University Of Sarajevo

Faculty of Political Science

Sarajevo, Bosnia and Herzegovina

Master Thesis

The importance of lifelong learning in media and information literacy for information security: information disorder during the COVID-19 pandemic

Author: Anes Mirojevic

Master Program in Information Security

June 2022

Abstract

KEY WORDS: Awareness, Human, Information, Literacy, Media

This paper analyzes the current situation related to massive information disorder related to information being shared on various internet-based platforms and its influence on the public perception of the same. As an example, we look into the information surrounding the COVID-19 pandemic, giving an overview of how information is served and consumed. Further on, the effects of the information disorder, based on information disorder and various conspiracy theories, has led to serious issues and threats related to the health of human beings even endangering their lives compared to others which are not trapped in the claws of misinformation. To uncover the reason why Information Disorder has such a strong effect, the human factor in information security, and how awareness training helps resolve issues, is analyzed. This paper moves on to identify Media and Information Literacy (MIL) as one of the forms to fight against the challenges faced in the digital world. In the end, the emphasis is given on the need for a well-defined awareness program for MIL, which should be based on experiences related to awareness trainings used to teach individuals of various organizations to protect against Information Security issues related to Social Engineering.

Acknowledgements

I would like to express my gratitude to my mentor and supervisor on this thesis Prof. Dr. Emir Vajzović, who has provided guidance in writing this thesis. I am mostly grateful for the patience, shared knowledge, and valuable assistance throughout every step of writing this thesis.

I would like further to express my thanks and gratitude to my professors at the University of Sarajevo, who have supported all of my efforts in completing the studies on the master's program in Information Security at the Faculty of Political Science, University of Sarajevo in Bosnia and Herzegovina.

The support of my parents, Aida and Sead, as they have provided support in my education all the way from primary school to this master's program for which I am, will always be in their debt.

Special tanks for the support of my fellow colleagues on the master's studies Selmir Ljevaković, Sadik Crnovršanin, Damir Softić, Dževad Mujadžić, and Ishak Kovacevic which was very important.

Special emphasis and thanks to the support given goes to my brother Faruk Mirojevic and my friend Jasmina Đikoli who have worked with me on all of the projects and activities of this master's program and have managed always to push me forward.

Finally, I would like to give my most gratitude to my wife Irma and our daughters Ajša and Đula, for being the inspiration in my life, making me always wanting to give my best and be the best person always. This would not be possible without your unconditional love and support.

Contents

1. Introduction

This chapter will focus on the matter related to the motivation for this thesis. This will include the presentation of the problem, solution method and the main contributions to the conducted research.

1.1 Background and motivation

Information security is evolving on a daily basis and has become an important aspect of every serious organization in the world in the recent couple of decades and the reason for this is the constant increase of dependencies related to the IT/ICT technologies and services which they provide in our everyday lives, both private and job related.

As the increase of use related to IT/ICT technologies is rising, the attacks on the same are rising even more. The attackers attempt to gain access to information, information systems and anything related to the mentioned in order of performing malicious activities, which can be related to financial gain, collecting personal data or intellectual property.

As identified by several studies like the ones by Parsons et al. (2010) and Yildirim, et al. (2011), humans are identified as the weakest link and the most frequent targets of malicious actors. The attacks on humans are mostly based on social engineering attacks, which intends to target the human nature and the basic human psychology to gain their attack goals (Mouton, et al. 2016). The identified best approach to protect against the attacks on the human factor is to educate humans on the threats and risks related to the mentioned attacks, by performing awareness trainings (He, et al. 2019).

All of the above mentioned refers to the corporate environment or within various organizations.

The research of this thesis explores the human factor, the attacks and social engineering used to exploit the human factor in everyday access to digital resources within and outside of an organization. The intention is to explore if awareness training performed to protect against the attacks within organizations, for the purpose of this paper referred to as the 'industry standard' in information security awareness, can be implemented to protect against exposure every human accessing the digital resources, predominantly the internet, is facing. The threat vector turns not only to the traditional social engineering exploits to gain information about an organization and individuals within, but also includes the voluntary information sharing on internet platforms such

as social media, web sites, content providers, that lead to consequences in the physical realm of our lives. Examples of such consequences are related to belief in scientific findings related to vaccination, calls for social unrest, electing officials and government in elections, abiding to law, respect for others etc.

1.2 Problem definition

So, to better understand the stance, we take a closer look at a human being, that is surrounded by the digital world we live in today, with a reflection on the many challenges faced. As humans are used to use their five senses to process the information, the need for the 'sixth sense' arises, in order to be able to tackle the challenges of information consumption today (Vajzović2, 2020). The constant growth of digital systems and information providers often feels overwhelming and challenging in the attempt to comprehend and function within the society.

On the other hand, what is an average human aware of in terms of the challenge the 'digital era' brings? What lies under the online platform accesses on a daily basis and what is the role individuals play in it?

The research is set to uncover what is the human perception and what information consumption can do by asking and answering to the question on how the human factor, incorporating all basic human capabilities, can be targeted, and manipulated by the digital platforms that are being used. Through analysis of events that were directly based on information consumption and the human perception of the same, we will analyze how the information has played its role in some major events with an example on how it played out in the first two years of the COVID-19 pandemic.

Further on, an attempt is made to see what the options are to protect against the identified threats to the various types of security (social, economic, health, state etc.). The information security aspect and its role in the game is based on identifying the forms used on online platforms and putting them into context of information security threat type.

Based on the results and examples, related to information security, we look at Media and Information Literacy (MIL), what it represents and how it can be used. Considering the mentioned, we have developed an awareness program, based on the identified threats of the 'new generation'. The program itself is based on the principles of MIL, incorporating the approach 'industry standard' awareness programs are developed and delivered. The goal is to test and

present a comparative analysis of what the developed awareness training can achieve in raising awareness in MIL.

In terms of the problem definition, this thesis intends to answer the following question:

Is it possible to establish a professional training program in the field of media and information literacy, which would lead to an increase in competences, the ability to recognize and defend against attacks from the domain of social engineering that are used for the purposes of creating information disorder, which lead to endangering health, social and general security in the organization?

1.3 Scope and method to provide solution.

As the main problem is defined and relevant issues mentioned including the efforts to address it, are represented in the steps taken in the research The steps are defined and depicted as shown in Figure 1. The Figure shows the steps that were taken to conduct this study. Four steps have been identified as the structure which was used to present the research.

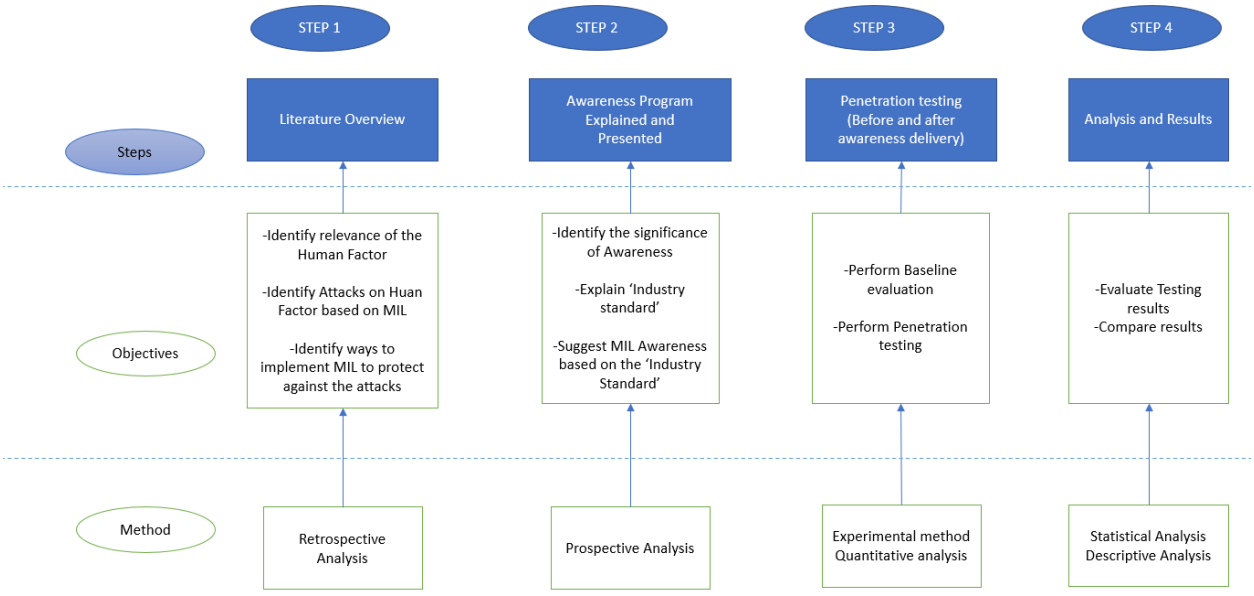

Figure 1 Research steps

1.4 Summary of contributions

The review and analysis of current approaches to resolving issues related to the human factor, including awareness training as one of the most common ways to resolve human factor related issues, provides the baseline or starting point. Based on the suggested resolution of the human factor related issues through awareness training, the new approach to address the reduction of risks related to the exploits of the human factor using information disorder in organizations in Bosnia and Herzegovina, is done through introducing Media and Information Literacy (MIL) awareness training.

The study relies on published academic papers which cover the subjects at hand and the actual tests conducted for the needs of this thesis.

The Contributions to this thesis are listed below:

I. Knowledge acquired from the literature review based on regulatory publications, professional articles, news segments and articles, and academic papers.
II. Development of the Media and Information Literacy awareness training, based on the 'industry' standards for information security awareness training approach as it is known to be a part of standard information security practices in organizations.
III. Penetration tests, based on Information Disorder Techniques and social engineering, conducted in an educational organization in Sarajevo, Bosnia and Herzegovina.
IV. Analysis of results based on the performed testing used to support the research needed for this thesis. Presentation and discussion based on the results which have shown that the Media and Information Literacy awareness training does improve the level of security in organizations.

1.5 The thesis structure.

The Thesis is structured as described below:

- Chapter 1 (Introduction) describes the background and the motivation to pursue the topics outlined in this thesis. It also defines the problem and the scope of the thesis.
- Chapter 2 (Literature Overview) provides a review of relevant literature which helps define the human factor, its significance from an information security perspective, what is needed to protect against human factor exploitations and social engineering as the main method of

exploiting the human factor. Further this chapter introduced Information Disorder identified and the source of the problem, and Media and Information Literacy as a way to fight information disorder. Furthermore, it goes to describe algorithms as amplifiers of information disorder and gives real-life examples of what consequences can be faced due to the information disorder techniques. In the last parts of this chapter, Media and Information Literacy awareness is suggested as the cure and response to the threats and weaknesses of the human factor.

- Chapter 3 (Applied research and testing methodology) explains the methodology used for the research, explains the objectives of the research, and sets the hypothesis of the thesis, based on the defined problem question. This chapter further goes to explain and define the Media and Information Literacy based awareness training, gives insight to the environment in which the testing was conducted and explains the testing process with actions taken to complete the tasks that are set in this thesis.
- Chapter 4 (Result analysis) analyzes and results of the performed testing. It also focuses on the results from the perspective of testing the outlined hypothesis of this thesis.
- Chapter 5 (Discussion) the result analysis is used as basis to discuss the end results and achievements of the preformed research.
- Chapter 6 (Conclusion) is a review of the conducted research, based on which the main conclusions are made, in relation to the hypothesis of this thesis. This chapter also discusses the limitations of the thesis and gives a suggestion to the potential additional research.
- Chapter 7 (Bibliography) displays the list of reference publications used in this thesis. It includes the relevant publications along with academic papers and research.

2. Literature Review

This chapter provides a review of relevant publications, academic papers, and relevant sources, which help define the human factor, its significance from an information security perspective, what is needed to protect against human factor exploitations and social engineering as the main method of exploiting the human factor. Further this chapter introduced Information Disorder as the source of the issues, Media, and Information Literacy as the response to deal with the issues, describes algorithms as amplifiers of information disorder and gives real-life examples of what consequences can be faced due to the information disorder techniques.

2.1 The Human factor

This part of the paper analyzes the Human Factor as the weakest link in information security and explains how targeting a human can cause serious information security issues and incidents.

Studying the human factor concentrates on researching of how humans behave physically and psychologically in relation to different environments, sources of information they have access to, various products, and/or services provided to them and/or provided by them. To give a high-level overview in terms of information security, the human factor is most commonly relating to the mishaps initiated by human interactions that could lead to creation of vulnerabilities to the IT systems used including the business-related applications (Parsons et al., 2010).

The influence of the human factor in cyber security is researched from many aspects taking different approaches towards the issue where the intention is to try to explain the problems associated to the human factor and possibly provide solutions to the same. There have been many studies in relation to how a human reacts in certain situations considering the psychology, the environment and atmosphere in which an individual operates and the personality of an individual in reflection to the effect that is left on information security. The main argument is that information security is not and should not only be based on the technologies used but is rather more related to the actual people that use the systems within their organizations (Bowen, et al., 2011).

In addition to the business environment in which human behavior is considered the weakest link, when it comes to online sources, the information consumption also affects the human as the weakest link and applies consequences on real-life events.

The human factor effect on the digital realm can be unintentional or intentional. Unintentional human factor interactions are all actions done by a human being without the real intention of creating a mishap and/or disrupting the overall information/cyber security in an organization.

Intentional human factor interactions are actions performed by a human that have the intention and goal of creating a mishap that is intended to disrupt the information/cyber security in an organization. Out of the two mentioned, at first sight more dangerous are the intentional interactions. (Hadlington, 2018).

From an organization's perspective, it is the result and the full effect of consequences related to mishaps caused by the human factor that is most important and is in most cases the only thing that matters, rather than the question, whether it is coming from an internal or external threat. Regardless of the source form which the malicious actor is coming from, it is known that many attacks are relying on the human factor to execute the malicious activity and/or to initiate such activities.

One additional argument made in reference to the human factor in organizations is based on the knowledge an individual and/or a group of individuals have in terms of information/cyber security. If an organization is equipped with the most sophisticated software and other equipment, the user is the one that still needs to use it properly to make sure the cyber security is not endangered. One simple example of email sending is given, where the user needs to have knowledge on what type of information is being sent, who is the recipient of the email and the knowledge required to ensure it is sent in accordance with the rules set by the organization.

As an example, sending an email with confidential data, and failing to use the company's encryption tool on their email solution. The user needs to be educated and aware that such confidential information must be encrypted before sending. (Hadlington, 2018)

One of the biggest challenges today in many organizations is related to the basic human nature and approach to life. A human has needs to be free and comfortable in the environment in which it operates and conducts its job-related tasks.

As one study shows, human beings have the need for both security and flexibility. In the attempt to resolve the issue, it has been concluded that the trick is to find the balance between these both. This is certainly the ultimate challenge when dealing with the human factor. A single approach has not yet been developed and might never be. So, to resolve the challenges, combinations of several approaches deployed, tested and confirmed to be able to select which approach is best applied to fit (Parsons et al., 2010).

Internet-based platforms as sources of information, and the sources of misinformation were not perceived as such serious issues until the information served online, resulted in real life effects, like the 2016 US Presidential elections. This directly played on the human psychological aspect of believing false information to be true. (Allcott et.al 2019)

One of the main differences between accessing digital and online resources within an organization and outside of the same, using private means and resources, is that generally organizations pay attention mostly to the assets and content on the devices and services owned and/or used to support the organizations processes, core business etc. The organizations are motivated to protect the business and do not necessarily consider the protection of individual especially when it comes to threats which are perceived as ones that will not have any impact on the security of the organization. Even though the knowledge acquired within the organization transcends to the 'home' environment, there is a need to develop a framework to consider information security in all environments at work and at home (Talib et al, 2010).

As any other aspect of information security, as long as information is considered, it can disrupt the security of humans as the weakest link. This also applies to receiving news. The society has gone and is still going through a digital transformation, which results in creating connected and plugged citizens. (Vajzović1, 2021). The connected citizens might be aware of the need to have an anti-virus, but there is no easily accessible and well distributed program to raise awareness of the threats that lurk in the digital realm. This is especially more worrisome since the growth of information consumers form online sources in in a constant and rapid growth. Analysis has shown that two thirds of adults in US read news and get informed on social media platforms (Wu et al, 2019).

The social Media platforms don't have editors, nor do they apply the concept of ethical editing to the content provided. Although some platforms tend to remove offensive content, misinformation, disinformation etc. overall it is not enough to protect the mentioned weakest link, the human factor, from being exploited. Basically, the situation is as such that social media platforms cannot be considered as an overall reliable source of information, and yet it is becoming the main source of news and information for more and more individual by the day. As similar situation resides in Bosnia and Herzegovina, where a lot of online content provides misinformation based on various motivations (i.e. Political gain, Click bait, bias information sharing etc.), the general perception of the public is that if the information is served online, it is trusted, even though most social media accounts and web sites that spread misinformation are coming from anonymous accounts and portals, which cannot be verified and/or held accountable (Cvjetićanin et al., 2019).

Methods of exploiting the human factor in information security is referred to as exploitation using 'Social Engineering'. Basically, social engineering attacks rely mostly on the human interactions to execute the attack. Social engineering attacks amongst other segments, include gaining knowledge of the individual being targeted, analysis of their behavior and their online habits. When considering the standard approaches to social engineering and the goals of the same, they included detection of inappropriate management of internal systems, which can be used to perform malicious activities and attacks on organizations. Today the social engineering targets the humanity of an individual in order to lure the victims to take certain actions, follow certain agendas, spread propaganda, purchase products, cause social unrest and many more.

2.2 Social Engineering

Social engineering methods are intended to lead to a deception of a human being with the intent to manipulate a single and/or multiple individuals into sharing and/or uncovering confidential information and/or taking action that can be related to personal and/or business sensitive information that can be used for malicious/fraudulent purposes (Mouton et al., 2016).
The usual Social Engineering attacks are identified, based on experiences and on the need to protect businesses or an industry from such attacks, which can be:

-Phishing: Phishing I s based on an attempt to obtain sensitive and/or confidential information by a fraudulent activity. These types of attacks usually are intended to uncover passwords, usernames, credit card numbers etc. The attack is usually executed by disguising as a trustworthy person using the means of electronic communication. **(**Terranova security, 2020**).**

Phishing activities might include email spoofing, instant messaging, false targeted advertising, and other forms of bait planting, using electronic channels and/or services. For example, phishing emails are attacks that use email massages as the point of contact and human interactions. The phishing emails usually contain malicious links, where the recipient is asked to click on the same, might have a malicious attachment asking the recipient to open the same and/or it might be giving instructions to the recipient to take certain action to enable the attack and/or share confidential information.

-Spear Phishing: It is the same approach as the phishing emails, but the attacker has defined a specific victim and/or group of victims to attack. The 'regular' phishing does not target anybody specifically or intentionally, but would aim to get anyone **(**Terranova security, 2020**).**

-Vishing: a name for a phishing method that uses voice communication or telephone to commit the attack. Voice phishing is mostly conducted by using a phone call. Combining the call, which nowadays can be done through other voice communication channels, with social engineering attempts to gain access to confidential and or/business sensitive information which can be private, financial and other relevant information which could lead to gaining the required access to a system within an organization. This type of an attack is called 'vishing' as a combination of words 'voice' and 'phishing' **(**Terranova security, 2020**).**

-Baiting, this type of an attack relies on the human nature to uncover and explore. Basically, an attacker plants a bait for the target(s), in a form of portable device (USB, CD, Hard Disk etc.). It often has a sticker and/or some form of marking on it baiting the target to check what is the content on it. The examples can be, 'list of employees suggested for a bonus', 'salary', 'confidential project' etc. Once the portable device is plugged in, it infects the computer with a malware intended to gain access and/or information **(**Terranova security, 2020**).**

-Pretexting: This type of an attack requires preparation as it is based on the creation and use of a made-up scenario (the pretext). When the target is attacked, the pretext is used to get in touch with the targeted individual(s) in such a way that increases the chance the target will share the required information and/or lead to actions to be performed by the targeted individual(s) which are usually not expected and would not be likely to happen without the use of the mentioned pretext **(**Terranova security, 2020**).**

2.3 Information disorder techniques (IDT):

Information Disorder Techniques (IDT) are methods of social engineering attacks that are based on information disorder. The IDTs are an addition to the more commonly known social engineering attacks where the targets are a broader audience with the intent to expand their influence on the society and drill down to an individual that operates in private and within an organization.

To better understand what IDTs are, below are the most common baselines to construct attacks using IDTs:

Misinformation/Propaganda: This type of activity is not usually perceived as an attack, even though in information security it can be regarded as such. Propaganda is a tool to spread skewed information to grow an ideological base. It is often based on facts, but the facts are used selectively without realistic context, which ends up being a misrepresentation of the actual situation. Misinformation is a method of presenting real data in such a way that it manipulates reality. It can often consist of fully inaccurate data as well (Poremba, 2021).

Although conventional social engineering techniques and approaches are perceived form an industry or business-related perspective, deception by serving misinformation with the intention of inciting action or lack of the same is a very common form of social engineering in the past three years. The perpetrator in this case uses means of deceiving the individuals consuming the information with the intention to sway their mindset into believing that the inaccurate information is actually to be trusted. (Poremba, 2021)

Disinformation: is false information designed to mislead and distort the truth. It is often used by nation-state actors to plant the seeds of false information on social media outlets, which is then spread until the false information is considered truth. (Poremba, 2021)
Propaganda/misinformation and disinformation are very similar but different in some ways. Disinformation tends to target the emotional component of an individual through representing correlated ideas and creating false narratives. For example, an ad shows that X number of people die as a result of terrorist attacks for every Y number of people in the U.S. that die from lack of health care. Both statements alone are true, but they have no connection to each other. Yet, the person who created the ad wanted to create outrage. (Poremba, 2021)

There are three common elements used to manipulate information for social engineering purposes:

-**Missing context.** Information is presented in a misleading way, or some vital facts are missing. This is commonly manifested on social media as presenting a photo that has nothing to do with

its caption. For example, a picture showing violence on a city street with the caption: “See what is happening across America today!” But the picture was, in fact, taken in a European city.

-**Deceptive editing.** Here, the threat actor is taking something that was once a genuine photo, video or illustration of a media story or event but editing key elements, so it distorts the reality to create a different message.

-**Malicious transformation.** This is the most serious of the three. Videos are altered through AI to create something fake that appears real. These are deep fakes. Threat actors use these techniques to push an agenda, whether it is a ransomware campaign for financial gain or to manipulate social outcomes like elections. (Poremba, 2021)

Social Engineering Attacks are carried out using an Attack lifecycle as demonstrated in Figure 2:

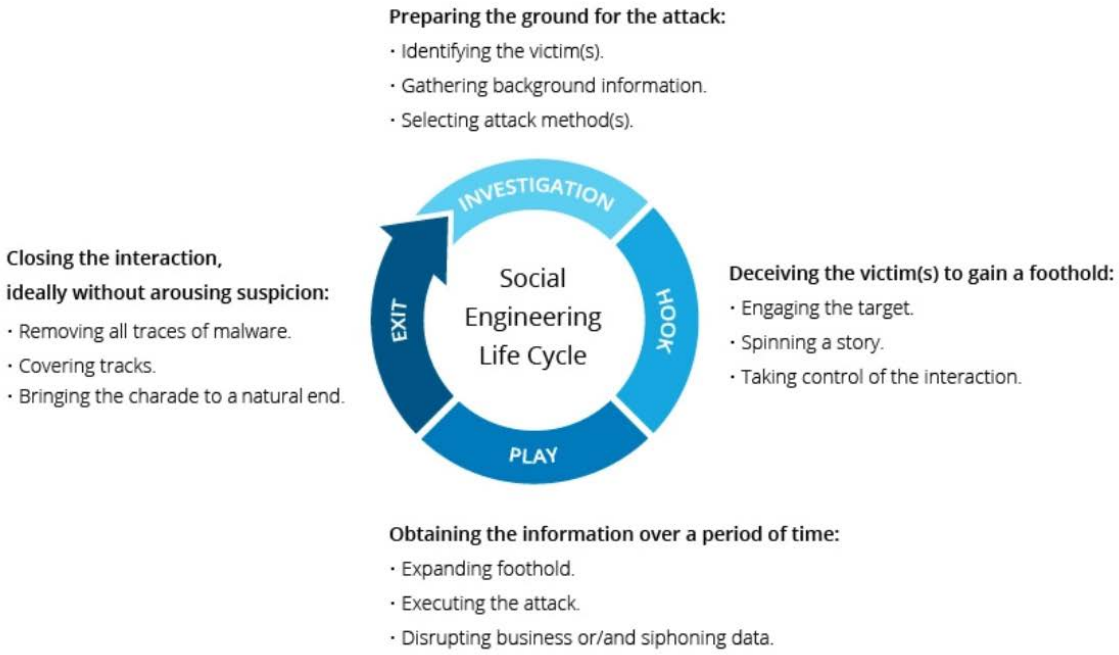

Figure 2 Social Engineering Life Cycle (Imperva.com, 2021)

The Social Engineering Life Cycle gives an overview of how attacks and actions based on Social Engineering are executed. A derived conclusion is that Social Engineering requires effort in preparation and execution of the steps, which says that Social Engineering like any other attack and/or malicious activity is premeditated.

The bottom line is that IDT gain momentum based on the human factor using psychology, living environment, and general atmosphere to influence humans into doing what the 'attacker' or a group of them has intended. When IDT is used, it can directly affect the perception of real-life events i.e., the Brexit vote in 2016, the Election of the U.S. president in 2016, election of the president in Trinidad and Tobago, the anti-vaccination movement, the pandemic caused by the SarsCov-2019 (COVID-19) and other.

2.4 Information Disorder, real life effect, a high-level overview

This part of the research will provide a high-level overview of how social engineering is used to cause information disorder through IDT and what real life effects and consequences come from the same. Apart for the effects and consequences that occur in the real realm that do not have consequences leading to the security of an individual, the society and state, we will be looking at the ones that caused consequences that influenced the lives of ordinary people and the way they go about their lives.

The Election(s)

It is known that on the political world information is crucial in obtaining support of voters in countries that have a democratic system. In the recent 15 years, there have been some very surprising election and referendum results that were directly influenced by the new technology residing on the internet. Although at first thought, some might say that sharing information on the internet is not forbidden, the actual actions taken are a part of much bigger effort to sway the results in a certain favor. To explain this in more detail, we turn to the three examples in Trinidad and Tobago, 2010, Brexit 2016 and the U.S. presidential elections of 2016.

During the campaign for the president in Trinidad and Tobago got the elections held in 2010, a company previously known as SCL Group, that changed their name into Cambridge Analytica (CaA), was hired by the Indian United National Congress (UNC) party that had the intentions to defeat their major opponent and at that time the president of the country coming from the Afro-Trinbagonian community and the leader of the People's National Movement party (PNM). The CaA had analyzed the structure of the voters, their overall beliefs and values and

had started a campaign to sway the elections in the favor of their clients. For the campaign itself they have used online platforms, predominantly Facebook and You tube.
The campaign was non-political and was calling all citizens from 18 to 35 years of age to avoid supporting any political party or program through a campaign called 'Do So'.
The result of the action was based on the premise that the younger voters that make up 40% of the total voters registered, should not participate in the elections. In the background, based on the analysis performed by the CaA, they anticipated that the Afro-Trinbagonian voters would not take part in the election, while the youth form the Indian community would do as they were told by their parents. In the strategy worked and the elected president was from the Indian community. (The Great Hack, 2019).

One of the biggest political upsets of the last decade occurred in Great Britain (UK) in 2016, when in June of the mentioned year the results of the referendum had led to UK to leave the European Union (EU). Apart from the real reasons some individuals have voted to leave, the so called 'undecided' voters were targeted into making their decision to leave on social media platforms, mostly Facebook, Twitter, and YouTube. The platforms displayed information, ads and videos that had spread misinformation to cause the recipients of the same to make their decision based on their emotionally driven impulses. The ads mostly contained information on how much money is spent on being a part of the EU, without anu insight to how much gain of the same is being returned overall. But the most significant emotionally based adds were talking about 'Taking Back Control" and preventing the increase of immigrants coming to the UK. The most efficient add was shared mostly on Facebook was stating that the county Turkey (Turkiye) will cause the flow of over 70 million new immigrants into UK once they join the EU. In some of the ads, the actual joining date was suggested to be within a few months after the referendum takes place (Cadwalladr, 2019).
In the movie that explains how the campaign had played out and what was being done in the campaign being managed by Dominic Cummings, it is clearly shown that the campaign has relied on targeting voter via online platforms. As it is explained in the movie, the goal was to target the eligible voters which have never and had no intention to take part in the election, actually are motivated to do so. In the movie the estimated 'undecided' voters were mentioned to be over 3 million (Movie Brexit, 2019).

In the post analysis phase conducted by journalists, they have found that the voters who voted to leave, were actually coming from the parts of UK, where the number of immigrants was the lowest, while in places like London which has one of the highest rates of immigrants, the vote was to stay within EU (Cadwalladr, 2019).

The Presidential election in the United States of America (US), have resulted in an unexpected result, having the republican candidate, Donald J. Trump getting elected. The reason why the unexpected result happened is based on the standard analysis of potential voters that had the Democratic candidate more favorable in the polls before the election actually took place. Among the verified voters when surveyed, the likeliness of voting for Trump were less by 3% on a national level (Doherty et al, 2018). Although the pols on many occasions can give a false sense, this was an indication that something was not right overall. As the analysis of what happened came to light, and it was clear that the majority of 'pro Trump' voters have not selected their candidate there because they support his agenda or the political party, but due to the fact that they have been swayed to vote in his favor.

Just like in the previous two cases, there were several influential factors coming from the online platforms, mostly the social media, Facebook, Twitter, YouTube and other.

In this specific campaign, we have the companies like Cambridge Analytica participate in spreading misinformation and targeting 'undecided' voters, but also a high surge of spreading disinformation and creating a platform for conspiracy theories in the process. The goal of all of the mentioned actions were strong enough to move the attention form the actual real and proven events that the Republican candidate has taken part in i.e. The 'All Access Hollywood' incident or the fact that the candidate was unfaithful to his wife days after she gave birth to their child.

On the other hand, a conspiracy that had no base or merit called 'Pizza Gate', alleging that the Democratic party along with their presidential candidate Hillary Clinton, were running a child sex-trafficking ring out of a pizzeria in Washington D.C., had the public swayed (Kang and Frenkel, 2020).

This specific 'Pizza Gate' conspiracy was amplified by a video of a pop star Justin Bieber where in the video he is touching his hat, while in the comments to the same he was asked to touch his hat if he had been a victim of the 'Pizza Gate'. The result of the popstar having a following

worldwide, in tens of millions of followers, prompted the information to be spread like an uncontrolled fire.

This example is one out of many which have developed over the time as the campaigning for the presidential election was going. There is no definite explanation to why such a strange story would take precedence over real scandals, but it is clear that the targeting of humans online has led to them deciding based on the information shared with them.

The Anti-Vaxxer movement

Over the past decade and some years before, in the general public discourse a question whether children should be vaccinated against measles, mumps and rubella (MMR) has been circulating. This idea comes mostly from the fake scientific study by Andrew Wakefield and other colleagues. The study form 1998 suggested that the use of MMR vaccines caused autism inf children. As the study was published it had created a worldwide anti vaccination movement. Having ground to support this idea, although the same study had not been replicated anywhere else, many parents ended up skipping on the vaccination of their children due to scares of their children being autistic as the cause of taking the vaccine.

As a result of the publications, the movement gained momentum in many countries and the number of children being infected by one of the three mentioned diseases was on the rise. In the UK it was recorded that a drop of children in vaccination form more than 96% had plummeted to only 87% being immune, while at least 95% of immune children is required for the 'heard immunity', which would protect the general public. In the 53 countries of the World Health Organization (WHO) for the region of Europe, has recorded 5273 cases of measles in 2016, while two years later in 2018 the number grew up to over 83,500 cases (Burki 2019).

The rise of infections and the global anti vaccination campaign was supported by the online platforms. This all comes after the study of Wakefield was proven to be false and after actual proof was shared to clearly indicate that the data in the study was forged. In 2011 the British Medical Journal has officially published a study showing that the research or Andrew Wakefield was fraudulent (Al Jazeera, 2011).

Even though the initial information surrounding the MMR vaccine was proven as false, even today we have a strong anti-vaccination campaign mostly present on the social media platforms

and conspiracy theory supporting websites. At this moment it is easy to find and explore more than a hundred of official studies trying to detect a link between the MMR Vaccine and autism, which confirm their safety, yet still many parents simply refuse to trust the same and don't allow their children to be vaccinated (DeStefano and Shimabukuro, 2019).

The basis of the decision not to vaccinate for MMR along with other types of vaccines is simply due to the fact that parents are emotionally affected by the wellbeing of their children. The construct is that the parents' belief that children do not benefit from the vaccines (Benoit and Mauldin, 2021).

To explain the phenomenon in more detail, several research studies, have indicated that the content on social media about vaccines and their effect on children uses information that cannot be proved, does not present any evidence, but shows specific cases of children adding media content like images, videos with narration and other tools to cause a human reaction (Ortiz-Sanchez et al, 2020). All of the content in the end creates confusion and leads to several parents refusing to vaccinate their children, which also endanger the ones who are not of age to receive the immunization and in the end has serious potential consequence on the society especially in the health security aspect of it.

The example and cases displayed above are just a small part of previous and ongoing attacks of the 'new generation' using IDT which target the human as the weakest link. As shown in the examples, the information disorder has led to real life consequences which affect all within the society, the state and even around the world. Further on we will turn the spotlight on the information disorder during the first two years of the COVID-19 pandemic, to give more insight on how humans are affected by the disorder in the digital realm.

2.5 The Information Disorder in the COVID-19 Pandemic

As the world was going through a pandemic caused by the COVID-19 virus the lives and health of human beings are directly affected by the information shared online. In this part of the paper, the goal is to analyze how IDT methods were used to spread wrong perceptions related to the pandemic and what effects it had on the consumers of information through IDTs related to the pandemic.

People all around the world have found themselves in a new situation caused by the spread of a virus that can seriously harm human health and in some cases even cause fatalities. Countries around the world have taken actions and very similar approaches to try to contain the spread of the virus among the public. The required actions for citizens of most countries included limitations of movement within the country and local area, limiting travel, and taking precautionary measures i.e., wearing protective masks, enforcing social distancing rules.

This new situation gave merit to malicious actors to spread misinformation which directly compromised health and security of countries. The information shared online from unreliable sources was taken as trusted sources of information. Initially before the vaccine had been discovered, the information shared consisted of information downplaying the effects of the virus, messages that depicted obligatory protective measures, promoting false remedies against the virus, and even promoting the notion that the virus does not even exist. (Hansson et al, 2021)

To better understand the evolvement of IDT techniques being used, we analyze how the approach of the same to the COVID-19 pandemic. Based on changes in representation of topics related directly to the new developments related to the information shared about the disease and other relevant discussions about the pandemic we can see the effect of the IDTs in full swing.

Initially after the pandemic was declared as such by the World Health Organization (WHO) on 11th March 2020 at least six types of harmful information were detected in the first three months of the COVID-19 pandemic:

- Messages that depicted recommended or obligatory protective measures (e.g., wearing a mask, using a hand sanitizer, observing lockdown) as either harmful or unnecessary. This was often done by appealing to fear and casting doubt on official requirements/advice.

  An example of such a case is information published on an Italian blog called 'Mondo Sporco' on May 30th, 2020, on which a picture of a protective medical mask was displayed with a message that 'The mask seriously damages your health'. Further on the explanations given claim that the masks prevent sufficient oxygen from being inhaled. (Hansson et al, 2021)

As an example of a real-life consequence related to the incorrect information, we were able to see several incidents that occurred all over the world, in which some individuals caused a disturbance and harmed the property of a store in which masks have been sold. More specifically a lady that witnessed falling into a 'rabbit hole' on the internet via media outlets and social media platforms which propagated that wearing masks should not be made mandatory. Melissa Rein Lively openly spoke about an incident she caused in a 'Target' store in Arizona U.S. Based on her testimony, she explained that due to the circumstances surrounding the restrictions due to the pandemic, with all the information she had been reading on social media platforms, had led her to destroy the racks in the store with face masks being sold. Her case escalated into having her family call the police as she continued to confront all due to the restrictions imposed because of the pandemic. At the moment she realized that she could lose her family, she decided to take action (Tomić 2021) (Cumberbatch 2019) and seek help in getting out of the mentioned 'rabbit hole' (CNN, 2021).

One additional case that caused issues in real life was the case of fake messages sent over Twitter, stating that hospitals in Lille, Brest and Nantes in France were overloaded and that there is no point in visiting the same as medical treatment is not possible at the mentioned medical facilities. The message was refuted by the mentioned medical facilities, but the message had influenced some potential patients to refrain from visiting them at the time the messages were being spread. (Hansson et al, 2021). There is no evidence on what the actual result of this message is as there in no proof of how many potential patients have suffered or even succumbed due to the IDT in this case, but it is clear that such information can directly affect the wellbeing of the individuals that depend on the mentioned healthcare institutions (Hansson et al, 2021).

- Messages promoting the use of false (or harmful) remedies against the virus – essentially giving scientifically unfounded medical advice.

  Such a message first time shared on 18th April 2020 with thousands of shares on Facebook in France, claimed that the famous French Roquefort cheese is a good remedy for the protection

from COVID-19. Scientists from the French National Scientific Research Center (CNRS) later refuted this message. (Hansson et al, 2021)

On other case recorded in Estonia had a false remedy published on Facebook, under the title 'Health with Natural Products-Coral Club' within the first few days of the pandemic, to be more precise on the 16th of March 2020. In the post the advertisement was promoting a colloidal silver product called 'Silver-Max' stating that it helps protect against the COVID infection. The ad was distributed and became known to many, even though there is no evidence that colloidal silver can help with any type of medical condition, while on the other hand a potential overdose is more likely, considering the fact that many people were anxious due to the unknown circumstances caused by the pandemic (Hansson et al, 2021).

- Misrepresentations of the transmission mechanisms of the coronavirus that could have made some people falsely believe that they were either immune to or unlikely to catch the coronavirus due to some personal factor such as their blood group, or practices such as smoking, healthy eating, or drinking hot beverages.

  On 18th March 2020 an Estonian news portal Lounaeestlane.ee, published a story titled 'New study: People with blood group A more susceptible to coronavirus' and mentioned that A is the most common blood group among Estonian population. Some other Estonian media later on shared the news. This information is actually not based on any real scientific analysis. On the other hand, lack of such studies reflected information shares as it was very hard to fact check such information. (Hansson et al, 2021)

- Messages that suggested that COVID-19 did not exist or was not severe, that the overall risk of catching it was low, and that the pandemic would end shortly, thereby potentially lowering the perceptions of health risk and discouraging cautious behavior. On 12th April 2020 a man claiming to be a veterinarian and understanding virology, on a Facebook post declared that viruses do not exist. This information among Facebook users in Estonia was share several hundreds of times. (Hansson et al, 2021)

  Even though verbally many individuals have heard that some don't believe that the pandemic is one at all, it is on the social media where many surprising and highly imaginative theories have developed. They are conspiracy theories that had spread very easily on all social media

platforms, personal podcasts and other IDT spreading platforms. Many people in the U.S had believed that the pandemic is caused by the government which they referred to as the 'deep state'. (BBC, 2020) Based on the mentioned misinformation and disinformation being spread, many individuals actually have let their guard down and, in the end, possibly have helped spread the virus among the society, directly compromising the health security of the same.

- Scammers exploited the uncertainty created by the pandemic to trick people into buying fake protection against the virus or into revealing their confidential information.
  As recorded on 9th March 2020, form several Facebook users in Norway, an ad appeared advertising 'high-tech masks' to better protect against the virus that caused COVID-19 pandemic. Although it is uncertain whether any users have actually received such a mask, it has been shown that several customers who ordered the mentioned masks, wasted their money believing that they would be a 100% safe form any infection if the 'high tech' mask was used.

  One additional case from Estonia recorded in March 2020, has shown that situation like the pandemic when using IDT can lead individuals into being scammed. While using Vishing, a man stated that due to the pandemic, the pension delivery would be stopped. The caller for that reason asked the pensioners to provide their personal account details to allow their pensions to be transferred to their accounts. Even though there was no further research to detect if anyone was actually harmed by this attempt, it is clear that individuals with malicious intent will use every opportunity for their own gain (Hansson et al, 2021).

- Certain individuals and groups were subjected to harassment/hate speech as the alleged spreaders of the virus. As it was reported on 21st April 2021, thousands of reports of harassment since the start of the COVID-19 pandemic were reported by people living in USA, but who are of Asian descent. (Ruiz et al, 2020)

  In Finland a case of harassment towards refugees was recorded. During the last days of April 2020, a group of refugees that were allowed to leave the refugee camp, which was in

quarantine due to several infections among the refugee community, were exposed to hostile behavior (Hansson et al, 2021).

As new information, developments and facts about the COVID-19 pandemic arose, new approaches and information shared through IDT emerged. As the vaccines for the protection against the COVID-19 virus started to get approved and put in use since December 2020, use of IDT created a whole new anti-vaccination movement.

The anti-vaccination movement related to the COVID-19 pandemic, is considered a very successful movements considering that it gained momentum and millions of followers around the world. Initially, the success of the IDT techniques was systematic in the form of taking a bottom-up form of media transmission, rather than the traditional top-to-bottom approach. This basically means that the information did not come from a verified realistic trusted source, but rather form anonymized and in many cases fake account on social media platforms. (Bajwa, 2021).

The process of the approach anti-vaccination actors taking the bottom-up approach is a form of a hybrid attack and can be explained as follows: First, it anonymizes the creators of the disinformation narratives and products, leaving other agents, such as blog creators to shape and influence it as it crosses networks. Such approach can be found on an example of how by ‘refraction’, which has been seen in use of Russian conspiracy networks in their intent to support the ‘Russian geopolitical objectives in Syria, by reproducing the same message through different political perspectives to create an illusion that there is an agreement form different sides and different even opposing narratives. (Bajwa, 2021)

The ‘refraction’ trend has been identified in the anti-vaccination campaigns on online platforms, coming from different sources seemingly creating and agreeable and ‘supported’ narrative.

One other technique known as ‘astroturfing’ (a deceptive ploy that aims to give the false perception of grassroots support to an issue for the purpose of misleading the public into believing that their opinion is shared by most people) (Longley, 2020), was used to support the

anti-vaccination movement. Basically, the Russian disinformation agents created tweets that were used in a form of creating grassroots advocacy for vaccine refusal. (Bajwa, 2021).

The vaccine hesitancy was recognized earlier by the WHO in terms of the analysis and approach to vaccines being administered to children, related to the mentioned MMR vaccines. In that respect the WHO has recognized the vaccine hesitation to be one of the top 10 threats to global health. The risks associated with vaccine hesitancy had shown its full form after the vaccines for COVID-19 were issued for emergency use by December 2020. The amplification in hesitancy in the general public was a direct result of the information being circled on social media and personalized custom content platforms i.e., pod casts. The mentioned platforms were able to deliver the message by using of a 'one-step-flow' approach to persuasion. This differs from the before known 'two-step-flow' persuasion. The difference is that in the 'one-step-flow' approach it was not a form of communication between the two sides, where one side would adopt the opinions of the other party and use the same to continue the process, but rather as a one-way street, the content providers were able to deliver the message on vaccination hesitancy just by serving their content (Hughes et al, 2021).

The examples discussed are an obvious pointer to the fact that the human is the weakest link susceptible to becoming persuaded only by being exposed to the content available on online platforms.

One other research hypothesizes that on the social media platforms, the outlined risks, based on misinformation, associated with vaccines appear more realistic and tangible than the identified risks associated with not being vaccinated. One of the sentiments is that the success of the vaccine is related to the absence of disease. The success of the anti-vaccine propaganda is based on the fact that it tends to be more focused on emotions and personal anecdotes with powerful imagery which contradicts the actual scientific empirical strategies used by pro-vaccination literature and platforms. The emotional approaches tend to be more appealing to social media users and are consistent with other content that tends to be shared on social media (Hughes et al, 2021). Just another argument in favor of human factor being the most exploitable point within the digital realm.

The ITD 'attacks' in the time of COVID-19 pandemic based on a social engineering approach were and are still being used misrepresented facts and information that creates the false and

wrong perception of the virus and the situation that evolved around it. This directly undermines the security of the general public, directly endangering their health.

One other problem that was created was the unrest of the public. Social Security is disrupted in several countries around the world. The people are against the rules enforced and especially the ones where vaccination is mandatory for individual to be allowed to go to work, access public spaces etc.

An example of such social disturbance are protests held in Brussels, Belgium in November 2021. The crowds were protesting COVID-19 related restrictions, mostly due to the outbreak of the Omicron variant, which was spreading faster than the previous variants of the COVID-19 virus. The peaceful protests had turned violent causing damage to the cars and shops in the part of the city where the protests were taking place (Guardian, 2021).

One other example of such social disturbance are protests held in Vienna in December 2021. Several thousands of protesters went out into the streets to protest against mandatory vaccination and rules that apply to the one who are not vaccinated. Like the ones in Belgium, they turned out to become violent, causing damage, endangering those who found themselves there and paralyzing the part of the city where the protests were being held. (Armstrong, 2021).

There were many other violent protests that occurred in several countries in Europe, U.S., Australia etc. The basis of the protests was amplified by the information shared on social media, which has led to a direct event in real life of individuals.

When it comes to B&H, the situation was not much better except that there were no violent protests recorded during the time restrictions were imposed. On the other hand, the information on social media platforms and even some mainstream media caused the perception of everything related to the pandemic to be perceived as some kind of a political game. Form wearing masks, maintaining social distance to the acquisition of vaccines and distribution of the same, was followed by a lot of misinformation and disinformation being spread on online platforms.

Although there are no official studies to show what kind of information was shared, it is based on several television pols held on the streets of B&H cities, that it is easy to see that the general public was not following the official health officials and their recommendations. Even though

many have obliged to the rules and had taken the vaccine due to other requirements not related to health, the general perception was that the trust in official expert instructions were not trusted. (O Kanal, 2022)

The general perception of the COVID-19 related information had a bad influence on the death rate from the virus in B&H. According to the data analysis by STATISTA, B&H ranked third in the world as of 26th April 2022 with a rate of fatalities as high as 4.8% (STATISTA, 2022).

Even though there is no official study to confirm the exact impact online platforms had in causing the percentage to be one of the highest in the world, it is safe to say that the total percentage was influenced even in the smallest percent and by correlation shows a direct influence online platforms can have to the security and wellbeing of humans in the real life.

2.6 Algorithms, bots and their amplification

One of the main ingredients to how our online experience is shaped is fully dependent on the algorithms that run in the background of all the online sources we access. Algorithms are defined as a procedure for solving mathematical problems in a finite number of steps that frequently involves repetition of an operation (Merriam-webster.com, 2022). Such operations, algorithms, use the computing power of modern computers to provide a more efficient calculation in order to get the required result quickly. This is the way initial search engines produced results and ranked websites when an online search was conducted.

Further to the development and increased use of algorithms that support online services, they have incorporated Artificial Intelligence (AI) into the picture. AI is commonly defined as "a system's ability to interpret external data correctly, to learn from such data, and to use those learnings to achieve specific goals and tasks through flexible adaptation" (Haenlein, Kaplan 2019). So, it is not only that the algorithms produce a computational result, but they also actually interpret the results of the computation and further on drive what users experience online. Just as an example every person has witnessed online, if you searched for something even after you stop, ads on websites and social media platforms will advertise a product or service you were interested in.

Initially the algorithms were perceived as a positive development incorporating AI into it as it can be put to good use. One of the examples is the use of IT and algorithms incorporating AI in delivering enhanced solutions for the management of infectious outbreaks like the recent pandemic, including tracking potentially and possibly infected individuals, tracing of contacts, the targeted delivery of healthcare and the ability to link across databases to elicit important patterns, such as health status and recent travel history (Green and Clayton, 2021).

All of the information we actually share using any digital system and especially the ones online consist of data, the actual components that in context form a set of information and the metadata. ‘Metadata describes other data. It provides information about a certain item's content. For example, an image may include metadata that describes how large the picture is, the color depth, the image resolution, when the image was created, and other data. A text document's metadata may contain information about how long the document is, who the author is, when the document was written, and a short summary of the document. Web pages often include metadata in the form of meta tags. Description and keywords meta tags are commonly used to describe the Web page's content. Most search engines use this data when adding pages to their search index’ (techterms.com, 2022).

In addition to the metadata that comes from files, websites, and the resources we interact with, we as users also generate metadata which is collected by the online resources we access: ‘When we are on the Internet, we are creating data that is being collected and analyzed. That's done through tracking cookies, which are stored in a user's web browser. Tracking cookies store data about what we do on a given website in a browser so that they can inform the website of our previous activities on subsequent visits. This kind of tracking is interesting for several purposes including the advertising industry because it allows them to get better access to Internet users’ (Wünsch 2013). In essence the biggest and largest internet companies i.e., Google, Facebook, Twitter, YouTube, and others gather data about our online navigation by collecting our metadata, which essentially is a part of our digital trail online.

However, anything that is managed by humans has the other side of the coin which is the negative side. More specifically the negative side of the use of AI and algorithms. To sum up what we have mentioned above, it is clear that we as users provide large data sets of metadata

which are processed by the mentioned algorithms and enhanced by the AI which serves the purpose of gaining control over our online experience without being able to actually notice what is happening. So, to better understand what is happening we need to consider the following: 'Big datasets are necessary inputs for training AI systems. They are also one of the key ingredients that can lead to bias, discrimination, and unequal outcomes. To understand data, we need to consider it within a bigger context. Data are abstracted elements or representations that seek to categorize and measure phenomena. While the data is often seen as conveying the facts, data is not purely objective. How data is collected, stored, and constructed, for whom and for what purposes, reflects the values of those compiling the data. Data are part of a "complex socio-technical system" that forms a "data assemblage" which includes "ideas, techniques, technologies, systems, people and contexts" that "evolve and mutate over time". All of this exists independent of AI; however, AI can amplify and codify these data agendas in ways that have profound societal implications.' (Ingram, 2021) In essence, the algorithms and AI within it is not a purely mathematical/computing in nature, with the intention of providing an end result. The algorithms are used to target users with a hidden agenda behind it, thus the negative side of the coin comes into the picture.

2.6.1 Voluntary informed consent-the golden standard (Algorithms, bots, and their amplification)

One important thing to remember is that initially all is done with the consent of the user, where the user accepts the 'terms and conditions' under which they access data is provided on the online systems and services we access. Consent is a mechanism based on which users grant access to their information and the 'golden standard' is the informed consent. In essence it is a voluntary agreement between parties with equal bargaining power in the physical realm, while online it is much more challenging. The challenge arises due to the fact that sovereignty of an individual and even a country does not apply as it applies to law regulations in the physical realm (Ingram, 2021). The second challenge arises from the fact the user online does not have real choice since access to the online resource requires acceptance of the informed consent through accepting cookies and/or terms and conditions, if they want to access the content considered: 'While we may legally "consent", we actually have no real choice'. 'Our data, including what we share publicly or within a selected online group, as well as the clicks, swipes, geo-location

information and other “meta data” becomes the cost to participate in modern society’ (Ingram, 2021).

After the data is collected is when the algorithms and AI kick in. As mentioned earlier AI and algorithms process a large number of data sets in a very quick time providing the wanted result for the large corporations, organizations and in some cases smaller entities.

The algorithms are actually a ‘black box’ for the user and the way they are constructed is often treated as a ‘trade secret’ by the owners of the same. However, sometimes even the owners are not exactly sure how the AI operates within the algorithm since it has a learning mechanism based on which results are provided: ‘AI cannot always explain its decisions. Part of the power of deep neural networks, which is the focus of much current AI research, is that they learn on their own, in a “black box” that even their creators can’t fully grasp’ (Ingram, 2021). In essence, the algorithms and AI are used by big technology companies (i.e., Google, Amazon, Facebook etc.) which have become the new ‘gatekeepers’ through the extensive use of their online platforms. This means that their use of such technologies enables them to control the content and type of information to which the information consumers are exposed to. Additionally, the technology itself is a ‘black box’ since access and review of the same is not possible. The corporations that developed and use the algorithms and AI, don’t allow any access to their technology in most cases under the premise that they represent a ‘trade secret’. (Vajzović2, 2020)

2.6.2 Algorithm bias (Algorithms, bots, and their amplification)

Apart from the fact that control over the algorithms cannot be established, the inability to fully understand and operate the algorithms also poses other challenges, like the legal aspect of taking responsibility of actions that are a direct result of the algorithm’s operations i.e., detection of potentially infected during the pandemic.

Amongst other problems of AI running within the algorithms, there are several other that directly influence the results:

- ***First, there may be concealed bias.***

- ***Second, algorithms cannot screen entirely autonomously, for a number of reasons. One is context. In English, for example, words can be modified by context or intonation and irony can turn a word into the opposite of its nominal meaning. Humans understand context and metaphor, but this is hard to encode. Another that words can be used to signify something that is obvious only to initiates.***
- ***A third is that language is fluid; English, for example, is spoken in many dialects and accents, which constantly evolve.***
- ***A fourth is that harmful misinformation can be presented in an acceptable form; spurious information about the dangers of vaccines can be presented in a pseudo-scientific manner that makes it appear credible.***
- ***A fifth is that it may be difficult to define when religion becomes political, and when an appeal for spiritual struggle is actually a call for extremist behavior.***
- ***A sixth problem is that terrorists can change platforms and spread different messages across multiple platforms, and terrorist organizations can morph into new forms, so that an algorithm may become increasingly inaccurate unless it is constantly retrained with new material.***
- ***A seventh problem is that there is a fundamental conflict between the business model of social media companies, which is based on advertising which is generated by viral content, and the idea that they should exclude posts that generate a lot of traffic.***
- ***An eighth potential problem is that the reliance on technology companies to use AI-based algorithms to moderate content amounts to the privatization of censorship. This would have mattered less in the past, but now that technology companies are, in effect, by far the largest media corporations in the world, it matters a great deal***. (Green and Clayton, 2021)

So, how does the algorithm and Ai incorporate with the IDT. The algorithms create bias and in addition influence the information we are served when online in the digital realm. In order to better understand the bias and the serving of information, we need to look at the fact that all AI is actually built by humans. Depending on which company is developing the AI, the intention is to drive the users to re-visit their platforms in order of obtaining larger numbers of ‘hits’ which directly affects their business model and potential advertisement revenue. When bias is added to

the formula, it can lead to the users receiving content which can drive them to be victims of IDT attacks.

The bias starts from the actual designers and developers of algorithms. Just to understand where it comes from consider the listed facts:

- According to a global AI talent survey, there are 22,400 people working in AI based on those who publish research, however 36,524 people self-report as AI specialists on Linked In. It is an elite community comprised primarily of highly educated white men.
- There are numerous studies and reports about the lack of gender diversity in the field of computing science. AI is even less representative with women making up only 16% of AI researchers.
- Decades of research also shows that who makes a technology impact what is constructed, and that gender plays a role in creating and reinforcing stereotypes through technology design.
- In Algorithms of Oppression, Safiya Noble challenges the notion that commercial search engines, fueled by powerful AI algorithms, provide an unbiased and value-free service. She presents numerous examples that highlight Google's racial and gender biased search results, particularly as they impact Black women. Search results seek to further entrench and protect the interests of those already in power, misrepresent marginalized groups and keep women and minorities locked out of participating in the creation of technology.
- Industry generally focuses on research with commercial applications which can mean that areas important to the public good, but not commercially viable, such as public health, may receive less attention.
- Universities are not only looking to corporate sources for funding, but they are also becoming increasingly interested in holding patents to commercialize research, leaving scientists with an escalating imperative to engage with market or economic considerations.
- In the algorithms themselves, the models are far from neutral, but rather "opinions embedded in mathematics". The choices made in designing AI models are another way that AI is socially constructed.

- Popular algorithms are also reused across applications, further amplifying, and encoding their reach. Underlying all of this, is tacit agreement that efficiency, prediction and optimization, the rationale to build AI, are worthy goals. Those higher order values both drive and are driven by the ability to harness big data. It enables a “data religion” which sees data as scientific, objective proof by which to explain the world. Like religion, the tribes within machine learning carry ideological perspectives that are encoded within the mathematical toolsets they choose to apply.
  (Ingram, 2021)

2.6.3 The filter bubble (Algorithms, bots, and their amplification)

Based on the listed facts which in essence create the based for bias and considering the fact that the math within is actually influenced to gain defined results. An example of such bias was proven on the hiring algorithm in amazon: ‘Amazon’s hiring algorithm, that discriminates against women for certain roles, was trained on datasets that privileged men’ (Ingram, 2021). Now, when combined with other IDT techniques, algorithms amplify the effects of the IDTs, create a very dangerous weapon online that can lead a user into a ‘rabbit hole’ on the internet by simply using subtle bias which is undetected by an average human being.

So how does the bias created by algorithms influence the online experience of an average user? To fully grasp the buildup, we first need to understand that there are two ways of how users get caught up in what is called a filter bubble. A filter bubble is a situation in which someone only hears or sees news and information that supports what they already believe and like, especially a situation created on the internet as a result of algorithms (= sets of rules) that choose the results of someone's searches (Cambridge dictionary, 2022). ‘The ‘filter bubble’ is a persistent concept which suggests that search engines and social media, together with their recommendation and personalization algorithms, are centrally culpable for the societal and ideological polarization experienced' when accessing online information (Burns, 2019). An average user gets into the filter bubble based on searches and content accessed online or by being targeted by a certain agenda. The actions that are performed online are not only limited to searches, but include other interactions like using the ‘like’ button, sharing content, making comments etc. One other important thing to remember is that the filter bubbles do not limit the possible search one can

perform online, but the content initially provided is based on suggestions given to the user on any of the platforms. The mentioned suggestions are the most likely content a user will interact with once one of online platforms are accessed (Tomlein, 2021).

2.6.4 Software bots (Algorithms, bots, and their amplification)

The algorithms actions are amplified by bots. 'Bots, shorthand for "software robots," come in a large variety of forms. Bots are typically automated in some fashion, either fully automated or human-in-the-loop' (Himelein-Wachowiak et al, 2021). When it comes to algorithms, the bots are the part of the AI component that amplify their effect on how a user experiences online interactions and helps maintain the filter bubble. This refers specifically to the automated bots mentioned in the definition above.

An example of such a case is the Brexit vote. One of the campaign managers for the 'vote leave' campaign Dominic Cummings, used the services of a tech company to target undecided voters and those who have never voted and lure them to vote for his campaign. Simply based on the metadata of users online, predominantly Facebook, they were able to use IDT in sharing information disorder, target the human emotions without them realizing it and influence their behavior in the real realm (Movie Brexit, 2019). The actions were amplified by the algorithms developed by the tech company whose services were used.

One other example of algorithms with AI capabilities was used to spread conspiracy theories about the US presidential elections in 2020, which on January 6th, 2021, led to a mass of people to storm and attack the US Capitol hill. All started on Facebook with conspiracy theory groups like 'QANON' and 'Proud Boys' which were being banned on the social media platform, but moved to other online platforms i.e., 'GAB', where they spread their misinformation and conspiracies. In the end it all led to a violent outbreak. (TRT World, 2021)

The same amplification influenced what people believed about the COVID-19 pandemic. Many users were led to obtain information from conspiracy theories and simple misinformation and disinformation instead of the actual facts and credible information. As an example, some studies

of information online showed that, when people were to enter the search for ‘COVID’ in the combination of the word ‘Truth’, they were more likely to receive information based on IDTs instead of the facts (CBSN, 2021).

We have seen that the algorithms that operate online are actually based on the agenda of the creators or companies that develop the same. The development of the algorithms and the AI within is built by mostly men considered to be of those who are ‘highly educated white males. The math within the algorithms can be and is influenced to produce a specific outcome and when combined with IDTs, creates a dangerous environment that targets the human factor and produces actions in the physical realm that have real and significant consequences. The algorithms do not actually produce the information, but they do amplify the ability of IDTs to reach the users online and based on this fact, the algorithms become one of the mentioned IDTs. The biggest challenge in the attempts to fight against the algorithms is the inability to legally hold the creators and owners of the algorithms legally accountable due to the facts that the internet does not recognize physical borders and because all users without a real choice accept the ‘terms and conditions’ under informed consent and in the end share their metadata.

2.6.5 Social Bots (Algorithms, bots, and their amplification)

Social bots are the non-automated actions that have the intention to drive the online experience of a user. One other definition states that ‘bots on social media are algorithms designed to communicate with humans and they are created in such a way that they replicate human behavior’ (Željeznik, 2021). However, the analysis of activities online has also led to defining human’s that perform agenda-based action online under avatar and alternate names as ‘bots’. These human bots are mostly politically, and agenda motivated to comment, share, distribute and conduct other actions online with the intent of discrediting, attacking, shaming and conduct other malicious activities against specific target(s), often ideological and political rivals on online platforms (Dan u životu jednog bota, 2021).

Bots in general are created to spread disinformation and misinformation, but also tend to lure the users to click on certain links and/or download certain files which have malicious software (malware) on them (Željeznik, 2021). However, in terms of IDT, the social bots, which are in our

focus, are both human and software-based bots that base their activity on attacking the human factor on online platforms. Just to clarify the difference form the AI bots, the software-driven bots are the ones that do not learn on experience, but are driven by a human bot, to direct users to the goals of their current agenda. Basically, depending on the 'current' goals defined, the human bot configures the activities of the software to support the mentioned goals, but does not include learning mechanisms powered by AI. The social bots operate through applications that replicate certain defined actions i.e., clicking like or dislike on post, comment etc.

When social bots are considered most of the targets are chosen based on the agenda of the operators behind the scenes which tend to apply fully devised plans and scenarios to achieve their goals. This specific form of social bots can be seen around the world, but their activities and IDTs are known to be used very often in the countries of ex-Yugoslavia and in that in Bosnia and Herzegovina (Dan u životu jednog bota, 2021).
It is not uncommon that the activities of the social bots are followed by agenda-based media organizations. These types of organizations are usually funded and supported by specific political parties and as such they are required to present the information in the way the political party needs it to be. Some of the activities in example can be glorifying their activities, attacking their opponents, or tending to drive the public focus on the topics and agenda they want the public to be steered in (Dan u životu jednog bota, 2021).

The social bots are usually a part of a larger network which devises a plan of action depending on the goal defined. On social media platforms the use specific methods to promote their stance is done by using hashtags, generating likes on a specific content and even on the comments shared. As it is known, the most liked content would be visible on any of the platforms, and it is usually the goal to achieve. On the other hand, it is also the case that negative tags are left on the content in the attempt to seem realistic as bots don't want to be caught by an average user on the online platforms, so if the can they give it a minus, thumbs down or use some other method to achieve their goal (Dan u životu jednog bota, 2021).

It is very clear that actions of the social bot networks get their action plan form an individual or a small group of leaders. The actions are disbursed to drive the information in the direction they

need it to be. Methods used aim to get the user's emotional triggers and basic human instinct based on which the public view would be formed. It is also not unusual that individuals that are a part of social bot groups, user their family and friends for their agenda. It is usually done by asking the family and friends to install certain applications on their devices, which the bots then use without the knowledge of the device owner to have multiple points of interaction in support of their activities. Sometimes the app is installed without the family member or friend even knowing it has been done so (Dan u životu jednog bota, 2021).

One of the ultimate goals of social bots is to get real users taking actions supporting their agenda, through their own accounts with their real names. This help additionally create a loop in convincing users that real people stand behind the agenda (Dan u životu jednog bota, 2021).

The social bot networks as IDTs have a very strong influence on the users that consider the content shared, comments made and other actions taken as a reference to making their own opinions on certain topics, accepting the view of other individuals, and applying the same as their own when considering political views. It has been recorded by research conducted by the Oxford Institute for the Internet, that social bot networks have been active and have had a direct impact on selection of political representatives in more than seventy countries all around the world. (Bradshaw, 2021). One of the most popular addon IDT used by social bot networks I 'astroturfing'. This method helps create false credibility that is perceived by the users on online platforms (Thomass et al, 2021).

The major difference social bots bring to the table when compared to other IDTs is the ability to take action anonymously. Being anonymous adds to the potential vulgarity and directness used when the bots are in action. The bots are usually committed to their activities full 'working hours' and are usually financially compensated for their actions. In many cases as recorded in Serbia and Bosnia and Herzegovina, the bots are hired as consultants or interns within an institution or political party, and their only job is to be a bot online. In the process as they do it under false names and avatars, with hiding methods like use of proxy and/or Virtual Private Networks (VPN) to hide their location, the bots have no limits in their vocabulary. As witnessed by several individuals that were targeted by social bots, they have seen very vulgar, personal,

appearance based, sexually offensive, ethnic based attacks on themselves. Although it is based on IDT, the victims of such attacks have seen and witnessed attacks by people online and on the street, believing that the content they read on the online platforms is actually, factually true (Dan u životu jednog bota, 2021).

2.6.6 Summary (Algorithms, bots, and their amplification)

Now that we have seen what IDT techniques can do and how they can directly influence the lives of humans, societies, the state and the world, the question is what can be done to 'fight back'? One of the ways that the retaliation can be achieved is through Media and Information Literacy (MIL)

MIL provides solutions on how the general public can be educated to process the information with critical thought, choose the information source, verify the information received and be on a constant lookout for the potential IDT threats. In the next chapter, we introduce MIL in more detail and conduct an analysis of MIL, which is intended to explain what MIL is and how it can confront the Social Engineering attacks performed using IDT.

2.7 Media and Information Literacy

MIL's importance from an information security perspective grows by the day in the world and societies today. However, the concept of MIL and what it represents is yet to reach the public in its full capacity as a response to the IDT techniques and overall information security.

'Media literacy, critical media literacy, media and information literacy, digital and media literacy, film literacy, screen literacy and so forth. All these subjects are based on a recognition that mediated societies require media-literate citizens. 'Literacy' has often been used synonymously with 'education'–media education, media literacy education, film education, screen education, etc. In the following review 'media literacy' is used for the most part; 'Media and information Literacy' (MIL) is the current term.' (Carlsson, 2019)

Initially the term 'media literacy' was coined in 1955 by a Professor Louis Forsdale, who had the intention of teaching his students on how to cope in the raise of media presence in their lives and in the process to gain the media literacy to be fully capable of processing the information received at that time through radio, television, and newspaper. Although there were several even

successful programs and lessons about media literacy over the years, it is in 2004 the UNESCO among other, defined literacy as ‘the ability to identify, understand, interpret, create, communicate, and compute, using written and printed (and visual) materials associated with varying contexts. Literacy involves a continuum of learning to enable an individual to achieve his or her goals, to develop his or her knowledge and potential and to participate fully in the wider society’. The emphasis on MIL form UNESCO was a response to the rapid growth of globalization, increase of internet usage, development of new information platforms, that were showing potential to mislead the recipients of information (Carlsson, 2019).

By 2007, UNESCO officially introduced the Media and Information Literacy as a term as the level of literacy required in the world of multiple information sources we witness today. In 2011 UNESCO has started projects that produced curriculum for MIL. The basic concepts of the MIL have evolved and have enhanced the efforts of other members of the academic community to get engaged in developing and promoting the mentioned concepts (Carlsson, 2019).

So, to better understand what MIL is and what it represents refer to the definition: “Media and Information Literacy consists of the knowledge, the attitudes, and the sum of the skills needed to know when and what information is needed; where and how to obtain that information; how to evaluate it critically and organize it once it is found; and how to use it in an ethical way. The concept extends beyond communication and information technologies to encompass learning, critical thinking, and interpretative skills across and beyond professional and educational boundaries. Media and Information Literacy includes all types of information resources: oral, print, and digital. Media and Information Literacy is a basic human right in an increasingly digital, interdependent, and global world, and promotes greater social inclusion. It can bridge the gap between the information rich and the information poor. Media and Information Literacy empowers and endows individuals with knowledge of the functions of the media and information systems and the conditions under which these functions are performed" (IFLA, 2011).

In the Declaration on the Importance of Media and Information Literacy in Bosnia and Herzegovina from 2019, the definition of MIL is refined as follows: “Media and information literacy refers to cognitive, technical and social skills and abilities citizens to approach, critically evaluate, use and contribute to information and media content through traditional and digital information and media platforms and technologies, with an understanding of how these

platforms and technologies work, how to occasion their use is governed by their own rights and respected the rights of others, how to recognize and avoid harmful facilities and services, to use information purposefully, media content and platforms to satisfy their own communication needs and interests as individuals and as members of their communities, and to practice actively and responsible participation in the traditional and digital public sphere and in the democratic processes." (DECLARATION ON THE IMPORTANCE OF MEDIA AND INFORMATION LITERACY IN BOSNIA AND HERZEGOVINA - 2019)

The curriculums, programs and educations that evolved in 'spreading the word' on MIL is in most cases based on the 'life skills' needed to participate in the information overflow individuals are exposed to by today's media and online platforms. The skills can be narrowed down and explained as follows:

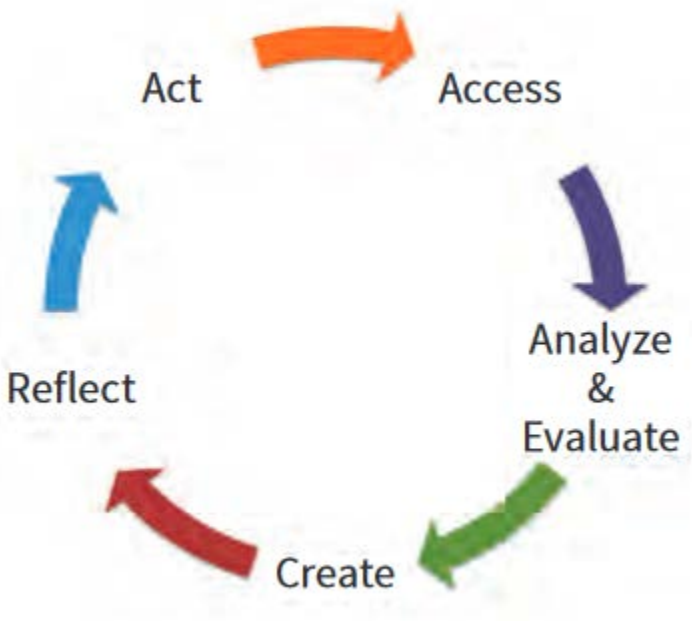

Figure 3 MIL Cycle of skills (Carlsson, 2019).

The MIL cycle starts with accessing the information, analysis of the same, potential creation of content in terms of further distribution and/or contribution, reflection on the information and

interaction based on which the final point is to act. The cycle is constantly ongoing and should add value to the mentioned life skills that is provided by MIL:

- Make responsible choices and access information by locating and sharing materials and comprehending information and ideas.
- Analyze messages in a variety of forms by identifying the author, purpose, and point of view, and evaluating the quality and credibility of the content.
- Create content in a variety of forms, making use of language, images, sound, and new digital tools and technologies.
- Reflect on one's own conduct and communication behavior by applying social responsibility and ethical principles.
- Take social action by working individually and collaboratively to share knowledge and solve problems in the family, workplace, and community, and by participating as a member of a community.

  (Carlsson, 2019).

Apart form the competencies of an individual, MIL requires the application of the concepts form the information delivery side as well. In that context the MIL intend to teach an enforce the listed principles:

- Higher level of media quality-requesting form media sources to provide quality, verified and reliable information.
- Enhanced quality in the public sphere-engaging MIL competent individuals an encourage them to take part in the public discourse through providing input, participating in debates, presenting information based on standing arguments, not only opinions and personal views.
- Enhanced quality in social debates-base the conclusions on the strength on the argument presented from both sides of the aisle, the society, and the media.
- Enhanced quality in the political community-ensure that political decisions are based on the previously mentioned principles which would include quality media, competent participants, quality information and quality debating.

(Vajzović 4, 2020)

In the attempt to understand the challenges of MIL, we need to be able to grasp the information sources and their overall effect on the information consumers.

All generations who were mature enough to receive information before the expansion of the Internet and online sources, were used to do so from what we define as traditional sources of information like Television, Radio, and newspapers. Other forms of information sources were based on written/printed information, verbal communication either directly or over telecommunication technologies that existed before cell phones or even more before smart phones were put in use. The mentioned generations are referred to as the 'Digital Migrants' (Vajzović1, 2021). Younger generations who were born into the era of smart phones, extensive use of the internet and other digital sources, can be referred to as the 'digitally inborn' generation. The digitally inborn generation uses the digital sources as primary sources of information as a given, opposed to the 'traditional' information sources. In the study comparing the efforts of parents of children that fall into the category of 'digitally born' there is an increase in the attempt to communicate the threats of internet usage that went form 81% in 2018 to 85% in 2020 based on results recorded in United Kingdom. However, the parents still feel less secure and confident that their knowledge of MIL is sufficient to educate their children (Vajzović 4, 2020). Regardless of the fact whether a person belongs to the group of 'digital migrants' or the 'digitally inborn', today's lifestyle and the fact that access to online platforms is within immediate grasp for information consumers, the mentioned platforms have already overtaken the 'traditional' information source or are in rapid growth to overtake them.

The access to online sources used as primary sources of information is dependent on the country, level of internet use, traditional views, and other factors. In the United States, a study has shown that than 80% of adult Americans receive their news from digital sources with the use of their smartphone, computer, or tablet (Shearer, 2021):

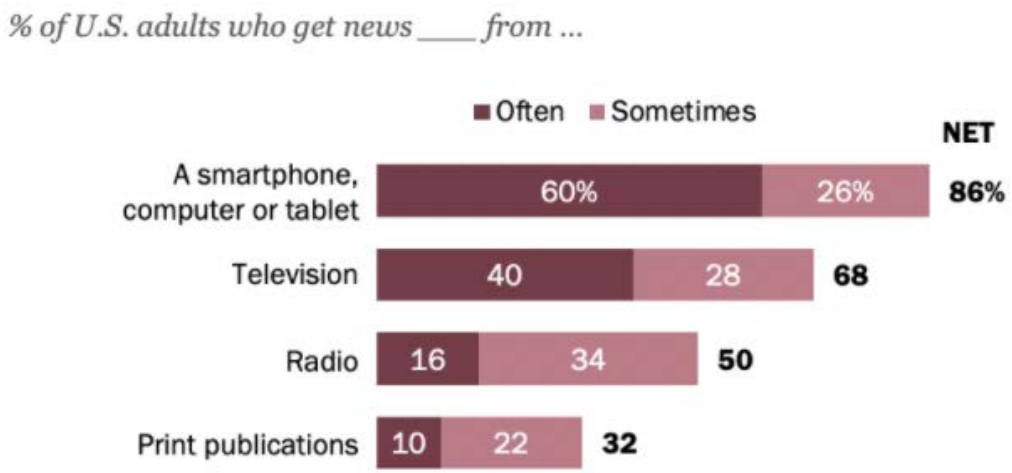

Figure 4 Survey of U.S. Adults on what they use as primary news source, Aug 31-Sep07 2020 (Shearer, 2021)

A Study in the UK on the other hand shows the difference between the 'Digital Migrants' and 'digitally inborn' news consumers form a different angle. In the report analyzing which news sources are predominant among information consumers, it is visible that the age and relevant habits affect the way news information is received: 'TV sources represent six of the top 20 most-used news sources- the most of any platform (the top 20 news sources also include four social media sites, five newspaper titles (print or digital format), three radio stations and two websites/apps). While TV is the most-used platform for news overall, there are some exceptions; for example, 16- 24s are still more likely to use the internet for news than TV (79% vs. 49%)' (Jigsaw Research, 2020)

In Bosnia and Herzegovina (B&H), the overall population mostly relies on the 'traditional' information sources. Although there is a constant rise in the numbers related to the use of online media, nearly three thirds of surveyed citizens of B&H chose TV as the primary source of news (71%). Half of the surveyed claimed that they use social media on a daily basis (50%) as their source of news comparing to one third (33%) of the surveyed who have confirmed that they use online media sources (i.e., media websites, online magazines etc.) as their everyday source of news. The survey also shows that when compared to previous years, online news source use is on

the rise while ‘traditional’ news sources are on the decline mostly among younger population. (Sokol, 2021).

From a MIL perspective the growing numbers of information sources is a constant challenge. This mostly applies to online platforms. The online platforms are the ones on which monitoring is not sufficient or does not exist at all. Those sources are most commonly social networks, blogs, and web sites (Cvjetićanin et al., 2019).

On the other hand, to further widen the issue, even though some online platforms and social networks like Facebook are monitored, the lack of ethical response creates a completely new dimension to the effects digitally consumed information is affecting the ones participating in the information exchange. The ‘Facebook Papers’ have revealed that the largest social media platform was aware of information using IDT being shared but failed to act upon it. The whistleblower in the Facebook Papers, Frances Haugen who is a former Facebook product manager, provided evidence that the social media giant was able to detect IDT and other forms of malicious content in forms of information, but decided to give prevalence to profit rather than protecting their consumers or taking actions to ban the detected IDTs. (Associated Press, 2021)

Development of new technologies and solutions has also helped in managing IDTs. The use of algorithms and artificial intelligence (AI), provides automated management of IDTs used by almost all social media platforms and often web sites, services and advertising companies and online ad agents.

An example of how algorithms and AI are used, adding to the previous section of this paper about algorithms, the documentary film ‘The Social Dilemma’ shows how social media companies manipulate users. The algorithms managed by AI are able to use IDTs to encourage addiction to the social media platforms, by collecting the metadata along with the information provided by users. These technologies are used to ‘hook’ the users to the platforms while targeting them with specific ads, which are based on algorithm and AI calculations. (Barnet, Bossio, 2020)

In the case of Edward Snowden, and the information he leaked to the public in 2013, it was revealed that the United States of America (US) government in their efforts to detect and collect information about all citizens and some world leaders used algorithms and AI. The program

called ‘Prism’ collected information from servers from at least nine major Internet companies which included Apple, Facebook, and Google. (History.com, 2018)

Considering all the challenges outlined, the goal of MIL is to start from a single individual through groups of individuals all the way to the highest level of societies, state, and group of states (i.e., European Union) and ‘train’ all included to be able to know what individuals as consumers of information are being served. MIL needs to enable the mentioned individuals to build the ‘sixth sense’ that would enable the individual to make decisions based on the analysis of the information being served, including: context in which it is presented, political motive, effects on the society, the effect on an individual level etc., considering the variety of all types of security (social, economic, health etc.).

In information security, the protection against all forms of social engineering is battled by awareness trainings. Awareness training as a form of an industrial standard in educating employees/members of various organizations, is used constantly. The main goal is to educate employees/members of an organization how to recognize and what to do in case of an event that could potentially lead to jeopardizing the security of an organization. Awareness in relation to MIL could help overcome the challenges related to IDTs as a form of social engineering.

2.8 Awareness training and potential of awareness using MIL.

As defined by the National Institute of Standards and Technology-U.S. Department of commerce (NIST) awareness trainings are: ‘Programs which set the stage for training by changing organizational attitudes to realize the importance of security and the adverse consequences of its failure. Awareness is a training in which people are taught the skills that will enable them to perform their jobs more effectively’ (NIST, 2021). It is also a requirement of the NIST 800-50 framework to provide the employees and/or members of an organization to attend awareness programs related to Information security as a standard practice (NIST 800-50, 2003).

The ISO27001 standard has a requirement under the clause 7, managing human resource and more specifically under the clause 7.2.2, to provide awareness trainings to organization members and outsource contractors and other relevant business partners. Any organization seeking certification based on the mentioned standard has to fulfil the awareness requirement to be compliant and receive the certification (ISO 27001:2017).

Although there are many organizations that do not seek full framework and/or standard implementation, or certification, it has become an industry standard to have an awareness program within the organization that teaches the members and other relevant individuals about information security what is referred to as the 'industry' standard. The concepts of MIL are not considered as a part of the average awareness programs even though there are some topics being covered within information security that fall into the MIL concepts.

The awareness training has shown to be the best approach when countering the human factor based vulnerabilities in an organization. Several papers and studies by professionals and the academic society have discussed this topic in order to analyze the importance of awareness training. In the overview of relevant literature, it can be concluded that awareness training in the context of battling the human factor in information security is ongoing for several years.

As most researchers suggest, the employees of an organization should always be considered for awareness training. The training should start when an employee joins an organization and from then on receive awareness training on a regular basis (Parsons, et al., 2010). Another study has shown that awareness, when delivered in a proper way, can be a powerful means of empowering the ones who attend the same (Abawajy, 2014). This means that the preparation and the delivery of the awareness training can be a powerful way to resolve human factor related issues in information security. Providing cyber security training to employees is considered a standard and has often shown to be essential. Even though some employees might have some knowledge in cyber security, this does not mean they have the experience in dealing with information/cyber security incidents, and therefore they need to be educated on a regular basis (He et al., 2019).

MIL is a form of awareness that is intended to teach an individual how to process the information being served in order of making decisions which can a do have impact on the security aspects of an individual's life. However, MIL is not an 'industrial standard' and as such is not present in the form of awareness trainings which are used for educating members/employees of various organizations. MIL is very vaguely present in educational systems and education programs. This also applies to B&H, where MIL as a concept or some parts of it are mentioned in state level strategies and initiatives. (Cvjetićanin et al., 2019). MIL is being studied in various academic programs like the master's program in Information Security at the Faculty of Political Science at the University of Sarajevo.

On the other hand, MIL is not reaching the general public, and many are not aware of the concept, what it means and how to use it. An analysis of the level of MIL in Europe related to the IDTs used to spread information on COVID-19 has shown that education is required to battle to effect IDTs have on the information consumption. (Lessenski, 2021). This study has shown that B&H is at the bottom of the list taking the 34th (19 points) place in 2021. Just a comparison, the best score was obtained by Finland, ranking as the first (78) points.

Another report shows what are the effect and missing parts when it comes to teaching MIL as a standardized program in schools in Europe. The competences that MIL can provide were defines as follows:

- Access: the ability to find and use media skillfully and to share suitable and valuable information with others (including browsing, searching, filtering, and managing data, information and digital content).
- Analysis and evaluation: the capacity to comprehend messages and use critical thinking and understanding to analyze their quality, veracity, credibility, and point of view, while considering their potential effects or consequences.
- Creation: the capacity to create media content and confidently express oneself with an awareness of purpose, audience, and composition techniques.
- Reflection: the capacity to apply social responsibility and ethical principles to one's own identity, communication, and conduct, to develop an awareness of and to manage one's media life.
- Action/agency: the capacity to act and engage in citizenship through media, to become political agents in a democratic society. (McDougall, J. et al, 2018)

The report has shown that the recipient of MIL education is lacking proper knowledge and that work needs to be done. The key finding is defined as follows:

- There is an urgent but ongoing need for media literacy educators and stakeholders to document their best practice in the form of empirical classroom research, and to address enduring disconnects between theory and practice, conceptual frameworks and pedagogic practice, and educational/political policy and classroom practices.

- The integration of digital literacy in the maturation phase, specifically into science education, is flourishing as a research area. The field of digital literacy is in general moving away from competence models and protectionist approaches to more robust research that embraces the complexity of 'dynamic literacies'.
- Successful implementation of media literacy education at the school level is facilitated by approaches to pedagogy that combine and/or cross boundaries between spaces and roles — the classroom and the extended 'third space', teachers and students working in partnership to co-create learning, and professional development in hybrid combinations of physical and virtual networks. This work also speaks to the need for media educators to be confident in accepting the need for the concept of 'Building' itself to change, as opposed to thinking of digital media as merely contributing to it as a stable entity.
- There is a wealth of evidence of more formal, funded, partnership engagements between media literacy educators and media industries, literacy organizations, NGOs, and other stakeholders at the level of resource production and single events. However, empirical evidence of the conditions for successful partnership and impacts at the school level are likely to be in the public domain within two to three years, as many relevant projects are ongoing. (McDougall, J., et al, 2018)

For the purpose of analyzing a standardized MIL educational program which is properly advertised, promoted, and made accessible to the general public, more precisely outside of the academic community and some education programs, no evidence was found in a satisfactory form, which would mean that the approach is insufficient in its approach on how MIL education is brought to the general public. One additional issue is that most of the current workforce and members of organizations around the world that have missed the topics in school that included them recently and are not in touch with the academic community, get left behind in terms of enhancing their capabilities and competencies in MIL.

In the attempt to resolve the defined issues, the MIL concepts need to be made more publicly available, they need to be made more appealing to the general public and the MIL concepts need to be introduced in the awareness programs within organizations in the same manner in which the information security has become a standard of the industry. The approach would also need to

include a public discussion on MIL and its effect on the lives of society and how it drives the perception of information consumption in the forms that are present today.

All the evidence found mostly referring to the academic society and academic professionals, missing the capabilities to present the importance of MIL awareness to the general public. As an example of the same is the response of undergraduate students of Political Sciences at the Faculty of Political science at the University of Sarajevo, where students reactions to the viewing of the documentary 'The Social Dilemma' had the following general reaction as paraphrased: 'We know that meta data an information is gathered about our online activities, but we were not aware of the gravity related to the same as it is presented in the documentary' (Vajzović3, 2021)

The awareness program needs to be built into the Information Security Awareness programs in organizations emphasizing the interests is to protect the organization. An average individual needs to be educated firstly on the importance of his/her role in the digital world and how it reflects on their lives, society, and all forms of security. The individuals need to be informed that the information consumption, distribution, and use can directly affect how all types of security can be affected and what every individual can do to 'protect' their interest and security.

The challenges to motivate individual to educate themselves in MIL can be resolved by clearly presenting what is the current status, how it is affecting their security and why it is important for them to be engaged. As it is a basic human reaction to protect the security of oneself and loved ones (Smajić, 2021), this needs to be the baseline to gain motivation and engage the general public into obtaining awareness for MIL.

The awareness needs to be effective and developed to make the MIL understandable to all who attend the awareness. It should include the listed as a minimum:

- Clear detail on how information is received

- What the information is about and how does it reflect on one's life

- What is the environment in which the information is consumed (Political, social, economic etc.)

- What are the methods to apply MIL in information consumption and distribution processes.

3. Applied research and testing methodology.

Based on the information gathered in the previous chapter and the performed literacy review, this chapter will explain the methodology used to conduct the research. After the Literature overview moving forward the goal is to test the effects of IDTs on an organization, prior and after the selected organization members have undergone training in awareness on MIL, to measure results and effectiveness of the proposed solution. This chapter explains the methodology, processes and actions taken to perform the testing required for this master thesis.

Figure 5 Research Methodology

STEP1 Literature review

The main idea behind the performed literature review was to present four main topics of this thesis and give an insight to them. The four topics are, information security, the human factor as the weakest link within information security, information disorder as a new fast developing aspect of information security which targets the human factor and MIL and awareness training in MIL as a form of education which help prevent serious consequences caused by information disorder.

STEP2 MIL Awareness developed and explained.

This step gives an insight to the awareness training for MIL, which is based on the concepts of industry information security awareness training. The approach is based on an in-class session where the participants are introduced to the information disorder, IDT techniques targeting humans and how to protect against the same.

STEP3 Testing before and after awareness training.

This step explains the tests that will be conducted on participants in a selected organization. It also gives insight to how they are performed, what are the goals of the same. The tests conducted for the need of this thesis are based on attacks using IDT methods as a sub section of social engineering attacks, which are in a broader sense penetration test on the participants. The test results will be collected for further analysis and compared to give a more detailed insight to the results before and after the participants receive awareness training in MIL to see if any improvements are made due to the training conducted.

STEP4 Analysis and results

Based on the results collected through the performed tests, full analysis and results will be presented. Initially, the results of the first tests conducted before the participants go through the awareness training. Then the results of the second test after participants complete the awareness training in MIL will be collected and presented. In further analysis, the results of the test will be compared based on the nature of the test performed. This means that type of IDT's that are similar or same in execution will be compared to give more precise insight into the responses which should give more details on the effects of the awareness training performed. The result presentation and comparisons will be done on a one-to-one basis.

3.1 Objectives

The research of this thesis focuses on the effects of MIL awareness training delivery in an institution in Bosnia and Herzegovina. It tends to prove that the education provided to members of the organization on MIL in order to raise the awareness can be a main point of reducing risks related to information security incidents caused by MIL including the potential reputational and/or financial loses for the organization. When approaching the members of the organization and considering the individual contribution of every participant to the human factor, it should not be taken lightly and with the intention of just delivering the general message. The focus should be shifted on the actual functional position an individual or a group of individuals have in an organization and deliver the education through awareness training, tailored to produce the best possible result.

3.2 Hypothesis

This thesis in its entirety is intended to review known cases and acquired knowledge in information security presented through the literature review, attempting to give insight to what MIL as a concept that can be taught to protect against information disorder and asymmetric hybrid attacks. Therefore, the problem set in this thesis is the question: ***Is it possible to establish a professional training program in the field of media and information literacy, which would lead to an increase in competences, the ability to recognize and defend against attacks from the domain of social engineering that are used for the purposes of creating information disorder, which lead to endangering health, social and general security in the organization?***

Based on the performed research, problem question and the conducted testing, the hypothesis of this master's thesis is:

The awareness-raising program in the field of MIL through the professional training program in organizations leads to a reduction in the risk of harming the security of the individual, the community, and the state, which are caused by information disorder and hybrid asymmetric attacks.

3.3 The MIL Awareness training

The basis of the awareness session is a known approach for in-class information security training. The training is delivered in a classroom, using a presentation with topics and most important information outlined on the slides.

During the process of delivery of the training, the participants are encouraged to engage in discussions, give their own examples and ask questions. The topics covered in the training in more details are:

- Introduction to Media and Information Literacy.
- Basic definitions of MIL and defining expressions used in MIL.
- Information disorder, what it is and how it exists online and around us in real life.
- The goals of information disorder
- Social engineering and the human factor/Link between social engineering and information disorder.

- Types of social engineering methods and attacks
- Defining information disorder techniques (IDTs) and hybrid asymmetric attacks.
- What types of IDTs have been recorded in recent history.
- Reflecting on effects of IDT through known cases (Antivaxxers, Brexit, Elections in various countries etc.)
- Amplification using AI algorithms and bot networks.
- How to protect against IDTs using MIL.
- Discussion sessions Q&A

The Training delivery details:

NOTE: The delivery details reflect how the actual training session was delivered as a general overview with some details.

i. The first part of the training was an introduction to the concept MIL and Information disorder. This part of the training explained all types of Information disorder techniques that exist today. The participants were introduced to MIL as a concept that fights against the information disorder.

ii. The second part of the training covered the Human Factor as the weakest link in the chain of information protection. The participants were introduced to the concept of social engineering and how it is used to exploit the human factor.

iii. The third part explained the social engineering attacks referred to as the IDT's (Phishing, Smishing, Vishing, Baiting, Missing context, deceptive editing, malicious transformation, emotion-based scamming and etc.). In this part of the awareness program, the participants were also introduced to the automated systems; algorithms that support the spread of information disorder and the execution of IDT attacks. Within the same part of the awareness training, the participants were introduced methods and already recorded exploitation that have been witnessed in B&H and the world.

iv. The fourth part of the training has introduced the participant to the real-life examples of information disorder and IDS's and what are their consequences.

v. In the fifth part, the participants were introduced to the MIL concepts of protection against information disorder and IDS's within their workplace and in their private affairs.
vi. The last part of the training was a fruitful discussion on the topics covered within the Q&A session.

3.4 The testing environment, process, and methodology.

This segment of the thesis gives insight to the environment in which the testing was conducted, provides relevant information on the participants form the selected organization. Further, the focus is turned to the actual test executions and the collection of results.

3.4.1 The testing environment

Initially, we have selected an organization whose participants are selected by the relevant organization representative. All of the participants have entered the testing process with the listed requirements, by the testers and responsible point of contact form the organization, have agreed on:

a. None of the participants have taken part in any kind of official and organized MIL awareness training.
b. None of the participants are allowed to be aware of any penetration testing which is to be conducted using IDT methods, before and after the awareness training is delivered to them.
c. None of the participants' personal information will be shared before, during or after the testing is conducted.
d. The organization itself is an educational institution within the University of Sarajevo. A total of six participants who are employed within the organization have been selected by the organization's point of contact.

The participants have provided some details about themselves in the entry survey which gives more insight to the selected participant and those details are:

a. There is a total of six participants out of which four are female and two males.

b. One participant is in the age range between 20-30. One participant in the range from 30-40. Four participants in the age range 40-50 and one participant from the range age of 50+ years.
c. Three of the participants have professional careers in the range of 4 to 7 years, while three of the participants have their professional career for more than 7 years.
d. All of the participants are working in Sarajevo, Bosnia and Herzegovina.

3.4.2 The testing process and methodology

The testing methodology has the goal to test how an awareness program on MIL can affect the members of an organization that has the intent of protecting its members and itself from IDT attacks on the same.

The whole testing process is planned in the below defined steps:

a. Obtaining the required permissions to carry out the testing.
b. Getting the list of participants and their initial contact details, which is their, full names and their email addresses.
c. Getting the initial responses through an entry survey related to MIL, prior to the awareness training.
d. Executing IDT attacks on the participants, before the awareness training.
e. Executing the awareness training.
f. Asking the participants to fill out the outgoing test-survey, one week after the training was conducted and conducting a second round of IDT based penetration tests.
g. Gathering testing results throughout the duration of the testing phase.
h. Analysis of the results and presentation of the same within this document.

Phase I consists of actions listed above, under points a, b, c and d. Phase II consists of only the action listed under point e. Phase III consists of actions listed under the above listed points under f and g. The last Phase IV consists of the actions listed under points h and i, from the above-listed action items.

The testing is intended to measure the responses and behavior of the selected participants before and after the MIL awareness training is conducted. The goal is to measure the response levels in

order to measure the efficiency of the IDT attacks and general understanding of MIL before and after the awareness training on MIL was delivered.

Once the required approvals to conduct testing in the organization with the code name XYZ have been provided, the list of participants was provided along with their full names and email addresses. In the first week of the testing, a full analysis was performed on the participants, analyzing their presence on social networks (Facebook, Instagram, LinkedIn, and Twitter). Seven days before the awareness training session was conducted, penetration testing within Phase I was initiated along with the entry survey being sent to the participants three days after the penetration testing was initiated.

The attacks performed were as follows:

-TEST 1-1: Engaging with all participants on social media platforms. The goal of this attack was to try and connect to the participants and gain their trust. After the trust was gained, by using false and misleading information, the goal of the test was to exfiltrate private information from the participants, especially the ones that cannot be retrieved just from publicly accessible information sources.

-While using a fake LinkedIn account, the participants were approached by a fake headhunter from Germany. The goal was to get the participants to interact and share their private information via LinkedIn communication and through the fake email that was setup for the fake headhunter account on LinkedIn.

-TEST 1-2: Using phishing by calls-vishing and impersonation. The goal was to try and get information that can be used to take control of an account within the organization that belongs to a student.

-Posing as one of the existing students within the educational organization, by creating a fake case of losing credentials, the goal is to get the student office to support our attempt with the respective IT team, to reset credentials for the student account and get them sent to a different ‘private’ email address of the impersonated student

-TEST 1-3: Getting participants to join a fake Facebook group/Page

-By creating a fake Facebook page, that promotes false information, the goal was to get the participants to share some of the false information on their Facebook accounts.

-TEST 1-4: Asking participants to fill out the entry survey on MIL.

-The entry survey was intended to measure the knowledge of participants related to MIL, IDS's in order to set a baseline before they take part in the MIL awareness training.

After the survey responses were gathered, the participants attended a three-hour training session on MIL which was a part of Phase II.

Seven days after the awareness session was delivered to the participants, the testing continued within the scope of Phase III as follows:

-TEST 2-1: Additional interaction through LinkedIn.

-The goal was to continue the already established communication posing as a headhunter, to see if any additional, even more confidential personal information would be shared by participants after they have completed the MIL awareness training (i.e., document numbers, unique personal identification number etc.).

-TEST 2-2: Second Vishing and impersonation case.

-The goal of this test was to impersonate a Social Health service representative, enquiring on the details about COVID-19 vaccinations among the staff of the XYZ organization. The goal was to get personal information and contact details of organization members which did not take part in the awareness training, which was to be used with the IT support to take over an account used by an employee.

-TEST 2-3: Additional interaction on Facebook.

-The goal was similar to the test conducted in 1-3, and that was to get the participants to believe in false information and to get them to distribute them further among their Facebook contacts using their personal accounts.

-TEST 2-4: Asking participants to fill out the outgoing test-survey on MIL.

-The outgoing test-survey was an actual test conducted on the participants, where they were tested on their ability to identify IDTs presented through misinformation, while planting

true information in between the cases presented to the participants The aim was to see how would the participants react to the news shared with them and check if they would use the knowledge gained in the awareness session when responding to the questions asked about the cases..

After the last test was completed, a move was made to complete the last phase, Phase IV.

4. Result Analysis

This chapter presents the results of the tests performed and the analysis of the same. The goal is to present what was captured during the execution of the testing and to show the results in more details.

4.1 Results of the first round of tests.

This part of the thesis analyzes the test results of the penetration tests and the entry survey conducted before the delivery of the MIL awareness training to the participants.

I. Results of the TETS 1-1:

A fake LinkedIn Account was set up along with the respective email address. The interaction was conducted with one participant, by approaching the individuals and offering a possibility of getting a job in Germany. Before the interaction started, the fake account had added several contacts from B&H and the western Balkans region. This was done to ensure that the engaged person would perceive the account as legitimate, and trusted since the contacts added were also contacts known to the participants.

During this test, it was observed that only two participants have active accounts on LinkedIn. One of the individuals was actively using the mentioned platform while the other was not active for more than a year at the time the test was conducted.

The decision was made to interact with a single individual in an attempt to exfiltrate personal information.

The Individual that was approached, had accepted an invite to connect and has immediately accepted the attempt to engage into a conversation for the fake job opportunity. To ensure

interaction would continue, the job potentially offered was in line with the type of work of the targeted individual.

This attempt was fully successful in getting the individual to trust the fake account and to share a CV along with some personal details during the course of the interaction. Considering that other participants were not using the LinkedIn platform, this test was considered 100% success.

This test allowed us to get all the personal information we needed and have set as the goal of this test:

-Personal information (Full name, private and business phone number, home and work address, DOB)

-Personal job preferences

-Personal professional experience details (full history of previous experiences, publications, interests, capabilities etc.)

Table 1 Results of TETS 1-1:

TEST 1-1	Success rate
While using a fake LinkedIn account, the participants were approached by a fake headhunter from Germany. The goal was to get the participants to interact and share their private information via LinkedIn communication and through the fake email that was set up for the fake headhunter account on LinkedIn.	100%

II. Results of the TETS 1-2:

As we were able to identify that one of the participants works in the student's office within the organization, we have prepared a scenario to target this specific individual.

The test was performed by using Vishing and impersonation. The impersonation was of an existing female student, which has lost access to private email account, which is registered in the student office, and in the end did not allow the 'student' to reset the credentials when accessing the organizations Learning Management System (LMS).

Using the above-mentioned pretext, we were successful in getting the trust and instruction on what steps to take to have the credentials reset and sent to a new email address which we have setup for this test.

After the process was followed (making sure that the initial targeted individual was copied in all of the emails), we were able to get the new set of credentials, after which we have taken over the account of a student, which was the goal of this test.

Table 2 Results of TETS 1-2:

TEST 1-2	Success rate
Using phishing by calls; vishing and impersonation. The goal was to try and get information that can be used to take control of an account within the organization that belongs to a student.	100%

III. Results of TEST 1-3:

This test was intended to initiate false information spreading using Facebook. During the test, none of the participants have accepted the interactions. This is due to the fact that during the time the testing was conducted, none of the participants have been active on Facebook, based on the publicly available information on user activity was observed for the targeted accounts.

It is important to note that four of the participants have accounts on Facebook. All others were not found using the information gathered on the mentioned participants.

Table 3 Results of TETS 1-3:

TEST 1-3	Success rate

Getting participants to join a fake Facebook group/Page	0%

IV. Results of TEST 1-4:

This test was actually a baseline information gathering survey, where the intention was to get some details on the participants and get their responses on how they evaluate their knowledge of information security and MIL. This survey will not be used to evaluate the success of the MIL awareness training but is used as a reference point of getting to know about the participants before the awareness training is delivered.

e. the response was as follows:

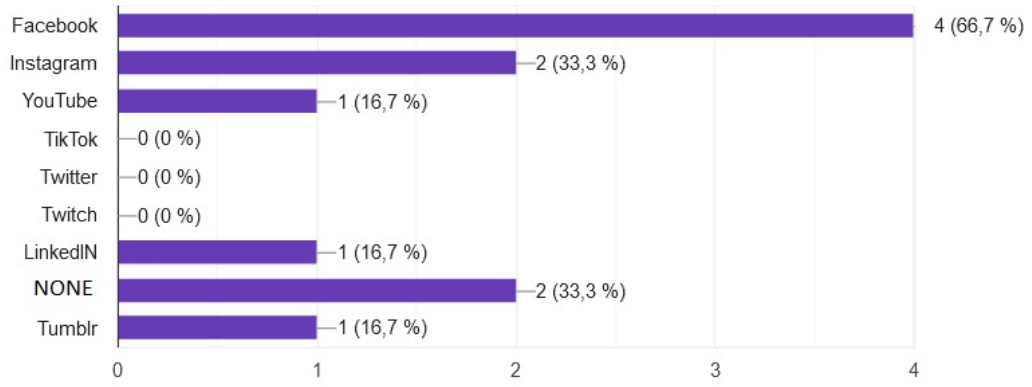

Figure 6 Participants respond to the question on which online platform they use and accounts for

As it can be seen in the above image, most participants have Facebook accounts, while two of the participants don't use any of the online platforms that were offered as a response. When it comes to TikTok, Twitter and Twitch, none of the participants have accounts for the same.

f. When it comes to daily access to the mentioned platforms and online access the responses were as follows:

NOTE: The question was: How often do you access the listed online platforms (Facebook, Instagram, TikTok, YouTube, Online news portals and other websites and Twitter.)

The options to choose from were: 'Daily', 'Once in two days', 'Rarely' and 'Never'.

i. Three participants access Facebook 'Daily', two selected 'Never', and one of the participants has not selected an option.
ii. Only one participant selected 'Daily' as their preference on accessing Instagram, four participants selected 'Never', while one participant did not select any of the provided options.
iii. When it comes to TikTok, two participants selected 'Rarely' as their choice, three responded with 'Never', while one participant did not provide an answer.
iv. When it comes to YouTube, four participants selected 'Daily' as their option, one participant selected 'Never', and one participant did not provide their response.
v. In the end when it comes to Online news portals and other websites, three selected 'Daily', one participant selected 'Once in Two Days', one participant selected the option 'Never' and one participant did not provide their response.
vi. In regard to accessing Twitter, five participants selected 'Never' as their option, while one participant did not select any of the offered options.

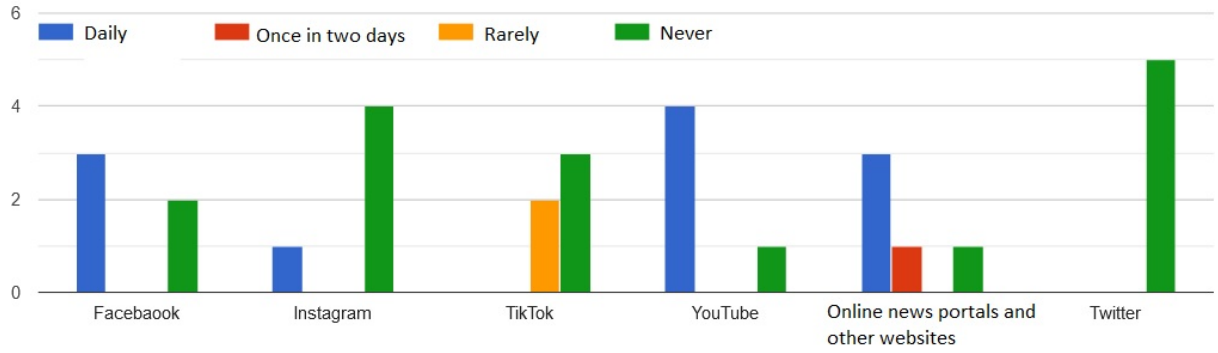

Figure 7 Participants respond to the question on how often they access online platforms.

g. The next series of questions are related to the competencies and understanding the participants personally feel they have in the domain of Information Security.

NOTE: The questions asked were:

Q1: The concept of Information Security is known to me.

Q2: Information Security is important for the organization I am currently employed at.

Q3: Information Security Awareness has a direct impact on the way I conduct my daily work tasks.

Q4: The competencies I have in Information Security, which I have gathered in my line of work are being used during on and off work activities which includes my private activities.

The possible responses to the questions were:

Strongly Agree (Blue), Agree (Red), Do not agree, or disagree (Orange/Amber), Disagree (Green), and Strongly Disagree (Purple/Violet)

The responses to the questions were as follows:

Q1: Three participants selected Agree, one participant selected Do not agree or disagree, one participant selected Disagree and one participant selected Strongly Disagree.

Q2: Two participants selected Strongly Agree and four participants selected the option Agree.

Q3: Two participants selected the option Agree, two participants selected the option Do not agree or disagree and two participants selected the option Disagree.

Q4: One participant selected Strongly Agree, three participants selected the option Agree, one participant selected Do not agree or disagree and one participant selected Disagree.

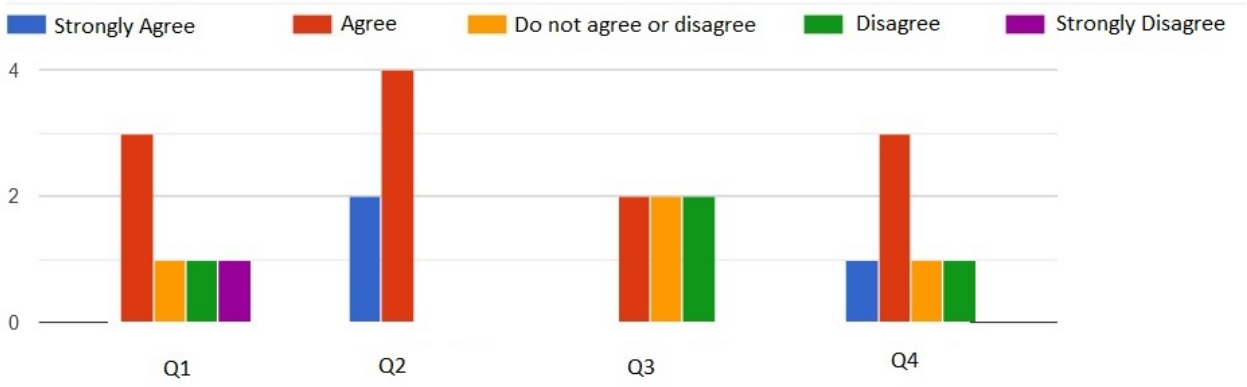

Figure 8 Responses of the participants on their Information Security competencies

h. The next series of questions are related to the competencies and understanding the participants feel they have in their ability to recognize information security threats.

NOTE: The questions asked were:

Q1: I know what a Phishing attack and I is am able to recognize such types of attacks.

Q2: I am able to protect sensitive business information within the organization I work for and outside when I am not at work.

Q3: The Security of my access credentials i.e., my password for work, is a very important part of information security.

Q4: I have personal experiences where knowledge about information security practices has helped in my line of work.

Q5: I have attended educations session in information security (awareness training) while working for organization in which I am currently employed at.

Q6: I am aware of the steps I am supposed to undertake in cases when information security is threatened, within the scopes of work I have within the organization in which I am currently employed at.

Q7: I have heard about and understand the concept of social engineering with regards to information security.

Q8: I believe I am capable of recognizing a social engineering attack.

The possible responses to the questions were:

Strongly Agree (Blue), Agree (Red), Do not agree or disagree (Orange/Amber), Disagree (Green), and Strongly Disagree (Purple/Violet)

The responses to the questions were as follows:

Q1: Two participants Agree, one participant Disagrees, and three participants Strongly Disagree.

Q2: One participant Strongly Agrees, three participants Agree, one participant selected Don not agree or disagree, and one participant selected the option Disagree.

Q3: Two participants selected Strongly Agree, three participants selected Agree and one participant selected Don not agree or disagree.

Q4: One participant selected Strongly Agree, three participants selected Agree, one participant selected Do not agree or disagree and one participant selected Disagree.

Q5: Two participants selected Agree, two participants selected Disagree and two participants selected Strongly Disagree.

Q6: Four participants selected Agree, one participant selected Do not agree or disagree and one participant selected Disagree.

Q7: Three participants selected Agree, one selected Disagree and two participants selected Strongly Disagree.

Q8: Two participants selected Agree, one participant selected Do not agree or disagree, one participant selected Disagree and one participant selected Strongly Disagree.

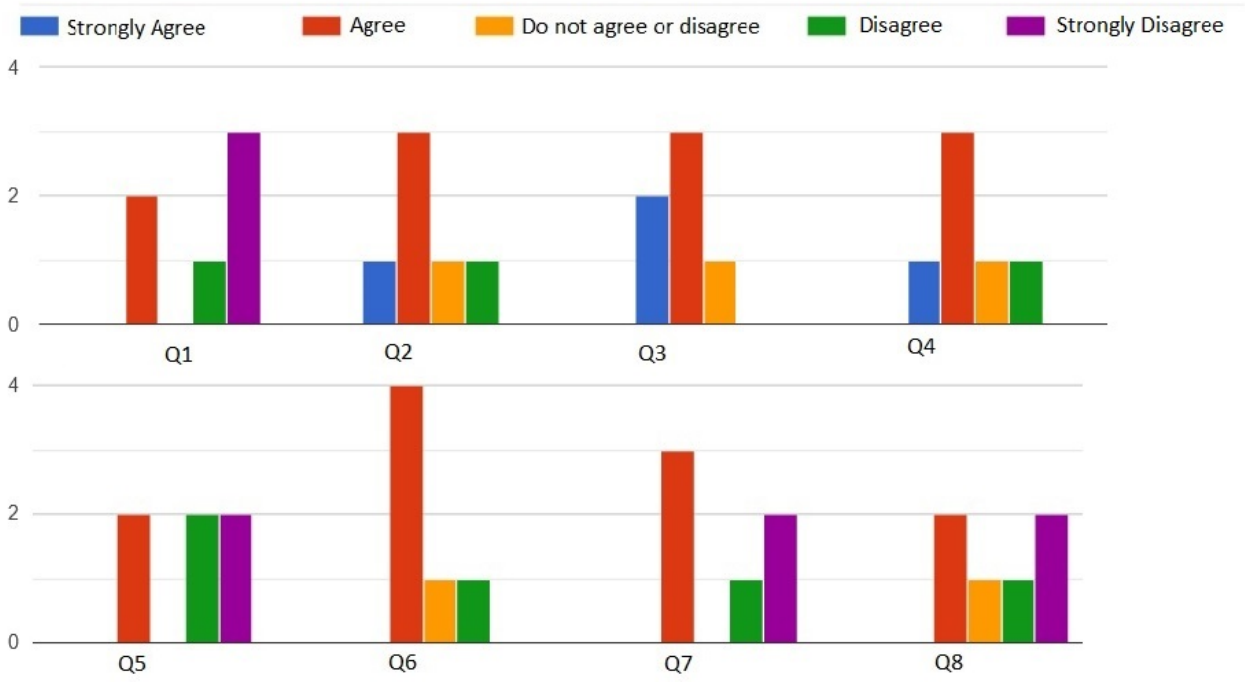

Figure 9 Participants respond to questions about their ability to detect Information Security threats.

i. The next series of questions are related to the competencies and understanding the participants feel they have in the domain of Media and Information Literacy.

NOTE: The questions asked were:

Q1: I know and understand the concepts of MIL.

Q2: MIL enhances the development of critical thinking within the organization in which I am employed at.

Q3: Knowledge, competencies and capabilities that are gained from MIL are very important in the world we live in today.

Q4: In our organization we have an existing program that teaches about the importance of MIL.

Q5: We require a change in the organization's awareness programs to ensure MIL receives mor attention in the organization where I am currently employed at.

The possible responses to the questions were:

Strongly Agree (Blue), Agree (Red), Do not agree or disagree (Orange/Amber), Disagree (Green), and Strongly Disagree (Purple/Violet)

The responses to the questions were as follows:

Q1: Two participants selected Strongly Agree, two participants selected Agree, one participant selected Do not agree or disagree and one participant selected Strongly Disagree.

Q2: One participant selected Strongly Agree, three participants selected Agree, one participant selected Do not agree or disagree and one participant selected Disagree.

Q3: Four participants selected Strongly Agree and two participants selected Agree.

Q4: Two participants selected Strongly Agree, three participants selected Agree and one participant selected Disagree.

Q5: One participant selected Strongly Agree, four participants selected Agree and one participant selected Disagree.

j. The next set of questions are asking the participant on their habit of accessing and using online provided information along with questions asking about their trust in the same.

NOTE: The questions asked were:

Q1: In my line of work, I use online information by browsing the Internet.

Q2: I am capable of recognizing a fully credible source of information from the Internet.

Q3: My primary source of news and information about global developments are online platforms and Internet sources.

Q4: I am capable of recognizing false and misleading information found online.

Q5: Information found on social networks are always credible and reliable.

Q6: Information found on corporate social networks i.e., Posao.ba and LinkedIn are always credible and reliable.

Q7: Information found on social networks are always credible and reliable if the sources are famous and/or known individuals.

Q8: The information found on online new portals are always credible and reliable.

Q9: Information is only credible and reliable if they are coming form known media organizations.

Q10: I only trust the information that is delivered via other media channels like radio, TV and newspapers.

The possible responses to the questions were:

Strongly Agree (Blue), Agree (Red), Do not agree or disagree (Orange/Amber), Disagree (Green), and Strongly Disagree (Purple/Violet)

The responses to the questions were as follows:

Q1: Two participants selected Strongly Agree and Four participants selected Agree.

Q2: One participant selected Strongly Agree and five participants selected Agree.

Q3: Two participants selected Strongly Agree and four participants selected Agree.

Q4: One participant selected Strongly Agree, three participants selected Agree and two participants selected Do not agree or disagree.

Q5: One participant selected Agree, two participants selected Disagree and thee participants selected Strongly Disagree.

Q6: Two participants selected Agree, two participants selected Do not agree or disagree and two participants selected Strongly Disagree.

Q7: One participant selected Agree, one participant selected Disagree and four participants selected Strongly Disagree.

Q8: One participant selected Agree, one participant selected Disagree and four participants selected Strongly Disagree.

Q9: One participant selected Agree, one participant selected Do not agree or disagree, one participant selected Disagree and three participants selected Strongly Disagree.

Q10: Two participants selected Agree, two participants selected Disagree and two participants selected Strongly Disagree.

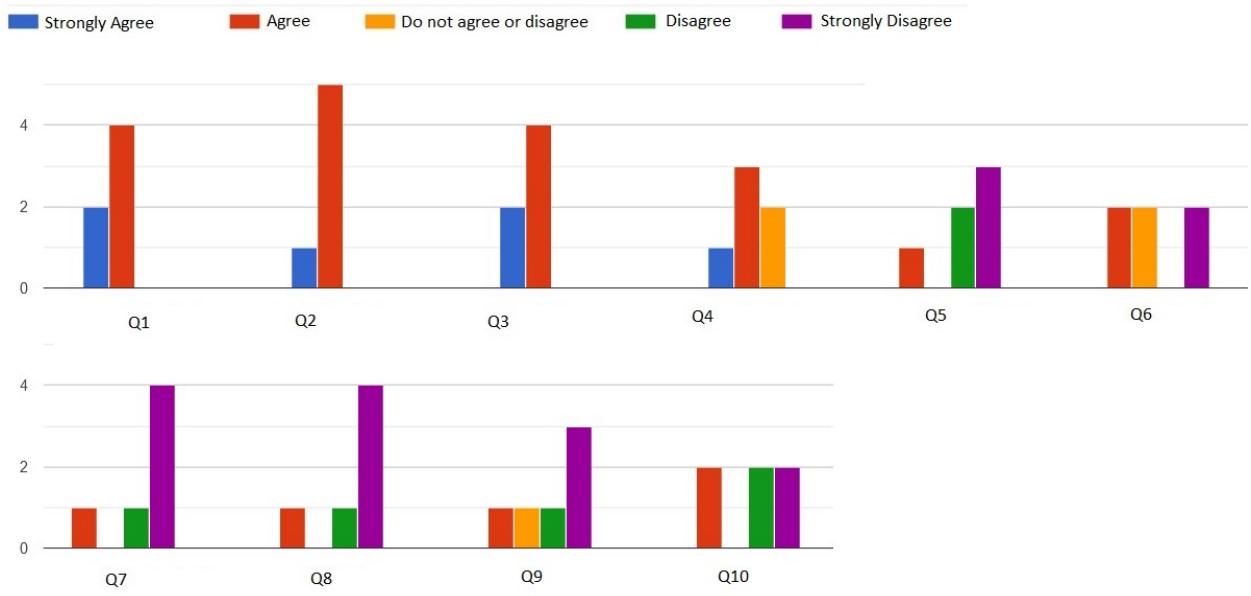

Figure 10 Participants habits on accessing and using online provided information.

4.2 Results of the Second round of tests.

This part of the thesis analyzed the test results after the participants have taken part in the MIL awareness training.

I. Results of TEST 2-1:

This test was the first conducted penetration test conducted after the participants have undergone the awareness training in MIL. The goal of this test was to gather more valuable personal information, like documents numbers (passport, ID etc.).

As the interaction continued, once the selected individual was asked to share information and copies of ID and passport, we received a positive response. However, the individual had blacked-out all of the document's information, which we were aiming to get.

The passport scan received had the following information blacked-out:

-Place of birth

-Details about the document issuer

-Passport number

-Unique personal identification number

The individual has emphasized in the email message that they have blacked-out the above listed information, as the same is personal and can be misused if the same is shared with anyone with malicious intentions.

Table 4 Results of TETS 2-1:

TEST 2-1	Success rate
Additional interaction through LinkedIn and email	0%

II. Results of TEST 2-2:

This test was performed on two participants. The intention was to re-test the person from the test 1-2 and to attempt the same approach with an additional participant.

In our Vishing and impersonation with pretext attempt, we tried to ask for details on individuals who were vaccinated for COVID-19 within the organization. In the intent to make the call seem fully legitimate, we have impersonated an existing representative of the Cantonal Social and Health Insurance, whose details are publicly available on LinkedIn. In the process we were trying to get contact details from the individuals in our attempt to use the information further to try and take over an employee account, in a similar way it was done under the TEST 1-2.

The responses from both targeted participants were as follows:

-The first participant asked us to contact the student's office as they were not allowed to share any information by phone and with an unknown individual

-The second participant refused to share any information and has instructed us to send a formal request for information and to address it to the legal department of the organization for review.

Table 5 Results of TETS 2-2:

TEST 2-2	Success rate

Second Vishing and impersonation case.	0%

III. Results of the TEST 2-3:

The goal was similar to the test conducted in 1-3, and that was to get the participant to believe in false information and to get them to distribute them further using their Facebook accounts.

This attempt has failed and no response from any of the participants who have Facebook accounts was recorded.

Table 6 Results of TETS 2-3:

TEST 2-3	Success rate
Additional interaction on Facebook.	0%

IV. Results of the TEST 2-4:

The goal of this particular test was to see if the participants would use their knowledge acquired during the MIL awareness session when receiving news. In the awareness session, the participants were instructed to always double check information published online, especially the information that could cause damage to the society based on the emotional stir the information could potentially cause.

For this test, we have created five news titles with brief explanations of the news. Two out of the five cases were actually true pieces of information, while three were created using the misinformation approach. This means that the false information news cases were created in a way that utilizes actual facts but presents the information in fake and false manner. Some of the information was also created based on pure disinformation.

For all of the five cases, the participants were asked to respond to the questions asked about the news presented with a Yes or a No. All of the cases had the same set of questions for all cases which were:

Q1: I believe this news is true.

Q2: In order to believe in the news presented, I would need to check the information form several sources including the ones which have proved to be credible.

Q3: This news is a clickbait.

Q4: The news presented in the same manner as this specific one, can potentially have consequences on the security of the community, and society we live in.

Q5: This news sounds like a conspiracy theory.

Case 1:

NEWS: *Zoran Milanović openly praised Dodik: I want such a Serb as a neighbor. RS should be independent from BiH.*
The President of the Republic of Croatia praised Milorad Dodik and stated that it is better to have Dodik as a neighbor in the independent RS than in Dayton BiH.
Note: This news was fake.
Responses:
Q1: All of the participants selected No.
Q2: Four participants selected Yes, and two participants selected No
Q3: All of the participants selected Yes.
Q4: All of the participants selected Yes.
Q5: Two participants selected Yes, and four participants selected No.

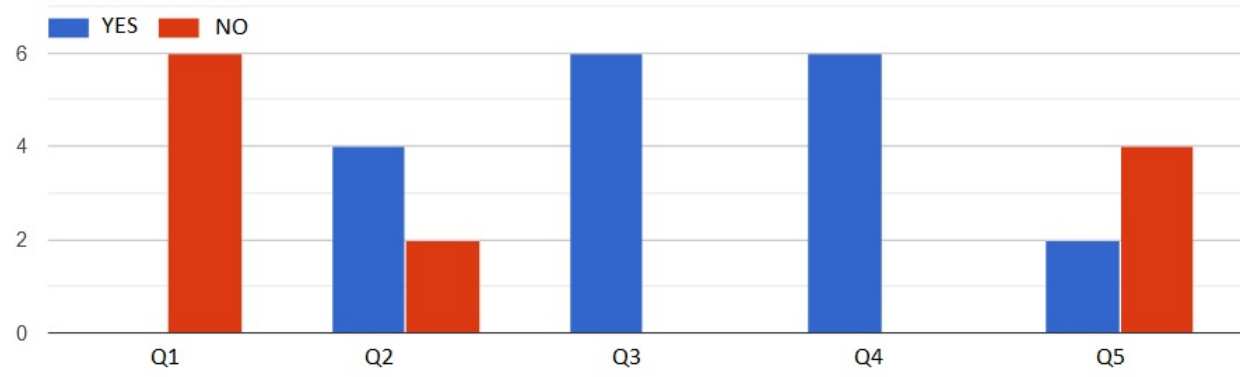

Figure 11 Participants responses to Case1

Case 2:

NEWS: *Milanović told officers from Bosnia and Herzegovina: Why are you coming to an enemy country?*

In a statement to the media during the swearing in of the 20th generation of cadets of the Croatian Military Academy "Dr. Franjo Tuđman" President of the Republic of Croatia Zoran Milanović suddenly declared: "Austrian soldiers are present in BiH, Croatia is not, because someone doesn't like it? I started to like that we were there" and then added: "There (pointing with his hand) are officers from Bosnia and Herzegovina, of Bosniak nationality, who come every year. Then why do they come to the enemy's country?"

Note: This news was true.

Q1: All of the participants selected No.

Q2: Five participants selected Yes, and one participant selected No.

Q3: All of the participants selected Yes.

Q4: Four participants selected Yes, and two participants selected No.

Q5: Three participants selected Yes, and three participants selected No.

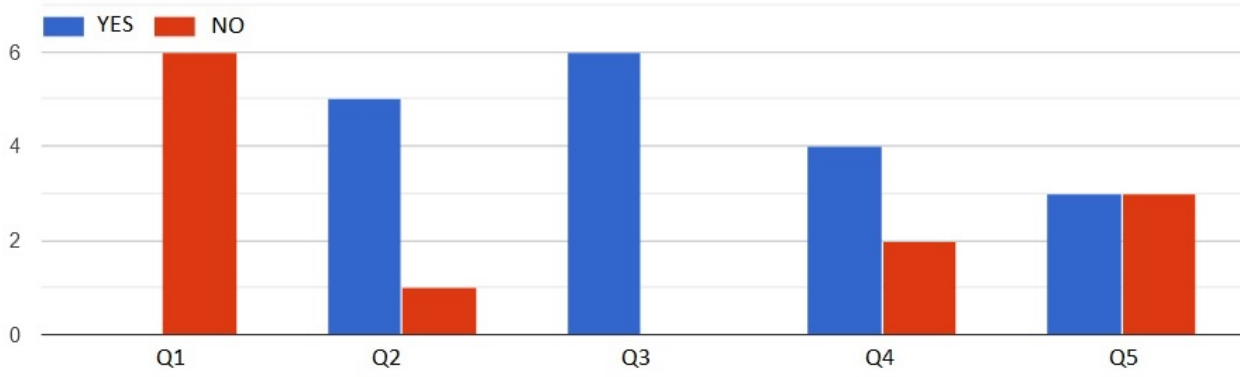

Figure 12 Participants responses to Case 2

Case3:

NEWS: *Shortage of diesel in the EU and the USA, a new rise in prices is expected.*

Although the price of crude oil has stabilized in the range of 80 to 90 dollars per barrel in the last two months, the latest problem could be the price of diesel. A shortage of diesel has appeared not only in the EU but also in the USA, and this will lead to an increase in prices. The consequences should be felt in B&H as well.

Note: This news was true.

Q1: Four participants selected Yes, and two participants selected No.

Q2: All of the participants selected Yes.

Q3: All of the participants selected Yes.

Q4: Two participants selected Yes, and four participants selected No.

Q5: One participant selected Yes, and five participants selected No

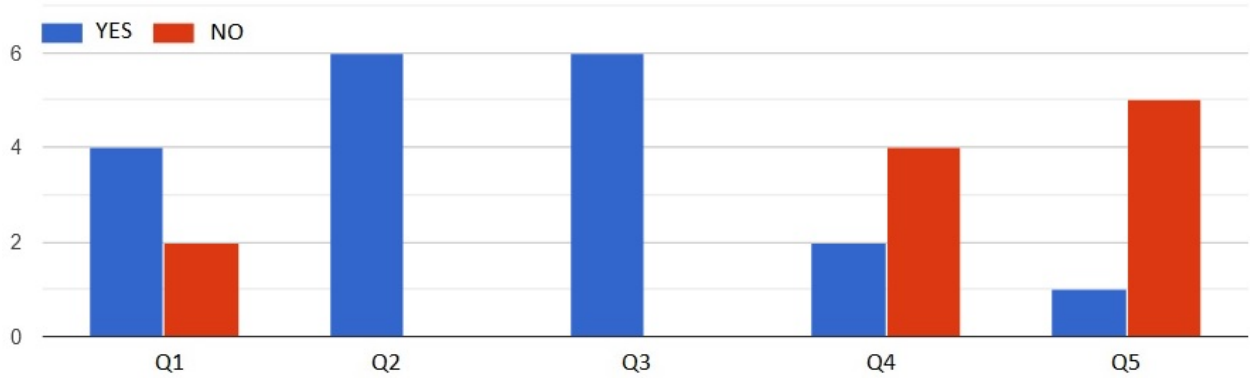

Figure 13 Participants responses to Case3

Case 4:

NEWS*: Instead of BiH, the RS national team will play a friendly match with Russia, confirmed Vico Zeljković, president of NS/FS BiH. After the decision to reject the friendly match between the football teams of Bosnia and Herzegovina and Russia, it was agreed to gather the informal team of the RS will play the match at the scheduled time.*

Note: This news was fake.

Q1: All of the participants selected No.

Q2: All of the participants selected Yes.

Q3: All of the participants selected Yes.

Q4: Three participants selected Yes, and three participants selected No.

Q5: Two participants selected Yes, and four participants selected No.

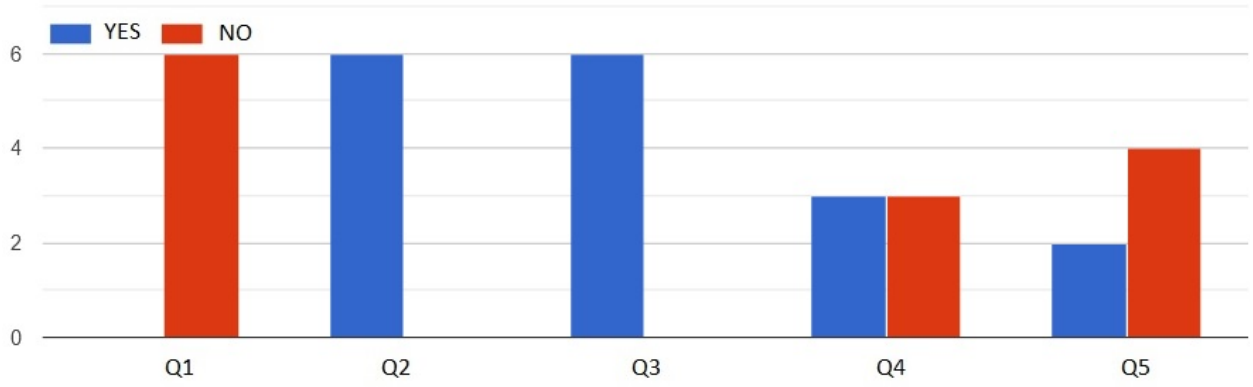

Figure 14 Participants responses to Case4

Case 5:

NEWS: An audio recording was leaked to the public, but Senad Hadžifejzović admitted that he said that Dodik should be "calmed down, killed". In an exclusive audio recording recorded with a mobile phone in the premises of FACE TV, Hadžifejzović admitted that he said "kill" instead of "pacify" and added "they should definitely be killed".

Note: This news was fake.

Q1: All of the participants selected No.

Q2: Five participants selected Yes, and one participant selected No.

Q3: All of the participants selected Yes.

Q4: Three participants selected Yes, and three participants selected No.

Q5: Two participants selected Yes, and four participants selected No.

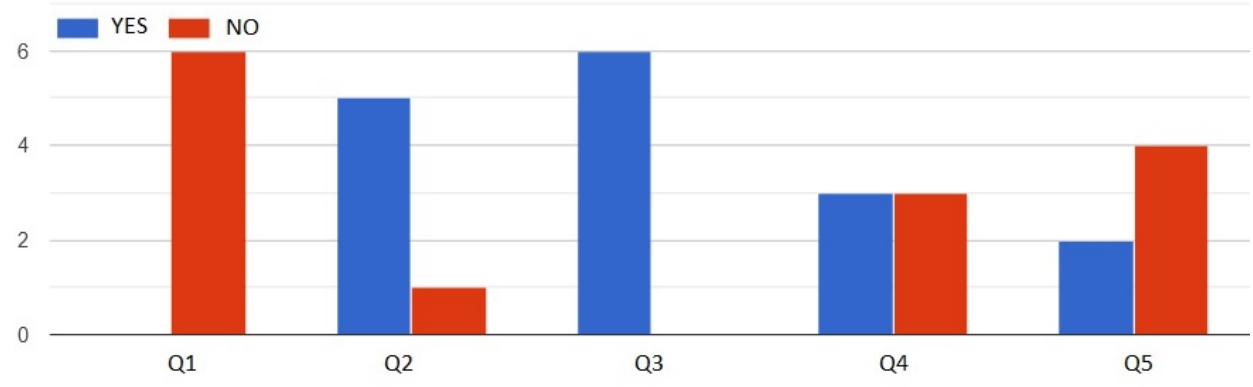

Figure 15 Participants responses to Case5

4.3 Results comparison between tests conducted before and after the participants attended the MIL awareness training.

Table 7 Table of comparison between attacks conducted before and after the awareness training was held for the participants:

	Details		
Attack performed Before the awareness session was delivered	Addressing participants through a fake LinkedIn account, using fraudulent approach	Pretexting through a phone call (Vishing) in the attempt to gain control of a student's account	Reaching out to participants to get them to join a Facebook page with the intent to have them share false information using their accounts
Attacks performed After the awareness session was delivered	Additional addressing participants through a fake LinkedIn account, using fraudulent approach	Pretexting through a phone call (Vishing) in the attempt to gain control of an employee's account	Reaching out to participants to get them to join a Facebook page with the intent to have

			them share false information using their accounts
Similarities between the two attacks	Both attacks performed on the same number of employees using a fake account and email	Both attacks conducted using same methodology and approach	Both attacks conducted using same methodology and approach
Differences between the two attacks	The type of information requested in the two attempts differed. On the fist attempt the goal was to gain trust of the participant and gather personal information. During the second attempt, the goal was to get more sensitive personal information like ID or Passport number.	Different participants were targeted in both cases and the goals were slightly different	Differences are in the information used and targeted fake information Facebook pages

Table 8 Attack success rates compared before and after the awareness training for the participants:

	TEST 1-1 vs TEST 2-1	**TEST 1-2 vs TEST 2-2**	**TEST 1-3 vs TEST 2-3**
Attack Success rate	100%	100%	0%
Attack Success rate	0%	0%	0%
Decrease in Attack Success	100%	100%	0%

I. TEST 1-1 vs TEST 2-1

Both of the tests were conducted using the same methodology and process. The goals differ slightly while in comparison TEST 1-1 was conducted with the goal of retrieving personal information enough to create a profile of the participant attacked along with information that is not publicly easily accessible i.e., personal address, personal phone number, personal preferences and capabilities related to work etc.

TEST 2-1 was an attempt to further gain trust of the attacked participant and try to lure them to share a scan of at least one personal document under the premise that the document is required for a job opportunity.

During the first test, the participant engaged was very enthusiastic about the possibility of working abroad and shared their CV and some other personal information via email, four days after the initial contact was made.

During the second test, the enthusiasm was still present and the response to share additional information (passport copy) was obliged within two days. However, when the passport scan was received, the most sensitive personal information was redacted in black. In the email message received the participant mentioned that the redacted information is blacked-out as they feel it is not crucial for applying for a job and if possessed by a malicious actor could be misused.

Therefore, it is derived that the awareness training had a positive impact on the participants who were tested in the mentioned tests with the same methodology. There was a clear sign of vigilance towards personal information sharing recorded after the MIL awareness session was delivered to the participants.

II. TEST 1-2 vs TEST 2-2

Both of the tests were conducted using the same methodology and process. The goals differ slightly while in comparison TEST 1-2 was conducted on one participant, TEST 2-2 was conducted on at least two participants including the participant in TETS 1-2.

During the first test, it was observed that the call recipient was willing to communicate without any prior checks of the person calling. The case based on a pretext was well received and further actions were explained to the caller in order of reaching the goal of resetting credentials of an existing student and taking over the account.

During the second test from the same type and methodology, it was observed that the call recipients were much more protective in regard to any information sharing. When this was noticed, we attempted to ease the conversation with communication being more emphatic to the call recipient. Even though the communication was with more ease, the recipients refused to share any type of information including any official information and has directed the caller to take more official approach by reaching out to the legal department of the organization with a formal request in writing.

Based on the experiences with TEST 1-2 and TEST 2-2, it was observed that the participants were more vigilant to potential fraudulent calls. Therefore, it is derived that the awareness training had a positive impact on the participants who were tested in the mentioned tests with the same methodology.

III. TEST 1-3 vs TEST 2-3

Both of the tests were intended to lure the participants to access a Facebook page and further share false information published on the mentioned page.

On both occasions, the participants did not take any action while they were asked to join the page. In both cases no action was recorded. During the follow-up check, we were not able to see any activity of any of the targeted participants. It is important to note that there was no attempt to befriend any of the targeted participants. Furthermore, in case the request to join any of the pages was declined, the attacker would not be able to see and record this action.

Since the second test on Facebook was left to stay for at least two weeks after it was initiated and since none of the targeted participants accepted to join any of the suggested Facebook pages with false information, it led to a potential conclusion that the targeted participants do not accept suggestions form unknown accounts. This is a sign of vigilance, which is a positive finding. Based on the above-mentioned both of the attempts were recorded as 0% success and cannot be compared to show the results before and after the participants took part in the MIL awareness training.

IV. TEST 2-4

The final test conducted on the participants was the test on how the participants perceive news published online after the awareness session in MIL was delivered. The objective was to see how the participants would react to true and fake news articles which based on their content are expected to cause an emotional response in addition to standard information receiving.

The results have shown that for four out of five cases presented, all of the participants selected NO as their response when asked if they believe if the news presented is true. Out of the mentioned four cases two were true and two were fake. For all of the mentioned four cases, all of the participants selected YES as their response when answering the question, is the news presented a clickbait.

When it comes to responding to the question whether the participants would have to verify the new presented form several sources including the ones know to be trustworthy, majority of the participants, no less than four out of six selected the option YES as their response. When asked if they believe the news presented could cause unrest within the community, no less than three participants selected YES as their response.

The only case on which the responses differ from the above-mentioned pattern of at least 50% participants giving cautious responses to the questions asked, is the case the participants perceived as the least to cause unrest within the community, and this is related to the Case 3, which is not related to a regional/local issue, but rather is a global issue regarding oil prices.

Based on the responses recorded it is safe to say that the participants do not believe in only one source when it comes to news that is potentially considered a danger to cause unrest within the community. All of the participants have shown a high level of vigilance related to multiple source information confirmations and have shown understanding that information is not to be trusted simply based on the fact that it is published online.

All of the findings and responses in the TETS 2-4, have clear indication that the awareness training in MIL has helped the participants in their attempt to verify and check the information given to them. Additionally, the participants have shown a more serious approach and satisfactory responses to the news they have perceived as potentially dangerous for the community and general society they live in.

5. Discussion

This chapter discusses the results of the research and the tests conducted in terms of the hypothesis outlined in this thesis, based on the literature overview, the MIL awareness training conducted, and, in the end, the tests executed in the XYZ organization. When all mentioned is taken into consideration, research has been conducted based on the hypothesis of this thesis, where the main goal is to show that MIL awareness training does actually increase the level of security within an organization in Bosnia and Herzegovina.

The results of a research by Siponen, et al. (2013) has used the survey responses of 669 employees from various companies in Finland, which shows that the delivery of messages related to awareness in information security significantly reduces the potential for the employees to miss-detect a malicious activity within their organization or, to put it in other words, awareness increases the levels of information security within an organization. Another research conducted by Bauer and Bernroider (2015), based on the survey responses of 183 bank employees working in banks in Germany, shows that information security awareness does have positive effects in increasing the security of the organizations they work for. These mentioned research results found in publications indicate that awareness training does have a positive effect in strengthening the posture of an organization. This is also expected since awareness strengthens the weakest link which is the human factor.

The same is expected to be valid when it comes to awareness training for MIL as an answer to the threats imposed by information disorder. The reason why the awareness effectiveness is compared to the standard ‘industry’ based awareness training in information security is due to the fact that MIL awareness trainings are not present in organizations as such. There might be some relevant topics, but none were found for the needs of this thesis.

The research and results of this thesis have shown that MIL awareness training as a subset of a general Information Security training program does increase the levels of security and higher levels of vigilance among members of the organization. Apart from the standard protection of job-related information and ensuring information is not leaked, the MIL training has increased the ability of that participants in processing the information served to them via online platforms i.e., news portals and/or social media.

When comparing the results, we have seen a significant increase in the way the participants responded after they attended the awareness training in comparison to the results of tests conducted before the awareness session. The results of TEST 1-1 vs TEST 2-1, which are similar in the method of execution have seen a 100% decrease in success. Although the sample is not large, it is clear that the targeted participants responded in a better way showing understanding of dangers that can evolve from trusting an individual which has not offered any kind of proof that they are who they represent them selves to be. While in the initial test before the awareness training, all of the information requested was shared without any questions asked. It is significant to note one more time that the response after the awareness training contained an email message mentioning that the information requested is not shared in full scope due to the concerns that the information might be misused.

In similar fashion TEST 2-2 has shown better responses, as it was conducted after the awareness session, to the results of TETS 1-2, which was a success in the vishing with pretext attack. During the execution of TEST 2-2, it was clear that the level of trust given to the caller was not the same as it was for TEST 1-2. The participants targeted in TEST 2-2, have shown to be more cautious and asking for an official communication to be sent before they disclose any information. This type of instruction was given to the participants during the discussion part of the MIL awareness session.

Then we have the TEST 1-3 and TEST 2-3, which were both unsuccessful. The participants were approached on Facebook, with the intent to get them to share false information via a Facebook page that shares misinformation. The targeted participants have shown that they will not blindly without any confirmation, accept such attempts. Even though we were not able to detect any major activities on the participants Facebook accounts, based on the information form the baseline survey, the participants have said that four out of six have Facebook accounts, while three of those four access the platform on a daily basis. This can indicate that they simply did not want to accept to be added to the Facebook page with false information, which is a positive response regardless of both tests of the same method being unsuccessful in its execution.

In the end the TEST 2-4, has given a lot of insight to the perception of the awareness training by the participant. Their responses to the five cases of news out of which two were true and three were false indicated a positive result in regard to the MIL awareness training the participants

have attended. Noting that no less than four of the participants have confirmed that they would check the information source and the news received from several sources before accepting it as the truth. Additionally, all of the participants chose not to believe in any of the news cases which they perceived as the ones most likely to pose a dangerous response within the community and the society, even though it was a mix of true and false information for the four cases they selected such options for. So, to sum up the analysis of TEST 2-2, the participants have used the knowledge form the awareness training to respond to the five cases and have shown a high level of understanding that information need to be verified from several sources, that it can cause unrest and pose danger to the organization, the community and the society they live in.

All of the results have a clear indication that MIL awareness training can be used to decrease the vulnerability of the human factor which is mostly targeted when it comes to information disorder and IDT attempts. The participants have shown understanding and the importance of MIL in its attempt to reduce the risks associated with information disorder.

6. Conclusion

The world we live in today is more and more significantly affected by the digital realm in which most of us complete more and more of our daily tasks both in our professional careers and in our private activities. As it is almost impossible to imagine our lives without the 'windows' into the digital realm though smartphones, tablets, laptops, and other devices, we as humans are constantly exposed to the content over which in most cases, we don't have control over. Although the intentions behind the development of the digital realm in our lives were good, the negative aspects are constantly showing that it affects our humane domain as we are prone to information receiving with an emotional response to the same. The human vulnerability referred to as the 'human factor', is what makes us humans the weakest link in information security and more prone to negative consequences caused by malicious intent behind information disorder and IDT's. To better understand the situation and threats the digital realm poses to humans we need to understand that the human factor is most commonly relating to the mishaps initiated by human interactions that could lead to creation of vulnerabilities to the ICT systems used to access the digital realm, including business-related applications (Parsons et al., 2010).

The effects of IDTs pose an ever-growing threat to the human factor, especially due to the fact that most humans are not aware of the IDTs. In many cases even when it is brought to their attention, individuals disregard the threats thinking it does not apply to them. The IDT techniques are fully in line with already known social engineering types of attacks which target the human factor.

The exposure to IDTs is amplified by bot networks and AI based algorithms that create information bubbles for individuals, which makes it even harder to detect when someone is exposed to IDTs. The Algorithms also referred to as the 'black boxes' of the internet, are known to be used by various online corporations to lure users to use their services and to be exposed to the content they want to serv. However, the same technologies are used in a malicious way even on the platforms owned by the mentioned large online corporations and digital media giants.

Digital media giants are using advanced technologies and AI to control which content is reaching us. Additionally, they are not doing much to protect the information consumers from malicious content by applying ethical-based restrictions on their platforms. It was shown that various actors from state level to individual are able to use the situation and sway the situations to support their own agenda. Such cases were recorded during the BREXIT vote, amplification misinformation for the antivaxxer movement, during the COCID-19 pandemic and many more. These real-life cases have caused significant political turmoil and have gone to the extent of endangering human lives.

The MIL as a response to the IDTs is a form which can help the fight against the issues faced. However, MIL is not very known to the general public and the situation requires the development of a different approach to resolve the issues at hand. For the purpose of proving that MIL training can help in organizations, we have developed, and awareness training based on MIL to see if educating individuals can lower the risks caused by exposure to IDTs. With proper knowledge the ability of an individual to recognize IDTs is much higher and the individual in more vigilant. Although MIL is known within the scientific spheres, it is not commonly known to an average individual and is not present as a part or a dedicated segment within an awareness training in organizations.

Utilizing the experiences of delivering awareness training in organizations related to the threats towards the information security of an organization, the MIL awareness training was developed. As McCormac, et al. (2010) state in their study that awareness training is the best approach to protect against social engineering and other potential attacks on an organization. The same study emphasizes that the awareness and education related to information security should start as soon as an employee joins the organization. Awareness needs to be performed on a regular basis for all employees as indicated in the paper by He, et al. (2019) and special attention needs to be given to the awareness which is one of the conclusions in a paper by Aloul (2012). The goal was to see if the positive effects of general information security awareness programs can be applied to MIL training, when it is done separately and fully dedicated to the topics related to IDTs and MIL as the response to the IDTs.

The XYZ organization was chosen along with specified participants. The ones chosen were tested prior to the awareness session and after the session. The results were compared to see if there are improvements due to the MIL awareness training. All in all, the research and have shown that the hypothesis of this thesis stands: **The awareness-raising program in the field of MIL through the professional training program in organizations leads to a reduction in the risk of harming the security of the individual, the community, and the state, which are caused by information disorder and hybrid asymmetric attacks.**

The results collected for the tests conducted on the participants prior and after the awareness training was delivered has shown a significant increase in the levels of vigilance, carefulness, skepticism, and critical thinking in all of the situations the participants have found the selves in. The approach conducted in this thesis needs to be applied in further research, and used as a reference to get support in the execution of more and more MIL based awareness training and other forms of education approaches within organizations and the general society, we live in.

6.1 Limitations of this thesis.

The attempts to protect organizations from malicious actors are constantly evolving every day based on new approaches to perform attacks on them. Apart from standard threats within information security, IDT techniques are an additional vector in most cases not considered by organizations as a spectrum of new threat vectors. Considering the human factor, it is always

interesting to find and identify the new approaches attackers and mass behavior influencers have, which are based on all elements which include information extractions, targeting private platforms, i.e., social media accounts and following daily routines and habits of targeted individuals, monitoring their online habits etc.

The testing for this thesis was not performed on large scale and was conducted only within one educational organization in Sarajevo, Bosnia and Herzegovina. The tests conducted were based on the agreed approach with the point of contact within the XYZ organization, which was limited in comparison to what real-life malicious actors potentially could achieve.

6.2 Open research questions and future work.

Future research and testing of the effects MIL provides as opposed to information disorder and hybrid asymmetric attacks, needs to be continued for different types of organizations and on a larger scale. Organizations need to acquire support of the top management in their attempts to conduct MIL based training, which can be approached in different ways i.e., full penetration tests using IDTs etc.

The awareness programs should be tested in different ways, different environments and in various locations. Some of the questions that could be asked for the purpose of academic research are:

- How do IDTs affect regions with state restricted/censored internet access?
- How does MIL awareness training resonate in different cultures of the world?
- Do different types of organizations require different approaches and selection of MIL related topics in their awareness programs?
- Can large media corporations be persuaded to support MIL on their platforms and how it can be done?
- What are the new challenges for MIL considering the recent use of AI software like chat GPT?

The above listed questions are just a few, that should be researched to determine the effectiveness of MIL as a whole and MIL based awareness trainings as a continuous improvement to the research conducted in this thesis.

Literature

1. (IEC), International Organization for Standardization (ISO) and the International Electrotechnical Commission. 2017. "ISO/IEC 27001 is an international standard to manage information security."

2. Abawajy, J. 2014. *User preference of cyber security awareness delivery methods, School of Information Technology.* Waurn Ponds, Australia: School of Information Technology, Deakin University.

3. Aloul, F., A. 2012. *The Need for Effective Information Security Awareness.* Sharjah, UAE: American University of Sharjah.

4. Armstrong, Mark. 2021. *Tens of thousands protest COVID-19 measures in Vienna.* December 12. Accessed December 15, 2021. https://www.euronews.com/2021/12/12/large-protests-in-vienna-against-covid-19-measures.

5. Bajwa, Aman. 2021. "INFORMATION DISORDER, THE TRIUMVIRATE, AND COVID-19: HOW MEDIA OUTLETS, FOREIGN STATE INTRUSION, AND THE FAR-RIGHT DIASPORA DRIVE THE COVID-19 ANTI-VACCINATION MOVEMENT." *The Journal of Intelligence, Conflict, and Warfare* 1 to 30.

6. Barbara Thomass, Martin Marinos, Jonila Godole, Lejla Turčilo, Orlin Spassov, Stjepan Malović, Remzie Shahini-Hoxhaj, Nataša Ružić, Aneta Gonţa, Marina Tuneva, Adina Marincea, Ana Milojević, Nikoleta Daskalova. 2021. "Three Decades Later: The Media in South East Europe after 1989." Konrad-Adenauer-Stiftung e.V. .

7. Bauer, S., Bernroider, E. W.N., & Chudzikowski, K. 2013. *End User Information Security Awareness Programs for Improving Information Security in Banking Organizations: Preliminary Results from an Exploratory Study.* Milano, Italy: Proceedings of the Eighth Pre-ICIS Workshop on Information Security and Privacy.

8. Bossio, Belinda Barnet and Diana. 2020. *Netflix's The Social Dilemma highlights the problem with social media, but what's the solution?* October 6. Accessed December 5, 2021. https://theconversation.com/netflixs-the-social-dilemma-highlights-the-problem-with-social-media-but-whats-the-solution-147351.

9. Bowen, B., M., Devarajan, R., & Stolfo, S. 2011. *Measuring the Human Factor in Cyber Security.* New York, USA: Department of Computer Science, Columbia University.

10. Brian Hughes, Cynthia Miller-Idriss, Rachael Piltch-Loeb, Beth Goldberg, Kesa White, Meili Criezis and Elena Savoia. 2021. "Development of a Codebook of Online Anti-Vaccination Rhetoric to Manage COVID-19 Vaccine Misinformation." *International Journal of Environmental Research and Public Health.*

11. Burki, Talha. 2019. "Vaccine misinformation and social media." *The Lancet (thelancet.com)* 258-259.

12. Burns, Axel. 2019. "Filter Bubble." *Internet Policy Review.*

13. Cadwalladr, Carole. 2019. "Facebook's role in Brexit — and the threat to democracy." *TED Talk.*

14. Camerota, Alisyn. 2021. *Former QAnon follower: I really believe it's a cult.* Video Report: https://www.youtube.com/watch?v=wYDMTCqZ0XU, CNN on Youtube.

15. Carlsson, Ulla. 2019. "Understanding Media and Information Literacy (MIL) in the Digital Age: A QUESTION OF DEMOCRACY." Department of Journalism, Media and Communication (JMG) University of Gothenburg.

16. Clayton, Cordel Green and Anthony. 2021. "Ethics and AI Innovation." *International Review of Information Ethics.*

17. Commerce, National Institute of Standards and technology (NIST) U.S. Department of. 2021. *Awareness, Training, and Education Controls.* Accessed December 14, 2021. https://csrc.nist.gov/glossary/term/awareness_training_and_education_controls.

18. 2019. *BREXIT.* Directed by Toby Haynes. Performed by Benedict Cumberbatch.

19. Dictionary, Cambridge. 2022. *dictionary.cambridge.org.* Accessed 2022. https://dictionary.cambridge.org/dictionary/english/filter-bubble.

20. Editors, History.com. 2018. *Edward Snowden discloses U.S. government operations.* June 26. Accessed December 15, 2021. https://www.history.com/this-day-in-history/edward-snowden-discloses-u-s-government-operations.

21. Editors, Imperva.com. 2021. *Social Engineering.* December. Accessed December 15, 2021. https://www.imperva.com/learn/application-security/social-engineering-attack/.

22. Elvira Ortiz-Sanchez, Almudena Velando-Soriano, Laura Pradas-Hernánde, Keyla Vargas-Román, Jose L. Gómez-Urquiza, Guillermo A. Cañadas-De la Fuente and Luis Albendín-García. 2020. "Analysis of the Anti-Vaccine Movement in Social Networks: A Systematic Review." *Environmental Research and Public Health.*

23. English, Al Jazeera. 2011. *UK autism-vaccine study was 'fraud'.* TV News Article, English, Al Jazeera.

24. Frenkel, Cecillia Kang and Sheera. 2020 . "'PizzaGate' Conspiracy Theory Thrives Anew in the TikTok Era." *The New York Times https://www.nytimes.com/2020/06/27/technology/pizzagate-justin-bieber-qanon-tiktok.html.*

25. Hadlington, L. 2018. *The Human Factor in Cyber Security: Exploring the Accidental Insider.* Leicester, UK: De Montfort University.

26. He, W., Anwar, M., Ash, I., Li, L., Yuan, X., Xu, L., & Tian, X. 2019. *Effects of Evidence Based Malware Cyber Security Training on Employees.* Cancun: Americas Conference on Information Security.

27. Hegarty, Stephanie. 2020. *QAnon: The conspiracy theory spreading fake news* . Video News report https://www.youtube.com/watch?v=u8Gd9MJsnnE, BBC Newsnight.

28. Hunt Allcott, Matthew Gentzkow and Chuan Yu. 2019. "Trends in the diffusion of misinformation on social media." *Research and Politics,* April-June.

29. IFLA, International Federation of Library Associations and Institutions. 2011. "IFLA Media and Information Literacy Recommendations." *IFLA Media and Information Literacy Recommendations.* Den Haag: IFLA, December 7.

30. Ingram, Katrina. 2021. "Constructing AI: Examining how AI is shaped by data, models and people." *International Review of Information Ethics.*

31. Julian McDougall, Marketa ZezulkovaBarry van Driel and Dalibor Sternadel. 2018. *Teaching media literacy in Europe: evidence of effective school practices in primary and secondary education.* Analytical, Luxembourg: Publications Office of the European Union.

32. Lessenski, Marin. 2021. *Media Literacy Index 2021.* March 14. Accessed December 1, 2021. https://osis.bg/?p=3750&lang=en.

33. Liang Wu, Fred Morstattery, Kathleen M. Carleyz, and Huan Liu. 2019. "Misinformation in Social Media: Definition, Manipulation and Detection." Arizona State University, Tempe, AZ,; USA USC Information Sciences Institute, Marina Del Rey, CA, USA; Carnegie Mellon University, Pittsburgh, PA, USA.

34. Limited, Guardian News & Media. 2021. "Violence breaks out in Brussels in protests against Covid restrictions." https://www.theguardian.com/world/video/2021/nov/22/violence-breaks-out-in-brussels-in-protests-against-covid-restrictions-video.

35. Longley, Robert. 2020. *What Is Astroturfing in Politics? Definition and Examples.* October 14. Accessed November 28, 2021. https://www.thoughtco.com/what-is-astroturfing-definition-and-examples-5082082.

36. Matus Tomlein, Branislav Pecher, Jakub Simko, Ivan Srba, Robert Moro, Elena Stefancova, Michal Kompan, Andrea Hrckova, Juraj Podrouzek and Maria Bielikova. 2021. "An Audit of Misinformation Filter Bubbles on YouTube: Bubble Bursting and Recent Behavior Changes." *In Fifteenth ACM Conference on Recommender Systems.* Amsterdam.

37. Mauldin, Stacu L Benoit and Rachel F. 2021. "The “anti-vax” movement: a quantitative." *BMC Public Health.*

38. McKenzie Himelein-Wachowiak, Salvatore Giorgi, Amanda Devoto, Muhammad Rahman, Lyle Ungar, Andrew Schwartz, David H Epstein, Lorenzo Leggio. 2021. "Bots and Misinformation Spread on Social Media: Implications for COVID-19." *JOURNAL OF MEDICAL INTERNET RESEARCH.*

39. Medve, Flora. 2022. *STATISTA.* Accessed November 5, 2022. https://www.statista.com/statistics/1129313/bosnia-and-herzegovina-covid-19-cases/.

40. Merriam-webster.com. 2022. "Merriam-webster Dictionary." https://www.merriam-webster.com/dictionary/algorithms.

41. Michael Haenlein, Andreas Kaplan, Chee-Wee Tan and Pengzhu Zhang. 2019. "Artificial intelligence (AI) and management analytics." *Journal of Management Analytics.*

42. Mouton, F., Leenen, L., & Venter, H.S. 2016. *Social Engineering Attack Examples, Templates and Scenarios.* Pretoria, South Africa: Department of Computer Science University of Pretoria.

43. Neil G. Ruiz, Khadijah Edwards and Mark Hugo Lopez. 2021. *One-third of Asian Americans fear threats, physical attacks and most say violence against them is rising.* April 21. Accessed December 5, 2021. https://www.pewresearch.org/fact-tank/2021/04/21/one-third-of-asian-americans-fear-threats-physical-attacks-and-most-say-violence-against-them-is-rising/.

44. News), CBSN (CBS. 2021. *Examining the spread of online misinformation during COVID-19 pandemic.* CBSN: https://www.youtube.com/watch?v=oZWpaA6XkI0.

45. 2019. *The Great Hack.* Directed by Jehane Noujaim.

46. Oslobođenje. 2022. "Komentari građana o vakcinaciji."

47. Oxford Online Learner's Dictionary. 2020. https://www.lexico.com/, .

48. Parsons, K., McCormac, A., Bautavicius, M., & Ferguson, L. 2010. *Human Factors and Information Security: Individual, Culture and Security Environment.* Edinburgh, Australia: Defense Science and Technology Organization of Australian Government, Department of Defense.

49. Pickle, Brian. 2022. Accessed Movember 5, 2022. https://techterms.com/definition/metadata.

50. Poremba, Sue. 2021. *Propaganda as a Social Engineering Tool.* July 13. Accessed December 1, 2021. https://securityboulevard.com/2021/07/propaganda-as-social-engineering-tool/.

51. Press, The Associated. 2021. *EXPLAINER: Just what are 'The Facebook Papers,' anyway?* October 25. Accessed December 5, 2021. https://apnews.com/article/what-are-the-facebook-papers-10e59530a699db5345ac3931509778b2.

52. Research, Jigsaw. 2020. *News Consumption in the UK: 2020.* Analytical, Ofcom .

53. Samantha Bradshaw, Hannah Bailey and Philip N. Howard. 2021. "Industrialized Disinformation: 2020 Global Inventory of Organized Social Media Manupulation." University of Oxford.

54. Shearer, Elisa. 2021. *More than eight-in-ten Americans get news from digital devices.* January 12. Accessed December 12, 2021. https://www.pewresearch.org/fact-tank/2021/01/12/more-than-eight-in-ten-americans-get-news-from-digital-devices/.

55. Shimabukuro, Frank DeStefano and Tom T. 2019. "The MMR Vaccine and Autism." National Library of Medicine https://www.ncbi.nlm.nih.gov/pmc/about/disclaimer/.

56. Shuhaili Talib, Nathan Clarke and Steven Furnell. 2010. "An Analysis of Information Security Awareness within Home and Work Environments."

57. Siponen, M., Mahmood, M., A., & Pahnila, S. 2013. *Employees adherence to information security policies: An exploratory field study.* Amsterdam, Netherlands: Elsevier Publishing Company.

58. Smajić, Prof. Dr. Mirza. 2021. *Zoom Lecture Information Security Masters program at the Faculty of Political Sciensce at the Sarajevo University.* Sarajevo, November 11.

59. Sokol, Anida. 2021. *Polarized Public trust in the media and social networks in Bosnia and Herzegovina (Original: Polarizirano povjerenje javnosti u medija i društvene mreže u Bosni i Hercegovini).* June 9. Accessed December 1, 2021.

https://www.media.ba/bs/magazin-novinarstvo/istrazivanje-gradani-u-bih-najvise-vjeruju-televiziji.

60. Sten Hansson, Kati Orru, Sten Torpan, Asta Bäck, Austeja Kazemekaityte,Sunniva Frislid Meyer, Johanna Ludvigsen, Lucia Savadori, Alessandro Galvagni and Ala Pigrée. 2021. "COVID-19 information disorder: six types of." *Journal of Risk Research* 1 to 16.

61. Terranova security. 2020. https://terranovasecurity.com/examples-of-social-engineering-attacks/.

62. Tijana Cvjetićanin, Emir Zulejhić, Darko Brkan, Biljana Livančić-Milić. 2019. "Dezinformacije u online sferi: slucaj BiH." Sarajevo: Udruženje građana ""Zašto ne".

63. 2021. *DAN U ŽIVOTU JEDNOG BOTA.* Directed by Bojan Tomić.

64. U.S. Department of Commerce, Mark Wilson and Joan Hash. 2003. *NIST Special Publication 800-50: Building an Information Technology Security Awareness and Training Program.*

65. University of Sarajevo/Consyltancy group for Media and Information Lirecy in Bosnai and Hezegovina. 2019. "Deaclaration about the importance of Media and Information Litrecy in Bosnia and Herzegovina (Original:DEKLARACIJA O ZNAČAJU MEDIJSKE I INFORMACIJSKE PISMENOSTI U BOSNI I HERCEGOVINI)." Sarajevo: University of Sarajevo, January 28.

66. Vajzović, Emir. 2020. *Medijska i Informacijska Pismenost: ISTRAŽIVANJE I RAZVOJ.* Fakultet političkih nauka Univerziteta u Sarajevu.

67. Vajzović1, Prof. Dr. Emir. 2021. *Zoom Lecture Media and Information Literacy Lecrute a part of Module1 of Information Security Masters program at Faculty of Political Science at the Sarajevo University.* Sarajevo, November 19.

68. Vajzović2, Prof. Dr. Emir. 2020. "Digitalna transformacija sigurnosti i algoritamska demokratija. Sarajevo Social Science." Sarajevo: Fakultet Politickih Nauka.

69. Vajzović3, Prof. dr. Emir. 2021. *Zoom Lecture Information Security Masters program at the Faculty of Political Science, Sarajevo University.* Sarajevo, December 09.

70. World, TRT. 2021. *The attack on Capitol Hill explained.* Video report: https://www.youtube.com/watch?v=tCTFhvVQG_Y, YouTube .

71. Wünsch, Silke. 2013. "Metadata reveals a lot." DW.com https://www.dw.com/en/metadata-reveal-much-about-what-you-do-online/a-16945685.

72. Yildirim, E. Y., Akalp, G., Aytac, S., & Bayram, N. 2011. *Factors influencing information security management in small and medium sized enterprises: A Case study from Turkey.* Bursa: Uludag University, Computer Technology and Programming Department, Bursa, Turkey.

73. Željeznik, Barbara. 2021. "Lažno informiranje na internetu." Undergraduate thesis / Završni rad Pula / Sveučilište Jurja Dobrile u Puli.

List of Figures:

List of Tables:

	FAKULTET POLITIČKIH NAUKA	Obrazac AR
UNIVERZITET U SARAJEVU – FAKULTET POLITIČKIH NAUKA IZJAVA o autentičnosti radova		Stranica **1** od **1**

Naziv odsjeka i/ili katedre: **Sigurnosne i mirovne studije**
Predmet: **Informacijska sigurnost**

IZJAVA O AUTENTIČNOSTI RADOVA

Ime i prezime:	Anes Mirojević
Naslov rada:	**Značaj cjeloživotnog učenja iz medijske i informacijske pismenosti za informacijsku sigurnost: informacijski nered u toku COVID-19 pandemije**
Vrsta rada:	Završni magistarski rad
Broj stranica:	97

Potvrđujem:

- da sam pročitao/la dokumente koji se odnose na plagijarizam, kako je to definirano Statutom Univerziteta u Sarajevu, Etičkim kodeksom Univerziteta u Sarajevu i pravilima studiranja koja se odnose na I i II ciklus studija, integrirani studijski program I i II ciklusa i III ciklus studija na Univerzitetu u Sarajevu, kao i uputama o plagijarizmu navedenim na web stranici Univerziteta u Sarajevu;
- da sam svjestan/na univerzitetskih disciplinskih pravila koja se tiču plagijarizma;
- da je rad koji predajem potpuno moj, samostalni rad, osim u dijelovima gdje je to naznačeno;
- da rad nije predat, u cjelini ili djelimično, za stjecanje zvanja na Univerzitetu u Sarajevu ili nekoj drugoj visokoškolskoj ustanovi;
- da sam jasno naznačio/la prisustvo citiranog ili parafraziranog materijala i da sam se referirao/la na sve izvore;
- da sam dosljedno naveo/la korištene i citirane izvore ili bibliografiju po nekom od preporučenih stilova citiranja, sa navođenjem potpune reference koja obuhvata potpuni bibliografski opis korištenog i citiranog izvora;
- da sam odgovarajuće naznačio/la svaku pomoć koju sam dobio/la pored pomoći mentora/ice i akademskih tutora/ica.

Mjesto, datum Doha, Katar, 29.05.2023

Potpis