complex world. *European Journal of Information Systems*, *22*.

Marôco, J. (2011). *Análise Estatística com o SPSS Statistics.: 5ª edição*. ReportNumber, Lda.

MyDataPrivacy (2020), *Categorias de Dados Pessoais*. Accessed in: April 5, 2020, in: https://mydataprivacy.eu/categorias-de-dados-pessoais/.

MyDataPrivacy (2020), *O que são dados pessoais para o RGPD?*. Accessed in: April 5, 2020, in: https://mydataprivacy.eu/o-que-sao-dados-pessoais-rgpd/.

MyDataPrivacy (2020), *O Regulamento*. Accessed in: April 5, 2020, in: https://mydataprivacy.eu/o-regulamento/.

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs*, *41*(1), 100-126. https://doi.org/10.1111/j.1745-6606.2006.00070.x

Oblinger, D., Oblinger, J. L., & Lippincott, J. K. (2005). *Digital Commons @Brockport Educating the Net Generation*. Retrieved from http://digitalcommons.brockport.edu/bookshelf%0Ahttp://digitalcommons.brockport.edu/bookshelf

Panigrahi, B. K., Abraham, A., & Das, S. (2010). Computational intelligence in powerengineering. Springer.

Paper, A. H. W. (2018). The Power of One : IBM + Hortonworks The Challenge : Driving Analytics Benefit from Today ' s Data, (May).

Pardal, L. & Correia, E. (1995). *Métodos e técnicas de investigação social*. Porto: Areal

Editores.

Pardal, L. & Lopes, E. L. (2011). *Métodos e técnicas de investigação social*. Porto: Areal Editores.

Parliament, E., & Council, E. (2016). Protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). (E. P. A. N. D. COUNCIL, Ed.). Official Journal of the European Union.

Peras, D., Mekovec, R., & Picek, R. (2018). Influence of GDPR on social networks used by omnichannel contact center. In *2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics, MIPRO 2018 - Proceedings* (pp. 1132–1137). Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.23919/MIPRO.2018.8400206

Petters (2020), *Data Privacy Guide: Definitions, Explanations and Legislation.* Accessed in: April 5, 2020, in: https://www.varonis.com/blog/data-privacy/.

Prensky, M. (2001). *Digital Natives, Digital Immigrants* (Vol. 9). MCB University Press.

Punch, K. F. (2000). *Developing effective research proposals*. London: SAGE Publications.

Punch, K. F. (2013). *Introduction to social research: Quantitative and qualitative approaches*. Sage.

Quivy, R. & Campenhoudt, L. (2018). *Manual de Investigação em Ciências Sociais*. Lisboa: Gradiva.

Robson, B. (2002). Alignment free methodology for rapid determination of differences between a test data set and known data sets. *U.S. Patent*, *6*, 434–488.

Shah, S. K., & Corley, K. G. (2006). Building Better Theory by Bridging the Quantitative– Qualitative Divide. *Journal of Management Studies*, 1821-1835.

Shapiro, B. J. J., & Hughes, S. K. (1996). *Information Literacy as a Liberal Art Enlightenment proposals for a new curriculum* (Vol. 31).

Shimko, J. (2020). A beginner's introduction to the ETL process. Accessed in: April 5, 2020, in: https://medium.com/@jennifershimko14/a-beginners-introduction-to-the-etl-process-7f5beb5c24fe

Tajel, H., & Turner, J. (2004). *The Social Identity Theory of Intergroup Behavior • 277 READING 16*. Retrieved from https://student.cc.uoc.gr/uploadFiles/B310/Tajfel & Turner 86_SIT_xs.pdf

Tankard, C. (2016). What the GDPR means for businesses. *Network Security*, *2016*(6), 5-8.

Tufekci, Z. (2008). Can you see me now? Audience and disclosure regulation in online social network sites. *Bulletin of Science, Technology & Society*, *28*(1), 20-36.

Vadakkanmarveettil, J. (2014). All You Ever Wanted to Know about Text Mining. Accessed in: April 5, 2020, in: https://www.jigsawacademy.com/all-you-ever-wanted-to-know-about-text-mining/?fbclid=IwAR3cebUaEj9ReYAxOErPVCB8EGvWbWMHALXdRq3VgXvcwZBtONfg6pohhwM#

What is value creation? definition and meaning - BusinessDictionary.com. (n.d.). Retrieved January 5, 2019, from http://www.businessdictionary.com/definition/value-creation.html

Young, A. L., & Quan-Haase, A. (2009, June). Information revelation and internet privacy concerns on social network sites: a case study of facebook. In *Proceedings of the fourth international conference on Communities and technologie* (pp. 265-274). New York, New York, USA: Association for Computing Machinery (ACM). https://doi.org/10.1145/1556460.1556499

Young, A. L., & Quan-Haase, A. (2013). Privacy protection strategies on Facebook: The Internet privacy paradox revisited. *Information, Communication & Society*, *16*(4), 479-500.

Annexes

Annex A – Questionnaire (In English) and Questionnaire (In Portuguese)

1- Age:

- □ < 18 years old
- □ 18 to 29 years old
- □ 30 to 44 years old
- □ 45 to 54 years old
- □ 55 to 65 years old
- □ > 65 years old

2- Gender:

- □ Male
- □ Female

3- What is your highest completed education level?

- □ Primary Education (Elementary+ Middle School)
- □ Secondary Education
- □ Bachelor Degree
- □ Master Degree
- □ PhD

4- Do you have academic training or develop/ developed some professional activity in the area of Information Technologies?

- □ Yes
- □ No

5- How often do you access social networks?

- □ several times a day
- □ once a day
- □ several times a week
- □ once a week
- □ several times a month
- □ once a month
- □ a couple of times a year

6- Regarding social networks, in average, what is your personal network size?

- □ Less than 250
- □ Between 250 and 500
- □ Between 500 and 1000
- □ Between 1000 and 3000
- □ More than 3000

7- What is your degree of concern about your privacy on social networks?

- □ never thought about it
- □ not concerned at all
- □ not too concerned
- □ somewhat concerned
- □ very concerned

8- What is your degree of concern about unwanted audiences accessing your profile, regarding the following items:

	never thought about it	not concerned at all	not too concerned	somewhat concerned	very concerned
a- The current and / or future employers and / or admission professionals of educational institutions use the personal information made available on social networks to assess whether it is suitable for their companies and / or educational institutions					
b- Police officers are using social networks to track underage drinking and other illegal activities					
c - Companies and / or educational institutions are monitoring social network's postings, personal information and images to identify violators of code of conduct and / or ethics (i.e., involvement in illegal activities)					
d - Employers and/or Educational Institutions are using social networks to monitor the extra-curricular activities of their employees or students					
e - Sexual predators use social network sites to track, monitor and locate potential victims					
f - Political parties have begun using social networks to target young professionals and students through the use of advertisements and data mining					

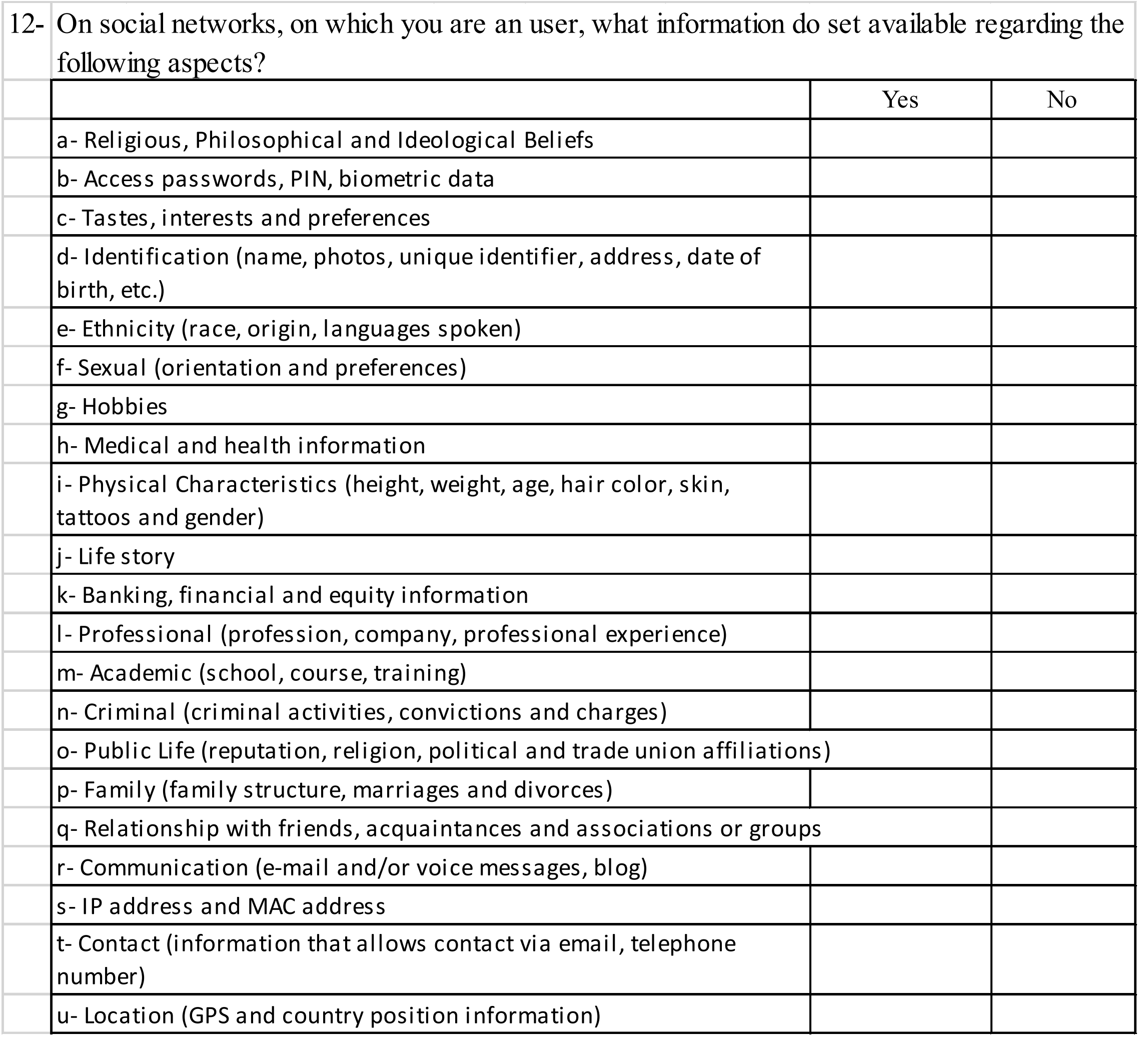
9- What is the degree of concern regarding access to your profile by future users, regarding the following categories:

	never thought about it	not concerned at all	not too concerned	somewhat concerned	very concerned
a- Employer and/or educational institutions					
b- Romantic partner					
c- Government					

10- To whom do you grant profile visibility access?

□ visible to only my friends
□ visible to some of my networks and all of my friends
□ visible to all of my networks and all of my friends
□ visible to anyone

11-	To what extent do you use the following data protection strategies?					
		never	rarely	sometimes	often	always
	a- Provide false or inaccurate information on social networks to restrict people I don't know from gaining information about me					
	b - Exclude personal information on social networks to restrict people I don't know from gaining information about myself					
	c - Send private email messages within social networks instead of posting messages to a friend's wall to restrict others from reading them message					
	d - Block former contacts from contacting me and accessing my social network profile					
	e - Certain contacts on my social network site only have access to my limited profile					
	f - Change my default privacy settings activated by social networks					
	g - Delete messages posted to my social network wall to restrict others from viewing/reading the message					
	h - Untag myself from images and/or videos posted by my contacts					



12-	On social networks, on which you are an user, what information do set available regarding the following aspects?		
		Yes	No
	a- Religious, Philosophical and Ideological Beliefs		
	b- Access passwords, PIN, biometric data		
	c- Tastes, interests and preferences		
	d- Identification (name, photos, unique identifier, address, date of birth, etc.)		
	e- Ethnicity (race, origin, languages spoken)		
	f- Sexual (orientation and preferences)		
	g- Hobbies		
	h- Medical and health information		
	i- Physical Characteristics (height, weight, age, hair color, skin, tattoos and gender)		
	j- Life story		
	k- Banking, financial and equity information		
	l- Professional (profession, company, professional experience)		
	m- Academic (school, course, training)		
	n- Criminal (criminal activities, convictions and charges)		
	o- Public Life (reputation, religion, political and trade union affiliations)		
	p- Family (family structure, marriages and divorces)		
	q- Relationship with friends, acquaintances and associations or groups		
	r- Communication (e-mail and/or voice messages, blog)		
	s- IP address and MAC address		
	t- Contact (information that allows contact via email, telephone number)		
	u- Location (GPS and country position information)		

13- To what extent do you consider yourself sufficiently informed about the General Data Protection Regulation (GDPR)?

If you answer ""Never heard of it"", in question 13, the questionnaire ends here. Thank you very much for participating!

□ Never heard of it

□ Not informed at all

□ Somewhat informed

□ Moderately Informed

□ Very informed

14- How did you change your behavior, when using social networks, after the implementation of the General Data Protection Regulation (GDPR), regarding the provision of personal data?

□ None

□ Very mild

□ Moderate

□ Very

□ Totally

15- In your opinion, to what extent have the social network management entities changed the way they operate, in order to comply with the provisions of the GDPR?

□ None

□ Very mild

□ Moderate

□ Very

□ Totally

1- Indique a sua idade:

□ <18 anos
□ 18 a 29 anos
□ 30 a 44 anos
□ 45 a 54 anos
□ 55 a 65 anos
□ >65 anos

2- Indique o seu género:

□ Masculino
□ Feminino

3- Qual o seu nível de escolaridade completo mais elevado?

□ Básico
□ Secundário
□ Licenciatura
□ Mestrado
□ Doutoramento

4- Tem formação académica ou desenvolve/desenvolveu alguma atividade profissional na área das Tecnologias de Informação (TI)?

□ Sim
□ Não

5- Qual a frequência com que acede às redes sociais?

□ Diversas vezes por dia
□ Uma vez por dia
□ Diversas vezes por semana
□ Uma vez por semana
□ Diversas vezes por mês
□ Uma vez por mês
□ Algumas vezes durante o ano

6- No que respeita às redes sociais, em média, com quantas pessoas se encontra conectado?

□ Menos de 250
□ Entre 250 e 500
□ Entre 500 e 1000
□ Entre 1000 e 3000
□ Mais de 3000

7- Qual o seu grau de preocupação com a sua privacidade nas redes sociais?

□ Nunca pensei nisso
□ Nada preocupado
□ Pouco preocupado
□ Moderadamente preocupado
□ Muito preocupado

8- Qual o seu grau de preocupação relativamente ao acesso ao seu perfil por utilizadores não desejados, em relação aos seguintes itens:

	Nunca pensei nisso	Nada preocupado	Pouco preocupado	Moderadamente preocupado	Muito preocupado
a- Os atuais e/ou futuros empregadores e/ou profissionais de admissão de instituições de ensino usarem a informação pessoal disponibilizada nas redes sociais para aferir se se é adequado para as suas empresas e/ou instituições de ensino					
b- Agentes da autoridade usarem as redes sociais para rastrear menores de idade a consumirem álcool e outras atividades ilegais					
c - As empresas e/ou instituições de ensino monitorizarem publicações nas redes sociais, informações pessoais e imagens para identificar violadores de código de conduta e/ou de ética (ou seja, envolvimento em atividades ilegais)					
d - As empresas e/ou instituições de ensino usarem as redes sociais para monitorizar as atividades extracurriculares dos seus funcionários e/ou alunos					
e - Predadores sexuais usarem as redes sociais para rastrear, monitorizar e localizarem potenciais vítimas					
f - Os partidos políticos usarem as redes sociais para se aproximarem de profissionais e/ou estudantes por meio de anúncios					

9- Qual o grau de preocupação relativamente ao acesso ao seu perfil por futuros utilizadores, em relação às seguintes categorias:

	Nunca pensei nisso	Nada preocupado	Pouco preocupado	Moderadamente preocupado	Muito preocupado
a- Empresas e/ou instituições de ensino					
b- Parceiros românticos					
c- Organismos do Estado					

10- A quem concede permissão para visualizar seu perfil?

□ Visível apenas para os meus amigos

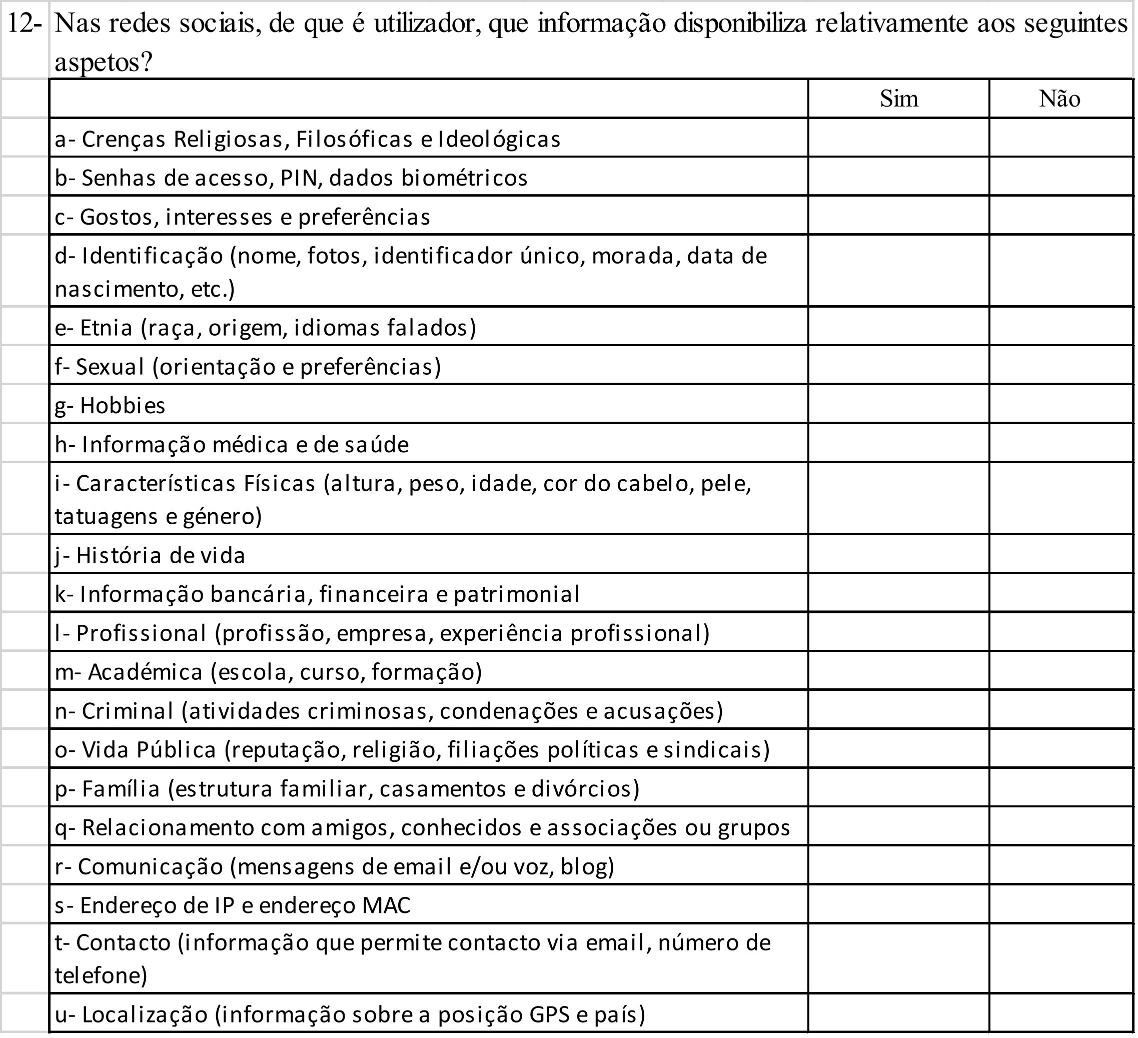
□ Visível para algumas das minhas redes e todos os meus amigos

□ Visível para todas as minhas redes e todos os meus amigos

□ Visível para todos

11- Em que medida utiliza as seguintes estratégias de proteção de dados?

	Nunca	Raramente	Algumas vezes	Frequentemente	Sempre
a- Fornecer informações falsas ou imprecisas nas redes sociais para restringir as pessoas que não conheço de obter informações sobre mim					
b - Excluir as informações pessoais nas redes sociais para restringir as pessoas que não conheço de obter informações sobre mim					
c - Enviar mensagens privadas nas redes sociais em vez de as publicar no mural de um amigo, para impedir que outras pessoas as leiam					
d - Impedir (bloqueando) antigos contatos de entrar em contato comigo e aceder ao meu perfil nas redes sociais					
e - Certos contatos nas minhas redes sociais só terem acesso limitado ao meu perfil					
f - Alterar configurações de privacidade nas redes sociais					
g - Excluir as mensagens publicadas no próprio mural nas redes sociais, para impedir que outras pessoas visualizem / leiam a mensagem					
h - Não proceder à identificação própria em imagens e / ou vídeos publicados pela rede de contatos					



12-	Nas redes sociais, de que é utilizador, que informação disponibiliza relativamente aos seguintes aspetos?		
		Sim	Não
	a- Crenças Religiosas, Filosóficas e Ideológicas		
	b- Senhas de acesso, PIN, dados biométricos		
	c- Gostos, interesses e preferências		
	d- Identificação (nome, fotos, identificador único, morada, data de nascimento, etc.)		
	e- Etnia (raça, origem, idiomas falados)		
	f- Sexual (orientação e preferências)		
	g- Hobbies		
	h- Informação médica e de saúde		
	i- Características Físicas (altura, peso, idade, cor do cabelo, pele, tatuagens e género)		
	j- História de vida		
	k- Informação bancária, financeira e patrimonial		
	l- Profissional (profissão, empresa, experiência profissional)		
	m- Académica (escola, curso, formação)		
	n- Criminal (atividades criminosas, condenações e acusações)		
	o- Vida Pública (reputação, religião, filiações políticas e sindicais)		
	p- Família (estrutura familiar, casamentos e divórcios)		
	q- Relacionamento com amigos, conhecidos e associações ou grupos		
	r- Comunicação (mensagens de email e/ou voz, blog)		
	s- Endereço de IP e endereço MAC		
	t- Contacto (informação que permite contacto via email, número de telefone)		
	u- Localização (informação sobre a posição GPS e país)		

13- Em que medida se considera suficientemente informado(a) sobre o Regulamento Geral de Proteção de Dados (RGPD)?

Caso responda ""Nunca ouvi falar"", na pergunta 13, o questionário termina aqui. Muito obrigado por participar!

□ Nunca ouvi falar

□ Nada informado

□ Pouco informado

□ Moderadamente informado

□ Muito informado

14- De que modo alterou o seu comportamento, na utilização das redes sociais, após a implementação do Regulamento Geral da Proteção de Dados (RGPD), relativamente à disponibilização de dados pessoais?

□ Nada

□ Pouco

□ Moderadamente

□ Muito

□ Totalmente

15- No seu entendimento, em que medida as entidades gestoras das redes sociais alteraram o modo de funcionamento destas, de forma a dar cumprimento ao estabelecido no RGPD?

□ Nada

□ Pouco

□ Moderadamente

□ Muito

□ Totalmente

Annex B – Information revelation (IT specialists and IT non-specialists) – Items

Descriptive Statistics

			Total				IT non-specialists				IT specialists			
			Freq	Perc	ValPerc	CumP	Freq	Perc	ValPerc	CumP	Freq	Perc	ValPerc	CumP
12a	Valid	No	468	77	77	77	335	78,6	78,6	78,6	133	73,1	73,1	73,1
		Yes	140	23	23	100	91	21,4	21,4	100	49	26,9	26,9	100
		Total	608	100	100		426	100	100		182	100	100	
12b	Valid	No	596	98	98	98	417	97,9	97,9	97,9	179	98,4	98,4	98,4
		Yes	12	2	2	100	9	2,1	2,1	100	3	1,6	1,6	100
		Total	608	100	100		426	100	100		182	100	100	
12c	Valid	No	147	24,2	24,2	24,2	107	25,1	25,1	25,1	40	22	22	22
		Yes	461	75,8	75,8	100	319	74,9	74,9	100	142	78	78	100
		Total	608	100	100		426	100	100		182	100	100	
12d	Valid	No	362	59,5	59,5	59,5	256	60,1	60,1	60,1	106	58,2	58,2	58,2
		Yes	246	40,5	40,5	100	170	39,9	39,9	100	76	41,8	41,8	100
		Total	608	100	100		426	100	100		182	100	100	
12e	Valid	No	409	67,3	67,3	67,3	288	67,6	67,6	67,6	121	66,5	66,5	66,5
		Yes	199	32,7	32,7	100	138	32,4	32,4	100	61	33,5	33,5	100
		Total	608	100	100		426	100	100		182	100	100	
12f	Valid	No	491	80,8	80,8	80,8	350	82,2	82,2	82,2	141	77,5	77,5	77,5
		Yes	117	19,2	19,2	100	76	17,8	17,8	100	41	22,5	22,5	100
		Total	608	100	100		426	100	100		182	100	100	
12g	Valid	No	261	42,9	42,9	42,9	192	45,1	45,1	45,1	69	37,9	37,9	37,9
		Yes	347	57,1	57,1	100	234	54,9	54,9	100	113	62,1	62,1	100
		Total	608	100	100		426	100	100		182	100	100	
12h	Valid	No	577	94,9	94,9	94,9	405	95,1	95,1	95,1	172	94,5	94,5	94,5
		Yes	31	5,1	5,1	100	21	4,9	4,9	100	10	5,5	5,5	100
		Total	608	100	100		426	100	100		182	100	100	
12i	Valid	No	525	86,3	86,3	86,3	369	86,6	86,6	86,6	156	85,7	85,7	85,7
		Yes	83	13,7	13,7	100	57	13,4	13,4	100	26	14,3	14,3	100
		Total	608	100	100		426	100	100		182	100	100	
12j	Valid	No	501	82,4	82,4	82,4	354	83,1	83,1	83,1	147	80,8	80,8	80,8
		Yes	107	17,6	17,6	100	72	16,9	16,9	100	35	19,2	19,2	100
		Total	608	100	100		426	100	100		182	100	100	
12k	Valid	No	601	98,8	98,8	98,8	422	99,1	99,1	99,1	179	98,4	98,4	98,4
		Yes	7	1,2	1,2	100	4	0,9	0,9	100	3	1,6	1,6	100
		Total	608	100	100		426	100	100		182	100	100	
12l	Valid	No	196	32,2	32,2	32,2	145	34	34	34	51	28	28	28
		Yes	412	67,8	67,8	100	281	66	66	100	131	72	72	100
		Total	608	100	100		426	100	100		182	100	100	
12m	Valid	No	128	21,1	21,1	21,1	98	23	23	23	30	16,5	16,5	16,5
		Yes	480	78,9	78,9	100	328	77	77	100	152	83,5	83,5	100
		Total	608	100	100		426	100	100		182	100	100	
12n	Valid	No	590	97	97	97	411	96,5	96,5	96,5	179	98,4	98,4	98,4
		Yes	18	3	3	100	15	3,5	3,5	100	3	1,6	1,6	100
		Total	608	100	100		426	100	100		182	100	100	

12o	Valid	No	542	89,1	89,1	89,1	380	89,2	89,2	89,2	162	89	89	89
		Yes	66	10,9	10,9	100	46	10,8	10,8	100	20	11	11	100
		Total	608	100	100		426	100	100		182	100	100	
12p	Valid	No	403	66,3	66,3	66,3	275	64,6	64,6	64,6	128	70,3	70,3	70,3
		Yes	205	33,7	33,7	100	151	35,4	35,4	100	54	29,7	29,7	100
		Total	608	100	100		426	100	100		182	100	100	
12q	Valid	No	244	40,1	40,1	40,1	171	40,1	40,1	40,1	73	40,1	40,1	40,1
		Yes	364	59,9	59,9	100	255	59,9	59,9	100	109	59,9	59,9	100
		Total	608	100	100		426	100	100		182	100	100	
12r	Valid	No	473	77,8	77,8	77,8	329	77,2	77,2	77,2	144	79,1	79,1	79,1
		Yes	135	22,2	22,2	100	97	22,8	22,8	100	38	20,9	20,9	100
		Total	608	100	100		426	100	100		182	100	100	
12s	Valid	No	589	96,9	96,9	96,9	414	97,2	97,2	97,2	175	96,2	96,2	96,2
		Yes	19	3,1	3,1	100	12	2,8	2,8	100	7	3,8	3,8	100
		Total	608	100	100		426	100	100		182	100	100	
12t	Valid	No	487	80,1	80,1	80,1	343	80,5	80,5	80,5	144	79,1	79,1	79,1
		Yes	121	19,9	19,9	100	83	19,5	19,5	100	38	20,9	20,9	100
		Total	608	100	100		426	100	100		182	100	100	
12u	Valid	No	489	80,4	80,4	80,4	344	80,8	80,8	80,8	145	79,7	79,7	79,7
		Yes	119	19,6	19,6	100	82	19,2	19,2	100	37	20,3	20,3	100
		Total	608	100	100		426	100	100		182	100	100	

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

*12a * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12a	No	Count	335	133	468
		% within IT Specialists	78,6%	73,1%	77,0%
	Yes	Count	91	49	140
		% within IT Specialists	21,4%	26,9%	23,0%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	2,225[a]	1	,136	,142	,084	
Continuity Correction[b]	1,923	1	,166			
Likelihood Ratio	2,181	1	,140	,142	,084	
Fisher's Exact Test				,142	,084	
Linear-by-Linear Association	2,222[c]	1	,136	,142	,084	,027
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 41,91.

b. Computed only for a 2x2 table

c. The standardized statistic is 1,491.

*12b * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12b	No	Count	417	179	596
		% within IT Specialists	97,9%	98,4%	98,0%
	Yes	Count	9	3	12
		% within IT Specialists	2,1%	1,6%	2,0%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,142[a]	1	,706	,767	,494	
Continuity Correction[b]	,003	1	,953			
Likelihood Ratio	,147	1	,701	,767	,494	
Fisher's Exact Test				1,000	,494	
Linear-by-Linear Association	,142[c]	1	,706	,767	,494	,242
N of Valid Cases	608					

a. 1 cells (25,0%) have expected count less than 5. The minimum expected count is 3,59.

b. Computed only for a 2x2 table

c. The standardized statistic is -,377.

*12c * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12c	No	Count	107	40	147
		% within IT Specialists	25,1%	22,0%	24,2%
	Yes	Count	319	142	461
		% within IT Specialists	74,9%	78,0%	75,8%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,686[a]	1	,408	,411	,235	
Continuity Correction[b]	,525	1	,469			
Likelihood Ratio	,694	1	,405	,411	,235	
Fisher's Exact Test				,469	,235	
Linear-by-Linear Association	,684[c]	1	,408	,411	,235	,059
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 44,00.

b. Computed only for a 2x2 table

c. The standardized statistic is ,827.

*12d * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12d	No	Count	256	106	362
		% within IT Specialists	60,1%	58,2%	59,5%
	Yes	Count	170	76	246
		% within IT Specialists	39,9%	41,8%	40,5%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,182[a]	1	,670	,718	,368	
Continuity Correction[b]	,113	1	,737			
Likelihood Ratio	,181	1	,670	,718	,368	
Fisher's Exact Test				,718	,368	
Linear-by-Linear Association	,181[c]	1	,670	,718	,368	,065
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 73,64.

b. Computed only for a 2x2 table

c. The standardized statistic is ,426.

*12e * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12e	No	Count	288	121	409
		% within IT Specialists	67,6%	66,5%	67,3%
	Yes	Count	138	61	199
		% within IT Specialists	32,4%	33,5%	32,7%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,073[a]	1	,787	,850	,429	
Continuity Correction[b]	,031	1	,861			
Likelihood Ratio	,073	1	,787	,850	,429	
Fisher's Exact Test				,850	,429	
Linear-by-Linear Association	,073[c]	1	,787	,850	,429	,072
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 59,57.

b. Computed only for a 2x2 table

c. The standardized statistic is ,270.

*12f * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12f	No	Count	350	141	491
		% within IT Specialists	82,2%	77,5%	80,8%
	Yes	Count	76	41	117
		% within IT Specialists	17,8%	22,5%	19,2%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1,803[a]	1	,179	,216	,110	
Continuity Correction[b]	1,514	1	,219			
Likelihood Ratio	1,763	1	,184	,216	,110	
Fisher's Exact Test				,180	,110	
Linear-by-Linear Association	1,800[c]	1	,180	,216	,110	,036
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 35,02.

b. Computed only for a 2x2 table

c. The standardized statistic is 1,342.

*12g * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12g	No	Count	192	69	261
		% within IT Specialists	45,1%	37,9%	42,9%
	Yes	Count	234	113	347
		% within IT Specialists	54,9%	62,1%	57,1%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	2,667[a]	1	,102	,108	,061	
Continuity Correction[b]	2,383	1	,123			
Likelihood Ratio	2,686	1	,101	,108	,061	
Fisher's Exact Test				,108	,061	
Linear-by-Linear Association	2,663[c]	1	,103	,108	,061	,019
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 78,13.

b. Computed only for a 2x2 table

c. The standardized statistic is 1,632.

*12h * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12h	No	Count	405	172	577
		% within IT Specialists	95,1%	94,5%	94,9%
	Yes	Count	21	10	31
		% within IT Specialists	4,9%	5,5%	5,1%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,084[a]	1	,772	,841	,455	
Continuity Correction[b]	,008	1	,929			
Likelihood Ratio	,083	1	,773	,841	,455	
Fisher's Exact Test				,841	,455	
Linear-by-Linear Association	,084[c]	1	,772	,841	,455	,150
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 9,28.

b. Computed only for a 2x2 table

c. The standardized statistic is ,290.

*12i * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12i	No	Count	369	156	525
		% within IT Specialists	86,6%	85,7%	86,3%
	Yes	Count	57	26	83
		% within IT Specialists	13,4%	14,3%	13,7%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,089[a]	1	,766	,797	,428	
Continuity Correction[b]	,029	1	,866			
Likelihood Ratio	,088	1	,767	,797	,428	
Fisher's Exact Test				,797	,428	
Linear-by-Linear Association	,089[c]	1	,766	,797	,428	,097
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 24,85.

b. Computed only for a 2x2 table

c. The standardized statistic is ,298.

*12j * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12j	No	Count	354	147	501
		% within IT Specialists	83,1%	80,8%	82,4%
	Yes	Count	72	35	107
		% within IT Specialists	16,9%	19,2%	17,6%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,477[a]	1	,490	,561	,281	
Continuity Correction[b]	,330	1	,566			
Likelihood Ratio	,471	1	,493	,561	,281	
Fisher's Exact Test				,487	,281	
Linear-by-Linear Association	,476[c]	1	,490	,561	,281	,072
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 32,03.

b. Computed only for a 2x2 table

c. The standardized statistic is ,690.

*12k * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12k	No	Count	422	179	601
		% within IT Specialists	99,1%	98,4%	98,8%
	Yes	Count	4	3	7
		% within IT Specialists	0,9%	1,6%	1,2%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,564[a]	1	,453	,680	,351	
Continuity Correction[b]	,113	1	,737			
Likelihood Ratio	,529	1	,467	,680	,351	
Fisher's Exact Test				,433	,351	
Linear-by-Linear Association	,563[c]	1	,453	,680	,351	,227
N of Valid Cases	608					

a. 2 cells (50,0%) have expected count less than 5. The minimum expected count is 2,10.

b. Computed only for a 2x2 table

c. The standardized statistic is ,750.

*12l * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12l	No	Count	145	51	196
		% within IT Specialists	34,0%	28,0%	32,2%
	Yes	Count	281	131	412
		% within IT Specialists	66,0%	72,0%	67,8%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	2,112[a]	1	,146	,156	,086	
Continuity Correction[b]	1,846	1	,174			
Likelihood Ratio	2,144	1	,143	,156	,086	
Fisher's Exact Test				,156	,086	
Linear-by-Linear Association	2,109[c]	1	,146	,156	,086	,027
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 58,67.

b. Computed only for a 2x2 table

c. The standardized statistic is 1,452.

*12m * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12m	No	Count	98	30	128
		% within IT Specialists	23,0%	16,5%	21,1%
	Yes	Count	328	152	480
		% within IT Specialists	77,0%	83,5%	78,9%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	3,263[a]	1	,071	,082	,043	
Continuity Correction[b]	2,882	1	,090			
Likelihood Ratio	3,380	1	,066	,082	,043	
Fisher's Exact Test				,082	,043	
Linear-by-Linear Association	3,257[c]	1	,071	,082	,043	,017
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 38,32.

b. Computed only for a 2x2 table

c. The standardized statistic is 1,805.

*12n * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12n	No	Count	411	179	590
		% within IT Specialists	96,5%	98,4%	97,0%
	Yes	Count	15	3	18
		% within IT Specialists	3,5%	1,6%	3,0%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1,557[a]	1	,212	,298	,162	
Continuity Correction[b]	,973	1	,324			
Likelihood Ratio	1,735	1	,188	,298	,162	
Fisher's Exact Test				,298	,162	
Linear-by-Linear Association	1,554[c]	1	,213	,298	,162	,104
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,39.

b. Computed only for a 2x2 table

c. The standardized statistic is -1,247.

*12o * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12o	No	Count	380	162	542
		% within IT Specialists	89,2%	89,0%	89,1%
	Yes	Count	46	20	66
		% within IT Specialists	10,8%	11,0%	10,9%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,005[a]	1	,945	1,000	,523	
Continuity Correction[b]	,000	1	1,000			
Likelihood Ratio	,005	1	,945	1,000	,523	
Fisher's Exact Test				1,000	,523	
Linear-by-Linear Association	,005[c]	1	,945	1,000	,523	,112
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 19,76.

b. Computed only for a 2x2 table

c. The standardized statistic is ,069.

*12p * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12p	No	Count	275	128	403
		% within IT Specialists	64,6%	70,3%	66,3%
	Yes	Count	151	54	205
		% within IT Specialists	35,4%	29,7%	33,7%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1,903[a]	1	,168	,190	,099	
Continuity Correction[b]	1,654	1	,198			
Likelihood Ratio	1,928	1	,165	,190	,099	
Fisher's Exact Test				,190	,099	
Linear-by-Linear Association	1,900[c]	1	,168	,190	,099	,029
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 61,37.

b. Computed only for a 2x2 table

c. The standardized statistic is -1,379.

*12q * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12q	No	Count	171	73	244
		% within IT Specialists	40,1%	40,1%	40,1%
	Yes	Count	255	109	364
		% within IT Specialists	59,9%	59,9%	59,9%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,000[a]	1	,994	1,000	,534	
Continuity Correction[b]	,000	1	1,000			
Likelihood Ratio	,000	1	,994	1,000	,534	
Fisher's Exact Test				1,000	,534	
Linear-by-Linear Association	,000[c]	1	,994	1,000	,534	,072
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 73,04.

b. Computed only for a 2x2 table

c. The standardized statistic is ,007.

*12r * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12r	No	Count	329	144	473
		% within IT Specialists	77,2%	79,1%	77,8%
	Yes	Count	97	38	135
		% within IT Specialists	22,8%	20,9%	22,2%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,264[a]	1	,607	,670	,344	
Continuity Correction[b]	,166	1	,684			
Likelihood Ratio	,266	1	,606	,670	,344	
Fisher's Exact Test				,670	,344	
Linear-by-Linear Association	,264[c]	1	,608	,670	,344	,075
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 40,41.

b. Computed only for a 2x2 table

c. The standardized statistic is -,513.

*12s * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12s	No	Count	414	175	589
		% within IT Specialists	97,2%	96,2%	96,9%
	Yes	Count	12	7	19
		% within IT Specialists	2,8%	3,8%	3,1%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,446[a]	1	,504	,611	,330	
Continuity Correction[b]	,171	1	,679			
Likelihood Ratio	,430	1	,512	,611	,330	
Fisher's Exact Test				,611	,330	
Linear-by-Linear Association	,445[c]	1	,504	,611	,330	,154
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 5,69.

b. Computed only for a 2x2 table

c. The standardized statistic is ,667.

*12t * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12t	No	Count	343	144	487
		% within IT Specialists	80,5%	79,1%	80,1%
	Yes	Count	83	38	121
		% within IT Specialists	19,5%	20,9%	19,9%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,156[a]	1	,693	,739	,385	
Continuity Correction[b]	,081	1	,777			
Likelihood Ratio	,155	1	,694	,739	,385	
Fisher's Exact Test				,739	,385	
Linear-by-Linear Association	,156[c]	1	,693	,739	,385	,081
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 36,22.

b. Computed only for a 2x2 table

c. The standardized statistic is ,394.

*12u * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
12u	No	Count	344	145	489
		% within IT Specialists	80,8%	79,7%	80,4%
	Yes	Count	82	37	119
		% within IT Specialists	19,2%	20,3%	19,6%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	,095[a]	1	,758	,823	,419	
Continuity Correction[b]	,038	1	,845			
Likelihood Ratio	,094	1	,759	,823	,419	
Fisher's Exact Test				,823	,419	
Linear-by-Linear Association	,094[c]	1	,759	,823	,419	,084
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 35,62.

b. Computed only for a 2x2 table

c. The standardized statistic is ,307.

Annex C – Concern for Internet Privacy (IT Specialists and IT Non-specialists)

			IT Specialists		
			No	Yes	Total
Concern for Internet Privacy	Never thought about it	Count	1	3	4
		% within IT Specialists	0,2%	1,6%	0,7%
	Not concerned at all	Count	6	4	10
		% within IT Specialists	1,4%	2,2%	1,6%
	Not too concerned	Count	45	19	64
		% within IT Specialists	10,6%	10,4%	10,5%
	Somewhat concerned	Count	282	111	393
		% within IT Specialists	66,2%	61,0%	64,6%
	Very concerned	Count	92	45	137
		% within IT Specialists	21,6%	24,7%	22,5%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	5,447[a]	4	,244	,241		
Likelihood Ratio	4,995	4	,288	,334		
Fisher's Exact Test	5,273			,244		
Linear-by-Linear Association	,186[b]	1	,666	,693	,356	,047
N of Valid Cases	608					

a. 3 cells (30,0%) have expected count less than 5. The minimum expected count is 1,20.

b. The standardized statistic is -,431.

Annex D – Concern for Unwanted Audiences

Descriptive Statistics

	Total					*IT non-specialists*					*IT Specialists*				
	N	Mi	Ma	M	SD	N	Mi	Ma	M	SD	N	Mi	Ma	M	SD
8a	608	1	5	3,5	1,21	426	1	5	3,42	1,25	182	1	5	3,55	1,090
8b	608	1	5	2,9	1,27	426	1	5	2,87	1,3	182	1	5	3,04	1,211
8c	608	1	5	3,2	1,27	426	1	5	3,16	1,31	182	1	5	3,23	1,181
8d	608	1	5	3,6	1,29	426	1	5	3,52	1,34	182	1	5	3,61	1,192
8e	608	1	5	4,5	1,03	426	1	5	4,48	1,07	182	1	5	4,59	0,922
8f	608	1	5	3,6	1,33	426	1	5	3,57	1,36	182	1	5	3,7	1,230
Valid N (listwise)	608					426					182				

Mi=Minimum; Ma=Maximum; M=Mean; SD= Standard Deviation

Concern for Internet privacy (IT Non-specialists and Specialists)

		IT Specialists		Total
		No	Yes	
Never thought about it	Count	1	3	4
	% within IT Specialists	0,20%	1,60%	0,70%
Not concerned at all	Count	6	4	10
	% within IT Specialists	1,40%	2,20%	1,60%
Not too concerned	Count	45	19	64
	% within IT Specialists	10,60%	10,40%	10,50%
Somewhat concerned	Count	282	111	393
	% within IT Specialists	66,20%	61,00%	64,60%
Very concerned	Count	92	45	137
	% within IT Specialists	21,60%	24,70%	22,50%
Total	Count	426	182	608
	% within IT Specialists	100,00%	100,00%	100,00%

*IT Specialists * 8a Crosstabulation*

			Never thought about it	Not concerned at all	Not too concerned	Somewhat concerned	Very concerned	
IT Specialists	No	Count	40	73	67	158	88	426
		% within IT Specialists	9,4%	17,1%	15,7%	37,1%	20,7%	100,0%
		Standardized Residual	1,5	-,1	-,9	,0	,0	
	Yes	Count	5	32	40	67	38	182
		% within IT Specialists	2,7%	17,6%	22,0%	36,8%	20,9%	100,0%
		Standardized Residual	-2,3	,1	1,4	,0	,0	
Total		Count	45	105	107	225	126	608
		% within IT Specialists	7,4%	17,3%	17,6%	37,0%	20,7%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	10,453[a]	4	,033	,033		
Likelihood Ratio	11,862	4	,018	,019		
Fisher's Exact Test	11,281			,023		
Linear-by-Linear Association	1,483[b]	1	,223	,226	,119	,014
N of Valid Cases	608					

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 13,47.

b. The standardized statistic is 1,218.

*IT Specialists * 8b Crosstabulation*

			8b					
			Never thought about it	Not concerned at all	Not too concerned	Somewhat concerned	Very concerned	Total
IT Specialists	No	Count	64	138	72	93	59	426
		% within IT Specialists	15,0%	32,4%	16,9%	21,8%	13,8%	100,0%
		Standardized Residual	1,2	-,2	,3	-1,1	,3	
	Yes	Count	15	62	27	56	22	182
		% within IT Specialists	8,2%	34,1%	14,8%	30,8%	12,1%	100,0%
		Standardized Residual	-1,8	,3	-,5	1,7	-,5	
Total		Count	79	200	99	149	81	608
		% within IT Specialists	13,0%	32,9%	16,3%	24,5%	13,3%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	9,411[a]	4	,052	,051		
Likelihood Ratio	9,660	4	,047	,048		
Fisher's Exact Test	9,400			,051		
Linear-by-Linear Association	2,353[b]	1	,125	,126	,067	,009
N of Valid Cases	608					

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 23,65.

b. The standardized statistic is 1,534.

*IT Specialists * 8c Crosstabulation*

			8c					
			Never thought about it	Not concerned at all	Not too concerned	Somewhat concerned	Very concerned	Total
IT Specialists	No	Count	47	116	65	119	79	426
		% within IT Specialists	11,0%	27,2%	15,3%	27,9%	18,5%	100,0%
		Standardized Residual	1,0	,0	-,5	-,6	,5	
	Yes	Count	11	50	34	60	27	182
		% within IT Specialists	6,0%	27,5%	18,7%	33,0%	14,8%	100,0%
		Standardized Residual	-1,5	,0	,8	,9	-,8	
Total		Count	58	166	99	179	106	608
		% within IT Specialists	9,5%	27,3%	16,3%	29,4%	17,4%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	6,351[a]	4	,174	,175		
Likelihood Ratio	6,640	4	,156	,159		
Fisher's Exact Test	6,400			,170		
Linear-by-Linear Association	,426[b]	1	,514	,531	,269	,022
N of Valid Cases	608					

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 17,36.

b. The standardized statistic is ,653.

*IT Specialists * 8d Crosstabulation*

			8d					
			Never thought about it	Not concerned at all	Not too concerned	Somewhat concerned	Very concerned	Total
IT Specialists	No	Count	42	73	59	125	127	426
		% within IT Specialists	9,9%	17,1%	13,8%	29,3%	29,8%	100,0%
		Standardized Residual	1,2	-,3	,0	-,8	,4	
	Yes	Count	8	35	25	66	48	182
		% within IT Specialists	4,4%	19,2%	13,7%	36,3%	26,4%	100,0%
		Standardized Residual	-1,8	,5	,0	1,2	-,6	
Total		Count	50	108	84	191	175	608
		% within IT Specialists	8,2%	17,8%	13,8%	31,4%	28,8%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	7,413[a]	4	,116	,116		
Likelihood Ratio	7,950	4	,093	,096		
Fisher's Exact Test	7,609			,106		
Linear-by-Linear Association	,600[b]	1	,439	,452	,230	,020
N of Valid Cases	608					

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 14,97.

b. The standardized statistic is ,775.

*IT Specialists * 8e Crosstabulation*

			8e					
			Never thought about it	Not concerned at all	Not too concerned	Somewhat concerned	Very concerned	Total
IT Specialists	No	Count	18	21	19	49	319	426
		% within IT Specialists	4,2%	4,9%	4,5%	11,5%	74,9%	100,0%
		Standardized Residual	,9	-,2	,2	,3	-,3	
	Yes	Count	3	10	7	18	144	182
		% within IT Specialists	1,6%	5,5%	3,8%	9,9%	79,1%	100,0%
		Standardized Residual	-1,3	,2	-,3	-,5	,5	
Total		Count	21	31	26	67	463	608
		% within IT Specialists	3,5%	5,1%	4,3%	11,0%	76,2%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	3,246[a]	4	,518	,523		
Likelihood Ratio	3,618	4	,460	,474		
Fisher's Exact Test	3,125			,539		
Linear-by-Linear Association	1,582[b]	1	,208	,213	,111	,016
N of Valid Cases	608					

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 6,29.

b. The standardized statistic is 1,258.

*IT Specialists * 8f Crosstabulation*

			8f					
			Never thought about it	Not concerned at all	Not too concerned	Somewhat concerned	Very concerned	Total
IT Specialists	No	Count	50	53	66	117	140	426
		% within IT Specialists	11,7%	12,4%	15,5%	27,5%	32,9%	100,0%
		Standardized Residual	1,1	-,3	,0	-,5	,0	
	Yes	Count	11	26	28	58	59	182
		% within IT Specialists	6,0%	14,3%	15,4%	31,9%	32,4%	100,0%
		Standardized Residual	-1,7	,5	,0	,8	-,1	
Total		Count	61	79	94	175	199	608
		% within IT Specialists	10,0%	13,0%	15,5%	28,8%	32,7%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	5,321[a]	4	,256	,257		
Likelihood Ratio	5,718	4	,221	,225		
Fisher's Exact Test	5,501			,239		
Linear-by-Linear Association	1,237[b]	1	,266	,271	,140	,014
N of Valid Cases	608					

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 18,26.

b. The standardized statistic is 1,112.

Annex E – Concern about future audiences

Descriptive Statistics

	Total					*Não Especialistas*					*Especialistas*				
	N	Mi	Ma	M	SD	N	Mi	Ma	M	SD	N	Mi	Ma	M	SD
9a	608	1	5	3,1	1,11	426	1	5	3,2	1,13	182	1	5	3,1	1,07
9b	608	1	5	3,1	1,26	426	1	5	3,1	1,27	182	1	5	3,1	1,22
9c	608	1	5	3,3	1,21	426	1	5	3,2	1,23	182	1	5	3,3	1,16
Valid N (listwise)	608					426					182				

Mi=Minimum; Ma=Maximum; M=Mean; SD= Standard Deviation

*9a * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
9a	Never thought about it	Count	22	6	28
		% within IT Specialists	5,2%	3,3%	4,6%
	Not concerned at all	Count	123	56	179
		% within IT Specialists	28,9%	30,8%	29,4%
	Not too concerned	Count	108	55	163
		% within IT Specialists	25,4%	30,2%	26,8%
	Somewhat concerned	Count	116	44	160
		% within IT Specialists	27,2%	24,2%	26,3%
	Very concerned	Count	57	21	78
		% within IT Specialists	13,4%	11,5%	12,8%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	3,038[a]	4	,552	,554		
Likelihood Ratio	3,090	4	,543	,546		
Fisher's Exact Test	2,892			,577		
Linear-by-Linear Association	,247[b]	1	,619	,633	,324	,028
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 8,38.

b. The standardized statistic is -,497.

*9b * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
9b	Never thought about it	Count	41	11	52
		% within IT Specialists	9,6%	6,0%	8,6%
	Not concerned at all	Count	111	59	170
		% within IT Specialists	26,1%	32,4%	28,0%
	Not too concerned	Count	104	47	151
		% within IT Specialists	24,4%	25,8%	24,8%
	Somewhat concerned	Count	86	30	116
		% within IT Specialists	20,2%	16,5%	19,1%
	Very concerned	Count	84	35	119
		% within IT Specialists	19,7%	19,2%	19,6%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	4,792[a]	4	,309	,311		
Likelihood Ratio	4,898	4	,298	,302		
Fisher's Exact Test	4,702			,319		
Linear-by-Linear Association	,121[b]	1	,727	,751	,377	,026
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 15,57.

b. The standardized statistic is -,348.

*9c * IT Specialists Crosstabulation*

			IT Specialists		
			No	Yes	Total
9c	Never thought about it	Count	33	7	40
		% within IT Specialists	7,7%	3,8%	6,6%
	Not concerned at all	Count	106	48	154
		% within IT Specialists	24,9%	26,4%	25,3%
	Not too concerned	Count	99	39	138
		% within IT Specialists	23,2%	21,4%	22,7%
	Somewhat concerned	Count	108	55	163
		% within IT Specialists	25,4%	30,2%	26,8%
	Very concerned	Count	80	33	113
		% within IT Specialists	18,8%	18,1%	18,6%
Total		Count	426	182	608
		% within IT Specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	4,401[a]	4	,354	,356		
Likelihood Ratio	4,695	4	,320	,324		
Fisher's Exact Test	4,401			,354		
Linear-by-Linear Association	,850[b]	1	,357	,361	,188	,019
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 11,97.

b. The standardized statistic is ,922.

Annex F – Profile visibility (IT Specialists and IT Non-specialists)

Descriptive Statistics

	Total				IT Non-specialists				IT Specialists			
	Freq	Perc	ValPer	CumP	Freq	Perc	ValPer	CumP	Freq	Perc	ValPer	CumP
a	444	73	73	73	317	74,4	74,4	74,4	127	69,8	69,8	69,8
b	75	12,3	12,3	85,4	47	11	11	85,4	28	15,4	15,4	85,2
c	34	5,6	5,6	91	26	6,1	6,1	91,5	8	4,4	4,4	89,6
d	55	9	9	100	36	8,5	8,5	100	19	10,4	10,4	100
Total	608	100	100		426	100	100		182	100	100	

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	3,555[a]	3	,314	,314		
Likelihood Ratio	3,496	3	,321	,323		
Fisher's Exact Test	3,541			,313		
Linear-by-Linear Association	,671[b]	1	,413	,430	,219	,026
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 10,18.

b. The standardized statistic is ,819.

Annex G – Privacy Protection Strategies (IT Non-specialists and Specialists)

Descriptive Statistics

	Total					*IT non-specialists*					*IT specialists*				
	N	Mi	Ma	M	SD	N	Mi	Ma	M	SD	N	Mi	Ma	M	SD
11a	608	1	5	2	1,3	426	1	5	2	1,26	182	1	5	2,2	1,39
11b	608	1	5	3,6	1,27	426	1	5	3,6	1,28	182	1	5	3,6	1,26
11c	608	1	5	3,7	1,21	426	1	5	3,7	1,23	182	1	5	3,8	1,17
11d	608	1	5	3,2	1,31	426	1	5	3,2	1,33	182	1	5	3,3	1,29
11e	608	1	5	3,2	1,33	426	1	5	3,2	1,36	182	1	5	3,3	1,28
11f	608	1	5	3,6	1,26	426	1	5	3,5	1,26	182	1	5	3,8	1,24
11g	608	1	5	3	1,29	426	1	5	2,9	1,32	182	1	5	3,1	1,23
11h	608	1	5	3,2	1,24	426	1	5	3,1	1,26	182	1	5	3,3	1,2
Valid N (listwise)	608					426					182				

Mi=Minimum; Ma=Maximum; M=Mean; SD= Standard Deviation

*11a * IT specialists Crosstabulation*

			IT specialists		
			No	Yes	Total
11a	Never	Count	226	82	308
		% within IT specialists	53,1%	45,1%	50,7%
	Rarely	Count	81	45	126
		% within IT specialists	19,0%	24,7%	20,7%
	Sometimes	Count	61	20	81
		% within IT specialists	14,3%	11,0%	13,3%
	Often	Count	27	13	40
		% within IT specialists	6,3%	7,1%	6,6%
	Always	Count	31	22	53
		% within IT specialists	7,3%	12,1%	8,7%
Total		Count	426	182	608
		% within IT specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	8,190[a]	4	,085	,085		
Likelihood Ratio	8,002	4	,091	,095		
Fisher's Exact Test	8,132			,085		
Linear-by-Linear Association	3,236[b]	1	,072	,076	,040	,005
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 11,97.

b. The standardized statistic is 1,799.

*11b * IT specialists Crosstabulation*

			IT specialists		
			No	Yes	Total
11b	Never	Count	30	13	43
		% within IT specialists	7,0%	7,1%	7,1%
	Rarely	Count	70	24	94
		% within IT specialists	16,4%	13,2%	15,5%
	Sometimes	Count	98	42	140
		% within IT specialists	23,0%	23,1%	23,0%
	Often	Count	92	44	136
		% within IT specialists	21,6%	24,2%	22,4%
	Always	Count	136	59	195
		% within IT specialists	31,9%	32,4%	32,1%
Total		Count	426	182	608
		% within IT specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1,260[a]	4	,868	,870		
Likelihood Ratio	1,281	4	,865	,866		
Fisher's Exact Test	1,272			,870		
Linear-by-Linear Association	,343[b]	1	,558	,578	,292	,023
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 12,87.

b. The standardized statistic is ,586.

*11c * IT specialists Crosstabulation*

			IT specialists		
			No	Yes	Total
11c	Never	Count	31	12	43
		% within IT specialists	7,3%	6,6%	7,1%
	Rarely	Count	44	13	57
		% within IT specialists	10,3%	7,1%	9,4%
	Sometimes	Count	97	38	135
		% within IT specialists	22,8%	20,9%	22,2%
	Often	Count	115	61	176
		% within IT specialists	27,0%	33,5%	28,9%
	Always	Count	139	58	197
		% within IT specialists	32,6%	31,9%	32,4%
Total		Count	426	182	608
		% within IT specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	3,566[a]	4	,468	,470		
Likelihood Ratio	3,593	4	,464	,469		
Fisher's Exact Test	3,446			,486		
Linear-by-Linear Association	,791[b]	1	,374	,382	,197	,020
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 12,87.

b. The standardized statistic is ,889.

*11d * IT specialists Crosstabulation*

			IT specialists		
			No	Yes	Total
11d	Never	Count	46	16	62
		% within IT specialists	10,8%	8,8%	10,2%
	Rarely	Count	103	41	144
		% within IT specialists	24,2%	22,5%	23,7%
	Sometimes	Count	96	46	142
		% within IT specialists	22,5%	25,3%	23,4%
	Often	Count	82	36	118
		% within IT specialists	19,2%	19,8%	19,4%
	Always	Count	99	43	142
		% within IT specialists	23,2%	23,6%	23,4%
Total		Count	426	182	608
		% within IT specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	1,087[a]	4	,896	,898		
Likelihood Ratio	1,097	4	,895	,896		
Fisher's Exact Test	1,065			,904		
Linear-by-Linear Association	,359[b]	1	,549	,567	,286	,022
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 18,56.

b. The standardized statistic is ,599.

*11e * IT specialists Crosstabulation*

			IT specialists		
			No	Yes	Total
11e	Never	Count	60	20	80
		% within IT specialists	14,1%	11,0%	13,2%
	Rarely	Count	78	35	113
		% within IT specialists	18,3%	19,2%	18,6%
	Sometimes	Count	103	44	147
		% within IT specialists	24,2%	24,2%	24,2%
	Often	Count	86	46	132
		% within IT specialists	20,2%	25,3%	21,7%
	Always	Count	99	37	136
		% within IT specialists	23,2%	20,3%	22,4%
Total		Count	426	182	608
		% within IT specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	2,989[a]	4	,560	,561		
Likelihood Ratio	2,988	4	,560	,563		
Fisher's Exact Test	2,944			,569		
Linear-by-Linear Association	,147[b]	1	,701	,715	,363	,025
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 23,95.

b. The standardized statistic is ,384.

*11f * IT specialists Crosstabulation*

			IT specialists		
			No	Yes	Total
11f	Never	Count	28	11	39
		% within IT specialists	6,6%	6,0%	6,4%
	Rarely	Count	74	19	93
		% within IT specialists	17,4%	10,4%	15,3%
	Sometimes	Count	96	34	130
		% within IT specialists	22,5%	18,7%	21,4%
	Often	Count	100	44	144
		% within IT specialists	23,5%	24,2%	23,7%
	Always	Count	128	74	202
		% within IT specialists	30,0%	40,7%	33,2%
Total		Count	426	182	608
		% within IT specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	9,296[a]	4	,054	,054		
Likelihood Ratio	9,464	4	,050	,052		
Fisher's Exact Test	9,242			,055		
Linear-by-Linear Association	7,162[b]	1	,007	,008	,004	,001
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 11,67.

b. The standardized statistic is 2,676.

Nota: Este resultado pode ser considerado significativo

*11g * IT specialists Crosstabulation*

			IT specialists		
			No	Yes	Total
11g	Never	Count	72	22	94
		% within IT specialists	16,9%	12,1%	15,5%
	Rarely	Count	109	39	148
		% within IT specialists	25,6%	21,4%	24,3%
	Sometimes	Count	103	54	157
		% within IT specialists	24,2%	29,7%	25,8%
	Often	Count	74	40	114
		% within IT specialists	17,4%	22,0%	18,8%
	Always	Count	68	27	95
		% within IT specialists	16,0%	14,8%	15,6%
Total		Count	426	182	608
		% within IT specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	5,854[a]	4	,210	,211		
Likelihood Ratio	5,889	4	,208	,211		
Fisher's Exact Test	5,806			,214		
Linear-by-Linear Association	1,983[b]	1	,159	,161	,085	,010
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 28,14.

b. The standardized statistic is 1,408.

*11h * IT specialists Crosstabulation*

			IT specialists		
			No	Yes	Total
11h	Never	Count	44	15	59
		% within IT specialists	10,3%	8,2%	9,7%
	Rarely	Count	112	34	146
		% within IT specialists	26,3%	18,7%	24,0%
	Sometimes	Count	96	53	149
		% within IT specialists	22,5%	29,1%	24,5%
	Often	Count	101	47	148
		% within IT specialists	23,7%	25,8%	24,3%
	Always	Count	73	33	106
		% within IT specialists	17,1%	18,1%	17,4%
Total		Count	426	182	608
		% within IT specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	6,211[a]	4	,184	,184		
Likelihood Ratio	6,306	4	,177	,181		
Fisher's Exact Test	6,218			,183		
Linear-by-Linear Association	2,082[b]	1	,149	,155	,080	,010
N of Valid Cases	608					

a. 0 cells (0,0%) have expected count less than 5. The minimum expected count is 17,66.

b. The standardized statistic is 1,443.

Anexx H – General Data Protection Regulation (GDPR) and Behaviors

Degree of information on the GDPR

	Total				IT non-specialists				IT specialists			
	Freq	Perc	ValPer	CumP	Freq	Perc	ValPer	CumP	Freq	Perc	ValPer	CumP
Never heard of it	7	1,2	1,2	1,2	6	1,4	1,4	1,4	1	0,5	0,5	0,5
Not informed at all	13	2,1	2,1	3,3	11	2,6	2,6	4	2	1,1	1,1	1,6
Somewhat informed	113	18,6	18,6	21,9	89	20,9	20,9	24,9	24	13,2	13,2	14,8
Moderately informed	347	57,1	57,1	78,9	247	58	58	82,9	100	54,9	54,9	69,8
Very informed	128	21,1	21,1	100	73	17,1	17,1	100	55	30,2	30,2	100
Total	608	100	100		426	100	100		182	100	100	

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

Change in behavior while using social networks, after the implementation of the GDPR

		Total				IT non-specialists				IT specialists			
		Freq	Perc	ValPer	CumP	Freq	Perc	ValPer	CumP	Freq	Perc	ValPer	CumP
Valid	None	107	17,6	17,8	17,8	82	19,2	19,5	19,5	25	13,7	13,8	13,8
	Very mild	176	28,9	29,3	47,1	123	28,9	29,3	48,8	53	29,1	29,3	43,1
	Moderate	216	35,5	35,9	83	157	36,9	37,4	86,2	59	32,4	32,6	75,7
	Very	84	13,8	14	97	49	11,5	11,7	97,9	35	19,2	19,3	95
	Totally	18	3	3	100	9	2,1	2,1	100	9	4,9	5	100
	Total	601	98,8	100		420	98,6	100		181	99,5	100	
Missing	System	7	1,2			6	1,4			1	0,5		
Total		608	100			426	100			182	100		

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

Perception of changes in the way social networks operate, by the management entities, having in mind the compliance with the established on the GDPR

		Total				Não Especialista				Especialista			
		Freq	Perc	ValPer	CumP	Freq	Perc	ValPer	CumP	Freq	Perc	ValPer	CumP
Valid	None	34	5,6	5,7	5,7	28	6,6	6,7	6,7	6	3,3	3,3	3,3
	Very mild	213	35	35,4	41,1	156	36,6	37,1	43,8	57	31,3	31,5	34,8
	Moderate	243	40	40,4	81,5	158	37,1	37,6	81,4	85	46,7	47	81,8
	Very	95	15,6	15,8	97,3	66	15,5	15,7	97,1	29	15,9	16	97,8
	Totally	16	2,6	2,7	100	12	2,8	2,9	100	4	2,2	2,2	100
	Total	601	98,8	100		420	98,6	100		181	99,5	100	
Missing	System	7	1,2			6	1,4			1	0,5		
Total		608	100			426	100			182	100		

Freq=Frequency; Perc=Percent; ValPerc=Valid Percent; CumP=Cumulative Percent

*AltFuncionEntGest * IT specialists Crosstabulation*

			IT specialists		
			No	Yes	Total
AltFuncionEntGest	None	Count	28	6	34
		% within IT specialists	6,7%	3,3%	5,7%
	Very mild	Count	156	57	213
		% within IT specialists	37,1%	31,5%	35,4%
	Moderate	Count	158	85	243
		% within IT specialists	37,6%	47,0%	40,4%
	Very	Count	66	29	95
		% within IT specialists	15,7%	16,0%	15,8%
	Totally	Count	12	4	16
		% within IT specialists	2,9%	2,2%	2,7%
Total		Count	420	181	601
		% within IT specialists	100,0%	100,0%	100,0%

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)	Point Probability
Pearson Chi-Square	6,589[a]	4	,159	,158		
Likelihood Ratio	6,815	4	,146	,154		
Fisher's Exact Test	6,428			,166		
Linear-by-Linear Association	2,093[b]	1	,148	,159	,082	,014
N of Valid Cases	601					

a. 1 cells (10,0%) have expected count less than 5. The minimum expected count is 4,82.

b. The standardized statistic is 1,447.

Annex I – Result of hierarchical multiple linear regression (Test of hypotheses 1, 2 and 3)

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,027[a]	,001	-,001	,15450	,001	,448	1	606	,504
2	,163[b]	,026	,023	,15263	,026	15,972	1	605	,000
3	,249[c]	,062	,058	,14992	,036	23,053	1	604	,000

a. Predictors: (Constant), FreqAcessoRS

b. Predictors: (Constant), FreqAcessoRS, DimConexões

c. Predictors: (Constant), FreqAcessoRS, DimConexões, PreocPrivRS

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,011	1	,011	,448	,504[b]
	Residual	14,466	606	,024		
	Total	14,477	607			
2	Regression	,383	2	,191	8,215	,000[c]
	Residual	14,094	605	,023		
	Total	14,477	607			
3	Regression	,901	3	,300	13,361	,000[d]
	Residual	13,576	604	,022		
	Total	14,477	607			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), FreqAcessoRS

c. Predictors: (Constant), FreqAcessoRS, DimConexões

d. Predictors: (Constant), FreqAcessoRS, DimConexões, PreocPrivRS

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,296	,013		23,120	,000		
	FreqAcessoRS	-,006	,009	-,027	-,669	,504	1,000	1,000
2	(Constant)	,248	,017		14,195	,000		
	FreqAcessoRS	-,006	,009	-,027	-,669	,504	1,000	1,000
	DimConexões	,021	,005	,160	3,997	,000	1,000	1,000
3	(Constant)	,432	,042		10,305	,000		
	FreqAcessoRS	-,008	,009	-,036	-,920	,358	,998	1,003
	DimConexões	,019	,005	,149	3,766	,000	,996	1,004
	PreocPrivRS	-,044	,009	-,190	-4,801	,000	,994	1,006

a. Dependent Variable: InformDisp

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,086[a]	,007	,002	,16146	,007	1,341	1	180	,248
2	,105[b]	,011	,000	,16161	,004	,653	1	179	,420
3	,163[c]	,027	,010	,16078	,016	2,857	1	178	,093

a. Predictors: (Constant), FreqAcessoRS

b. Predictors: (Constant), FreqAcessoRS, DimConexões

c. Predictors: (Constant), FreqAcessoRS, DimConexões, PreocPrivRS

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,035	1	,035	1,341	,248[b]
	Residual	4,692	180	,026		
	Total	4,727	181			
2	Regression	,052	2	,026	,996	,372[c]
	Residual	4,675	179	,026		
	Total	4,727	181			
3	Regression	,126	3	,042	1,623	,186[d]
	Residual	4,602	178	,026		
	Total	4,727	181			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), FreqAcessoRS

c. Predictors: (Constant), FreqAcessoRS, DimConexões

d. Predictors: (Constant), FreqAcessoRS, DimConexões, PreocPrivRS

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,323	,023		14,002	,000		
	FreqAcessoRS	-,019	,016	-,086	-1,158	,248	1,000	1,000
2	(Constant)	,302	,035		8,549	,000		
	FreqAcessoRS	-,018	,016	-,084	-1,125	,262	,999	1,001
	DimConexões	,008	,010	,060	,808	,420	,999	1,001
3	(Constant)	,413	,075		5,531	,000		
	FreqAcessoRS	-,017	,016	-,076	-1,020	,309	,994	1,006
	DimConexões	,006	,010	,044	,583	,561	,981	1,019
	PreocPrivRS	-,027	,016	-,126	-1,690	,093	,978	1,022

a. Dependent Variable: InformDisp

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,002[a]	,000	-,002	,15137	,000	,001	1	424	,971
2	,198[b]	,039	,035	,14855	,039	17,264	1	423	,000
3	,302[c]	,091	,085	,14465	,052	24,106	1	422	,000

a. Predictors: (Constant), FreqAcessoRS

b. Predictors: (Constant), FreqAcessoRS, DimConexões

c. Predictors: (Constant), FreqAcessoRS, DimConexões, PreocPrivRS

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,000	1	,000	,001	,971[b]
	Residual	9,715	424	,023		
	Total	9,715	425			
2	Regression	,381	2	,190	8,633	,000[c]
	Residual	9,334	423	,022		
	Total	9,715	425			
3	Regression	,885	3	,295	14,105	,000[d]
	Residual	8,830	422	,021		
	Total	9,715	425			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), FreqAcessoRS

c. Predictors: (Constant), FreqAcessoRS, DimConexões

d. Predictors: (Constant), FreqAcessoRS, DimConexões, PreocPrivRS

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,284	,016		18,289	,000		
	FreqAcessoRS	,000	,012	,002	,036	,971	1,000	1,000
2	(Constant)	,228	,020		11,217	,000		
	FreqAcessoRS	1,666E-5	,011	,000	,001	,999	1,000	1,000
	DimConexões	,026	,006	,198	4,155	,000	1,000	1,000
3	(Constant)	,462	,052		8,944	,000		
	FreqAcessoRS	-,007	,011	-,028	-,594	,553	,985	1,015
	DimConexões	,025	,006	,194	4,176	,000	1,000	1,000
	PreocPrivRS	-,055	,011	-,230	-4,910	,000	,985	1,015

a. Dependent Variable: InformDisp

Annex J – Results of simple linear regressions (Test of hypothesis 4)

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,104[a]	,011	,009	,15373	,011	6,603	1	606	,010

a. Predictors: (Constant), UtiNãoDesej

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,156	1	,156	6,603	,010[b]
	Residual	14,321	606	,024		
	Total	14,477	607			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), UtiNãoDesej

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,349	,024		14,418	,000		
	UtiNãoDesej	-,017	,007	-,104	-2,570	,010	1,000	1,000

a. Dependent Variable: InformDisp

Model Summary

					Change Statistics				
		R	Adjusted R	Std. Error of the	R Square	F			Sig. F
Model	R	Square	Square	Estimate	Change	Change	df1	df2	Change
1	,045[a]	,002	-,004	,16189	,002	,366	1	180	,546

a. Predictors: (Constant), UtiNãoDesej

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,010	1	,010	,366	,546[b]
	Residual	4,718	180	,026		
	Total	4,727	181			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), UtiNãoDesej

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,331	,053		6,294	,000		
	UtiNãoDesej	-,009	,014	-,045	-,605	,546	1,000	1,000

a. Dependent Variable: InformDisp

Model Summary

					Change Statistics				
		R	Adjusted R	Std. Error of the	R Square	F			Sig. F
Model	R	Square	Square	Estimate	Change	Change	df1	df2	Change
1	,132[a]	,017	,015	,15005	,017	7,517	1	424	,006

a. Predictors: (Constant), UtiNãoDesej

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,169	1	,169	7,517	,006[b]
	Residual	9,546	424	,023		
	Total	9,715	425			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), UtiNãoDesej

Coefficients[a]

		Unstandardized		Standardized			Collinearity	
		Coefficients		Coefficients			Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,355	,027		13,155	,000		
	UtiNãoDesej	-,020	,007	-,132	-2,742	,006	1,000	1,000

a. Dependent Variable: InformDisp

Annex K – Results of simple linear regressions (Test of hypothesis 5)

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,230[a]	,053	,051	,15043	,053	33,724	1	606	,000

a. Predictors: (Constant), PermiVisPer

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,763	1	,763	33,724	,000[b]
	Residual	13,714	606	,023		
	Total	14,477	607			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), PermiVisPer

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,233	,011		20,370	,000		
	PermiVisPer	,037	,006	,230	5,807	,000	1,000	1,000

a. Dependent Variable: InformDisp

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,316[a]	,100	,095	,15375	,100	19,987	1	180	,000

a. Predictors: (Constant), PermiVisPer

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,472	1	,472	19,987	,000[b]
	Residual	4,255	180	,024		
	Total	4,727	181			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), PermiVisPer

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,220	,021		10,275	,000		
	PermiVisPer	,052	,012	,316	4,471	,000	1,000	1,000

a. Dependent Variable: InformDisp

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,187[a]	,035	,033	,14871	,035	15,289	1	424	,000

a. Predictors: (Constant), PermiVisPer

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,338	1	,338	15,289	,000[b]
	Residual	9,377	424	,022		
	Total	9,715	425			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), PermiVisPer

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,239	,014		17,728	,000		
	PermiVisPer	,030	,008	,187	3,910	,000	1,000	1,000

a. Dependent Variable: InformDisp

Annex L - Results of simple linear regressions (Test of hypothesis 6)

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,007[a]	,000	-,002	,92785	,000	,027	1	606	,868

a. Predictors: (Constant), DimConexões

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,024	1	,024	,027	,868[b]
	Residual	521,705	606	,861		
	Total	521,728	607			

a. Dependent Variable: EstProtDados

b. Predictors: (Constant), DimConexões

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	3,194	,082		38,822	,000		
	DimConexões	-,005	,032	-,007	-,166	,868	1,000	1,000

a. Dependent Variable: EstProtDados

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,009[a]	,000	-,005	,91006	,000	,015	1	180	,903

a. Predictors: (Constant), DimConexões

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,012	1	,012	,015	,903[b]
	Residual	149,078	180	,828		
	Total	149,090	181			

a. Dependent Variable: EstProtDados

b. Predictors: (Constant), DimConexões

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	3,296	,161		20,512	,000		
	DimConexões	-,007	,056	-,009	-,122	,903	1,000	1,000

a. Dependent Variable: EstProtDados

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,021[a]	,000	-,002	,93421	,000	,193	1	424	,661

a. Predictors: (Constant), DimConexões

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,169	1	,169	,193	,661[b]
	Residual	370,045	424	,873		
	Total	370,213	425			

a. Dependent Variable: EstProtDados

b. Predictors: (Constant), DimConexões

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	3,178	,097		32,927	,000		
	DimConexões	-,017	,039	-,021	-,440	,661	1,000	1,000

a. Dependent Variable: EstProtDados

Annex M – Results of simple linear regressions (Test of hypothesis 7)

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,018[a]	,000	-,001	,15454	,000	,198	1	606	,657

a. Predictors: (Constant), InfRGPD

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,005	1	,005	,198	,657[b]
	Residual	14,472	606	,024		
	Total	14,477	607			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), InfRGPD

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,275	,033		8,317	,000		
	InfRGPD	,004	,008	,018	,445	,657	1,000	1,000

a. Dependent Variable: InformDisp

Model Summary

					Change Statistics				
		R	Adjusted R	Std. Error of the	R Square	F			Sig. F
Model	R	Square	Square	Estimate	Change	Change	df1	df2	Change
1	,042[a]	,002	-,004	,16191	,002	,323	1	180	,571

a. Predictors: (Constant), InfRGPD

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,008	1	,008	,323	,571[b]
	Residual	4,719	180	,026		
	Total	4,727	181			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), InfRGPD

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,261	,070		3,703	,000		
	InfRGPD	,010	,017	,042	,568	,571	1,000	1,000

a. Dependent Variable: InformDisp

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,003[a]	,000	-,002	,15137	,000	,003	1	424	,953

a. Predictors: (Constant), InfRGPD

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	,000	1	,000	,003	,953[b]
	Residual	9,715	424	,023		
	Total	9,715	425			

a. Dependent Variable: InformDisp

b. Predictors: (Constant), InfRGPD

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	,286	,038		7,619	,000		
	InfRGPD	-,001	,010	-,003	-,059	,953	1,000	1,000

a. Dependent Variable: InformDisp

Annex N – Results of simple linear regressions (Test of hypothesis 8)

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,243[a]	,059	,057	1,002	,059	37,570	1	599	,000

a. Predictors: (Constant), InfRGPD

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	37,696	1	37,696	37,570	,000[b]
	Residual	601,007	599	1,003		
	Total	638,702	600			

a. Dependent Variable: AltCompPós

b. Predictors: (Constant), InfRGPD

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	1,122	,237		4,742	,000		
	InfRGPD	,359	,059	,243	6,129	,000	1,000	1,000

a. Dependent Variable: AltCompPós

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,185[a]	,034	,029	1,065	,034	6,358	1	179	,013

a. Predictors: (Constant), InfRGPD

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	7,210	1	7,210	6,358	,013[b]
	Residual	202,978	179	1,134		
	Total	210,188	180			

a. Dependent Variable: AltCompPós

b. Predictors: (Constant), InfRGPD

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	1,501	,491		3,053	,003		
	InfRGPD	,295	,117	,185	2,522	,013	1,000	1,000

a. Dependent Variable: AltCompPós

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,250[a]	,063	,060	,971	,063	27,956	1	418	,000

a. Predictors: (Constant), InfRGPD

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	26,377	1	26,377	27,956	,000[b]
	Residual	394,385	418	,944		
	Total	420,762	419			

a. Dependent Variable: AltCompPós

b. Predictors: (Constant), InfRGPD

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	1,065	,271		3,932	,000		
	InfRGPD	,361	,068	,250	5,287	,000	1,000	1,000

a. Dependent Variable: AltCompPós

Annex O – Results of simple linear regressions (Test of hypothesis 9)

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,276[a]	,076	,074	,993	,076	49,272	1	599	,000

a. Predictors: (Constant), AltFuncionEntGest

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	48,545	1	48,545	49,272	,000[b]
	Residual	590,157	599	,985		
	Total	638,702	600			

a. Dependent Variable: AltCompPós

b. Predictors: (Constant), AltFuncionEntGest

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	1,668	,132		12,617	,000		
	AltFuncionEntGest	,322	,046	,276	7,019	,000	1,000	1,000

a. Dependent Variable: AltCompPós

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,259[a]	,067	,062	1,047	,067	12,850	1	179	,000

a. Predictors: (Constant), AltFuncionEntGest

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	14,078	1	14,078	12,850	,000[b]
	Residual	196,110	179	1,096		
	Total	210,188	180			

a. Dependent Variable: AltCompPós

b. Predictors: (Constant), AltFuncionEntGest

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	1,758	,280		6,270	,000		
	AltFuncionEntGest	,342	,095	,259	3,585	,000	1,000	1,000

a. Dependent Variable: AltCompPós

Model Summary

					Change Statistics				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	R Square Change	F Change	df1	df2	Sig. F Change
1	,278[a]	,077	,075	,964	,077	34,965	1	418	,000

a. Predictors: (Constant), AltFuncionEntGest

ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	32,480	1	32,480	34,965	,000[b]
	Residual	388,282	418	,929		
	Total	420,762	419			

a. Dependent Variable: AltCompPós

b. Predictors: (Constant), AltFuncionEntGest

Coefficients[a]

		Unstandardized Coefficients		Standardized Coefficients			Collinearity Statistics	
Model		B	Std. Error	Beta	t	Sig.	Tolerance	VIF
1	(Constant)	1,647	,148		11,130	,000		
	AltFuncionEntGest	,306	,052	,278	5,913	,000	1,000	1,000

a. Dependent Variable: AltCompPós

Anexx P – Results of t-student test (Test of hypothesis 10)

Group Statistics

	IT specialists	N	Mean	Std. Deviation	Std. Error Mean
AltFuncionEntGest	No	420	2,71	,909	,044
	Yes	181	2,82	,818	,061

Independent Samples Test

		Levene's Test		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
AltFuncionEntGest	Equal variances assumed	6,602	,010	-1,448	599	,148	-,114	,079	-,268	,040
	Equal variances not assumed			-1,511	377,056	,132	-,114	,075	-,262	,034