Surveillance in the Digital Era

Master`s Thesis

Master in Development and International Relations

Student: Razvan-Stefan Strajeru

Supervisor: Malayna Raftopoulos

Hand in date: 31 May 2017

- Aalborg University -

Abstract

One of the first writings about privacy was published in 1980 in the Harvard Law Review. The writers of the article, Louis Brandeis and Samuel Warren defined Privacy as the Right to be let alone. Following this, in 1948 the United Nations published the Universal Declaration of Human Rights, which entitled every individual with the Right of Privacy. Furthermore, once with the appearance of the Internet and with the evolution of the technology, human beings were able to communicate with each other and to do different daily tasks with the use of electronic devices. Any occurred communication, any task achieved, any search made online, creates data that leaves behind an electronic footprint. As a result, in 1990 the United Nations published the Regulation of Computerized Personal File document which establishes guidelines for the protection of personal data. Personal data is considered to be the property of the individual who creates it. One of the biggest invasions of privacy and data is represented by the act of surveillance.

This thesis aims to analyze how, by invading individuals` privacy, key information about their lives can be obtained. With the use of the obtained information, it is possible to establish a personal profile. Moreover, it also examines what could be achieved after a personal profile is established. In order to analyze this situation a case study will be applied.

According to the Macmillan dictionary, surveillance represents the process of carefully watching a person who is involved in a criminal activity. The field of surveillance can be split into two categories: Targeted surveillance and Mass surveillance. The targeted surveillance is deployed against a specific individual, while the mass surveillance is not concentrated on the observation of a specific individual. Mass surveillance evolved from the Panopticon model built by Jeremy Bentham and with the use of technology is being deployed through two methods: Video surveillance and Dataveillance. The Video surveillance method is being deployed through closed-circuit television cameras and automated license plate reader system. The Dataveillance method is formed by an assemblage of steps that are including the use of different techniques. With the use of these steps, and by putting together different pieces of information obtained about an individual, a mosaic of his life is created.

What makes this situation so controversial is that, with the use of the obtained mosaic every individual can be influenced or manipulated to take different decisions, in order to achieve the goal of the person who obtained the mosaic.

Abbreviations

ALPR	Automated License Plate Reader
ARPA	Advanced Research Project Agency
BCCLA	British Columbia Civil Liberties Association
CCTV	Closed-Circuit Television Cameras
DOJ	Department of Justice
FBI	Federal Bureau of Investigation
FOIA	Freedom of Information Act
GCHQ	British Government Communications Headquarters
GPS	Global Positioning System
GSM	Global System for Mobile Communications
HEW	Health Education Welfare Advisory Committee
IMSI	International Mobile Subscriber Identity
ISP	Internet Service Provider
KDD	Knowledge Discovery in Databases
MIT	Massachusetts Institute of Technology
NSA	National Security Agency
OECD	Organization for Economic Co-operation and Development
RCMP	Royal Canadian Mounted Police
RFID	Radio-Frequency Identification
SCL	Strategic Communication Laboratories
UCLA	University of California Los Angeles
UDHR	Universal Declaration of Human Rights

Table of Content

Abstract

Abbreviations

1. **Introduction…………………………………………………………………………1**

2. **Methodology……………………………………………………………………...5**

2.1 Synopsis………………………………………………………………………….5

2.2 Research design………………………………………………………………….6

2.3 Research method…………………………………………………………………8

2.3.1 *Data Collection*…………………………………………………………...9

2.3.2 *Limitations*………………………………………………………………..9

3. **Theories and Background…………………………………………………………10**

3.1 Privacy……………………………………………………………………………..10

3.2 Mass Surveillance…………………………………………………………………17

3.2.1 *Internet and Technology*…………………………………………………17

3.2.2 *Surveillance*……………………………………………………………...20

3.2.3 *Data collection*…………………………………………………………..22

3.2.4 *Big Data*…………………………………………………………………27

3.2.5 *Data mining*……………………………………………………………...29

4. **Analysis…………………………………………………………………………….32**

4.1 Superpanopticon…………………………………………………………...32

4.1.1 *Data Collection*…………………………………………………………….35

4.1.1.1 *User generated data*……………………………………………………...35

4.1.1.2 *Data collected without user`s agreement*…………………………………39

4.1.2 *Data mining*……………………………………………………………………...43

4.1.3 *Individual Profiling*………………………………………………………………...46

4.1.4 *Case study: Strategic Communication Laboratories*……………………………….49

5. **Conclusion………………………………………………………………………….55**

6. **Bibliography……………………………………………………………………….58**

1. **Introduction**

Over the last years, the debate over Privacy and Surveillance has captured the world`s attention, thus becoming a hot topic. When we are thinking about surveillance, we are thinking at the importance of privacy, which means to be free from observation. While it is true that surveillance can represent a means to protect privacy such as “biometric identification and video cameras that film those with access to sensitive data” (Marx 2016:23) and that privacy on its turn can protect surveillance, for example, the police men who go undercover and need to use “fake IDs and call forwarding to protect their identity” (ibid), surveillance can also represent an invasion of privacy.

The world`s attention was triggered when Edward Snowden released “top-secret documents showing that the National Security Agency was spying on American citizens” (Franceschi-Bicchierai 2014).

We live in a world where almost every person is using a device with an Internet connection and this is why the reveals of Edward Snowden through which he showed the world the real purpose of surveillance, made the public aware of the violation of their privacy. He revealed that, every telephone company in America was providing the NSA – which is the world`s largest surveillance organization - with “their customers` phone records” (ibid). Furthermore, he revealed that the NSA “intercepts 200 million text messages every day worldwide through a program called Dishfire” (ibid) and that they are not spying only American citizens but also on foreign countries and world leaders, for example German Chancellor Angela Merkel, Brazil`s President Dilma Roussef and Mexico`s former President Felipe Calderon. According to Snowden, NSA even have a tool called “XKeyscore” which is used “to search nearly everything a user does on the internet through data it intercepts across the world” (ibid). The NSA describes this tool as the “widest-reaching system to search through the Internet data” (ibid).

Edward Snowden opened the chapter of the mass surveillance. While the surveillance is likened to the Panopticon idea of Jeremy Bentham, the mass surveillance is represented by the indiscriminate monitoring of the population. In the past, surveillance existed too, but it did not have the same impact on people`s privacy as the new mass surveillance has. In the past, it was just about national data bases which included techniques such as: “censuses registering the subject of a kingdom, ID documenting individuals and tattoos marking them, and numbering and categorizing humans” (Privacy International- What is mass surveillance?).

The technologized mass surveillance involves way more techniques which represent an invasion of privacy. It is based on using the technology and the internet and it evolves with each passing year. In public spaces mass surveillance is being deployed through: closed-circuit television cameras (CCTV) which enable “people`s movements to be tracked and stored for later analysis” (ibid), facial recognition software which identifies individuals, automated license plate reader system (ALPR) which identifies the movement of automobiles and through Stingray devices which are collecting data from every smartphone in a given geographic location, and with computer software’s that are monitoring the communications and extracting data. Economical aspects of our lives can be tracked too, when using a bank card.

With the use of technology and by deploying acts of mass surveillance, the authorities “can now have access to information concerning the entirety of an individual`s life: everything they do, say, think, send, buy, record and obtain, everywhere they go and with whom” (ibid).

For an individual, the idea that he is under surveillance, may have striking results which may lead to changes in his behavior, this depending from individual to individual. People might change the way they act, the way they speak or communicate, in other words they will change their entire behavior because of the fear of knowing that someone is watching their private life. He or she will be more reticent in expressing ideas or opinions, they will also be more reserved when it will come to associate with certain groups or to express different opinions in the political field. All of these behavioral changes are seen as the “chilling effect” (Ahmed 2016) of surveillance.

For the citizens, the use of mass surveillance was explained as a tool of fighting against terrorism. Even though this was the official explanation, a member of the White House review panel on NSA surveillance, Geoffrey Stone, said that, the mass collection of data represented a “logical program from the NSA`s perspective” (Isikoff 2013), but in fact they did not find any evidence that this mass collection of data had actually stopped a single terror attack.

The first words about privacy were written by Louis Brandeis and Samuel Warren, in an article, named The Right to Privacy. The article was published on 15 December 1890 in the Harvard Law Review. In the publication, Brandeis and Warren defined privacy as “the right to be let alone” (Warren, Brandeis 1890: 193).

The most important declaration which recognizes the importance of privacy is the Universal Declaration of Human Rights from 1948 issued by the United Nations. According to this Declaration, every individual is entitled to privacy. This Declaration contains an article that talks about the importance of privacy. According to this article, namely Article 12 of the Declaration, "everyone has the right to the protection of the law" (Universal Declaration of Human Rights 1948:4) against arbitrary interference within his private life or attacks against his privacy.

Later on, more exactly in 1980, the Organization for Economic Co-operation and Development known as OECD, published more guidelines in order to help the protection of privacy. It is important to mention that the "OECD guidelines are not legally binding" (OECD) and those guidelines involve trans-border flow of personal data. The basic principles talk about the importance of personal data.

According to the principles, there should exist limits regarding the amount of personal data collected and it should be obtained "by lawful and fair means" (ibid) and the reason why the data is being collected should be "specified no later than at the time of data collection" (ibid) and most importantly, "the personal data should not be disclosed except with the consent of the data subject; or by the authority of law" (ibid). Furthermore, one of the principles states that the personal data that is collected should be protected against risks such as "destruction, use, modification or disclosure of data" (ibid). Another important principle, states that, the individual should have the right to know from the data controller, if they have information related to him and also, he should be able to "challenge data relating to him and if the challenge is successful to have the data erased, rectified, completed or amended" (ibid).

Ten years later, on 14 December 1990 the United Nations General Assembly adopted the Guidelines for the Regulation of Computerized Personal Files. These guidelines provide ten principles "concerning the minimum guarantees that should be provided in national legislations" (UN General Assembly 1990: 1). Among the ten principles, the principle 8 named Supervision and Sanctions, states that every country should establish the authority that should be responsible for supervising "observance of the principles" (ibid) which concern the minimum guarantees of the protection of personal data. Some of the principles are inspired from the 1980 OECD guidelines.

In the United States of America, the most important writing on privacy and computerized personal files is The Code of Fair Information Practices which was established in 1973 by the Health, Education, Welfare (HEW) advisory Committee on automated data systems. The Code of Fair Information Practices is based on five principles. Among these five principles, the first one clearly states that, there should not exist any secret system that keeps record of personal data, while the other principles state that an individual should have means to find out what information is in his record and how it is being used. Furthermore, the individual should also have means to prevent the information from being used or disclosed without his consent (1973: The Code of Fair Information Practices: 1). These five principles were enforced in the Privacy Act and passed into law in 1974.

As for the European Union, the European Parliament together with the Council of European Union adopted the Directive 95/46/EC of 24 October 1995. It refers to the protection of individuals "with regard to the processing of personal data and on the free movement of such data" (Directive 95/46/EC of the European Parliament and of The Council 1995:1). According to this Directive: "whereas data-processing systems are designed to serve man; whereas they must, whatever the nationality or residence of natural persons, respect their fundamental rights and freedoms, notably the right to privacy" (ibid). Furthermore, according to this Directive, the personal data has to be collected for a specific purpose and it has to be proceeded according to the law. In 2002, the European Parliament and the European Council have adopted the Directive on privacy and electronic communications. This Directive refers to "the processing of personal data and the protection of privacy in the electronic communications sector" (Directive 2002/58/EC of the European Parliament and of The Council 2002:1).

Although there exist all these legislations and directives both in the United States of America and in The European Union, personal data is still being collected. Therefore, the right for privacy is being abused on a daily routine. This subject seemed worth studying to me because the issue of the data collection has become a hot and also an important topic in the International arena and through it, experts may be able to analyze human behavior in detail. Therefore, I have decided to set as the research question of my thesis the following:

How is individual profiling achieved through surveillance and data collection and why is it so controversial?

2. **Methodology**

In this chapter I will describe the research design of my thesis, the procedure of the data collection and the types of data that were used in this thesis. Furthermore, the limitations that I have faced while writing the thesis, will be presented. Moreover, this chapter includes a Synopsis through which I will provide a short briefing about the chapters.

2.1 Synopsis

The first chapter, namely Introduction offers insights over the debate between Privacy and Surveillance. It also presents the reveals made by Edward Snowden and the chilling effect of the Surveillance. Furthermore, a couple of regulations regarding Privacy are presented. Lastly, the research question of the thesis is stated.

The Methodology of the thesis is presented in the second chapter. This chapter presents the research design, research method and data collection of the thesis. Moreover, it also describes the limitations faced when writing the thesis.

The third chapter is called Theories and Background. This chapter is divided into two subchapters. In the first subchapter, namely Privacy, the definition of the word 'Privacy' is presented along with the first writings about Privacy and how the norms and regulation regarding the field of Privacy have evolved through time. Furthermore, it also covers the Panopticon theory framed by Jeremy Bentham and it provides information about how the theory was modernized by Michel Foucault. The second subchapter, namely Mass Surveillance it is split into five subchapters. The first two subchapters, are covering how The Evolution of Internet and Technology have redefined the field of Surveillance. The following subchapters describe the tools used for the Data Collection process, and the meanings of Big Data and Data Mining. The Mosaic theory is explained in the last subchapter.

Analysis is the fourth chapter. The main subchapter of the analysis is named Superpanopticon. This subchapter is divided into four subchapters which are analyzing the steps of the Superpanopticon assemblage. The first three subchapters are analyzing: Data collection, Data Mining and Individual Profiling. Why the Individual Profiling is so controversial is analyzed in the fourth subchapter by using the case study of the Strategic Communications Laboratories.

The fifth and the sixth chapters are the Conclusion and the Bibliography chapter, respectively.

2.2 Research design

This thesis uses the case study as the research design. Robert K Yin stated that: "A case study is an empirical inquiry that investigates a contemporary phenomenon within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident" (Yin: 13), while Peter Swanborn defines the case study as the study of a social phenomenon "in which the researcher focuses on process-tracing : the description and explanation of social processes that unfold between persons participating in the process, people with their values, expectations, opinions, perception, resources, controversies, decisions, mutual relations and behavior, or the description and explanation of processes within and between social institutions" (Swanborn 2010:13).

A case study has several advantages and disadvantages and according to C.R. Kothari some of the advantages are:

"Case study method enhances the experience of the researcher and this in turn increases his analyzing ability and skill.

Case study method has proved beneficial in determining the nature of units to be studied along with the nature of the universe. This is the reason why at times the case study method is alternatively known as mode of organizing data.

The method facilitates intensive study of social units which is generally not possible if we use either the observation method or the method of collecting information through the schedules. This is the reason why case study method is being frequently used, particularly in social researches" (Kothari 2004,1990,1985: 115).

Furthermore, as Kothari states some of the disadvantages are:

“The danger of false generalization is always there in view of the fact that no set of rules are followed in the collection of the information and only few units are studied.

It consumes more time and requires lot of expenditure. More time is needed under case study method since one studies the natural histories cycles of social units and that too minutely.

Case study method is based on several assumptions which may not be very realistic at times and as such the usefulness of case data is always subject to doubt” (Kothari 2004,1990,1985: 116).

Furthermore, Robert K. Yin states that in order to determine the quality of a case study one has to follow three simple principles. These three principles are:

“a) Using multiple, not just single, sources of evidence;

b) Creating a case study database;

c) Maintaining a chain of evidence;” (Yin 2009: 101).

This are the principles which are going to be used throughout my thesis. For instance, I will be using multiple sources of evidence such as government reports, literature and articles from newspapers around the world and through these sources, a database of the thesis will be created. By maintaining a chain of evidence in this thesis, the reader is able to trace back the steps of this thesis. For instance, the reader can follow the steps from the research question to conclusion or from the conclusion to the research question.

The reason of choosing case study as the type of research design is linked with the research question of the thesis. For example, the research question of the thesis starts with “how” and it is related to a social phenomenon and as Yin stated: “case studies are the preferred strategies when how or why questions are being posed when the investigator has little control over events and when the focus is on a contemporary phenomenon within some real-life context” (Yin: 1). Furthermore, he also states that if you are trying to find out “how or why the program had worked or not, you would lean toward a case study” (Yin: 7). On the other hand, research questions like “where”, “who” and “what” are favored for using surveys or economic research.

In order to answer the thesis research question, I have chosen to use the case study of Strategic Communication Laboratories. The current actions of the Strategic Communication Laboratories are considered to be part of ongoing phenomenon, which in my opinion needs to be better understood. Taking into consideration that in the analysis chapter I will use the case study of Strategic Communication Laboratories, this thesis will be regarded as a single case. I have chosen the single case study to find out how individual profiling can be achieved and what makes it so controversial based on the real situation of Strategic Communication Laboratories. The advantage of using a single case study is represented by the fact that it: "can be intensively examined when the research resources at the investigator`s disposal are relatively limited" (Lijphart 1971: 691).

2.3 Research method

This thesis uses as a research method the qualitative approach and it is based on qualitative secondary data. The qualitative research methods are "often regarded as providing rich data about real life people and situations and being more able to make sense of behavior and to understand behavior within its wider context" (Research Methodology- Qualitative Research). According to Robert K. Yin, the qualitative research is described by five features. These features are: "Studying the meaning of people`s lives, under real-world conditions; Representing the views and perspectives of the people in a study; Covering the contextual conditions within which people live; Contributing insights into existing or emerging concepts that may help to explain human social behavior; and Striving to use multiple source of evidence rather than relying on a single source alone" (Yin2011: 7-8) Moreover, Robert K. Yin also argues that: "The study`s conclusions are likely to be based on triangulating the data from the different sources. This convergence will add to the study`s credibility and trustworthiness" (Yin2011: 9).

On the other hand, the quantitative research is "entailing the view of relationship between theory and research as deductive, a predilection for natural science approach, and as having an objectivist conception of social reality" (Research Methodology- Quantitative Research). This thesis also uses some quantitative secondary data in order to provide a better understanding of the findings. By using the quantitative secondary data, the credibility of the findings has been strengthened.

2.3.1 *Data collection*

As I have not collected data by myself, I have not used primary data. This thesis is built on a high amount of secondary data sources, data which was collected, written and analyzed by authors and journalists. I am aware of the risks regarding the secondary data sources such as: reliability, suitability or adequacy.

The qualitative secondary data sources used are: books such as Windows into the Soul by Gary T. Marx; newspapers such as The Guardian, government reports such as British House of Lords and United Nations documents such as The Universal Declaration of Human Rights.

2.3.2 *Limitations*

The most important limitation that I have faced while writing this thesis was the gathering of primary data. I was not able to travel to The United States in order to conduct interviews to obtain information from the National Security Agency about the surveillance tolls. Moreover, I was not able to get in touch with the Data Collection Companies or with the Data Mining Companies.

3.**Theories and Background**

The theories and background chapter is divided into two subchapters. The first subchapter, provides information about the regulations of Privacy, the first writings about it as well as the theory that seemed to be the most suitable for this subchapter. The second subchapter offers insights about Mass Surveillance and through its subchapters, it provides information about the techniques of Mass Surveillance and about the theory that seemed to be the most suitable for the subchapter.

3.1Privacy

Every human being feels the need for privacy. As Jan Holvast stated, the subject of privacy dates back in time and it can be seen "in the writings of Socrates and other Greek philosophers" (Holvast 2009: 15) where a distinction is being made between "the outer and the inner, between public and private, between society and solitude" (ibid). While it is true that from time to time private was perceived as an antisocial behavior, there have also been times when the "periods of retirement normally were accepted" (ibid).

"In the sixteenth century, Queen Elizabeth declared that she did not want to make windows into men`s hearts and secret thoughts. Her actions bolstered liberty and sharpened distinctions between the public and the private that became central to our ideas of the good society and the dignity of the person". (Marx 2016:1)

One of the first publications regarding Privacy was written by Samuel Warren and Louis Brandeis, and it was published in Harvard Law Review in 15 December 1890 and it was titled "The Right to Privacy". According to Warren and Brandeis the term privacy is as old as the common law and it states that an individual should have protection in person and in property but this concept should be defined and updated from time to time. The old meaning of the right to life was represented by the fact that life was assimilated with freedom, and the right to property was represented by the fact that individual`s properties were secured. Later on, this concept was updated and the scope of these rights was broadened. The right to life, took a new meaning in which it meant “the right to enjoy life-to be let alone” (Warren, Brandeis 1890: 193) and property started to “comprise every form of possession” (ibid).

The Constitution of the United States of America, explained in the 19th century the concept of privacy as an individual right with private property forming the basis of the right. The concept evolved and in 1948 the United Nations published the Universal Declaration of Human Rights. In the article 12 of the Universal Declaration of Human Rights it is stated that: “No one shall be subjected to arbitrary interference with his privacy, family, home or correspondence, nor attacks upon his honour and reputation. Everyone has the right to the protection of the law against such interference or attacks” (Universal Declaration of Human Rights 1948:4).

In 1967, Alan Westin published Privacy and Freedom where he defined privacy as “The claim of individuals, groups, or institutions to determine for themselves when, how and to what extent information about them is communicated to others. Viewed in terms of the relation of the individual to social participation, privacy is the voluntary and temporary withdrawal of a person from the general society through physical or psychological means, either in a state of solitude or small-group intimacy or, when among larger groups, in a condition of anonymity or reserve” (Holvast 2009: 13-16). After this definition was published, it was used in almost any publication related to privacy.

On 8 April 1988 at the Thirty-Second Session of the Human Rights Committee, the article 17 was adopted, which was titled as The Right to Respect of Privacy, Family and Correspondence and Protection of Honor and Reputation.

The field of privacy can be split into two parts: relation privacy and informational privacy. The relation privacy refers to the relation that an individual has with others such as who can enter in one`s house or who “is allowed to touch one`s body” (Holvast 2009: 16). The informational privacy is related to “the collection, storing and processing of (personal) data.” (ibid) In this thesis the focus will be on the informational privacy.

Alan Westin established a description for the functions of privacy that can be implemented in our days. He stated that 4 of these functions are: “The need for a personal autonomy; privacy as a form of emotional release; self-evaluation and decision making and the need for a limited and protected communication” (Holvast 2009: 17). With the need for a protected communication

Westin argues that, “the value of privacy recognizes that individuals require opportunities to share confidences with their family, friends and close associates” (ibid). Here we can say that, privacy gives an individual the opportunity to express his thoughts without any fear.

The term privacy as defined by the Macmillan dictionary refers to “the freedom to do things without other people watching you or knowing what are you doing” (Macmillan English Dictionary). According to the Cambridge Dictionary, the noun privacy means “someone`s right to keep their personal matters and relationships secret” (Cambridge Dictionary).

The term privacy has evolved throughout time containing different meanings and it was adapted to the ages representative icons. To have a better understanding of what privacy represents in our days, we should focus on the relation between privacy and the digital age.

The rapid development of the technology enables humans all over the world to use advanced information and communications devices for a better life. At the same time, “technology is enhancing the capacity of governments, companies and individuals to undertake surveillance, interception and data collection” (Brown 2016) which is a direct abuse against human rights, more exactly against the right for privacy.

On 18 December 2013, The United Nations General Assembly, adopted the resolution 68/167. The resolution is titled the Right to Privacy in the Digital Age. In this resolution, the United Nations reaffirmed the resolution from 1948 and added that they are:

> “Emphasizing that unlawful or arbitrary surveillance and/or interception of communications, as well as unlawful or arbitrary collection of personal data, as intrusive acts, violate the rights to privacy and to freedom of expression and may contradict the tenets of a democratic society” (UN General Assembly 2014: 2) and that they are:
>
> “Deeply concerned at the negative impact that surveillance and/or interception of communications, including extraterritorial surveillance and/or interception of communications, as well as collection of personal data, in particular when carried out on a mass scale, may have in the exercise and enjoyment of human rights” (ibid).

This resolution expresses concern related to the personal data collection, the interception of the communications and points out the mass surveillance deployed around the world, which can be called as a daily issue. It calls for the “strengthening of prevention of and protection against such violations” (Brown 2016).

Nowadays, more than 100 countries around the world have “adopted comprehensive data protection/privacy laws to protect personal data held by private bodies” (Banisar 2016). These laws are being applied in relation to “personal information held in both electronic and physical form” (ibid) held by governments or private companies. These laws, that should be respected by both the public and private organizations involved in the collection of personal data, are designed around some basic principles. These principles state that:

“There should be limits to what is collected” (Privacy International- What is Data Protection?). Limits should be created regarding the collection of personal data. Also, the collection of personal data should be “obtained by lawful and fair means, with the knowledge or consent of the individual” (ibid).

“There must be no secret purposes” (ibid). The collection of personal data has to be done with a specific purpose. The collected data should be used only for the specified purpose.

“The information should be correct” (ibid). The personal data collected has to be up to date, complete and accurate. Also, it has to be relevant for the specified purpose.

“There must be no creeping purposes. Personal information can only be disclosed, used, or retained for only the original purposes, except with the consent of the individual or under law” (ibid). Moreover, the data should be destroyed when it is not needed anymore for the purpose that it was collected for.

“No secret organization, sources or processing” (ibid). Every individual should know about the collection, the use and the purpose of the collection of his data. Moreover, everyone who has his data collected has to know about the organization that is the “data controller” (ibid).

“Individuals have rights to be involved” (ibid). Everyone has the right to access his information, and should be able to challenge the information and to “seek its deletion, rectification, completion or modification” (ibid).

“Organizations must be held to account” (ibid). All of the organizations involved in the process of the collection of personal data, must be accountable for respecting the principles inside the laws.

Most of the countries that have adopted the laws based on these principles, have established data protection commissions to “oversee and enforce the laws” (Banisar 2016). The capabilities of these commissions but also their independence from Government differs from country to country. These commissions can conduct investigations, “act on complaints and impose fines when they discover that an organization has broken the law” (Privacy International- What is Data Protection?).

Although, there exist regulations and declarations regarding the individual right to Privacy, and norms and guidelines regarding the protection of the personal data, these norms and regulations are considered to be violated on a daily basis. In order to analyze the purpose behind these violations, I will use the Panopticon theory framed by Jeremy Bentham and modernized by Michel Foucault.

Jeremy Bentham was born in London, on 15 February 1748. He was a philosopher and a jurist who developed the idea of Panopticon. The term Panopticon comes from the Greek language, and it is defined as the ‘all-seeing’. The idea of an inspection-house or Panopticon was developed by Bentham after a trip to his brother, Samuel, who was working in Russia for “Prince Potemkin” (UCL -The Panopticon). Samuel developed a method which enabled him to supervise his unskilled workers. He placed his desk in the middle of the factory and arranged the workers in a circle around his desk. By doing this, he could supervise all the workers.

Jeremy took the principle of Samuel`s idea and applied it to design a prison in which “a number of persons are meant to be kept under inspection” (Foucault 1995: 204). The proposed prison was “a circular building, with the prisoners` cells arranged around the outer wall and the central point dominated by an inspection tower. From this building, the prison`s inspector could look into the cells at any time, but the inmates themselves would never be able to see the inspector himself” (UCL- The Panopticon).

The watch tower is emitting a bright lighting and because of it, the inmates could not figure out if they are being observed and when. As a consequence, they were living under the constant idea that they are always watched.

The idea is to assume that an all-powerful inspector is standing in an inspection tower which is placed strategically to always watch the inmates. “Bentham expected that this new mode of obtaining power of mind over mind” (ibid) will make certain the fact that the prisoners will change their behavior and moreover, they will work harder, so as not to get punished. The idea of constantly being under surveillance is unpleasant, but even though it is unpleasant, Bentham argued that “the Panopticon and its central inspection principle” (ibid) have its benefits such as “morals reformed, health preserved, and all by a simple idea in Architecture” (ibid).

Michel Foucault, who was born on 15 October 1926 in Poitiers, France was a philosopher, social theorist and literary critic. In 1975, he wrote the book “Discipline and Punish” in which he modernized the idea of Panopticon and transformed it into a theory of power. He used it as a way to explain the predisposition of societies to subjugate its citizens. Foucault explains that the prisoner of Panopticon who is seen by the observer but who cannot see the observer, is as he stated when framing the theory, the: “object of information, never a subject in communication” (Foucault 1995: 200).

Foucault is one of the social theorists who always believed that knowledge is a form a power. Furthermore, he explained that knowledge can be obtained from power by “producing it, not preventing it” (Mason). According to the Panopticon theory, the power was in the hands of the man who was sitting in the center of the prison, the one who was watching the prisoners. In his own view, knowledge means power and knowledge is obtained through observation.

The Panopticon theory “functions as a kind of laboratory of power” (Foucault 1995: 204). Due to its capabilities of observation, it is able to achieve such an efficiency that it makes it able to invade someone`s privacy. Therefore, the idea is to discover “new object of knowledge over all the surfaces” (ibid) on which power could be applied.

According to Foucault, the theory of Panopticon has to be understood as a “generalizable model of functioning” (Foucault 1995: 205). In other words, it is a method of defining power with regard to everyday life of individuals. In his own view, the Panopticon can be applied to more than just one situation. For instance, it can be used to “supervise workers, to reform prisoners or to instruct school children” (ibid). Whatever the situation is that needs the application of the Panopticon theory, it makes it possible to exercise the power. Furthermore, this mechanism that is applied straight on individuals, it provides the idea stated by Bentham of “power of mind over mind” (Foucault 1995: 206).

The Panopticon theory, represents a way of achieving from the power obtained through knowledge a new instrument that can be used by the government or companies. The political sphere is not the only field where it can be used but it can also be used in other fields such as education, medical treatment or production. This mechanism works as a way of making the relations of power to act as a function and it further makes this function to work through these power relations. That is to say, the Panopticon organizes things in such a manner that the power which is exercised from the outside it is not felt like a heavy constraint “to the function it invests” (ibid) but, instead, it is felt as it is slightly present into the functions in order to increase their efficiency.

Using Bentham`s Panopticon idea together with Foucault`s modernized theory of Panopticon, into our days, the circular building and the observer in the middle are the methods of surveillance and the data collection methods are the functions which provide power through knowledge. The prisoners who were being watched are the citizens whose data is being collected and whose human rights are being abused. In Foucault`s words, the citizens from nowadays, represent “the object of information” (Foucault 1995: 200).

3.2 Mass Surveillance

In this subchapter, I will provide information about how the evolution of Internet and Technology have put their mark on the field of surveillance by providing the necessary means for a new type of surveillance, namely mass surveillance. Moreover, I will present how the process of mass surveillance is being conducted by introducing the Data Collection, Big Data and Data Mining subchapters. In the end, I will present the theory that seemed to be the most suitable for the presented information.

3.2.1 *Internet and Technology*

“On October 4, 1957, the Soviet Union launched the world`s first manmade satellite into orbit. The satellite known as Sputnik, did not do much: It tumbled aimlessly around in outer space, sending blips and bleeps form its radio transmitters as it circled the Earth.” (History: The Invention of the Internet)

The launch of the Sputnik satellite triggered alarms in the United States. They have realized that their scientists and engineers were designing cars and TV sets when their focus should be in other different fields like science and technology. So, they have started to invest in “scientific research and development” (ibid). “The Federal Government itself formed new agencies such as the National Aeronautics and Space Administration (NASA) and the Department of Defense`s Advanced Research Project Agency (ARPA), to develop space-age technologies such as rockets, weapons and computers” (ibid).

Being in the ages of the Cold war, it meant that everybody was living under a constant fear of a different kind of attack. United States` scientists and military experts were afraid that the national telephone system, that made the long-distance communication possible could be destroyed with a single rocket. J.C.R Licklider a scientist from MIT who was working for ARPA proposed a solution to this problem in 1962: “a galactic network of computers that could talk to one another” (ibid).

Packet Switching, that was proposed in 1965 by a MIT student was the way of sending information between computers. “Packet switching breaks data into blocks, or packets, before sending it to its destination" (ibid). The packet switching idea enabled the creation of Arpanet which in 1969 delivered the first message between two computers.

While in the end of the 1969 there were just four computers connected to the Arpanet, in the end of 1979 the number of computers rose significantly. This means that, the number of computer networks rose too, but it was hard to have all of them into a single network and this represented a real problem. This problem was solved when a computer scientist, Vinton Cerf, developed the Transmission Control Protocol and upgraded it with a supplementary protocol called Internet Protocol. What this means, is that, all the computers connected on networks can communicate with each other. One writer describes Cerf`s protocol as "the handshake that introduces distant and different computers to each other in a virtual space" (ibid). These protocols created the worldwide network.

In 1991, Tim Berners-Lee who was a computer programmer in Switzerland modified the meaning of the internet of that time, from a way of sending information between two computers, to the internet of our days, by introducing the World Wide Web. This means that, the Internet evolved to "a web of information that anyone could retrieve" (ibid).

From the introduction of the World Wide Web, to our days, we are the witnesses of one of the fastest and significant evolutions known by the mankind, namely the Internet. We have seen that once with the evolution of the Internet, the Technology industry started to evolve significantly. Nowadays, due to the evolution of the Internet and Technology, the market is full of devices that have an Internet connection such as: laptops, smartphones, smartwatches, tablets, smart TV`s and so on.

These devices are supposed to make our life better and easier. With them, we can take pictures using the camera in order to immortalize a nice moment in our lives, we can search online for the information we need, we can listen to music, we can navigate to a place easier by using the GPS and the maps application, we can speak with our families or friends by calling them or we can use free applications to text, call or video call such as WhatsApp, Viber or Skype. We create our contact list and save it in the memory of the device in order to reach easier a person we want to communicate with. Basically, big parts of our life are organized and stored with and in these devices. Essential information such as: what our interests are, are associated with the things we search for online, memories that we want to keep are associated with our photos and notes, also the persons who we get in touch with are associated with our contact lists.

Edward Snowden stated: “every part of a private life, today, is found on someone’s phone. We used to say a man`s home is his castle; Today, a man`s phone is his castle” (Edward Snowden interviewed by Shane Smith for Vice on HBO). One would like to keep all this information for himself. Now, just imagine for one second, that all the information you have in these devices for which you have paid, is no longer only yours, private. Someone else can see your memories, the list of persons who you get in touch with, your texts, your fields of interests, the places you have visited, where you are, your schedule and so on. How would you feel about that? Would you have a strange feeling that you have been robbed by someone who is invisible to you but somehow steals something that means a lot to you? Would you feel like your privacy has been violated?

3.2.2 *Surveillance*

“We live in a golden age of convenience enabled by technology. So, that means that, you and I can be on other sides of the planet and we can have a conversation in real time for no money. Technology has enabled convenience of communication, but also convenience of surveillance.” (Edward Snowden interviewed by Shane Smith for Vice on HBO)

Surveillance has the potential to invade privacy. Privacy is an important condition for an individual right of freedom and by its invasion the democracy system of our days is being undermined.

“The English noun surveillance comes from the French verb *surveillir*. It is related to the Latin term *vigilare* which means to keep watch, with its hint that something vaguely sinister or threatening lurks beyond the watchtower and town walls”. (Marx 2016: 15) One might ask what does surveillance mean. According to the Macmillan Dictionary it represents the process of carefully watching a person or a place that is involved in a criminal activity, but other thesaurus and dictionaries define surveillance as the activity of: looking, observing, watching, supervising, controlling, gazing, viewing, examining, scanning, monitoring, tracking, following and spying (ibid).

The field of surveillance is split into two broad types of surveillance: Targeted surveillance and Mass surveillance. The targeted surveillance can be described as the traditional process of watching a suspected person who is involved in a crime, which represents a direct action against an individual that involves the use of specific powers and can be carried out “overtly or covertly” (House of Lords 2009:12) with the implication of human agents. The methods used to keep under observation a suspect in the targeted surveillance are: “the interception of communications and visual surveillance devices” (ibid).

The improvement of writing, language and the appearance of “more different forms of social organization involving larger political entities” (Marx 2016:17) led to the creation of more and more complex methods of surveillance that are developed through technology. These methods of surveillance are making the traditional targeted surveillance definition to be outdated, because in the new mass surveillance, watching a suspected person means watching “everyone in a given group” (ibid).

Mass surveillance it is further known as “passive or undirected” (House of Lords 2009:12) surveillance. It is not concentrated on the observation of an individual, but it is designed to collect “images and information for possible future use” (ibid). It depends on the evolution of technology and it is deployed through the use of devices that were created for and adapted to this purpose. The methods of mass surveillance are: Video surveillance and “Dataveillance” (ibid)

Video surveillance is being deployed with the use of closed-circuit television cameras (CCTV) that are using the facial recognition software which is able to recognize and track an individual’s movement. Video surveillance is also being deployed through the automated license plate reader system (ALPR), which uses cameras that are placed on roads and highways to track the movement of vehicles.

In the online world, every individual “leaves electronic footprints behind with the click of mouse, making a phone call, paying with payment card, using joined up government services, searching online” (House of Lords 2009: 15) which are considered to be personal data. “The combined term dataveillance covers the systematic use of personal data systems in the investigation or monitoring of the actions or communications of one or more persons” (House of Lords 2009: 12). All the devices that we are using are enabling the Online Surveillance. The “Dataveillance” (ibid) is deployed through different software`s and devices. Devices such as Stingray`s and IMSI catchers and Software`s such as PRISM and XKeyscore. “Dataveillance” (ibid) is based on data collection.

3.2.3 *Data collection*

"The King no longer has to come to data. The data come to him". (Marx 2016:48)

The new modern surveillance can be characterized as the extraction of information known as data, through the use of technology from different persons or groups of persons. It involves an agent (most often someone who works for intelligence agencies of police forces) who opens the doors of someone`s personal data, by extracting and peeking into it with the use of technology. In this case, the agent is seen as the watcher, and the person that has its personal data collected is seen as the surveillance subject.

The extracted data can reveal personal information of the subject such as: real name, nicknames, passwords, gender, age, education, occupation, memberships, e-mail and contacts. Also, it can reveal information about location such as: the places he visited, residence address and for how long the subject stayed in a place. It can provide as well business information, communication information, information about the economic behavior, beliefs and interests. This information is easily connectable to the subject and can answer to questions such as who and where.

Once with the evolution of the technology, we have started to spend more and more time in our digital life. We were accustomed with the landline telephones after which, we have started to use the desktop computers. Due to the technological evolution, we have entered into a portable era which implies the use of portable computers (laptops) and portable phones (mobile phones). After this, we have discovered that it is far more easily and also more convenient to carry in our pockets a combination between a computer and a mobile phone, which is the smartphone.

According to PCmag a specialized magazine, a smartphone is "a combination cellphone and handheld computer that created the greatest tech revolution since the Internet. A smartphone can do everything a personal computer can do, and because of its mobility, much more" (PCMag). It also states that a smartphone combines a "cellphone with e-mail and Web, music and movie player, camera and video recorder, GPS navigation, voice dictation for messaging and a voice search for asking questions about anything. A lot more personal than a personal computer, a smartphone is generally within reach no matter where you are." (ibid)

Due to its capabilities and to the portability that it has, we have it in our pockets or around us, 24 hours per day. This means that, we spend a lot of time on a daily basis interacting with it by speaking with our family or friends, or by searching online for something that we want to know or by taking pictures and so on. The amount of data that we have created and stored on these devices represents a big part of our life.

Every smartphone has what is called an IMSI. IMSI stands for International Mobile Subscriber Identity and it is a unique number that it is "associated with the Global System for Mobile Communications (GSM) and with the Universal Mobile Telecommunications System (UMTS). The IMSI is a unique number identifying a GSM subscriber" (Techopedia- International Mobile Subscriber Identity). The IMSI has three components such as: a mobile country code and a mobile network code which are the location area identifier and a mobile subscriber identification number which "identifies the mobile subscriber and is assigned by the operator" (ibid). All the IMSI details are stored in the SIM, inside the phone and are sent by the phone to the nearest network. "When a mobile is affiliated, a temporary IMSI is allocated and used to identify the subscriber in future exchanges. This is embedded in the SIM of the mobile equipment and is provided anytime the network is accessed" (ibid).

In order to have access to the network carrier, the smartphone connects to the nearest cell phone tower in the area. In this way, you can make a call or send a text message or use the internet connection. Once with the evolution of the technology not only the computers and portable devices have evolved, but also the tools of surveillance. The appearance of the smartphone led to the appearance of IMSI catcher`s also called Stingray.

IMSI catchers are also known as Stingrays. According to the American Civil Liberties Union they are" invasive cell phone surveillance devices that mimic cell phone towers and send out signals to trick cell phones in the area into transmitting their locations and identifying information. When used to track a suspect`s cell phone, they also gather information about the phones of countless bystanders who happen to be nearby" (ACLU). According to the Guardian "they are the size of a suitcase, they work by pretending to be a cellphone tower in order to strip metadata" (Woolf 2016). Furthermore, these "devices are also capable of listening to phone calls" (ibid).

One might ask what metadata is and what can these devices take from your smartphone. According to Edward Snowden these devices can take a lot of information from a smartphone. "Everything in your contacts list, every SMS messenger that you use, every place that you were, the place where the phone is physically located. Even if you`ve got GPS disabled, because they can see which wireless access points are near you" (Edward Snowden interviewed by Shane Smith for Vice on HBO).

Metadata is formed by descriptive metadata and structural metadata. Descriptive metadata relates to piece of information and gives you actual details about it. For example, for a photo that was taken with a device, the descriptive metadata will tell the location where the picture was taken, the date and time when the picture was taken and the exposure time of the camera. Structural metadata relates to many things and tells you how similar bits of information are stored and what they mean. For example, for the same picture the structural metadata will stay the same but it will help you to organize and interpret data by describing what the fields mean. According to Edward Snowden, "metadata is the fact that the communication occurred" (ibid).

The IMSI catchers or Stingray`s are not targeting only a suspect`s phone, but instead they are taking data from everybody with a smartphone in a specific area. What this means is that, if you are walking, driving or living in a specific area where the police forces or the intelligence agencies are placing these devices, the data that it is in your smartphone will be extracted.

One might say that, in our days the use of these devices is certainly a method used by the intelligence agencies and the police forces against civilians but, as these devices are not projected to catch the data of a suspected person in a geographic area, they are instead collecting the data from the smartphones of all the innocent persons that are in that area.

Another way to collect the content of a smartphone is through a software developed and used by the United States National Security Agency called "Dishfire" (Ball 2014). The NSA program, which the agency refers to it as the "SMS Text Messages: A Goldmine to Exploit" (ibid) has the capability to "collect pretty much everything it can" (ibid). According to The Guardian, by using this program, the NSA "collected almost 200 million text messages a day from across the globe, using them to extract data including location, contact networks and credit card details" (ibid).

Individual`s data with no suspicion of any illegal activity has been collected and stored. The data collected through the "Dishfire" (ibid) program can reveal the text messages sent and received, missed calls alerts, details of border crossing, the electronic business cards and financial details. The NSA allowed the British intelligence and security organization GCHQ to access the servers of the collected data in order to search for the UK phone numbers.

A representative of the British GCHQ stated that the program contains a "large volume of unselected SMS traffic" (ibid) this making it "particularly useful for the development of new targets, since it is possible to examine the content of messages sent months or even years before the target was known to be of interest". (ibid)

According to The Guardian, NSA developed another program called "PRISM" (Greenwald, MacAskill 2013) with which they have obtained direct access to the servers of the most important U.S based IT companies. The Washington Post newspaper obtained a document whose authenticity was verified by The Guardian too. The document containing a "41-slide PowerPoint presentation – classified as top secret with no distribution to foreign allies- which was apparently used to train operatives on the capabilities of the program. The document claims collection directly from the servers" (ibid) of the companies.

The PRISM program was launched in 2007 and started to collect data from the Microsoft servers. The servers of Yahoo were added in 2008 while in 2009 were added the servers of Google, Facebook and PalTalk on the PRISM data collection list. In 2010 and 2011 were added the servers of YouTube, Skype and AOL respectively. The next year, namely 2012, meant the addition of Apple`s servers into the PRISM program.

For the users of Microsoft, Yahoo, Google, Facebook, PalTalk, YouTube, AOL and Apple it meant that their data was no longer private. This meant that their search history, file transfer, videos, audio conversation, along with their "video chats, photographs, e-mails, documents as well as connection logs" (Gellman, Poitras 2013) were collected by the NSA. This program allows NSA to collect the communications in which they are interested in without requesting them from the service providers and without any court orders.

One of the “widest reaching system developing intelligence from computer networks” (Greenwald 2013) is called XKeyscore. The software which is one of the NSA`s most important tools was leaked by Edward Snowden in the documents that he released. The XKeyscore software is capable of collecting data in real time and to automatically store the collected data. The XKeyscore database is fed with “a constant flow of internet traffic from fiber optic cables that make up the back of the world`s communication network, among other sources, for processing” (Weissman 2015).

Also, this tool gathers data such as: Skype sessions, file uploads, “pictures, documents, voice calls, webcam photos, web searches, intercepted username and password, e-mail and social media traffic” (ibid). XKeyscore can be used to track an individual`s online activity, to monitor who access a specific webpage, to search in documents for a specific language which is not used commonly in a country, to search for people who are using VPN services and so on.

I will provide a brief understanding of one of the NSA`s most important tools, namely XKeyscore with all of its attributions in the Analysis chapter.

3.2.4 *Big Data*

The quantity of computer generated data is growing with each passing day. Every business, every field of activity and even almost each of us uses technologized devices and creates data. In order to improve their work, "retailers are building vast data bases of recorded customer activity, organizations working in logistics, financial, health care and other sectors are also capturing more data. Even public social media is creating vast quantities of digital material" (Banerjee et al., 2015: 591). As more smart devices that are using an internet connection are launched every day, big data is being expanded by the "Internet of things" (Meola 2016).

Big data is the term that characterizes, the large amount of data that is produced on a daily basis. It is the act of collecting and storing large data sets from traditional and digital sources to identify trends and patterns. Companies are using this information in order to improve what they know about their customers; more exactly about what customers want and need.

Big data can be classified into three categories: structured data, semi-structured data and unstructured data. Structured data is the data that has a proper format associated to it. It resides in a specific field in a record or file. For example, it is the data that is present in the databases, store receipts or web traffic. Structured data "has the advantage of being easily entered, stored, queried and analyzed" (Beal). The semi-structured data is the data that does not have a proper format associated with it. It is the data that is present in the e-mails, world documents or in the log files. The unstructured data is the data that does not have any format associated with it. It cannot "be so readily classified" (ibid) and requires special software to organize and understand it. For example, the image files, the audio files and the video files.

Big data is characterized with volume, velocity and variety. Volume refers to the "sheer size of the datasets" (Harris 2012) or to the amount of data. Volume represents the greatest challenge. Velocity refers to the speed at which the data is generated, acquired and used. It represents an issue because of the rate with which the data is flowing into the servers. The last type, variety, refers "to the different types of data that are available to collect and analyze in addition to the structured data found in a typical database" (ibid). There have been identified four categories of information that establish big data:

“1. Machine-generated data. This includes RFID data, geolocation data from mobile devices and data from monitoring devices such as utility meters.

2. Computer log data, such as clickstreams from websites.

3. Textual social media information from sources such as Twitter and Facebook.

4. Multimedia social and other information from Flickr, YouTube and other similar sites.” (ibid)

After we have seen how the evolution of the internet and technology enabled us to use the portable computers and the smart devices, how the surveillance evolved with the use of technologized devices, how the user`s data is being collected through different methods, and how these large data sets are being stored in data warehouses, it is the time to see that all of the above steps are connected and that they provide the analysts with the possibility to analyze the collected data through the process called data mining.

3.2.5 *Data mining*

Data mining can be pictured as the "result of the natural evolution of information technology" (Han, Kamber 2006:1). It is the process of "analyzing data from different perspectives and summarizing it into useful information" (Bornschlegl et al., 2016). It assumes the action of searching and analyzing big data sets from different perspectives and extracting useful information. Data mining is defined as the process "of extracting data, analyzing it from many dimensions or perspectives, then producing a summary of the information in a useful form that identifies relationship within the data" (Greiner 2011) It is also called "knowledge discovery" (Techopedia- Knowledge Discovery in Databases).

Data mining can be divided into two categories: Descriptive data mining and Predictive data mining. Descriptive data mining defines the general properties of the data that is held in the databases. Descriptive data mining is formed by different functions such as: "Class/ Concept description; Mining of Frequent Patterns; Mining of Associations; Mining of Correlations; Mining of Clusters" (Tutorialspoint 2014: 3). Predictive data mining is the mining of the data with the desire of "using business intelligence or other data to forecast or predict trends" (Techopedia- Predictive Data Mining). Predictive data mining analyzes the data with the aim of making predictions.

Data mining implies using various types of software`s like the analytics tools. Among the used software`s we can find: RapidMiner, WEKA, ORANGE, KNIME, NLTK and R-Programming. It can be both automated and "labor-intensive" (Techopedia- What is the difference between big data and data mining?). To go further, one software of data mining is able to search "through dozens of years of accounting information" (ibid) in order to find a specific information related to expenses or accounts for a specific period of time.

The process through which data mining is obtaining knowledge from the available data is done through specialized software. The process starts with the raw data which, through selection becomes target data. In the second step, the target data becomes preprocessed data. Through the process of transformation, the data is normalized and becomes transformed data which represents the third step. By applying data mining algorithms on the transformed data, the data will be divided into diverse patterns. In the end, the knowledge is obtained by interpreting the patterns.

Data mining is used for different fields, for example in the fields of market in order to analyze it. In other words, data mining is used by the managers from the fields of market in order to find out “what kind of people buy what kind of products” (Tutorialspoint 2014: 2), this process being known as the Customer Profiling. It further helps the companies to determine which are the most suitable products for the different customers who usually buy from their shop. That is to say, through data mining the companies can learn about the factors that will attract customers, this process being called “Identifying Customer Requirements” (ibid). Furthermore, in the process of target marketing, data mining is being used to “find clusters of model customers who share the same characteristics such as interests, spending, habits, income etc.” (ibid). Not only that it helps to find clusters of model customers, but it also helps to create a pattern of every customer purchasing habits.

By using the process of data mining, the raw data from the large data sets becomes knowledge. We have seen in the market example, how the discovered knowledge can be used. There are a lot of fields that can be improved by mining the data related to them, but one might ask, what knowledge can be achieved after mining one`s personal data collected through different techniques?

In the digital era, different types of data are being created every minute. Different methods of data collection such as devices like Stingrays or IMSI catcher`s or specialized software`s such as Dishfire, XKeyscore or PRISM are used on a daily basis. All the collected data is stored and forms what is called, the large data sets. Following the process of mining the collected data, all the information can be put together through the Mosaic theory by the specialized analysts.

The mosaic theory was first used in the stock market. It is the method used by analysts to gather information about a company. It involves collecting non-material information, public and non-public information about a company in order to determine the “underlying value of the company`s securities” (Investopedia). This method allows the analysts to make recommendations to clients based on the information they have gathered about the company they were looking for.

After it was applied to the stock market, the theory was proposed to other fields such as intelligence agencies. The mosaic theory describes a basic precept of intelligence gathering: “Disparate items of information, though individually of limited or no utility to their possessor, can take on added significance when combined with other items of information. Combining the items illuminates their interrelationships and breeds analytic synergies, so that the resulting mosaic of information is worth more than the sum of its parts" (Pozen 2005: 630).

The Department of the Navy of The United States of America, defines the theory in its Freedom of Information Act regulations as "the concept that, apparently harmless pieces of information when assembled together could reveal a damaging picture" (ibid).

The framework of theory was used in a court too, in the United States. A suspected drug trafficker had his car bugged with a GPS surveillance device for 28 days. "Whenever the car was in motion, the GPS device used cell phone technology to broadcast signals of the car`s location to a government computer every seven seconds. The device produced over 2000 pages of location data over twenty-eight days". (Kerr 2012: 323). The assigned judge of the case was Mr. Ginsburg who stated that: "The collective sum of twenty-eight days of surveillance revealed more than the sum of its parts" (Kerr 2012: 325). The judge referred to the location collection as a non-search and stated that: "Many non-searches packaged together as a collective entity became a search because the individual pieces of the puzzle that seemed small in isolation could be assembled together like a mosaic to reveal the full picture of a person`s life" (ibid).

Here we can say that, by collecting small parts of information and by putting them together, it creates a mosaic that can be seen as a picture of someone`s life. "These types of information can each reveal more about a person than does any trip viewed in isolation. Repeated visits to a church, a gym, a bar tell a story not told by any single visit" (Kerr 2012: 326). We can say that, if a person goes to a church constantly he can be profiled as being religious. "The sequences of a person`s movements can reveal still more; a single trip to a gynecologist`s office tells little about a woman, but that trip followed a week later by a visit to a baby supply store tells a different story" (ibid).

4. **Analysis**

The main chapter of the Analysis is called Superpanopticon, which is formed by four subchapters and analyzes the steps of the mass surveillance process and the achieved result. The first subchapter is also divided into two subchapters that are analyzing the situation of the Data Collection. The second subchapter analyzes the process of Data Mining and the results obtained through the process. The third subchapter is the Individual Profiling, which is the result of the Superpanopticon model. What makes the Individual Profiling so controversial is analyzed in the fourth subchapter using the case study of the Strategic Communication Laboratories.

4.1 Superpanopticon

One of the oldest forms of surveillance has its roots in the birth of the nation-state and it can be seen as data collection, but it is based on the old way of documenting. According to Anthony Giddens, the nation-state kept official statistics that included: "births, marriages and deaths; […] moral statistics relating to suicide, delinquency, divorce and so on" (Allmer 2011: 5).

One of the of oldest ways of surveilling individuals was developed by Jeremy Bentham. His idea was to build a round building, which was divided in cells. In the middle of the building, Bentham placed a huge tower which was emitting light. A supervisor was standing in the tower while the prisoners were standing in the cells and because of the light the prisoners were unable to see inside the tower. As a result of the great architecture, the supervisor was able to see every prisoner without being seen and the observation was possible at any time, thus creating for the prisoners the feeling that they were under surveillance all the time. The principle of the architecture was described by Bentham as the Panopticon. Through the Panopticon, Bentham has created among the prisoners the feeling of being surveilled all the time with the aim that they will "discipline themselves out of fear of surveillance" (Allmer 2011: 3).

For Michel Foucault, the Panopticon architectural concept represents a symbol for the "modern disciplinary society" (ibid). In the book Discipline and Punish, were he transformed the architectural designed framed by Jeremy Bentham into a theory of power, Foucault states that: "On the whole, therefore, one can speak of the formation of a disciplinary society in this movement that stretches from the enclosed discipline, a sort of social quarantine, to an indefinitely generalizable mechanism of Panopticism" (ibid). In Foucault`s interpretation, the Panopticon theory "creates a consciousness of permanent visibility as a form of power" (ibid) where the old tolls of domination such as handcuffs and chains are not needed any more. Foucault evaluates surveillance in the perspective of the evolution of the modern disciplinary societies where he sees discipline as a "form of operational power relations and technologies of domination in order to discipline, control and normalize people" (ibid). In his own view, the symbol of the modern surveillance societies is represented by the Panopticon. According to Foucault`s theory of Panopticon, the prisoner of Panopticon who is seen by the observer but who cannot see the observer represents the: "object of information, never a subject in communication" (Foucault 1995: 200). Furthermore, according to the theory, the power was in the hands, of the one who was watching the prisoners. In Foucault`s view, knowledge means power and knowledge is obtained through observation.

With the evolution of the Internet and technology, the methods of surveillance have evolved too, through the use of different devices and software`s which were created to accomplish this purpose. Mark Poster, who was a Professor Emeritus of History explains according to Foucault`s tradition that, surveillance nowadays is a "major form of power in the mode of information" (Allmer2011: 11). According to him, the evolution of technology has caused not only new ways of surveillance, but also "an electronic Superpanopticon mode of information" (ibid). As he describes, the Superpanopticon, represents a process in which the mases are normalized and controlled. "Today`s circuits of communication and the databases they generate, constitute a Superpanopticon, a system of surveillance without walls, windows towers or guards" (Poster cited in Allmer2011: 11). He emphasizes the fact that the evolution of the information and the communication technologies have also led to the evolution of new methods of surveillance and as a result, it led to the creation of new forms of power.

Superpanopticon can be described as an advanced technology which implicates the process of collecting, processing and sharing of personal data about people and groups of people, data which is generated on a daily basis from their regular activities. That is to say, societies, nowadays have been transformed into "data, markets, and samples" (ibid). Moreover, Haggerty and Ericson pointed out that mass surveillance is able to put "different systems, practices and technologies together into a larger whole" (Allmer2011: 8). In this way, mass surveillance is seen as an assemblage.

Here we can say that, the complex process which generates power for those who own it, namely mass surveillance, can be seen as a Panopticon system, in which the prisoners from the old Panopticon, are represented by the nowadays citizens whose data is being collected. Moreover, with the continuous evolution of the personal devices, mass surveillance has adapted itself to the daily requirements and has become a super electronic system of surveillance, which is seen as a Superpanopticon. What is astonishing is the fact that, the Superpanopticon operates as an assemblage, formed by different strategies and techniques that are deployed step by step, and each deployed step enables the possibility for the next step to occur. By following and maintaining a chain of evidence for each step, in the end, the assemblage reveals key information about different individuals. The first step of this Superpanopticon model of mass surveillance is the data collection.

4.1.1 *Data collection*

Data is being created everywhere around the world, with the use of different smart devices. Data is being created when an individual who walks with a smartphone in his pocket, and a specialized software is tracking the number of steps, when making a phone call, when searching for something online or when accomplishing tasks that are regular on a daily basis. The amount of data created is huge and it covers an enormous number of fields. Data collection is the process of gathering the data through different techniques that include the use of specialized software`s and specialized devices. The process of data collection can be split into two categories:

- The data that the individuals provide being conscious or unconscious when using social media or installing different applications on their devices. This type of data is available at any time for collection and is called user generated data.
- The data that is being extracted by the authorities or private companies through specialized software`s and devices. This process can be seen as an intrusion of privacy or as a theft and it can be called data collection without user`s agreement.

4.1.1.1 *User generated data*

The first category to be analyzed is the category in which individuals are giving away data about themselves, by using social media or installing different applications on their devices. The most used social media networks are Facebook and Twitter. According to a website, specialized in statistics namely Statista, “in the first quarter of 2017, Facebook had 1.94 billion monthly active users” (The Statistics Portal: Number of monthly active Facebook users worldwide as of 1st quarter 2017) while in the same period of time, Twitter had “328 million monthly active users” (The Statistics Portal: Number of monthly active Twitter users worldwide as of 1st quarter 2017).

“If you are not paying for it, you`re not the customer; you`re the product being sold”
(Andrew Lewis)

Every day individuals around the world, while using Facebook are sharing pictures and videos, are updating their activities, are clicking the like button when they see a post they like, a company they prefer, are tagging their friends and are using the messenger service to speak with their families and friends. Moreover, in the description of individuals` profiles on Facebook, sometimes the user fills in his profile with information such as address, education, birthday, phone number and so on. This type of data is called the user generated data. This data can be seen and collected by everyone for any purposes.

In order for the company to work and to offer the user with the necessary means to enjoy his experience when using the provided services, Facebook has to pay for: the maintenance of the servers, employees, network infrastructure and so on. In 2015, Facebook reported "$2.52 billion in capital expenditures on data centers, servers, network infrastructure and office buildings" (Data Center Knowledge). So, what is the trick? Why would a company spend so much money to provide you with a service that you are not paying for? Actually your "private data pays for free Facebook" (Segall 2011).

When you sign up to use the service, you agree to different terms and conditions which are called the Facebook agreement. According to the terms of service, a photo or video posted on Facebook remains the "intellectual property of the user but Facebook`s terms give the company a worldwide, non-exclusive, royalty-free license, with the right to sublicense" (Smith 2013). Basically, Facebook, through its license has the right to use the content that you have uploaded, for example, photos and videos, in any way they want. Furthermore, Facebook has the ability to "transfer or sub-license its rights over a user`s content to another company or organization if needed" (ibid). To go further, if you delete your account it does not mean that Facebook license upon your content has ended. This license ends only when all other users such as people or companies/organizations that have interacted with your content "have also broken their ties with it" (ibid).

Moreover, according to Facebook`s terms, they collect different kinds of information from or about you. They are collecting the "things you do and information you provide" (Facebook). For instance, they retain the information about every time you log in into your account, the content of the messages and the people with whom you communicate with. They are also collecting data about what type of content you view, what is the frequency and the duration of certain activities you do through the platform. Furthermore, you are tracked through other people's movements, as Facebook described it, the "Things others do and information they provide" (ibid) related to you. Facebook collects the data about you, every time an individual shares a photo of you, send you a message or tags you in a photo.

Networks and connections represents another piece of data stored by Facebook. For instance, they collect information about every group you are connected with, and the content of the information that you share with the respective group. Also, if you make a purchase on Facebook, they will retain the information about the payment, such as the details of your card as well as "billing, shipping and contact details" (ibid).

Your devices represent another source of information for Facebook. That is to say, they collect the information about all your devices in which you have installed and accessed Facebook. After collecting the information, they associate the respective data across your devices. Facebook also offers some examples of the information that they collect from the devices: "The operating system, hardware version, device settings, files and software names and types, battery and signal strength, device identifiers, device locations including GPS or Wi-Fi signals and Connection information such as mobile operator, ISP and IP address" (ibid).

Although some digital-agency executives "described Facebook as a black hole" (Turow 2011: 138), apparently more and more individuals are agreeing to be tracked and monitored, to have their personal information and preferences stored and used for different proposes, by others. In the end, what appears to be as a free service, comes with a price. According to Peter Eckersley, who is a senior staff technologist for the Electronic Frontier Foundation, "the cost of being on Facebook is being data-mined" (Segall 2011).

In order for a device to function it has to have an operating system. According to The Verge, "99.6 percent of the new smartphones run Android or iOS" (Vincent 2017). Both operating systems have a built-in store, in the form of an application. For the Apple iOS, the store application is called App Store while for the Google Android OS, the store application is called Play Store. From these stores, individuals can buy different applications or install them for free. These applications that we are installing on our devices have different requirements. This means that, when you press install, you automatically agree with the application permission requirements.

One might think that the application permissions are required in order for the application to function properly. For instance, in the case of WhatsApp, permissions such as Camera, Contacts or Microphone are normally to be asked for, because through the application you can videocall with your contacts and for that to happen, the application needs to have access to those features. But what happens when an individual installs an application that requires for strange permissions?

According to The Guardian, “tens of millions of people have downloaded apps like Super-Bright LED Flashlight or Brightest Flashlight Free without reviewing the permissions” (Fox-Brewster 2014 flashlight). According to the Google Play store statistics, the Brightest Flashlight application has been installed more than 50 million times. This application requires permission to “delete applications, track location, tinker with Bluetooth connectivity, view call details and write additional software to the phone” (ibid). Additionally, it requires permission to the storage where all the files such as documents, pictures, videos of the device are being stored and permission to read Identity.

A security company, namely Bitdefender, which offers security products for different devices, stated that, the Flashlight application should only require a “limited set of permissions to run – for example, around use of a device`s camera, in order to use its flash to provide light” (ibid). As described by The Guardian, with the help of different applications with significant permission requirements that individuals install without reviewing, in our case, the Flashlight, “developers are often asking for far greater power over a user`s device in order to collect data” (ibid). Here, the data that is stored inside the device, is created or saved by the owner of the device, and is considered to be the user`s property. But, from the moment that an individual installs an application like the Flashlight with its permission requirements, the user agrees consciously or unconsciously to the collection of his personal data. With the individual`s agreement, the process of data collection can start and the collected that can be seen as user generated data, because of the user`s agreement to offer access to his information.

4.1.1.2 *Data collection without user`s agreement*

The second category of data collection to be analyzed, is the category in which the authorities or private companies are using software`s or specialized devices to collect data from individuals. The most used tools are the IMSI catcher`s also known as Stingray's or the XKeyscore software developed by the American National Security Agency.

The use of IMSI catchers or Stingray`s are spread across the globe with proofs available from Norway, The United States of America and Canada. This surveillance technique is expanding very fast, as more and more police stations and intelligence agencies are buying these devices.

Aftenposten which is a Norwegian newspaper, conducted an investigation in the Norwegian capital, Oslo in 2014. The Norwegian journalists have used for two months a "German made CryptoPhone 500" (Johansen 2014) to monitor and to "disclose a number of locations in the city with suspicious mobile activity" (ibid). They have actually found so many locations, that they were questioning themselves if their device, the CryptoPhone was working properly. To find the answer, Aftenposten cooperated with two security companies "Aeger Group and CEPIA Technologies" (ibid), companies that have the necessary equipment for this kind of mission. After two weeks of measurements the companies reached to the conclusion that surveillance equipment is used actively in the center of Oslo. The targeted areas were around the Prime Minister Office, around the Ministry of Defense office, the Parliament and in the Embassies area. The manager of CEPIA technologies, Kyre Sletsjoe, who has "a long experience from Norway`s intelligence services" (ibid) stated: "If we had made such findings for a private company, they would prompted a request for the authorities to begin an investigation" (ibid).

In August 2016, The Guardian newspaper published an article about the use of Stingray in The United States of America. According to The Guardian, "court documents ordered released by a judge in Oakland, California, have revealed rare insights into how local police and the FBI use a sophisticated surveillance device known as Stingray" (Woolf 2016). The revealed documents showed that in a search for a suspect called Purvis Ellis, the city police department used "its Stingray device for several hours" (ibid). The documents also revealed that after the local police department failed to locate the suspect, they "asked the FBI for help. The FBI using their own Stingray located Ellis and brought him into custody" (ibid).

The use of these devices is heavily increasing in The United States of America and according to the American Civil Liberties Union, “at least 66 state and federal agencies are now known to use the devices, including the IRS, as well as dozens of state and local police departments” (ibid).

Another country where the authorities officially confirmed that they have used the Stingray devices is Canada. According to the British Columbia Civil Liberties Association (BCCLA), the Vancouver Police Department (VPD) used the surveillance device Stingray, despite the fact that, they do not own any. The Vancouver Police Department borrowed the device from the Royal Canadian Mounted Police (RCMP) and used it in Vancouver. After an exchange of letters between the BCCLA and the VPD, the BCLLA sent a letter where they asked if the VPD used the RCMP`s Stingray and if they would do it again. The answer came shortly: “Yes they have used an RCMP`s Stingray and yes, they would do so again” (Vonn 2016).

Here we can say that, in all of the three cases, in the United States of America, in Canada and in Norway, the data of individuals being in a given geographic area has been extracted from their smartphones. Although that in the case that occurred in the USA the authorities were looking for a specific individual, they have collected and analyzed data from everyone in the area in order to locate the suspect`s phone. As these devices are not designed to target a specific smartphone, they are stripping the data of every smartphone that is connected to the fake cell tower. This technique of data collection can be seen as an invasion of individual`s privacy and as a technique of mass surveillance.

Among Edward Snowden`s reveals, it was a Pdf document containing over 30 pictures about a software called XKeyscore. The capabilities of this software were described in many ways and according to the NSA`s own interpretation is the “widest-reaching system to search through the Internet data” (Franceschi-Bicchierai 2014). As the Intercept magazine describes, this surveillance system “boasted approximately 150 field sites in the United States, Mexico, Brazil, United Kingdom, Spain, Russia, Nigeria, Somalia, Pakistan, Japan, Australia, as well as many other countries, consisting of over 700 servers” (Marquis-Boire et al., 2015).

The servers in the above countries are storing “full-take data” (ibid) from the websites that are accessed in a country. This means that, the servers are capturing all the data that it is transmitted and received in every specific country. The ability of these servers is not only to allow the data traffic flow, but for a short period of time, the used data is stored. As the data continuously flows, the servers are not able to keep the data for a long period, usually the data is saved for “3 to 5 days” (ibid). Moreover, before rewriting the data with new data, the metadata is being stored and kept for “30 to 45 days” (ibid). Moreover, XKeyscore is able to access these servers and to search inside them, as NSA describes: “It is a fully distributed processing and query system that runs on machines around the world” (ibid). With this process, the XKeyscore software collects: “emails, chats and web browsing traffic, pictures, documents, voice calls, webcam photos, web searches, advertising analytics traffic, social media traffic, Skype sessions and more” (ibid).

Furthermore, the capabilities of the XKeyscore software are not only represented by the collection of web traffic. According to an internal document of NSA, namely “VoIP Configuration and Forwarding Read Me” (ibid) that was revealed by Edward Snowden, the software was absorbing data from up to 75 servers about “700.000 voice, fax, video and tag files per day” (ibid).

Today, when an individual or a company wants to keep its files safe, they are using the process of encryption. When a file is encrypted, it automatically creates a key for decryption, thus making the user able to access the file again by using the decryption key. According to the Intercept, in what is described as the “Great sim heist” (Scahill, Begley 2015), the NSA and the British GCHQ, by using the XKeyscore software, have hacked into the “internal computer network of the largest manufacturer of SIM cards in the world” (ibid), namely Gemalto.

This multinational company from the Netherlands is making the “chips used in mobile phones and next-generation credit cards” (ibid). Gemalto produces around 2 billion SIM cards per year. Among its customers there are companies such as Verizon, T-Mobile, AT&T and more than 400 wireless network providers around the world. With the hack, the NSA and the GCHQ managed to steal the “encryption keys used to protect the privacy of cellphone communications across the globe” (ibid). This hack provided the agencies with the potential to “monitor a large portion of the world`s cellular communications, including both voice and data” (ibid). The greatest result of the hack was the fact that the intelligence agencies can monitor the mobile communications without needing approval from the telecommunication companies or from foreign governments.

Moreover, the XKeyscore software is able to show activities and interest of people based on location, website visits or nationality. For instance, in one of the slides that were in the document revealed by Edward Snowden it was presented how the software searched for “germaninpakistan” (Marquis-Boire et al., 2015). The search results were containing details for all “individuals in Pakistan visiting specific German language” (ibid) websites, or chatting in the German language. Furthermore, it also revealed documents that were located in computers in Pakistan, written in the German language. The software has the ability to search in a given area for a specific language and it also has the ability to search for a language that pops out to be unusual in a given area.

Through the XKeyscore software the intelligence agencies are capable of collecting a large amount of data from different sources. When trying to figure out which data was whose, the agencies were facing a problem, the same problem that was faced by the Internet companies which found the solution for the problem. The problem was solved by tracking the internet users with the use of “identifiers that are unique to each individual, often in the form of browser cookies” (ibid). These cookies are sent from a website when an individual access it and stored in the user’s browser and are in the form of small pieces of data. These cookies are used for different purposes such as: “authenticating users (cookies make it possible to log in to websites), storing preferences and uniquely tracking individual even if they`re using the same IP address as many other people” (ibid).

The NSA`s ability to “piggyback off of private companies tracking their own users” (ibid) is a critical tool that enables the agency to track back the collected data to individual users. Once with the establishment of this path, if an individual switch to another Wi-Fi network or uses a VPN service he will still be tracked as long as he is “using the same browser and fail to clear the cookies” (ibid). In the same way, through the XKeyscore software the smartphones are tracked, but with only one the difference: On the smartphones, the cookies are the “unique tracking identifiers that application developers use to track their own users” (ibid). For instance, in May 2015, CBC news together with the Intercept magazine, have revealed that the XKeyscore software was used “to track smartphone connections to the app marketplace run by Samsung and Google” (ibid).

With the use of XKeyscore, the NSA is able collect a large amount of data and through the collected data to track individuals. Moreover, the most astonishing thing is, as Edward Snowden stated: “I, sitting on my desk, I could wiretap anyone, from you or your accountant to a federal judge” (ibid). For certain the use of XKeyscore software is an invasion of individual`s privacy and the capabilities of the software represent the necessary means for the act of mass surveillance.

4.1.2 *Data Mining*

The data collected from different sources and through different techniques is stored and organized in large data files in databases, which coins the term Big Data. Furthermore, by organizing the data in large data files, it enables the possibility for the analysis of the data. The process of analyzing data, through which knowledge is discovered, is referred to as Data Mining, which is the second step of the Superpanopticon.

“Data mining is an automated analysis of data, using mathematical algorithms, in order to find new patterns and relations in data” (Custers 2013: 7). The process of data mining consists in only one step which is called Knowledge Discovery in Databases. This process is seen as the “nontrivial process of identifying valid, novel, potentially useful and ultimately understandable patterns in data” (ibid). The KDD process is formed by five successive steps.

The first step is the data collection. This step has been presented in the above subchapter.

The second step in the KDD process is called data preparation. In this step, the data from the databases is arranged and ordered. In order to obtain a better result from this step, the data should be aggregated. For instance, “zip codes may be aggregated into regions or provinces, ages may be aggregated into five-year categories” (Custers 2013: 8).

The third step of the process is the actual data-mining. In this step, the data is analyzed in order to find “patterns or relations” (Custers 2013: 9) and the process is done through mathematical algorithms. “Data mining is different from traditional database techniques or statistical methods because what is being looked for does not necessarily have to be known” (ibid). Moreover, data mining is used to find new patterns or to confirm “suspected relationships” (ibid).

Although there are different data mining techniques, the most common “discovery algorithms with regard to group profiling are clustering, classification and, to some extent, regression” (ibid). Regression is used to describe data “with a mathematical function” (ibid) while classification map the data into “several predefined classes” (ibid). Moreover, clustering is used to “describe data by forming groups with similar properties” (ibid).

In data mining, a pattern is a “statement that describes relationships in a (sub)set of data such that the statement is simpler than the enumeration of all the facts in the (sub)set of data” (ibid). If a pattern in data is “interesting and certain enough for a user, according to the user`s criteria, it is referred to as knowledge” (ibid). In order for a pattern to be seen as knowledge, a “particular certainty” is required. Certainty can involve “several factors, such as the integrity of the data and the size of the sample” (ibid).

The fourth step in the Knowledge Discovery in Databases process is the interpretation of the results obtained in the third step, the data mining. Many of the results are statistical and they must be transformed into “understandable information such as graphs, tables, or casual relations” (ibid).

The last step of the KDD process “consists of determining corresponding actions” (Custers 2013: 10) and is called “Acting upon Discovered Knowledge” (ibid). This last step is about the decision of the analyst who observes the discovered knowledge and needs to take a decision about how the discovered knowledge will be used.

The KDD process has proved to be very useful by mining the data related to different fields such as market or health. Using the KDD process, different companies were able to increase their revenue, because once their data was mined, it has revealed patterns in the shopping behavior of their customers. Furthermore, the obtained knowledge about a customer’s shopping behavior can be seen as a piece of information that helps the company to increase the revenue by providing the customer with his preferred merchandise. But, as a big company has a lot of customers, the revenue of the company will be for certain increased when the company will be able to generate a common solution for the customers. Furthermore, by using the KDD process the companies were able to find patterns in the shopping behavior of all of their customers. Once the company had pieces of information about every customer, they were able to put all of the discovered knowledge head to head, and through this process they were able to form a bigger picture of their customers` habits, this leading to the increase of their revenue by applying a common solution for their customers.

One example which was presented in the Business News Daily magazine, tells the story of how the KDD process is used. According to this magazine, the example was given by a professor of UCLA to his students. The story is as follows: “A grocery store chain discovered that when men bought diapers on Thursdays and Saturdays, they also tended to buy beer” (Brooks 2014). Furthermore, they have also discovered that these customers were typically doing their weekly shopping “on Saturdays and only bought a few items on Thursdays” (Data Mining). According to the professor, the grocery store chain used the discovered information “in various ways to increase revenue, such as moving the beer display closer to the diaper display and making sure that beer and diapers were sold at full price on those days” (Brooks 2014). This example clearly describes that, by putting together pieces of information about customer`s shopping habits, it reveals a bigger picture of all of the customers` shopping habits, picture that is helping the company to increase its revenue.

Furthermore, the companies are using the Knowledge Discovery in Databases process not only to create a bigger picture of customers` habits from the discovered information, which is seen as general because the provided solution applies to more than one customer, but to create a bigger picture of an individual`s shopping habits using his personal information discovered, which is seen as specific, because it offers personalized solutions.

For example, the “Blockbuster Entertainment mines its video rental history database to recommend rentals to individual customers. American Express can suggest products to its cardholders based on analysis of their monthly expenditures” (Data Mining Techniques 2010).

In both cases, for the general and for the specific solution, the process through which the solution was found, can be seen as an application of the Mosaic Theory, because as the Mosaic Theory states, by putting together pieces of information it can reveal a bigger picture, a mosaic.

The companies are storing data about their customers` shopping habits. With the use of their databases, the companies are able to perform the KDD process and through it, they have the possibility to form a personal profile related to the shopping habits of every individual that is a customer. For instance, in the customer profile created by the Blockbuster Entertainment, the company has added every video rental of a customer to its profile and by doing this, they were able to establish the preferences of the customer related to the rented movies and to make some recommendations. Here we can say that, by having a database filled with the customers` information and by using the process of data mining, the companies were able to form a personal profile of their customers as a method of increasing their revenue.

4.1.3 *Individual Profiling*

The third and the last step of the Superpanopticon model is achieved after the first two steps were completed meaning that, the collected data was mined. The result of the Superpanopticon is formed by a file, a dossier of every individual that had his data collected, and it can be seen as a profile of an individual.

According to the Webster dictionary, profiling is defined as "the act or process of extrapolating information about a person based on known traits or tendencies, e.g. consumer profiling" (Ferraris et al.: 6). Moreover, it adds that it is "the act of suspecting or targeting a person on the basis of observed characteristics or behavior, e.g. racial profiling" (ibid).

In an article wrote in 1993, Roger Clarke introduced profiling as a "dataveillance technique" (ibid) and refers to it as the "process of creating and using a profile" (ibid). Moreover, Clarke describes profiling as "a technique whereby a set of characteristics of a particular class of person is inferred from past experience and data-holding are then searched for individuals with a close fit to that set of characteristics" (ibid). According to Clarke`s definition, profiling is seen as a process of construction of a series of information (a profile), "which is then applied to something or someone by techniques of data elaboration" (ibid).

"A personal profile is a property or a collection of properties of a particular individual" (Custers 2013: 13). An example of a personal profile is "the personal profile of Mr. John Doe (44) who is married, has two children, earns 25.000 Euro per year and has two credit cards and no criminal record. He was hospitalized only twice in his life, once for appendicitis and last year because of lung cancer" (ibid).

According to the Bloomberg magazine, ThorpeGlen Ltd. is a company that provides "electronic surveillance solutions for homeland security, organized crime and major frauds" (Bloomberg). In 2008, Vincent Barry who is the Vice President of sales & marketing of the company, held a seminar which had as theme "Solutions to Current Challenges" (ThorpeGlen 2008). During the seminar, he explained the company's ability to identify and profile individuals. The main tools used to track and profile an individual, are his electronic devices. According to the company, these electronic devices forms what is called the "electronic DNA" (ibid). Moreover, by collecting the data from the "electronic DNA" (ibid), the company has the ability to analyze massive amounts of data and through the analysis, it is able to profile the individual whose data was collected, to identify his SIM card, how much does he use his device and it can track in real time the location of the device. Furthermore, the company claims that it has the ability to detect "change of SIM and change of handset" (ibid). If the person changes his devices along with his SIM card, after his profile was established, he can be tracked again. This is possible by comparing his Telecommunications behavior or device data with his historical profile which is stored in the company`s database.

In 2006, in The United States of America the Bush Administration demanded all the search data from the major search engines such as: Google, AOL, MSN and Yahoo. The reason for demanding the search information was to "help revive a child protection law" (Sullivan 2006). The answers were as follows: "Google says No; AOL MSN and Yahoo Said Yes" (ibid).

In the case of AOL, the company agreed to provide the list of searches to the administration but, with one condition. The condition was to anonymize the searches. AOL released a statement in which they claimed that: "We gave the DOJ a generic list of aggregate and anonymous search terms" (ibid). With the anonymization process, they wanted to make sure that the users` privacy was protected and they have replaced the users` usernames with some random numbers.

According to the Constitution of the United States of America, once you turn over something to the Administration unless it is a secret hearing, it becomes a public record. Once the documents were given to the administration, the New York Times newspaper made a FOIA request to obtain a copy of all of the provided documents from the AOL search engine.

In the documents about the searches made on the AOL search engine, searches that were collected and provided to the Bush Administration by the AOL company, there were buried in a list of "20 million web search" (Barbaro, Zeller 2006) the searches of the user "No. 441779" (ibid).

The AOL user "No. 4417749 conducted hundreds of searches over a three-month period on topics ranging from 60 single man" (ibid) to "landscapers in Lilburn, Ga" (ibid), "Arnold" (ibid), "homes sold in shadow lake subdivision gwinnett county Georgia" (ibid) and "dog that urinates on everything" (ibid).

The journalists from the New York Times followed the data trail and "search by search, click by click the identity of user No. 4417749 became easier to discern" (ibid). The journalists discovered “a 62-year old widow” who lives in the Shadow Lake subdivision, in Lilburn Georgia named Thelma Arnold who has three dogs. When one of the reporters read to her the searches that were on the paper, Ms. Arnold replied: “Those are my searches” (ibid). Moreover, she added: “My goodness, it`s my whole personal life” (ibid) and she argued the fact that: “I had no idea somebody was looking over my shoulder” (ibid).

Furthermore, she also had searches such as: “swing sets” or “dry mouth” or “competitive market analysis of homes in Lilburn” (The New York Times 2008). Ms. Arnold agreed to discuss with the reporters about her searches and described them one by one. “I was thinking about my grandchildren” (ibid) when I was searching for “swing sets” (ibid), Ms. Arnold stated. When she was searching for “market analysis of homes in Lilburn” (ibid) she argued that she “wanted to find out what my house was worth” (ibid).

According to the New York Times newspaper, Ms. Arnold`s searches “are a catalog of intentions, curiosity, anxieties and quotidian questions” (Barbaro, Zeller 2006). Furthermore, based on her searches Ms. Arnold was profiled as being 62 years old, living in Lilburn Georgia, widow or single – based on her “60 single man” (ibid) search, as a grandmother based on her “swing sets” (The New York Times 2008) searches and as a pet lover based on her “dog that urinates on everything” search.

This example illustrates perfectly how from simple searches made online, normal people get to be profiled. It further illustrates the assemblage of the Superpanopticon model, in which every achieved step enables the possibility for the next one to occur. In the case of Ms. Thelma Arnold, the data that was collected by the AOL company represents the process of data collection. Furthermore, the process of the de-anonymization succeeded by the New York Times journalists represents the process of data mining through which they have obtained knowledge about Ms. Arnold. As the Mosaic Theory states, by putting together pieces of information, a bigger picture, a mosaic of someone`s life will be obtained. In this case, the Mosaic Theory was applied by putting together the searches of Ms. Thelma Arnold. That is to say, a mosaic of her life was obtained and through it, she was tracked by the journalists and later on she was profiled.

4.1.4 *Case study: Strategic Communication Laboratories*

One might ask, what makes individual profiling so controversial? Is it the fact that if an individual is profiled as a smoker based on his shopping habits, he will have to pay more for his health insurance? Is it the fact that individuals can be turned off when applying for a job because of their established profile?

Or is it the fact that once the individual profile is established, the individual can be manipulated based on his behavior and personality traits? If this situation occurs, then, the statements - made by Michel Foucault - when modernizing the Panopticon theory – in which he sees the individual as an "object of information" (Foucault 1995: 200) from which knowledge can be extracted through observation, and that the extracted knowledge means power - are considered to be true and the Panopticon theory to be applied. Moreover, in the Panopticon model, the power was in the hands of the observer who was standing in the middle of the prison, while in this situation, the power is in the hands of the one who establish the individual profile, the one who uses the power to manipulate the individual.

Strategic Communication Laboratories (SCL) is the parent of a group of companies among which Cambridge Analytica company can be found. The SCL`s website states that: "We are the premier election management agency" (Grassegger, Krogerus 2017). Cambridge Analytica has captured the world`s attention and differentiated itself from the rest of the Big Data driven companies, with their revolutionary approach in the data field, by using psychometrics also known as psychographics techniques.

Psychometrics is "a data-driven sub-branch of psychology" (ibid) whose goals are "measuring psychological traits, such as personality" (ibid). Two teams of psychologists managed to develop in the 1980s a model which had the aim "to assess human beings based on five personality traits" (ibid). This model was called the Big Five. According to the Big Five model, the personal traits needed to asses a human being are: "openness (how open you are to new experiences?), conscientiousness (how much of a perfectionist are you?), extroversion (how sociable are you?), agreeableness (how considerate and cooperative you are?) and neuroticism (are you easily upset?)" (ibid). According to this model, with the help of these traits it is possible to make an assessment of the type of person who stands in front of us, including the person`s "needs and fears" (ibid) and how he might behave. This Big Five model which is also known as OCEAN, “has become the standard technique of psychometrics” (ibid). For a long period of time, the main problem that this model had, was the process of data collection. Back then, psychometrics had to fill out a personal questionnaire and it was a complicated procedure.

The solution to the data collection problem was provided by Michal Kosinski when he was accepted to do his “PhD at the Psychometrics Centre” (ibid) at the Cambridge University. Kosinski together with another student, David Stillwell launched a Facebook application called “MyPersonality” (ibid). Through this application, the Facebook users were filling out “different psychometric questionnaires, including a handful of psychological questions from the Big Five personality questionnaire” (ibid). After filling in the questionnaire, based on the Big Five values the users were receiving “a personality profile” (ibid).

In the beginning, both PhD students were expecting the questionnaire to be filled in by a couple of friends, but with the time passing by, more and more people completed the questionnaire. When “millions of people had revealed their innermost convictions” (ibid), Kosinski and Stillwell became the owners of the “largest dataset combining psychometric scores with Facebook profile ever to be collected” (ibid). Over the years, by using the questionnaire dataset to calculate the “personal Big Five values” (ibid) of every individual and by comparing the obtained results with the data gathered online about individuals, such as: “what they liked, share or posted on Facebook, or what gender, age, place of residence they specified” (ibid), the researchers were able to create an image about individuals` life. “While each piece of such information is too weak” (ibid) to produce an accurate personal profile, when “tens, hundreds, or thousands of individual data points are combined” (ibid) the resulted individual profile becomes accurate.

Moreover in 2012, Kosinski proved that, with the use of his model he could establish from the available data if the users consume drugs, cigarettes or alcohol and what their “religious affiliations” (ibid) are. He also proved that, “on the basis of an average of 68 Facebook likes by a user, it was possible to predict their skin color (with 95 percent accuracy) and their affiliation to the Democratic or Republican Party (85 percent)” (ibid). The model created by Kosinski, which is called Kosinski`s method, is able to analyze phycological traits of the individuals based on their online activity. That is, to create a profile of every individual who was analyzed and to store it. Through the created profile, the one who owns it, could establish what decisions are most likely to be taken by an individual. Moreover, by storing these profiles, it provides the possibility to search for people who have the same interest and who are likely to take the same decision, for example, if we are thinking retrospectively, the individuals who are likely to vote for the Republican nominee in the race for the White House, Donald J. Trump.

The method of analyzing and profiling individuals developed by Kosinski which can be used for future purposes was implemented by the Cambridge Analytica company. This company captured the world`s attention in 2016 when it was hired by Donald J. Trump, to help him with his electoral campaign, to become the president of the United States of America. While other campaigners have relied on demographics as usual, Cambridge Analytica used psychometrics.

Furthermore, Alexander James Ashburner Nix, who is the CEO of Cambridge Analytica, described in a presentation held in 2016 at the Concordia Summit, the steps which were going to be used for Mr. Trump`s campaign. First of all, Mr. Nix established "five different faces, each face corresponding to a personality profile" (ibid). Those five personality profiles are the Big Five also known as the OCEAN model. Moreover, as Mr. Nix stated: "At Cambridge, we are able to form a model to predict the personality of every single adult in the United States of America" (ibid). His company technique is based on an assemblage formed by three elements: "behavioral science using the OCEAN model, Big Data analysis, and ad targeting" (ibid).

The first step deployed by the company in Mr. Trump`s campaign, was to buy personal data from different sources such as: "like land registries, automotive data, shopping data, bonus cards, club memberships, what magazines you read, what churches you attend" (ibid) and so on. The second step was to combine the data that was collected and bought from others with the "electoral rolls of the Republican Party" (ibid) and with the data collected by the Cambridge Analytica from the online world, thus resulting in calculating a Big Five personality profile of every individual whose data was collected. Through this process, which is seen as a description of the Superpanopticon assemblage, all the digital footprints left behind by individuals become "real people with fears, needs, interests and residential addresses" (ibid). Both steps used in this process, namely Data Collection and Data Mining had as a result the third step of the Superpanopticon model, which is Individual Profiling. The achieved result is considered to be valuable, because as Mr. Nix stated: "We have profiled the personality of every adult in the United States of America-220 million people" (ibid).

After the company created the individual profiles of the American citizens and categorized the voters, it was the time for Cambridge Analytica to try to influence individuals, having as goal the victory of Donald J. Trump. In order to achieve the desired goal, the company tried to influence the potential Clinton voters such as: “left-wingers, African-Americans and young women” (ibid), “to suppress” (ibid) their votes which lead us to the fourth step of the process used by the Cambridge Analytica. This step is split into two parts: Testing and Delivery. The company decided that they have better chances to influence individuals by providing them with advertisings on the Facebook platform. The company “tested 175.000 different ad variations for his arguments, in order to find the right version” (ibid). Moreover, the messages “differed for the most part only in microscopic details, in order to target the recipients in the optimal psychological way: different headings, colors, captions, with a photo or video” (ibid). The delivery of the messages did not represent a problem, as Mr. Nix stated: “We can address villages or apartment blocks in a targeted way. Even individuals” (ibid).

In the delivery part of the fourth step, the company targeted individuals in different States, by sending them “sponsored news-feed-style ads in Facebook timelines that can only be seen by users with specific profiles” (ibid). For example, in the Miami district of Little Haiti, Cambridge Analytica provided the locals on Facebook with “news about the failure of the Clinton Foundation following the earthquake in Haiti” (ibid). In Detroit, Michigan where lives the biggest population of African-Americans in the United States of America, the adds contained video “in which Hillary Clinton refers to black men as predators” (ibid). The use of Facebook to send advertising to different individuals has proved to be the best choice as Mr. Nix stated: “Facebook proved to be the ultimate weapon and the best election campaigner” (ibid).

The fifth step of the process is designed to help the campaigners that are campaigning in the old way, by going from door to door and speaking with individuals. From July 2016, the campaigners “were provided with an app with which they could identify the political views and personality type of the inhabitants of a house” (ibid). In order to achieve the best result and to be able to speak with as many individuals as possible, the campaigners rang at the “houses that the app rated as receptive to his messages” (ibid). Furthermore, the application was providing the campaigners with “guidelines for conversations tailored to the personality type of the resident” (ibid).

Following the work of Cambridge Analytica and the work of others who were working for the election of Mr. Trump, on 8th November 2016, Donald J. Trump was elected as the 58th President of the United States of America. On 9th November 2016, the CEO of Cambridge Analytica sent out a press release stating: "We are thrilled that our revolutionary approach to data-driven communication has played such an integral part in President-elect Trump`s extraordinary win" (ibid).

In the end we can say that, by deploying each step of the Superpanopticon assemblage, it enabled the possibility for the next step to occur. In the case of the Cambridge Analytica, the steps involved in the process which was used to help Mr. Donald J. Trump to become president were:

The Data collection, which represents the first step of the process. Through this step, they have bought and collected data about American citizens. The second step of the process, is the step in which they have combined the collected data with the "electoral rolls of the Republican party" (ibid) which provided them pieces of information about individuals. This step represents the process of Data Mining. The third step of the process was possible to occur after the discovered pieces of information were put together and formed as a mosaic the individual profile of 220 million American citizens.

What makes the creation of the individual profile so controversial is represented by the fourth and fifth steps of the process and by the obtained result. In the fourth step, individuals were targeted with adds tailored by the findings about themselves in their personal profile in order to be influenced not to vote for Hillary Clinton. In the fifth step, with the use of an application that was using the established individual profile to determine the "political views and personality type of the inhabitants of a house" (ibid), the campaigners were addressing only to individuals who were living in the houses that were rated to be receptive by the application and with questions created to suit their personality.

The fact that an individual can be influenced by others, using his or her personal profile as a psychological weapon, in order to achieve their desired goal, - in our case the wining of the presidential election by Donald J. Trump with the help of Cambridge Analytica – makes the whole process of the Superpanopticon assemblage to be controversial. The fact that the future of a country can be influenced by an assemblage of techniques of mass surveillance is considered to be a controversial issue and this is the reason why Cambridge Analytica has captured the world`s attention after Donald J. Trump was elected as the President of the United States of America.

In this case, where a company has influenced the final outcome of the United States of America Presidential Election, basing themselves on an old strategy of surveillance, the Panopticon theory, and furthermore by developing an assemblage of techniques of mass surveillance, which is seen as a Superpanopticon, they were able to use the American citizens as an "object of information" (Foucault 1995: 200), from which knowledge was extracted through observation, – by using methods of Data Collection and the process of Data Mining – and the discovered knowledge was used to form individual profiles - by applying the Mosaic Theory - in order to manipulate and to influence individuals, we can say that it is true that knowledge represents power, and that the power was in the hands of the company. Furthermore, we can say that the statements made by Michel Foucault are true, and the whole process used by the company to influence or manipulate individuals based on their personal profile is constructed upon a modernization of the Panopticon theory. Moreover, we can also say that the Panopticon Theory and the Mosaic Theory have applied.

5. **Conclusion**

Throughout the human history privacy has played an essential role in respecting individuals` private lives. From the writings of the Greek philosopher, Socrates, who perceived privacy as a normal period of retirement from the public life, to the declaration of Queen Elizabeth who stated that: “she did not want to make windows into men`s hearts and secret thoughts” (Marx 2016: 1). According to Gary T. Marx, her actions emphasized the idea that a distinction should be made between the public and the private life. Moreover, her actions became the starting point of our society, of our conception about what a good society represents. Furthermore, in December 1890, for the first time in the human history, the words right and privacy were used in the same sentence. Samuel Warren and Louis Brandeis published an article in the Harvard Law Review, titled The Right to Privacy. According to their own view, The Right to Privacy was seen as “the right to enjoy life-to be let alone” (Warren Brandeis 1890 :193).

Moreover, with the evolution of the society clear regulations about privacy were required. As a result, in 1948 the United Nations published the Universal Declaration of Human Rights, which entitled every individual with the Right of Privacy. The appearance of the Internet and the continuous evolution of the technology provided human beings with different electronical devices that are used to accomplish different daily tasks. Every single interaction with these electronical devices produces personal data. As the use of these devices has increased from day to day, clear personal data regulations were required. As a result, in December 1990 the United Nations General Assembly adopted a resolution regarding the Regulation of Computerized Personal Files. While the use of these devices has enabled the convenience of communication, in the same time it has also enabled the convenience of surveillance, which is deployed through a process of invading individuals’ privacy.

Through the analysis chapter I have tried to answer the research question of this thesis, by analyzing the complex process of mass surveillance, that enables the possibility for the creation of individual profiles. Moreover, I have also analyzed what can be achieved with the use of the obtained profiles. I was able to establish that, by modernizing the Panopticon model framed by Jeremy Bentham -that assumes an architectural model in which an observer stands in the middle of the prison and observes the behavior of all of the prisoners without being seen- with the implication of technology, it results a complex process of mass surveillance, called

Superpanopticon. The Superpanopticon model of mass surveillance operates as an assemblage, formed by different strategies and techniques that are deployed step by step, and each deployed step enables the possibility for the next one to occur.

In order to create an individual profile, there are required pieces of information that mixed together reveal a mosaic of an individual`s life. The first step for achieving this goal, is represented by the collection of the data. This process can be done at a targeted level, for example, by collecting information from social media about an individual or by buying the data that it was already collected by data companies about the respective individual. This process can also be done at an untargeted level, where by using specialized software's and devices, the data of the individuals that are in a given geographical area is collected. In order for the second step of the process to occur, the collected data needs to be stored in large datasets and prepared for analysis, which coins the term Big Data. The second step of the process is called Data Mining. In this step, the collected data is analyzed through the Knowledge Discovery in Databases process in which Knowledge is discovered. Knowledge represents key information about individuals and offers the possibility for the third step to occur. In the third step, the goal of the process is reached, by applying the Mosaic Theory, which states that by putting together the discovered pieces of information about an individual, it creates a mosaic of his life, in our case the mosaic represents, the Individual Profile.

With the use of the profiles, individuals can be manipulated or influenced to take different actions, actions that are required in order to achieve the goal of the person or the company that deployed the necessary steps in order to establish the profile. In the case of the Strategic Communication Laboratories, the parenting company of Cambridge Analytica, that was hired to help Donald J. Trump to become the president of the United States of America, the proposed solution that brought the victory in the electoral campaign, was psychometrics. After the company has created the personal profile of 220 million American citizens, using the same process as presented above, the company has targeted individuals with personalized advertising tailored from their profiles. The company has used the citizens as an - "object of information" (Foucault 1995: 200) from which knowledge about themselves was extracted through observation. According to Foucault, knowledge represents power and with the use of power the company manipulated the American citizens - towards the company`s goal, not to vote for Hillary Clinton. In the case of the

Strategic Communications Laboratories both theories, namely the Panopticon Theory and the Mosaic Theory were applied.

All the steps of the Superpanopticon assemblage including the result, namely individual profiling, and what can be achieved with it, represent a clear invasion of individual's privacy. While there are norms and regulations regarding the privacy of an individual and the protection of personal data, these norms are not respected and are being violated on a daily basis. With the passing of time and with the continuous evolution of the technology, the techniques of mass surveillance are merging together to cover every corner of human life and human behavior. One might say that the society of our times is developing against the principles that have "bolstered liberty and sharpened distinctions between the public and the private that became central to our ideas of the good society" (Marx 2016:1) towards a society governed by a desperate need for making "windows into the men`s hearts and secret thoughts" (ibid). The issue of invading individuals` privacy, which clearly represents a problem nowadays, raises concerns more and more often and at a more advanced level. We should ask the real questions nowadays, because if we take into account the direction from where our society is coming, and how the electronic assemblage of mass surveillance has evolved in the last years, we notice the direction towards the electronic assemblage of mass surveillance is pushing our society in the near future, will be one in which, according to Foucault, the observers will have the power in their hands. If the power is in the wrong hands, not only that the Right to Privacy will suffer a more serious deterioration, but also the Democratic Values will be undermined.

Tony Porter, the British Government`s surveillance camera commissioner stated that: "I`m worried about the overt surveillance becoming much more invasive because it is linked to everything else. You might have a video photograph of somebody shopping in Tesco. Now it is possible to link that person to their pre-movements, their mobile phone records, any sensor detectors within their house or locality. As smart cities move forward, these challenges are so much greater for people like myself. And members of the public need to decide whether they are still happy with this. The nightmare scenario is that there is a lack of understanding about how big and effective this can be. This technology is there to protect us, but there needs to be informed and consent about what is capable of doing" (Weaver 2017).

6. **Bibliography**

ACLU. *"Stingray Tracking Devices"* in *American Civil Liberties Union Website*. Available at <https://www.aclu.org/issues/privacy-technology/surveillance-technologies/stingray-tracking-devices> [Accessed at 17 March 2017]

Ahmed, Nafeez. "*'Chilling Effect' of Mass Surveillance is Silencing Dissent Online, Study Says"* in *Motherboard Vice Website*, 17 March 2016. Available at <www.motherboard.vice.com> [Accessed at 6 March 2017]

Allmer, Thomas. "*Towards a Critical Theory of Surveillance Studies"* edited by the Unified Theory of Information Research Group (UTI) Vienna, Austria.The Internet & Surveillance – Research Paper Series, 3 November 2011

Ball, James. *"NSA collects millions of text messages daily in 'untargeted' global sweep"* in *The Guardian Website*, 16 January 2014. Available at <www.theguardian.com> [Accessed at 22 March 2017]

Banerjee, Sagnik, Subhadip Basu and Mita Nasipuri. *"Big Data Analytics and Its Prospects in Computational Proteomics".* In "*Information System Design and Intelligent Applications: Proceedings of Second International Conference INDIA 2015, Volume 2"* ed. by Mandal, J.K, Suresh Chandra Satapathy, Manas Kumar Sanyal, Partha Pratim Sarkar, Anirban Mukhopadhyay. India: Springer, 2015

Banisar, David. *"National Comprehensive Data Protection/Privacy Laws and Bills 2016"* in *Social Science Research Network Website*, 31 October 2016. Available at <https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1951416> [Accessed at 13 March 2017]

Barbaro, Michael, Tom Zeller Jr. *"A Face Is Exposed for AOL Searcher No. 4417749"* in *The New York Times Website*, 9 August 2006. Available at <www.nytimes.com> [Accessed at 13 May 2017]

Beal, Vangie. *"Structured data"* in *Webopedia Website*. Available at <www.webopedia.com> [Accessed at 23 March 2017]

Bloomberg. *“Company Overview of ThorpeGlen Ltd.”* in *Bloomberg Website*. Available at <https://www.bloomberg.com/research/stocks/private/snapshot.asp?privcapId=29688915> [Accessed at 5 May 2017]

Bornschlegl, Marco X., Kevin Berwind, Michael Kaufmann, Felix C. Engel, Paul Walsh, Matthias L. Hemmje and Ruben Riestra. *“IVIS4BigData: A Reference Model Advanced Visual Interfaces Supporting Big Data Analysis in Virtual Research Environments”.* In *“Advanced Visual Interfaces Supporting Big Data Applications: AVI 2016 Worksop, AVI-BDA 2016 Bari, Italy, June 7-10, 2016”* Revised Selected Papers ed. by Marco X. Bornschlegl, Felix C. Engel, Raymond Bond, Matthias L. Hemmje. Springer International Publishing: 2016

Brooks, Chad. *“What is Data Mining”* in *Business News Daily Website*, 19 February 2014. Available at <www.businessnewsdaily.com> [Accessed at 28 April 2017]

Brown, Deborah. “*New UN resolution on the right to privacy in the digital age: crucial and timely”* in *Internet Policy Review Journal on internet regulations Website*, 22 November 2016. Available at <https://policyreview.info/> [Accessed at 13 March 2017]

Cambridge Dictionary *“Privacy”* in *The Free Online English Dictionary.* Available at <http://dictionary.cambridge.org/> [Accessed at 11 March 2017]

Custers, Bart. “*Data Dilemmas in the Information Society: Introduction and Overview”*. In “*Discrimination and Privacy in the Information Society. Data Mining and Profiling in Large Databases* ed.by Bart Custers, Toon Calders, Bart Schermer, Tal Zarksy. Berlin: Springer, 2013

Data Center Knowledge. *“How Much Does Facebook Spend on Its Data Center?”* in *Data Center Knowledge Website*. Available at <http://www.datacenterknowledge.com/the-facebook-data-center-faq-page-three/> [Accessed at 2 April 2017]

“Data Mining” in *California State University Northridge Website*. Available at <http://www.csun.edu/~sh111280/Hassas.html> [Accessed at 3 May 2017]

"Data Mining Techniques", Data Mining Seminar Report, 30.04.2010, in *Collegelib Website*. Available at <http://www.collegelib.com/t-data-mining-seminar-report.html> [Accessed at 3 May 2017]

"Directive 95/46/EC of the European Parliament and of The Council" in the *European Commission Website*, 23.11.1995. Available at <http://ec.europa.eu/justice/policies/privacy/docs/95-46-ce/dir1995-46_part1_en.pdf> [Accessed 9 March 2017]

"Directive 2002/58/EC of the European Parliament and of The Council" in the *European Union Law Website*, 31.7.2002. Available at <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32002L0058&from=en> [Accessed at 9 March 2017]

Facebook. "*Data Policy"* in *Facebook Website*. Available at <https://www.facebook.com/policy.php> [Accessed at 7 April 2017]

Ferraris,V., AMAPOLA F. Bosco, G. Cafiero, E. D`Angelo, Y.Suloyeva. *"Working Paper Defining Profiling"* in *United Nations Interregional Crime and Justice Research Institute Website*. Available at <http://www.unicri.it/special_topics/citizen_profiling/WP1_final_version_9_gennaio.pdf> [Accessed at 5 May 2017]

Foucault, Michel. *"Discipline & Punish. The Birth of the Prison"*. New York: Vintage Books a Division of Random House, Inc: 1995

Fox-Brewster, Tom. *"Check the permissions: Android flashlight apps criticized over privacy"* in *The Guardian Website*, 3 October 2014. Available at <www.theguardian.com> [Accessed at 10 April 2017]

Franceschi-Bicchierai, Lorenzo. *"The 10 Biggest Revelations from Edward Snowden`s Leaks"* in *Mashable Media Website*, 05 June 2015. Available at <www.mashable.com> [Accessed 5 March 2017]

Gellman, Barton, Laura Poitras. "*U.S., British intelligence mining data from nine U.S. Internet companies in broad secret program"* in *The Washington Post Website*, 7 June 2013. Available at <www.washingtonpost.com> [Accessed at 22 March 2017]

Grassegger, Hannes, Mikael Krogerus. *"The Data That Turned the World Upside Down"* in *Motherboard Vice Website*, 28 January 2017. Available at <https://motherboard.vice.com/en_us> [Accessed at 18 May 2017]

Greenwald, Glenn, Ewen MacAskill. "*NSA Prism program taps in to used data of Apple, Google and other"* in *The Guardian Website*, 7 June 2013. Available at <www.theguardian.com> [Accessed at 22 March 2017]

Greenwald, Glenn. *"XKeyscore: NSA tool collects 'nearly everything a user does on the internet'"* in *The Guardian Website*, 31 July 2013. Available at <www.theguardian.com> [Accessed at 23 March 2017]

Greiner, Lynn. *"What is Data Analysis and Data Mining?"* in *Database Trends and Applications Website*, 7 January 2011. Available at <http://www.dbta.com/Editorial/Trends-and-Applications/What-is-Data-Analysis-and-Data-Mining-73503.aspx> [Accessed at 26 March 2017]

Han, Jiawei, Micheline Kamber. "*Data Mining. Concepts and Techniques"* Second Edition. Amsterdam, Boston, Heidelberg, London, New York, Oxford, Paris, San Diego, San Francisco, Singapore, Sydney, Tokyo: Morgan Kaufmann Publishers: 2006

Harris, Drew. "*Analytics and Big Data Glossary"* in *Data Informed Website*, 27 April 2012, last updated on 16.3.2017. Available at <www.data-informed.com> [Accessed at 25 March 2017]

History. "*The Invention of the Internet"* in *History Website*. Available at <www.history.com> [Accessed at 16 March 2017]

Holvast, Jan. "*History of Privacy"* in Matyáš V., Fischer-Hübner S., Cvrček D., Švenda P. (eds):*" The Future of Identity in the Information Society"*, IFIPT AICT 298, pp.13-42. Berlin, Heidelberg, New York; Springer, 2009

House of Lords. "*Surveillance: Citizens and the State"*. Select Committee on The Constitution 2nd Report of Session 2008-09, Volume I Report, published on 6 February 2009. London: The Stationery Office Limited in the *United Kingdom Parliament Website*. Available at <https://www.publications.parliament.uk/pa/ld200809/ldselect/ldconst/18/18.pdf> [Accessed at 16 March 2017]

Investopedia. *"Mosaic Theory"* in *Investopedia Website.* Available at <www.investopedia.com> [Accessed at 28 March 2017]

Isikoff, Michaell. "*NSA program stopped no terror attacks, says White House panel member"* in *NBC News Website*, 20 December 2013. Available at <www.nbcnews.com> [Accessed at 6 March 2017]

Johansen, Per Anders. *"Secret surveillance of Norway`s leaders detected"* in *Aftenposten Website*, 13 December 2014, Updated on 16 December 2014. Available at <http://www.aftenposten.no/norge/Secret-surveillance-of-Norways-leaders-detected-71828b.html> [Accessed at 22 April 2017]

Kerr, Orin S. "*The Mosaic Theory of the Fourth Amendment"*, in the *Michigan Law Website*, Michigan Law Review, Volume 111, Issue 3, 2012 Available at <http://repository.law.umich.edu/> [Accessed at 28 March 2017]

Kothari, C.R. "*Research Methodology Methods and Techniques"* Second Revised Edition. New Delhi, New Age International (P) Limited Publishers, 2004,1990,1985

Lijpart, Arend. "*Comparative Politics and the Comparative Method"* in *The American Political Science Review,* Vol.65, No3 (Sept., 1971), 682-693

Macmillan English Dictionary "*Privacy"* in *The Free Online English Dictionary from Macmillan Publishers.* Available at <http://www.macmillandictionary.com/> [Accessed at 11 March 2017]

Marquis-Boire, Morgan, Glenn Greenwald, Micah Lee. *"XKEYSCORE NSA`s Google for the World`s Private Communications"* in *The Intercept Website*, 1 July 2015. Available at <https://theintercept.com/2015/07/01/nsas-google-worlds-private-communications/> [Accessed at 26 April 2017]

Marx, T.Gary. "*Windows into the Soul. Surveillance and Society in an Age of High Technology"*. Chicago and London: The University of Chicago Press, 2016

Mason, K. Moya.*" Foucault and His Panopticon"* in *Moya K. Mason Research Website*. Available at <http://www.moyak.com/papers/michel-foucault-power.html> [Accessed at 14 March 2017]

Meola, Andrew. *"What is the Internet of Things (IoT)?"* in *The Business Insider Website*, 19 December 2016. Available at <http://www.businessinsider.com/what-is-the-internet-of-things-definition-2016-8?r=US&IR=T&IR=T> [Accessed at 23 March 2017]

OECD. "*OECD Guidelines on the Protection of Privacy and Transborder Flows of Personal Data"* in *OECD Website.* Available at <www.oecd.org> [Accessed at 7 March 2017]

PCMag. *"Definition of: Smartphone"* in *PCMag Magazine Website*. Available at <http://www.pcmag.com/encyclopedia/term/51537/smartphone> [Accessed at 17 March 2017]

Pozen, E. David. *"The Mosaic Theory, National Security, and the Freedom of Information Act"* in *The Yale Law Journal Website*, 2005. Available at <http://www.yalelawjournal.org/> [Accessed at 28 March 2017]

Privacy International. *"What is mass surveillance?"* in *Privacy International Website.* Available at <https://www.privacyinternational.org/node/52> [Accessed at 5 March 2017]

Privacy International. *"What is Data Protection?"* in *Privacy International Website*. Available at <https://www.privacyinternational.org/node/44> [Accessed at 14 March 2017]

Research Methodology. *"Qualitative Research"* in *Research Methodology Website*. Available at < http://research-methodology.net/research-methods/qualitative-research/ > [Accessed at 23 May 2017]

Research Methodology. *"Quantitative Research"* in *Research Methodology Website*. Available at <http://research-methodology.net/research-methods/quantitative-research/> [Accessed at 23 May 2017]

Scahill, Jeremy, Josh Begley. *"The Great Sim Heist. How Spies Stole the Keys to the Encryption Castle"* in *The Intercept Website*, 19 February 2015. Available at <https://theintercept.com/2015/02/19/great-sim-heist/> [Accessed at 26 April 2017]

Segall, Laurie. *“Your private data pays for ‘free’ Facebook and Google”* in *CNN Money Website,* 28 January 2011. Available at <http://money.cnn.com/2011/01/28/technology/google_data_privacy_day/> [Accessed at 5 April 2017]

Smith, Oliver. *“Facebook terms and conditions: why you don`t own your online life”* in *The Telegraph Website*, 4 January 2013. Available at <www.telegraph.co.uk> [Accessed at 5 April 2017]

Sullivan, Danny. *“Bush Administration Demands Search Data; Google Says No; AOL, MSN & Yahoo Said Yes”* in *Search Engine Watch Website*, 19 January 2006. Available at <www.searchenginewatch.com> [Accessed at 11 May 2017]

Swanborn. G.Peter. *“Case Study Research: What, Why and How?”.* Los Angeles, London, New Delhi, Singapore, Washington DC: SAGE Publications, 2010

Techopedia. *“International Mobile Subscriber Identity (IMSI)”* in *Techopedia Website*. Available at <www.techopedia.com> [Accessed at 17 March 2017]

Techopedia. *“Knowledge Discovery in Databases (KDD)”* in *Techopedia Website*. Available at <www.techopedia.com> [Accessed at 26 March 2017]

Techopedia. “*Predictive Data Mining”* in *Techopedia Website*. Available at <www.techopedia.com> [Accessed at 27 March 2017]

Techopedia. “*What is the difference between big data and data mining?”* in *Techopedia Website*. Available at <www.techopedia.com> [Accessed at 27 March 2017]

The New York Times. *“What Revealing Search Data Reveals”* [Picture] in *The New York Times,* 8 August 2006. Available at <http://www.nytimes.com/imagepages/2006/08/08/business/09aol-graphic.html> [Accessed at 13 May 2017]

The Statistics Portal. *“Number of monthly active Facebook users worldwide as of 1st quarter 2017 (in million)”* in *The Statistics Portal Website.* Available at <www.statista.com> [Accessed at 2 April 2017]

The Statistics Portal. *“Number of monthly active Twitter users worldwide as of 1st quarter 2017 (in million)”* in *The Statistics Portal Website.* Available at <www.statista.com> [Accessed at 2 April 2017]

ThorpeGlen. *“ISS- Webinar – 13th may 2008 Identification of Nomadic Targets”,* presented by Vincent Barry, 13th May 2008. Available at <https://assets.documentcloud.org/documents/1217046/1361-thorpeglen-presentation-identification-of.pdf> [Accessed at 9 May 2017]

Turow, Joseph. *“The Daily You. How the New Advertising Industry is Defining Your Identity and Your World”*. New Have & London: Yale University Press, 2011

Tutorialspoint. *“Data Mining data pattern evaluation”* in *TutorialsPoint Simply Easy Learning Website,* 2014. Available at <https://www.tutorialspoint.com/data_mining/data_mining_tutorial.pdf> [Accessed at 27 March 2017]

UCL. “*The Panopticon”* in the *UCL London Website*. Available at <https://www.ucl.ac.uk/Bentham-Project/who/Panopticon> [Accessed at 14 March 2017]

UN General Assembly. *“Guidelines for the Regulation of Computerized Personal Data Files”* in *Refworld Website,* 14 December 1990. Available at <http://www.refworld.org/docid/3ddcafaac.html> [Accessed at 7 March 2017]

United Nations General Assembly. “*Resolution adopted by the General Assembly on 18 December 2013*” in *NATO Cooperative Cyber Defence Centre of Excellence* Website, 21 January 2014. Available at <https://ccdcoe.org/> [Accessed at 13 March 2017]

“Universal Declaration of Human Rights” in *OHCHR Website.* Available at <http://www.ohchr.org/EN/UDHR/Documents/UDHR_Translations/eng.pdf> [Accessed at 6 March 2017

Vice on HBO. “*‘State of Surveillance’ with Edward Snowden and Shane Smith”* in *Vice News Website*, published on 8 June 2016. Available at <https://news.vice.com/article/state-of-surveillance-with-edward-snowden-and-shane-smith> [Accessed at 16 March 2017]

Vincent, James. *"99.6 percent of new smartphone run Android or iOS While BlackBerry`s market share is a rounding error"* in *The Verge Website*, 16 February 2017. Available at <www.theverge.com> [Accessed at 10 April 2017]

Vonn, Micheal. *"Stingray surveillance: more of the story"* in *British Columbia Civil Liberties Association Website*, 8 August 2016. Available at <www.bccla.org> [Accessed at 25 April 2017]

Warren d. Samuel, Louis D. Brandeis. "*The right to Privacy"* in the *Harvard Law Review*, Vol.4, No.5 (Dec 15, 1980), pp. 193-220

Weaver, Mattew. *"UK public faces mass invasion of privacy as big data and surveillance merge"* in *The Guardian Website*, 14 March 2017. Available at <https://www.theguardian.com/uk-news/2017/mar/14/public-faces-mass-invasion-of-privacy-as-big-data-and-surveillance-merge> [Accessed at 29 May 2017]

Weissman, Cale Guthrie. *"It turns out the NSA was collecting voice, calls, photos, passwords, documents, and much more"* in *The Business Insider Website*, 1 July 2015. Available at <http://www.businessinsider.com/nsa-xkeyscore-surveillance-program-details-revealed-in-new-snowden-documents-2015-7?r=US&IR=T&IR=T> [Accessed at 23 March 2017]

Woolf, Nicky. *"Stingray documents offer rare insight into police and GBI surveillance"* in *The Guardian Website,* 26 August 2016. Available at <www.theguardian.com> [Accessed at 22 March 2017]

Yin, K. Robert. *"Study Research Design and Methods"* Second Edition. London, New Delhi: SAGE Publications

Yin, K. Robert. *"Case Study Research Design and Methods"* Fourth Edition. United States of America: SAGE Publications, 2009

Yin, Robert K. *"Qualitative Research from Start to Finish".* New York, London: The Guilford Press, 2011

"1973: The Code of Fair Information Practices". Available at <http://simson.net/ref/2004/csg357/handouts/01_fips.pdf > [Accessed 7 March 2017]

--

Front page image source: Marquis-Boire, Morgan, Glenn Greenwald, Micah Lee. *“XKEYSCORE NSA`s Google for the World`s Private Communications”* in *The Intercept Website*, 1 July 2015. Available at <https://theintercept.com/2015/07/01/nsas-google-worlds-private-communications/> [Accessed at 26 April 2017]