CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Political Studies

Department of Security Studies

Master's Thesis

2022 **Bc. Andrea Tunysová**

CHARLES UNIVERSITY

FACULTY OF SOCIAL SCIENCES

Institute of Political Studies

Department of Security Studies

Predicting and Preventing Terrorism with Artificial Intelligence and Machine Learning: Implications for Security in Israel

Master's thesis

Author: Bc. Andrea Tunysová

Study programme: Security Studies

Supervisor: Mgr. Petr Špelda, Ph.D.

Year of the defence: 2022

Declaration

1. I hereby declare that I have compiled this thesis using the listed literature and resources only.
2. I hereby declare that my thesis has not been used to gain any other academic title.
3. I fully agree to my work being used for study and scientific purposes.

In Prague on May 1, 2022 Bc. Andrea Tunysová

References

TUNYSOVÁ, Andrea. *Predicting and Preventing Terrorism with Artificial Intelligence and Machine Learning: Implications for Security in Israel.* Praha, 2022. 78 pages. Master's thesis (Mgr.). Charles University, Faculty of Social Sciences, Institute of Political Studies. Department of Security Studies. Supervisor Mgr. Petr Špelda, Ph.D.

Length of the thesis: 163 639 characters with spaces

Abstract

The thesis examines the use of Artificial Intelligence and Machine Learning for predicting and preventing terrorism and the resulting security risks. At the conceptual level, the thesis examines the approach of predicting threats with a focus on predictive policing and presents risks associated with the use of predictive machine learning systems, which are then discussed within the context of counterterrorism. The paper aims to answer the question to which extent we can rely on machine learning systems used to predict and prevent terrorism and what are the implications of their use for security in Israel. The thesis points out that although the predictive tools seem to be faster and more precise than human analysts, they cannot be trusted to a full extent. If the results of these systems are used to employ strict measures such as the restriction of a suspect's liberty, it may lead to the violation of human rights. Therefore, in the case of counterterrorism in Israel, which is sometimes presented as the only democracy in the Middle East, it is necessary to bear in mind the risks associated with the limits of predictive machine learning systems together with the up-to-date practice of Israeli security agencies and Israeli historical-social context, indicating that it would be very difficult, if not impossible, objectively and precisely predict terrorist activities and identify terrorists before they commit an attack.

Abstrakt

Magisterská diplomová práce zkoumá využití nástrojů umělé inteligence a strojového učení pro předvídání a předcházení terorismu a s tím spojená bezpečnostní rizika. Na konceptuální úrovni se práce zabývá přístupem předvídání hrozeb se zaměřením na predictive policing a rozebírá jeho dopady na bezpečnost při použití metod strojového učení. Rizika, která jsou spojena s užitím prediktivních nástrojů a strojového učení jsou následně rozebírána v kontextu proti-terorismu. Cílem práce je odpovědět na otázku, do jaké míry lze důvěřovat prediktivním nástrojům strojového učení pro předvídání a předcházení terorismu, a jaké bezpečnostní dopady může mít jejich užití v Izraeli. Práce poukazuje na to, že ačkoliv se analýza dat prediktivních nástrojů může zdát rychlejší a přesnější než posouzení člověkem, nelze na tyto technologie spoléhat v plném rozsahu. Naopak, pokud výsledky těchto systémů povedou k razantním opatřením jako omezení svobody jedince, může tak docházet k porušení lidských práv. Proto i v případě Izraele, který je často prezentován jako jediná demokracie na Blízkém východě, je třeba brát v úvahu rizika spojená s limity nástrojů strojového učení v kombinaci s dosavadní praxí izraelských bezpečnostních složek a historicko-společenským kontextem Izraele, který napovídá, že je obtížné, ne-li nemožné objektivně a přesně předpovědět možné teroristické aktivity a identifikovat teroristy předtím, než spáchají útok.

Keywords

predictive policing, terrorism, artificial intelligence, machine learning, Israel

Klíčová slova

predictive policing, terorismus, umělá inteligence, strojové učení, Izrael

Název práce

Předpovídání a předcházení terorismu pomocí umělé inteligence a strojového učení:
Důsledky pro bezpečnost v Izraeli

Acknowledgement

I would like to express my gratitude to my supervisor Mgr. Petr Špelda, Ph.D. for his guidance, helpful advice and always very kind approach.

My gratitude goes also to my loving family who patiently kept supporting me throughout this seemingly never-ending study.

Table of Contents

Introduction

The development of Artificial Intelligence (AI) and Machine Learning (ML) has had a significant impact on today's society as their applications affect all spheres of our lives – healthcare, agriculture, financial sector, education as well as security. Although these technologies raise high expectations and seem to have the potential for improving our everyday lives, as the United Nations Secretary-General António Guterres stated in 2018, "they are not risk-free, and some inspire anxiety and even fear," and might "be used to malicious ends or have unintended negative consequences." (United Nations 2018, p. 8) This statement applies to all of the spheres in which AI and ML have been used, whether to a lesser or greater extent, and crime and terrorism prevention are not the exceptions. The predictive programs have raised hopes among law enforcement agencies' officials as they have promised to simplify and speed up the process of data analysis and to increase the accuracy and objectiveness of the results. The introduction of ML into predictive programs was expected to enhance their capabilities and improve their predictions based on gained experience. At the same time, however, the algorithms have been proven to be unfair and regarding ML, in some cases inexplicable, questioning its ethical and fair use as well as the accountability for such decisions.

National security strategists have similar thoughts about the potential benefit of these technologies, as they assume that predicting major security threats would allow them to take appropriate measures to avoid these threats. Terrorism is often represented as such a vital threat and although each government has a different definition for this phenomenon, in all cases the term arouses fears and thoughts of irreversible harm that might occur anytime and anywhere without warning. This uncertainty and the persistent possibility of huge damages made the concept of terrorism being used as a justification for the employment of exceptional measures by governmental agencies. The early prediction of a terrorist attack promises the prevention of such a disaster and loss of lives. Thus, it is not surprising that the intelligence services have also found great interest in these technologies.

The aim of this paper is to examine to which extent we can rely on the ML systems to predict and prevent terrorism and to assess the security implications of their use in Israel. The thesis will be concerned with the functioning of ML systems and their employment in predictive policing in order to identify the most significant risks the law enforcement agencies could face in their attempt to predict and prevent crime. The thesis aims to demonstrate that intelligence services will face similar problems in an effort to identify terrorists and reveal their intentions before they have a chance to realise them.

Conceptually, the thesis will first present a review of the existing literature focused on predicting and pre-empting harmful events. It will also explain the reasons behind the adoption of preventive approaches and describe the role of technology in preventing harm. The paper will then introduce the functioning of ML predictive policing in general and present its technical and ethical limits. Due to a limited amount of data on its application in counterterrorism, the thesis aims to explain the basic limits of the predictive systems through the predictive policing approach. Although the predictive counterterrorist measures cannot be examined entirely through the lens of predictive policing in the empirical sense, as terrorism and criminality are two distinct phenomena of incomparable nature, the technology that is used functions on the same principles and bears similar risks. Namely, the algorithm biases and lack of transparency and inexplicability accompanying predictive systems' results will be examined. Accordingly, the thesis will stress the issues of accuracy and the production of false negatives and false positives. The last technical concerns examined are the non-robustness and vulnerability of ML systems to adversarial actions, which might make the system produce incorrect answers. Finally, the impact of ML on surveillance and its ethical implications on human rights will be emphasized.

The conclusions will be drawn on the criteria elaborated in the first part. These criteria in the form of limits of ML in predicting crimes will be discussed within the context of counterterrorism. Accordingly, the thesis will examine how these shortcomings might be manifested in the efforts to predict and prevent terrorism through identifying terrorists, and what might be the security impacts in Israel. The security implications for Israel will be evaluated based on the Israeli approach to terrorism in past years and present practices. Israel is often claimed to be threatened by terrorism since the creation of the state. On the other hand, the country is also praised for being one of the most developed nations in terms of capabilities in cyberspace and artificial intelligence. Thanks to their position of a technologically advanced nation and successful fighter against terrorism, Israeli security agencies have become providers of education and consultation in the field worldwide. There is a sufficient body of literature on the Israeli approach to terrorism and its practices for detecting terrorists. However, less literature is concerned with the security implications of such practices when they are enriched by ML technologies. Although it is impossible to determine the concrete features and functioning of the systems the Israeli intelligence services work with, drawing on the literature on the approach to terrorism in Israel and sources concerned with predictive policing and the functioning of AI/ML in general, the thesis's research questions will be answered.

1. Review of the existing literature

There has been a considerable body of literature examining the efforts to anticipate criminal activity as a part of criminal analysis literature. These efforts have aimed to identify trends in past crimes, deploy patrols more effectively and thus prevent undesirable events in time. Such procedures involve crime mapping and offender profiling, but also the attempts to calculate under which conditions the crime is likely to occur. Therefore, the effort to identify future threats and to determine the strategies to thwart possible harmful events is not a new approach. Despite the absence of a consensus on whether and to which extent the earlier crime analysis techniques differ from today's predictive policing, it is clear that what has changed in past decades is the increasing amount of data the law enforcement agencies have at their disposal, the advanced technologies and the growing insistence on predicting and forestalling crime before it happens rather than on post-crime investigation (Egbert and Leese 2021, p. 19-25).

The authors concerned with ethical and practical implications of crime prediction and crime prevention usually build on pre-crime and precautionary principles that underline the necessity of undertaking the action to prevent harmful yet uncertain events from occurring. They often refer to the novella *Minority report* written by Philip K. Dick, in which *Precrime* is a term designed to name a law enforcement agency, whose members possess the ability to foresee future criminality. The agency is therefore capable to take appropriate measures to prevent it. The story illustrates how the predictions limit one's right to decide about the future (as the future is already pre-determined) and how they might become a self-fulfilling prophecy (Lennon 2015, Perry et al. 2013, McCulloch and Wilson 2016). The literature concerned with predictive and pre-emptive approaches to criminal offences mainly examines their impacts on criminology and justice (McCulloch and Wilson 2016, Egbert and Leese 2021, Perry et al. 2013). According to McCulloch and Wilson, the pre-crime approach is "ideally suited" to terrorism as the application of precautionary measures to something so ill-defined as terrorism allows governments to gain control over everything and everyone who poses a risk to their political interests (McCulloch and Wilson 2016, p. 71). While this literature is rich in ethical implications of the pre-crime approach, the role of technologies is often neglected and does not offer a connection with the most advanced predictive technologies.

On the other hand, literature on anticipation and prevention of terrorism through ML methods usually contains scientific articles from the field of information technology and deals with the technical aspects and functioning of big data analysis and ML tools and their potential for being used by security agencies and intelligence services. Such applications include the detection of promoting extremism and incitement to terrorism on social networks (Kaur, Kaur Saini and Bansal 2019, Ahmad et al. 2019, Ferrara et al. 2016), location of vulnerable areas (Ding et al. 2017, Zhang et al. 2018), profile construction of potential terrorist and terrorist behaviour recognition (Zheng et al. 2017), bank transfers analysis and financing terrorism detection (Domashova and Mikhailina 2021, Garcia-bedoya, Granados, and Cardozo Burgos 2020, Rocha-Salazar, Segovia-Vargas, Camacho-Minano 2021), biometrics data analysis (Chamieh et al. 2018), and predictive policing programs that analyse huge amounts of data to predict terrorist attacks and to generate alerts (Verma, Malhotra and Verma 2018, Kalaiarasi et al. 2019, Gohar, Butt and Qamar 2014). This literature often attempts to prove the potential usefulness of the AI /ML applications. Likewise, there is a significant amount of literature with a more critical view of the efficiency of advanced technologies and their use by law enforcement agencies, pointing to their moral and ethical implications (Ganor 2019, McKendrick 2019, Browning and Arrigo 2020). The examination of predictive policing usually deals with the description of the benefits and risks associated with these techniques, partly from an ethical point of view and from the point of view of the technical operation of AI/ML. However, the literature is less concerned with the use of machine learning in systems designed to predict and prevent terrorism and its possible limits. Although trying to anticipate crime and taking steps to prevent it from occurring, and the attempts to thwart theorist attacks and to stop the spread of terrorism differ in nature, the use of predictive means can lead to similar consequences in numerous cases, as it requires the use of the same technologies for a similar purpose. Therefore, this thesis aims to assess the functioning of ML predictive policing programs, their benefits and limits, and to point out the problems that might be encountered by intelligence services when attempting to predict terrorist attacks and to detect (potential) terrorists.

2. Predicting and preventing crime with machine learning

Since the future is by its nature uncertain and “imagined fears are unbounded,” the uncertainty has become the cornerstone of the need for risk prevention (McCulloch and Wilson 2016, p. 41-71, Zedner 2007, p. 275). One of the most significant moments that led the authorities and society to lean towards the prevention and pre-emption of harmful events was the terrorist attack on the World Trade Centre (McCulloch and Wilson 2016, p. 1-2, Lennon 2015, p. 45). The former US secretary of defence Donald Rumsfeld compared the uncertainty that leads the leaders to employ countermeasures to the “things we don’t know we don’t know” as “unknown unknowns” – the opposite of “known unknowns” – “things we know we don’t know." The attack of 9/11 was in his words "the most horrific single unknown unknown America has experienced" (Rumsfeld 2010). Security in this sense ensures the conditions under which the crime or terror are unthinkable. However, the feeling of insecurity is very often driven by anxieties that are not necessarily connected to criminal activity (Zedner 2007, p. 262-264).

In reaction to the same incident, George W. Bush called for a precautionary approach in the fight against terrorism and similar attacks: "if we wait for threats to fully materialize, we will have waited too long […] we must take the battle to the enemy, disrupt his plans, and confront the worst threats before they emerge." (The White House, 2002) As the organisers of this attack were not unknown to intelligence services, the incapability to anticipate the event and apprehend the terrorists led to the belief that any measure must be taken in order to prevent such events from happening again (Brayne 2017, p. 978). The attack on the World Trade Centre and similar terrorist attacks that claimed human lives encouraged national governments to employ pre-emptive countermeasures and underlined the importance of intelligence gathering and the necessity to increase surveillance (Zedner 2007, p. 264-265).

Risks associated with predicting and preventing harmful events

According to McCulloch and Wilson, the attempts to predict and prevent uncertain future events might result in false prediction, its misuse for political goals, arrestation and punishment of potential criminals, and most importantly may lead to "self-fulfilling prophecy," as we could see in the case of *Minority Report*. By taking steps to ensure that the event does not happen, we are essentially limiting the number of choices one possesses. Moreover, the crime that was averted is often seen and approached by institutions as if it had already happened. The precautionary principle has been applied to many spheres, including the environmental law and the finance sector. Contrary to crime and counterterrorism, however, the incentives to forestall harmful incidents employed in other spheres do not lead to the prosecution or punishment of the suspected culpable (McCulloch and Wilson 2016, p. 1-10, 41-42).

It is also important to note that with regard to terrorism, there is not any concrete and widely accepted definition and therefore many types of unwanted behaviour have been determined as being terrorist while providing the authorities with justification to apply exceptional measures. As anything can be labelled terrorist and anyone can be considered terrorist, anyone can be pursued under the pretext of preventing human casualties. Since counterterrorism is one of the priorities of national governments, the penalties for suspects are much stricter than for criminals that have already committed the crime (McCulloch and Wilson 2016, p. 72).

Role of technologies

Constantly evolving technologies represent hope for catastrophic scenarios to be predicted more quickly and accurately, which is doubly true for statistical and computational methods that provide a veil of science, precision, and objectivity. Hence, these methods are often considered to represent a solution to complex and once unresolvable problems. One should not forget that technological explanations and principally computational methods might provide mathematically accurate results, however, they do not take into account social, historical and cultural contexts, and answer (in the case of crime and counterterrorism) questions of when, where and by whom the dangerous activity is likely to be committed rather than why such actions occur. Accordingly, the solutions generated by statistical methods and technological tools provide evidence for detention and punishment instead of possibilities to remove the causes of these activities (McCulloch and Wilson 2016, p. 76-78, McKendirck 2019, p. 30).

To summarize this part, as the future is always uncertain, there will always be fears of possible yet ambiguous harm, requiring appropriate counteraction in advance. The technology, giving the impression of science, accuracy, and neutrality, has become a hope for improving the well-being of society in almost every way, including preventing crime and facing terrorism. The following section will briefly describe the operation of AI and ML in predictive policing programs and their potential risks, which prove that these technologies are very often overestimated and not as flawless as expected. They are man-made, work with data that reflect human society and are accompanied by contexts that these machines do not necessarily consider while producing outcomes.

2.1. Artificial Intelligence and Machine Learning

The aim of the following part is to explain key terms discussed in the thesis. Due to the extensive scope of the AI and ML topics, this section will briefly introduce the functioning of the technologies to the extent that enables us to understand the predictive policing methods and the risks involved in using AI and ML for crime and terrorism prediction and prevention. First, the operation of artificial intelligence and machine learning will be introduced. The second part deals in more detail with predictive policing and issues that law enforcement agencies may encounter.

The widely agreed approach to AI is absent among researchers but might also differ across disciplines. Researchers concerned with AI examine its technical operations and potential use, whereas policymakers usually perceive AI as being able of reasoning and conduct operations comparable to human capabilities and behaviour. Although both approaches have offered significant findings, the lack of a common definition raises further concerns and prevents the regulation of these systems due to the impossibility to determine which systems should be classified as AI and thus to which systems the regulations of these technologies are applicable (Krafft et al. 2020, p. 1-2). To illustrate, the OECD understands AI as a "machine-based system that is capable of influencing the environment by producing an output (predictions, recommendations or decisions) for a given set of objectives." Accordingly, OECD puts emphasis on AI systems' ability to work with a certain degree of autonomy (OECD 2021).

Machine learning then represents a subfield of AI. ML systems can learn and improve their accuracy in a given situation or task based on their experiences and the increasing amount of inserted data. They are enabled to do so by using computational algorithms whose repetition and modification make the system learn to predict output to a given input. The outputs are then compared to known outcomes and judged and adjusted to improve the ability of further prediction (Goodfellow, McDaniel and Papernot 2018, p. 56-57). McCulloch and Willson stress that technology has given the approaches to predict and prevent crime an impression of scientism, impartiality, and precision. This statement is valid for the ML methods as well, if we consider that they are based on mathematical formulas and computational methods. Another particularity of ML lies in the ability to adapt to situations based on its past experiences and learned knowledge. To put it in the context of terrorism, the CCTV camera can develop its definition of "normal behaviour" based on its acquired knowledge about usual and unusual behaviour, since the terrorist behaviour is deviating from the one it considers normal. Therefore, the camera is supposed to recognize a terrorist according to his unusual behaviour, although the person has not committed any terrorist offence yet (McCulloch and Wilson 2016, p. 82).

2.2. Predictive policing

The term predictive policing refers to the use of data analysis to predict criminal activity and to precisely deploy the resources to prevent such actions from occurring. The literature on predictive policing often emphasizes the use of many different types of data and its analysis through quantitative tools and advanced technology (Meijer and Wessels 2019, Perry et al. 2013, McCulloch and Wilson 2016). The authors often draw attention to its connection to the pre-emptive approach, implying that the counteraction to avert it is taken before the crime is committed (McCulloch and Wilson 2016, Meijer and Wessels 2019, Egbert and Leese 2021, Bennett Moses and Chan 2016). Finally, as the definitions indicate, the resulting predictions are based on data about past crimes. Perry et al. delineate predictive policing as the use of analytic methods, especially quantitative tools, to determine likely targets of crime in order to undertake appropriate measures to prevent the action from occurring, thereby to "solve past crimes by making statistical predictions." (Perry et al. 2013, p. 1-2) Meijer and Wessels define predictive policing as the collection of data on past crimes followed by the use of statistical methods to identify places where criminal activity is likely to occur or to determine the persons that are likely to commit such actions. Predictive policing is expected to facilitate the work of law enforcement agencies, and to invent strategies and concrete counter-interventions (Meijer and Wessels 2019, p. 1033).

Accordingly, the rapid analysis of large amounts of data that would otherwise require a longer time if they were analysed by humans, and its potential to increase the effectiveness of law enforcement agencies, especially with regard to the deployment of resources at concrete times and places, are believed to be the most significant advantages of predictive policing (Egbert and Leese 2021, p. 8, Meijer and Wessels 2019, p. 1033).

The definitions of predictive policing often operate with the assumption that predictive policing aims at reducing crime rates. However, the authors also point out the difficulty if not an impossibility to prove the efficiency and success of such measures (Egbert and Leese 2021, p. 164-185). Meijer and Wessels highlight that there is a lack of empirical evidence to affirm both the efficiency and inefficiency of predictive programs (Meijer and Wessels 2019, p. 1034). Although predictive policing methods have the potential to ameliorate security in particular regions (Mohler et al. 2015, Pearsall 2010), they cannot reduce all forms of crime as they are not even meant to mitigate other forms of crimes – for instance, the domestic violence (McCulloch and Wilson 2016, p. 11).

According to Amoore, predictive technology is based on "anticipatory logic," which does not necessarily exclude the possibility of the occurrence of a certain event, and is rather intended to "incorporate the very unknowability and profound uncertainty of the future into imminent decision." (Amoore 2013, p. 9). As the purpose of predictive policing is to intervene in the present to prevent crime from occurring in the future, the use of predictive policing programs is in a large number of cases pre-emptive. This offence is often approached and intervened in the present "as if it will have happened." According to Sheehey, predictive policing "uses and reinforces a racialized past to generate its preemptive power." (Sheehey 2018, p. 54-55).

After introducing the predictive policing approach, it is evident that predictive techniques usually need a large amount of data and quantitative methods to process them. The definitions of predictive policing also stress the employment of advanced technology to predict and thus reduce crime by acting pre-emptively. The advantage of using predictive programs is that the analysis is conducted faster and with perhaps better precision than if done by a human, and this benefit is enhanced with the incorporation of machine learning systems. However, some authors challenge predictive policing and methods that are supposed to calculate the conditions under which these events are likely to take place. The next part of the thesis focuses on the risks arising from using machine learning predictive programs, although in some cases, these are common to big data analysis.

3. The risks associated with the use of AI/ML in predictive policing

3.1. Algorithm bias

The algorithms are a set of formulas determining the way data are analysed and the correlations in data are discovered. Predictive policing algorithms are supposed to identify correlational patterns in a huge amount of data and accordingly predict for example where the police sources should be deployed to prevent likely crimes (O'Donnel 2019, p. 544-549).

Although we cannot recognize the exact ML model and algorithms the predictive policing work with as this is proprietary knowledge, O'Donnel suggests two types of models that are very likely to be used. The first one is KNearest Neighbors which is expected to detect new variables that resemble the most (the "nearest neighbours") to the ones the algorithm was programmed for, and to include them into the code to use them in the following decision-making processes. A well-known application of this algorithm is a consumer behaviour prediction – people whose behaviour was the same in the past are likely to behave similarly in the future. However, the human programmer is not always able to determine which variables exactly the algorithm incorporated. Programmers can only measure whether the predictions based on new algorithms are valid (O'Donnel 2019, p. 549-550).

The other method the predictive algorithm could rely on is neural networks. The artificial neural network receives training data that are subsequently proceeded by algorithms, that find correlational patterns and product outcomes, which are then evaluated. According to the evaluation of what is a "wrong" answer, the algorithm learns how it should not proceed, and on the other hand, what is "right" and how it should analyse the data in future tasks. According to acquired knowledge, the algorithm's code can be modified to analyse data in the subsequent operations. To change the model after its deployment, the human operator can initiate and operate re-training or fine-tunning. The algorithm is applied to the examined data set, in which it employs what it has learnt, and progressively learns to recognize further patterns from new data and again encode the new knowledge and apply it (O'Donnel 2019, p. 549-550, Guo et al. 2019, p. 1-2).

These algorithms have been long believed to represent an impartial method of data processing, as they are not influenced by the conditions the humans are, as the socio-economic situation, emotions, convictions and prejudices. Yet this assumption has been proven false. Since the algorithms were developed by man and the data comprise information about society, these conditions are built into the system and might produce results that are not always impartial (Green and Hu 2018).

To illustrate, the use of predictive policing models for profiling is assumed to result in further stigmatization of individuals and groups, as it reflects a continuation of the practices of previous policymakers or prejudices that are entrenched in society. When they are translated into data, they are incorporated into the algorithm's thinking. Perhaps the most often cited issue relates to racial discrimination and criminality when positive feedback loops teach the algorithm that it is right to associate race with criminality (Meijer and Wessels 2019, p. 1036, O'Donnel 2019, p. 544). It has been also proved that police databases do not contain data about all crimes, as a high number of crimes stays unreported, the information the report contains might be incomplete or influenced by the bias on the side of the law enforcement agency's officers. If it is the case, biased data are incorporated into the algorithm that subsequently produces biased outcomes, without being necessarily programmed to explicitly use race as a variable (Green and Hu 2018, O'Donnel 2019 p. 560). Green and Hu argue that such ML techniques are not able to detect social changes as well as do not possess common sense and assume that the correlations found in historical data apply to the new ones. This could lead to a distorted view of reality and society as well as to discrimination reinforcement (Green and Hu 2018, Meijer and Wessel 2019, p. 1036).

One of the principal triggers of the call for AI fairness has become Propublica's research from 2016 examining the COMPASS prediction program by Northpointe (today Equivant). The researchers argued that the system was racially biased against black defendants, who were assumed to be prone to recidivism. Although the scoring system did not explicitly include race as a criterion to rate whether the person will re-offend, the categories used were indirectly connected to race. The authors demonstrated that black defendants were considered 77% more likely to commit a violent crime in the future and 45% more likely to commit any type of crime. Overall, the system proved itself to not be as accurate as expected, as only 20% of people designed to commit the crime again in the future did so. The founder of Northpointe Tim Brennan contended that some elements that were associated with race – such as the jobless and social marginalization – could not be removed from the scoring system, otherwise, the algorithm's accuracy would be significantly reduced (Angwin et al. 2016). The different opinions on the objectivity and accuracy of COMPASS proponents and critics of the system also stemmed from a distinct understanding of AI and ML bias and fairness, as there is no common approach toward this issue (Lepri et al. 2018, p. 618).

Similarly, the biased algorithms can be found in other domains than security, as in Amazon's software that was used to rank job applications. The system was discovered to prefer men over women, as the model was taught to identify patterns in resumes received in the last past 10 years, where the majority of applications were submitted by men (Dastin, 2018).

3.2. Lack of transparency and inexplicability

The concept of transparency of AI and ML might refer to the transparency of the entire model as such, including all the model's components, but also to the transparency at the level of each component or the transparency of a concrete algorithm. It is usually the model of low complexity that can be fully explained. In some cases, the researchers refer to transparency as to the feature of the model's parameters or of the algorithm used in the training phase, however, without understanding entirely the whole model as such (Lepri et al. 2018, p. 619). As mentioned in the previous paragraphs, even the programmers might encounter difficulties when explaining why and how certain patterns were encoded into algorithms. In some cases, revealing how the system examined the data and what elements made the ML algorithm produce such results is almost impossible, making the ML systems being often compared to black boxes. Apart from the explanation of technical aspects of predictive programs, and therefore how such result was made by ML systems, the transparency of the prediction process implies also the ability to provide information about how the data are recorded, collected and stored, by which entities and for what uses, including the aim and the question to which the results are expected to answer. The transparency also suggests that the concerned individuals are aware that these techniques are used, as well as that they are informed about their potential risks. And finally, the transparency of the process could include also the transparency of the legal framework, which allows the use of predictive systems to analyse and prevent risks (McKendrick 2019, p. 28).

Taking into account the possible bias in data and the outcomes of ML analysis, it should be emphasized that predictive policing programs "hint at the possibilities, rather than providing any credible proof." (McKendrick 2019, p. 11) When the results of predictive systems are used as evidence, the notion of evidence takes on a new form, losing the reliability and authenticity that are commonly considered to be the essential features of adequate evidence. McKendrick talks about the "erosion of standards of proof," which, especially with regard to counterterrorism, has already fallen below the threshold of definite proof. The growing use of predictive ML systems shifts the boundary of criteria defining the evidence all the more, enabling the law enforcement agencies to act solely on the basis of the ML system's results, without asking how and why such result was produced (McKendirck 2019, p. 30).

Accordingly, the impossibility to fully explain the process and the results has the potential for human rights violations – in the case of security agencies and the prediction of criminal or terrorist activity, especially the right to due process, as it might not be always possible to provide reasons why a particular person was identified as the offender. In the future, therefore, the use of evidence generated by ML systems poses a challenge to law enforcement agencies and courts, necessitating them to adapt to these new technologies in a way they do not compromise human rights. The opacity and impossibility to define who should be accountable for the actions based on the ML system's results also suppress the right for the remedy for those who were unfairly treated (Meijer and Wessels 2016, p. 1036, McKendrick 2019, p. 30, UNICRI and UNOCT 2021b, p. 44).

Il follows that the lack of transparency of ML programs could undermine the public trust in technologies and lead to their refusal (Ben-Israel, Matania and Friedman 2020a, p. 131). The procedures to prevent these actions are a subject of research of explainable AI that focuses on the methods and tools whose purpose is to clarify the results of ML algorithms, concretely to fully explain the model, its potential consequences, including biases as well as to define the concepts of accuracy and fairness so they can be used for decision-making processes (IBM Corporation 2021). Although there are efforts by researchers to introduce these techniques into predictive policing and counterterrorism applications, they are not always employed in practice and if they are, they rather mitigate the issue than resolve it completely, and as the AI technologies become more advanced, their application becomes more complex.

Besides the general assumption that explainable AI seeks to introduce more transparent and interpretable systems enabling to explain and understand the results, Miller emphasizes that the researchers concerned with explainable AI usually do not work with the concept of explanation from social sciences, and when deciding about whether the explanation is useful, they are rather driven by their intuition. Therefore, it is essential to understand the way individuals themselves understand the concept of explanation and how do they create and evaluate it. In other words, this issue should not be examined separately only by computational science alone as it cannot moderate the distrust in AI systems. For a specific case of the technology, there are different questions we need to ask to understand, and the questions will be differently relevant or useful than others for understanding the particular model. Hence, according to Miller, a complex approach comprising "the intersection of artificial intelligence, social science, and human-computer interaction" needs to be employed and the literature on explainable AI should also clearly state what questions it is expected to explain (Miller 2018, p. 4-22).

3.3. Algorithms' accuracy and false positives/negatives

The accuracy of the algorithms and thus the exactness of the predictions are affected by the features of the data sets and the quantity of data used for training. Some phenomena such as the curse of dimensionality, class imbalance and spurious correlations that are connected to the data's quantity and quality represent another potential limit linked to the use of ML predictive systems, as they could influence algorithms' accuracy.

We might assume that increasing the number of dimensions in data sets will ensure more precise results. However, this is not always true, and contrary, it could slow down the learning process and decrease the performance quality. Indeed, the introduction of a higher number of features would in some cases require extending the data sets, which is not possible when we are looking for "needles in a haystack," about whom we usually do not have enough data for training. This trade-off between the complexity of dimensions and the algorithm's accuracy is referred to as the curse of dimensionality (Verhelst, Stannat and Mecacci 2020, p. 2978). The more dimensions are used to make predictions, the higher number of combinations we obtain, making it impossible to extract concrete and regular patterns from these combinations. In the case of crime prediction, the data from law enforcement agencies contain information from various datasets that might include more complex approaches to crime. As a result, we obtain an infinite number of possible combinations that are too complex to find sought correlations (Delgado et al. 2011, p. 5).

Second, the accuracy could be lowered when the data sets supposed to train the algorithm for prediction do not comprise a balanced number of positive and negative cases. This situation represents a very real-world scenario, as the minority group is often the target of the research. Not only with regard to the detection of non-terrorists and terrorists but also in health diseases detection or financial fraud detection (Johnson and Khoshgoftaar 2019). Although it is possible to change the threshold to control the false positive (situation when a result was incorrectly classified as positive) or false negative errors (situation when a result was classified incorrectly as negative) to optimize the results, the error will always occur to some extent, making the researchers obliged to opt to reduce either false negatives or false positives. To balance the data during the training process, a kind of "uniformized subsets" can be employed; the dimensions and features in training data are reduced or increased. If it is not possible, only a few subdomains can be made, which might result in overfitting. It means that although the machine learning algorithms produce perfect results in terms of accuracy during the training, the algorithm might not fit so well in the real-world scenario, in which it is used to find correlations on whole populations. For instance, when the algorithm with an accuracy of 99,9% is applied to a population of 1 million people, it will suggest 10 000 inaccurate results (Verhelst, Stannat and Mecacci 2020, p. 2972).

Spurious correlations could represent another type of potential limit. When the number of features in data is increased, some of the discovered correlations can be irrelevant for the research, as they are not causally dependent. These correlations might be encountered regardless of the balance of data and are therefore difficult to be avoided, decreasing the accuracy of the algorithm. When speaking of the effort to track down criminals, we might for instance find a correlation between the criminal and the size of the shoes, which in our case has no benefit for the prediction of criminal activity and the identification of future offenders (Verhelst, Stannat and Mecacci 2020, p. 2977).

Another assumption is that collected data are a reflection of reality. However, regarding data about crime, each police officer might have a different view on what constitutes a crime, so the classifiers used by one actor might not necessarily correspond to what other law enforcement officers see under those categories. It is also necessary to be aware that a large number of crimes is not reported. It may happen that the given type of crime will be neglected in the predictions. Therefore, the quality of the data used to train the models has a great influence on their accuracy (Strikwerda 2021, p. 432, Bennett Moses and Chan 2016, p. 4-5).

The changing conditions and situations will also be reflected in the prediction's accuracy. As the results of predictive programs are based on past data, the law enforcement agencies presume that patterns that have been identified in past will be equally relevant to future crimes. Consequently, the predicted future is rather a reflection of the expectation of continuity of current police practice (Strikwerda 2021, p. 431-432). The resulting police activity based on predictions generated by predictive programs may also lead to positive feedback loops, as the results of following policing will confirm the accuracy of the prediction. When the police force increases its activities in the area the predictive program identified as highly vulnerable to criminal activities, the police will accordingly obtain data about this particular spot and the predictions will become "self-affirming" and "self-perpetuating." (Bennett Moses and Chan 2018, p. 810)

Moreover, the assumption of continuity and reinforcement of the current practice in anticipating crimes may then incite potential criminals to invent a different strategy to escape the detection procedures. If the predictive system finds that there is a correlation between crime and individuals with a particular feature, the police might assume that people with these features are more likely to commit a crime and focus police on them. Nevertheless, the people who really intend to offend might bypass the system by avoiding having such a feature. To give a very simplified example, the predictive policing systems may find that there is a correlation between crime and individuals with a tattoo, making the police focus on people with tattoos. However, criminals that become aware of such assumptions might decide not to get a tattoo to not get revealed. It is, therefore, necessary to realize that crime, as well as terrorism, is a dynamic phenomenon that is constantly changing and adapting to current trends and technological progress (Bennett Moses and Chan 2016, p. 31).

3.4. Adversarial ML

Once the evidence exists in the digital form, it is always at the risk of being attacked or modified by actors with malicious intentions. This concerns also predictive programs – the adversary who accesses information about the functioning and characteristics of the system can manipulate its operation to make it produce incorrect results. The adversary might be able to access the system's structure, parameter values, or the data used to train the model and thus perturb the data directly in the model or even before they are input (UNICRI and UNOCT 2021b, p. 39, Goodfellow, McDaniel and Papernot 2018, p. 56-59).

To illustrate, concerning Neural Networks, the most powerful adversary could access information about the number of layers and their characteristics, about neurons' activation functions or information about weights and biases, as well as the training algorithm. However, the adversary could simulate the functioning of the network and cause significant damage also even if he knows only the network architecture, including the information about layers, activation functions and the weights and biases obtained during the training phase. Likewise, the malicious actor could be capable of revealing and attributing the outputs to given inputs to discover the functioning of the algorithms (Papernot et al. 2015, p. 4). According to the amount of knowledge about the system and thus the ability to simulate the system on one's own machine, the employment of adversarial machine learning could be divided into the white box and black box scenarios (Goodfellow, McDaniel and Papernot 2018, p. 59).

Goodfellow, McDaniel and Papernot distinguish the adversarial examples according to the adversary's goals. "Untargeted misclassification attack's" purpose is to modify the input data in any incorrect way, whereas "targeted attack" aims to produce a particular misclassification outcome. It is also possible to invade the training process itself – the adversary could be capable to input his own data from which the system will learn, meaning that it will produce distorted and modified results (Goodfellow, McDaniel and Papernot 2018, p. 58-60).

The reason why these attacks have become alarming is that, unlike a machine, human beings are not capable of noticing these changes. Probably the most discussed application of adversarial ML is image recognition. An adversary can produce even a slight perturbation of the image, which is perceptible only by a machine, but unrecognizable to a human eye that continues to see still the same image (McKendrick 2019, p. 22-23). The literature concerned with adversarial ML often refers to autonomous vehicles and the recognition of traffic signs. The perturbation of the sign STOP might confuse the recognition system and make it deduce that the sign represents the symbol allowing driving, which increases the risk of an accident. However, research on adversarial ML has been also conducted in the domains of audio recognition, summarization of texts or fraud detection (Cartella et al. 2021).

Therefore, the potential benefits that adversarial ML could bring to criminals are enticing, as the ability to deceive predictive systems might carry the opportunities in the image recognition of criminals, fraud detection and corruption or in disrupting the systems aimed at finding the correlations between suspected persons. For instance, the fuzzing technology could provide the adversary with the discovery of the system's vulnerabilities. The adversary can also operate in a black box scenario and cause the machine to produce wrong results even without knowing well the inner functioning of the systems (Solomon 2020b).

It should be noted that the adversarial environment is a subject of further research that has already yielded results regarding the vulnerability of models and how to make ML systems more resilient or how to attenuate the effects of adversarial attacks. However, these practices are not perfect and cannot be applied generally to all models. There have also been attempts to create defence systems against adversarial examples. Reddy Mekala, Yerramreddy and Porter divide them into two groups. Adversarial rectification enables the adversarial input to be corrected in such a way that it can generate correct results, and the adversarial detection capability can detect adversarial inputs and not let them enter the system. The adversarial input is recognized either based on the observation of the system's behaviour or by using the algorithm that was trained on real and adversarial data so it is able to distinguish between them. However, this technique works rather with small datasets and cannot be accurately applied to a big dataset containing cases from a real-world scenario (Reddy Mekala, Yerramreddy and Porter 2021, p. 1-8).

3.5. Increased surveillance and the role of private and state sector

The surveillance practices that are supposed to detect vulnerable areas where the crime is likely to occur or that are able to observe a typical characteristic of a criminal are not new, and as technological tools are being developed, these techniques are getting enhanced. What we are witnessing in AI and ML is a move from modern surveillance tools and big data analysis to a more advanced level, and that is predicting the future and scoring individuals according to the risk they pose to society's security, based on the data we gained about them through surveillance. These advanced surveillance technologies might reinforce the risks associated with the attempts to foresee threats and with the belief that predictive technology is the solution to these unpredictable risks (Miller 2014, p. 106).

What has changed with the advent of big data, automated analysis and machine learning capabilities, is first that a huge number of people can be monitored. This allows law enforcement agencies to keep under surveillance not only individuals who already have a criminal record, and are thus considered as likely to re-offend but every single citizen, even those who have not committed any action yet. They are monitored to make sure they are caught in the case they decide to offend, as the future is uncertain and always contains the possibility that they commit the crime (Brayne 2017, p. 977). Therefore, data are collected and stored "not for a specific criminal investigation but rather for an undetermined purpose, serving a mentality of 'nice-to-have' rather than 'must-have' intelligence." (Vogiatzoglou 2019, p. 4)

These days, an individual leaves his traces everywhere. When sending an e-mail, searching the Internet, making a phone call, purchasing goods either online or in a store using his payment card. All of these can be used to track criminals and predict their future activity. Regarding crime detection, machine learning has been also employed in surveillance cameras in combination with a database of images or other biometric data for the identification of citizens. In this way, it is possible to detect who is doing what in a given moment (Miller 2014, p. 111). Machine learning text analysis, on the other hand, helps understand private messages, without the need of a human reader, and is thus able to detect suspicious behaviour through messages, just as CCTV cameras can recognize unusual behaviour through video records (McKendrick 2019, p. 2-6).

According to the opponents of AI/ML technology for surveillance and accordingly for crime predictions, its use has two key effects. First, it violates individual privacy rights and the right to due process, transforming the democratic state into a "digital authoritarian state." (Ganor 2019, p. 5). Second, it may reinforce the inequalities embedded in society (Miller 2014, p. 106). The control over citizens has always been disproportionally focused on particular groups, institutions or places that are expected to represent a risk to society (Lyon 2007). According to Brayne, it is essential to "understand who is surveilled by whom, in what way, and for what purpose." Once we become conscious of who is surveilled, by which means and what is the aim of this control, we will be capable of understanding how differently and unequally the opportunities in society are distributed (Brayne 2017, p. 1003). Lyon offers an example of biometric ID card systems that match screened individuals with the identity in national databases and allow to treat members of a particular group differently (Lyon 2007, p. 112). At the same time, however, the introduction of big data and analysis tools allowed gathering and storing the data about every single citizen. The 9/11 attacks were considered to be caused by the failure of intelligence and information exchange, which made the authorities believe that it is necessary to gather data about everyone, including those, who do not have any significant links that would indicate they are prone to criminal activity. Surveillance, therefore, becomes intensified, automated and targeted to everyone, but on the other hand, also more invisible and imperceptible by the surveilled ones. This control aims to prevent unwanted events from occurring (Brayne 2017, p. 978-979).

McKendrick points to the tendency to abuse acquired data. As we discussed in the previous part, the ML systems can be misused by offenders, but also by non-democratic states to repress the citizens and impose total control over them. As the governments have a different understanding of what constitutes crime and terrorism, the data collected and analysed by ML surveillance tools can be used as proof that a citizen is involved in criminal activity, that is, however, beyond the definition of a crime in a democratic state. Such behaviour could be the opposition to the government and manifestations. Similarly, sensitive data can be misused to threaten and blackmail (McKendrick 2019, p. 31, G. Waters 2019, p. 575). This issue is also one of the main arguments of the opponents of the use of AI/ML surveillance – that the technology compromises the right to privacy, freedom of speech and freedom of expression. The behaviour of citizens is easily controlled through modern technologies, and any person condemned by the state authorities as uncomfortable and dangerous can be pursued if not physically harmed thanks to the information generated by ML surveillance systems (Ganor 2019, p. 4-5).

Each trace produced by individuals can be owned and stored by different actors and might not make much sense when considered separately. When the data are put together to analyse and find correlations, they might reveal a lot of details about one's private life. With regard to surveillance and big data manipulation, we can observe function creep, as the data that are gathered with a completely different intention may be inadvertently and unexpectedly transferred to actors who put the data together and use it for other purposes (McKendrick 2019, p. 13, Brayne 2017, p. 980). It is usually the private sector that collects and stores the data containing sensitive information. The private sector is therefore asked to participate in citizens' surveillance, transferring individuals' data to the state's security sector. Surveillance means are spreading and intersecting in all possible spheres, spilling over from private data holders into state surveillance, which oversees citizens' behaviour (Brayne 2017, p. 977). Although technological companies sometimes confirm that they secure the personal data of their users, it is often only a statement rather than a serious commitment. Similarly, while many companies have denied government agencies access to the data they have obtained, they may still access them through third parties and thus bypass the system. If the government does not have access to certain data by law, they can, in many cases, buy them legally from private companies (McKendrick 2019, p.17).

4. Methods

The thesis aims to answer two research questions: 1) To which extent can we rely on the AI/ML systems used in predicting and preventing terrorism? 2) What are the implications of their use for security in Israel?

The theoretical part was concerned with the conceptualisation of the predictive and preventive approach to crime that is expected to increase the efficiency and accuracy of law enforcement agencies. Yet it is difficult to prove whether they successfully contribute to the decrease in crime occurrence. Moreover, the use of ML predictive systems also poses several security risks. As the inferences are not based on criminal activity that fully reflects the reality, yet on the patterns retrieved from past policing that comprise of what training data we obtained about the incidents, rather the continuity of police practice is expected. Hence the stigmatisations embedded in society are likely to persist, if not to get reinforced when integrated into algorithms. The argument that compared to the analysis conducted by humans, algorithms are characterised by a higher level of neutrality, thus seems to be weakened. The expected continuity of current police practice may also make the criminals adapt to the system and circumvent it.

Therefore, the first part partially answered the question of to what extent we can trust AI and ML methods designed to make predictions, especially about criminal activity. The second part applies these limits to the field of counterterrorism. First, Israel's approach to terrorism and its technological advances will be briefly introduced, outlining why the Israeli case was chosen for the case study.

Risks discussed in the first part will be then examined in the context of the Israeli fight against terrorism. Concretely, the following chapters will examine (1) what are the implications of biased data and biased algorithms when the predictive programs are used to detect potential terrorists, how the algorithms could be biased considering the Israeli context and what are the possible impacts on security of those who are disadvantaged by these algorithms (2) what may be the risks associated with the opacity of results generated by predictive programs, what are the Israeli experiences with unclear evidence in decision making and sentencing persons for the sake of preventing terrorism, (3) how may be reflected the fact that these programs cannot be entirely accurate and that they run the risk of misjudging someone as a terrorist, or, on contrary, to overlook some of them, and what could be the consequences in the case of Israel, given their current practice with terrorism detection, (4) what impact adversarial examples can have on attempts to predict and prevent terrorism and how terrorists can exploit these vulnerabilities, (5) what are the effects of the ML surveillance conducted under the pretext of countering terrorists, to what extent it is practised in Israel, and what are the implications for the security of the population living on the territories controlled by Israeli authority.

To answer the abovementioned questions, the paper draws conclusions from academic literature focused on criteria that were defined in the theoretical part, as well as from the sources dealing with counterterrorism in Israel. The sources available are also considered as limits of this thesis – due to security reasons, the information about concrete systems used in Israel is not made public to such an extent as we could have seen in the case of predictive policing programs, since revealing such information could provide a manual for terrorists on how to breach these measures or how to bypass them, therefore, it is not possible to examine these measures and functioning of the systems the Israeli agencies use on the empirical level in such a depth. The lack of resources on counterterrorism is also the reason why the theoretical part used predictive policing as a proxy to discuss the limits associated with the prediction of crimes. The empirical part then draws mostly from electronic resources such as the official websites of involved institutions, Israeli companies and Israeli journals concerned with the current events in Israel.

After the presentation of the Israeli case, its technological development and fight against terrorism, common and differentiated features of predictive policing and terrorism prediction and prevention will be highlighted, and the significant challenges that may be encountered by the intelligence forces when deploying counterterrorism measures will be emphasized.

The main goal of this paper is to answer whether these ML predictive systems can be considered reliable since they are based on patterns that correspond to the history of terrorist attacks and since some of their results cannot be fully explained. The thesis aims to show that intelligence applications bear the same limits as predictive policing, although in the intelligence community in general, and in the case of Israel in particular, there is a very positive view of these technologies. Even though predictive tools may allow us to find the needles in a haystack, they are the needles we have seen before, while the new ones stay undiscovered. The same problem arises in the case of lone-wolf attacks that the individual commits impulsively without any previous connection to the terrorist network and thus for which there is not enough data to indicate that he decided to commit an attack.

5. Case study: Israel

Introduction to the Israeli case

Israel faces terrorism since the creation of the state, however, does not have any official counterterrorism strategy nor any national security strategy that would be approved by the Israeli parliament. Instead, Israel draws on its acquired experiences in countering terrorism and tries to adapt its strategies to the evolution of threats. The first intifada demonstrated that Israeli Defence Forces did not have an adequate response to insurgencies, and instead of striving to solve the issue by a "nuanced policy-style" strategy, employed a hard military approach. Therefore, Israel has learned from its experiences starting with a popular uprising, suicide bombing and missile operations and ending with knife attacks, car-ramming and lone-wolf offenders. It is through this evolution of threats and political situations that Israel attempts to adapt its capabilities to provide security. Israeli representatives are also aware that the source of terrorism is not located only within Israeli territories but also in foreign countries. The literature usually presents Arab-Israeli and namely Palestinians as a source of threat of terrorism, however, terrorist attacks have been committed also by Israelis, yet on a lower scale. As the following part will demonstrate, counterterrorism measures usually take into consideration national and ethnic identities. It is also noteworthy that Israel stresses preventive and deterrence measures that are supposed to thwart terrorist attacks before they occur. However, Israeli counterterrorism strategies lack the notion of deradicalization. It seems that the Israeli policymakers are convinced that their neighbouring states do not intend to coexist in peace, making the Israelis accept terrorism as a long-time threat that they should learn to live with (Kfir 2019, p. 227-236).

The agencies responsible for providing internal security are the Israeli Security Agency (ISA, also known as Shin Bet or Shabak) and the military organization Israeli Defence Forces (IDF) with the support of the Police. Recently, the ISA has undergone changes to be able to face new challenges. The agency tries to adapt to technological progress and develop its capacity to efficiently process textual, visual, and vocal data and to recognize relevant information that could help identify dangerous individuals. According to the former director of ISA Nadav Argaman, technologies play a supporting role in the agency's operations, as they enable the analysis of a huge amount of data to make the most accurate predictions possible (Argaman 2018, p. 4). In 2015, former chief of the ISA's Information Technology Division Ronen Horowitz stated that Israel had been developing big data analysis tools and machine learning capabilities for fifteen years and emphasized the success of their acquired know-how:

“I am telling you with certainty that quite a few terrorists are looking at us from the sky owing to Big Data capabilities [...] We are looking for a needle in a haystack – very weak signals, when the enemy is highly sophisticated.” (Rapaport 2015)

Around one-quarter of ISA’s employees are experts in technologies. According to ISA, cyber technologies and big data analysis are essential tools in the fight against terrorism and espionage. One of the ISA’s units, The Information Systems Technology division is responsible for the development of infrastructure and the innovations of such systems as speech recognition, computer vision, natural language processing and data mining systems. Reportedly, the division “has developed sophisticated technologies for highlighting information relevant to the prevention and disruption of terrorist activity,” that are at the same time “unique to ISA and considered state-of-the-art in the industry and the intelligence community,” and that represent “crucial weapons in ISA’s efforts to obtain important intelligence in real-time and disrupt terrorist intentions in advance.” (ISA 2020)

In 2018, the division was awarded by prime minister Benjamin Netanyahu for a significant contribution to the state’s security, concretely for thwarting numerous terrorist attacks, and for their achievements regarding the development of machine learning capabilities that also positively contributed to the operations (ISA 2020, Times of Israel 2018). The ISA managed to thwart 500 terrorist attacks in the planning process and 1 000 potential terrorist attacks thanks to social media analysis that identified individuals with typical characteristics of such offenders. ISA has been using social media such as Facebook to detect potential terrorists since 2015 when the attacks committed by lone-wolf attackers became more frequent. By employing algorithmic tools, ISA tracked the activity of suspects, including their posts and connections to other persons. The screening concerned the individuals considered to be to some degree likely to commit a terrorist attack, such as the family members of those who already took part in terrorism or the bereaved family members of those who were killed by IDF. Some of the arrested suspects reportedly affirmed that they planned or considered committing an attack (Jeremy Bob 2020).

The startup nation and security provider

Since the creation of the Israeli state in 1948, Israeli leaders emphasized the importance of science and technological progress that were supposed to go hand in hand with its national security imaginations (Spektor 2021, p. 108). Data processing and intelligence services thus represent an essential part of ensuring and operating security in Israel. The country is also often named the "startup nation," referring to the high number of established companies focusing on innovations, usually in the fields of cybersecurity, defence and AI (Keilaf 2020). Security reinforced by technological progress has therefore become an essential part of Israel's vision of the future, making technological innovations one of Israel's main interests (Spektor 2020, p. 103). Specifically, the Israeli policymakers believe AI to be "critical to the welfare, economy and security of Israel's citizens." Their goal is to "establish a holistic and sustainable secured AI ecosystem, driven by the private sector but in which government, private industry and academia all participate, and which supports the use of AI at all levels." (Ben-Israel, Matania and Friedman 2020a, p. 120) Israel is known for its extensive investments in technology and innovations related to security, but except for contributing to national security, these technologies are exported beyond the frontiers of the state. Most of these are surveillance technologies that are in demand in many countries, although they are often questioned because of the concerns over human rights violations (Spektor 2020, p. 103-110).

Antebi points to the "dual feeding" processes of the Israeli military and intelligence units. Such processes are characterized by the transfer of human resources between national security agencies and private companies concerned with technological innovations. One of these processes is the training of IDF's Unit 8200, which hires talented university graduates into military service. After they finish their military training, they are employed by various start-ups or establish their own companies. Thanks to their practical knowledge from military service, they are believed to work more effectively and resolve issues faster than recent graduates, who usually lack such practical experience. Thus, the research experiences of security forces play an important role in the Israeli security industry. Also, the communication and cooperation with academia, private sector and state security agencies are occurring at all levels. Accordingly, these three levels are interconnected and support each other in innovation and development. This transfer of know-how enables effective cooperation and maintenance of national security (Antebi 2021, p. 84-88).

Israeli security companies are often considered to represent a prime example of successful security providers not only in Israel but also abroad. Israeli security services have become popular both in the US and in Europe owing to the spread of fears of potential terrorist attacks as well as thanks to the popularity of Israeli security forces worldwide. This popularity stems from the assumption that Israel is a qualified and experienced country due to its history full of armed conflicts and efforts to preserve the integrity of the state. Israeli security providers benefit from Israel's experience in counterterrorism and are hired by clients from all over the world (Grassiani, 2017). Grassiani presents examples from the US, where the clients explicitly ask for Israeli experts to provide security services. IDF former members are also highly demanded in private security companies worldwide. The author uses the notion of "ISE brand" to illustrate how the Israeli knowledge and experience in security providing become commodities sold on the market. Israeli experts thus apply their knowledge and experience in providing security and counterterrorism to different contexts, as to other criminal activities. One of the characteristic elements of this ISE brand is also the ability to fight against terror and provide security seminars in this field. However, according to the author, these workshops are often focused on concrete identities and the knowledge provided is "very one-sided, if not to say racist and generalizing" (Grassiani, 2017). That is also why Israel is believed to represent an appropriate case study for this paper. Not only that Israel has a long history of countering terrorism and is one of the most advanced countries regarding AI technologies, but also because Israeli companies export their know-how abroad.

Recently, Israel drew attention to its fight against covid-19, enhanced by advanced capabilities in data analysis and intelligence. The government applied its intelligence techniques based on ML originally used by intelligence services to detect terrorists and thwart terrorist attacks, to slow down the spread of the virus. As in other countries, where surveillance technologies have been used to detect the virus by sensitive data about citizens, concerns about the violation of privacy rights have been raised. The former Prime Minister Netanyahu admitted that sacrificing a small degree of citizen's privacy was worthwhile to save lives, explaining that Israel was "one of the few countries with this capability," therefore, they took advantage of it. According to analysts, using these methods has been proved very effective in fighting the virus, as detecting terrorists and patients are very similar in principle – the goal is to identify a person's contacts based on where the person was, when and with whom (Hendrix, 2020).

5.1 Algorithm bias

Algorithm bias in predictive programs could lead to the incorrect determination of some individuals as terrorists based on the data used to develop and train the algorithms. Similar issues might arise when the algorithm is used in different contexts than in the one it was programmed for. As mentioned in the first section, predictive programs raised concerns regarding their use by law enforcement agencies, as some of them were proven to be biased against minorities. Similarly, such a problem may be encountered when trying to identify terrorists.

Although biased algorithms may manifest themselves in any sphere where the discrimination takes place, when the model is trained on the population with an ethnic majority, bias determined by cultural factors is intrinsic to these data (McKendrick 2019, p. 28). Therefore, regarding Israeli counterterrorism, this paper will focus on ethnonational identity as a possible source of bias and will pay attention to different approaches of institutions and society to minorities, namely Palestinians.

Ethnic profiling

The use of profiling method, such as ethnic profiling, raise concerns among human rights defenders worldwide. Except for doubts regarding its efficacy, the attention to a particular characteristic as nationality or religion carries a risk of discrimination based on the affiliations to a concrete group. These groups are usually perceived as a potential source of threat (Moeckli and Thurman 2011, p. 34).

The characteristic features and behavioural patterns of terrorists might change across the terrorist organisations, regions and religions, as well as they change with time and the evolution of accessible technologies. In a multicultural environment, a behaviour or feature considered unusual in one culture could be considered normal behaviour in another one. Indeed, the difficulties may arise when the system is trained on a different culturally diverse sample of data than the one to which it is later applied or when the systems learned the bias from the user involved in training. To illustrate, advanced video analysis relying on ML techniques is expected to recognize unusual behaviour and to develop the profiles of suspicious behaviour. When the system learns through receiving feedback from the user, the machine is taught to produce the results that answer the user's questions and aims. When the user providing the feedback is influenced by his own biases against particular groups, their incorporation into the system during training is inevitable (Moeckli and Thurman 2011, p. 34-37).

Ethnic identity has always been a constitutive element and one of the arguments to defend the foundation of the Israeli state. It also played an important role in the securitization of Israeli identity in the Israeli parliament after the first intifada. The proposals discussed between the political leaders were linked to the ethnic identity of Israelis and the Palestinian minority, and although almost none of these proposals was approved, this ethnonational element cannot be overlooked when examining the Israeli conception of security and counterterrorism. Understanding ethnic identity as a part of the threat in times of uncertainty and heightened fear contributes to the securitization process and the exacerbation of these fears (Oleskser 2014).

There is neither any direct evidence confirming Israel's employment of measures based on ethnic, national or religious affiliation, nor any official document that would explicitly confirm that in order to prevent terrorism, ethnic profiling is employed. Hasisi, Margalioth and Orgad (2012) and Hasisi et al. (2019) demonstrate that the security service at the Ben-Gurion Airport used ethnic profiling to detect terrorists carrying explosives in their luggage before the automated "Hold Baggage Screening" (HBS) system was introduced. Non-Israelis and Israeli-Arabs were obligated to undertake additional control of their luggage in more than 40% of cases compared to 9,8% of Israeli Jews (Hasisi, Margalioth and Orgad 2012, p. 542). The replacement of targeted screening by the automated HBS system in 2015 led to the improvement of passengers' impressions regarding the fairness of the procedures, and the number of reports about the humiliation of minority groups decreased (Hasisi et al. 2018, p. 417). However, today, discriminatory practices are still present in both institutions and Israeli society (Zussman, 2013). Human rights organisations such as Amnesty International point out that ethnical minorities, especially the Palestinians, are still treated differently in areas such as budget allocation, political participation or policing. As claimed by Adalah – The Legal Center for Arab Minority Rights in Israel, around 65 Israeli laws are discriminatory toward Palestinian citizens (Amnesty International 2021, Adalah 2017).

Therefore, the incorporation of ML systems into profiling methods might reinforce the existing discriminations and grievances among ethnonational groups. The case of Israel shows the importance of ethnic identity in conceptualising threats and designing the measures supposed to identify terrorists. Although it has not been explicitly expressed that the attention is targeted to the members of a particular group, the profiling practices indicate that they are watched on a larger scale than others. Considering Israel's role as an exporter of security knowledge and technological innovations, the efficacy of Israeli ML security systems might also produce less accurate results when applied to other contexts.

Surveillance focused on minorities

Another indication of possible discrimination based on already biased data is the analysis of social networks, which focuses on the accounts of individuals who are assumed to be likely to become extremists. When the number of lone-wolf terrorist attacks, to which people were called upon and get inspired on social networks, increased in 2014, Israeli security services launched a surveillance campaign of monitoring social networks. The attention was paid to young Palestinians, who were perceived as the most likely to succumb to these incitements. Subsequently, these individuals, who were evaluated as potential terrorists, were notified by telephone, or their parents were requested to talk to their children and eventually dissuade them from committing the attack (Hasisi et al. 2019, p. 415-416, Harel 2017).

Similarly, since the 1967 Six-Day War, Israel has been issuing different IDs for Palestinians and dividing them into groups according to the territory of their residence (Weitzberg 2020, p. 8). Accordingly, Palestinians are not allowed to move freely on Israeli territories as they need to be screened and controlled at checkpoints, whereas the Israeli Jews are not required to follow these procedures to move across the same territories (Tawil-Souri 2012, p. 153). The use of biometric systems and databases of photographs and fingerprints, as well as the control of Palestinians at checkpoints, are supposed to increase the security of Israeli citizens (Weitzberg 2021, p. 13). Tawil-Souri puts forward that the fact that Palestinians have different documents and their movement is controlled and recorded, demonstrates that Israel approaches their surveillance "differently and unevenly," thereby contributing to their discrimination (Tawil-Souri 2012, p. 160).

More recently, it was revealed that the Israeli military force has been using a smartphone application to verify whether the Palestinians passing the checkpoints represent a danger. The application is called Blue Wolf and is able to take pictures of the individuals, store them and match them with the person's identity in the database used by the Israeli military and intelligence forces. When the photo is associated with one's identity details, the application shows a particular colour, signifying either that the passing person is not dangerous and thus can pass or that the person represents a threat and should be detained or arrested. In 2016, the Israeli authority released similar technology called White Wolf, which enabled to scan Palestinian identification cards to verify whether the person should be allowed to enter the workplace. Former Israeli Army members reportedly compared Blue Wolf to "Facebook for Palestinians" and claimed that there had been even a competition among soldiers about who takes more photos. The soldiers were told that such systems were employed with the aim of preventing terrorism. Within two years, the Army collected thousands of captions of Palestinians (Roth 2021, Middle East Eye 2021). The data of Palestinians are also heavily collected in the occupied territory under the control of the Israeli military. Except for the limitations such as curfew and movement restrictions, Palestinians are surveilled by face recognition cameras that monitor their moves nonstop, and some of them are even installed in a way that they directly point into people's houses (Roth, 2021). Such practices as ethnic profiling and surveillance focused on a certain group are expected to lead to a self-fulfilling prophecy to some extent – the more we focus on members who belong to a certain group, the more offenders are found among them (Munk 2017).

Prediction based on facial features

Perhaps the most controversial issue has been the predictive program designed by Tel Aviv based company Faception. Faception claims it can predict personality and behaviour based on one's facial features. Faception's technology follows the assumption that DNA is linked to the personality, meaning that appearance can tell us what the character is. The company developed 15 classifiers, including the categories such as terrorist, poker player, individual with high IQ, academic researcher, extrovert or paedophile. The software analyses the video and image databases to associate these types of personalities to the recorded individuals. The algorithm encodes facial traits, face's width and height ratio, eyes' corners and the like. The company claims the accuracy of their result is 80%, praising that its software was able to identify 9 out of 11 terrorists involved in the 2015 Paris November attacks. Although we do not have precise information about the use of Faception in national security organisations, in 2016, Faception's CEO announced that it had signed a contract with the Israeli homeland security agency to facilitate the detection of terrorists (Haaretz 2016, Ring 2016, Faception, 2021). According to Faception's website, their technology is "objective like blood testing," without "discrimination by race, class, gender, age etc." The main advantage that makes it more effective than face recognition cameras, in Faception's words, is that it can focus on anonymous individuals, for whose identification only a picture of their face is needed to associate the facial traits to the categories of personality. To identify terrorists, face recognition programs usually work with pictures of already known criminals or terrorists that are not needed in the case of Faception, which is capable to identify anonymous individuals (Faception, 2021).

Thus, targeting particular individuals can have severe consequences if employed to prevent terrorism, namely to identify potential terrorists, and in the case of ML predictive analysis, can lead to the wrong identification of individuals as likely to be terrorists. The above-mentioned examples of ethnic profiling, targeted surveillance, and programs evaluating people according to their facial traits, aimed at demonstrating that the past and current practices involve an unequal approach towards minorities, which is expected to become intensified if incorporated into ML predictive systems. The bias in the case of counterterrorism does not include only facial recognition, which was stressed in this chapter but can be found also in voice recognition systems that may be biased against individuals with an accent and thus might affect ethnic and national minorities (Orbán and Ní Aoláin 2020, p. 25-27).

5.2. Lack of transparency and inexplicability

The non-transparency and inability to explain some of the processes and results of ML systems could be caused by the system's adjustment of its guidelines over time according to acquired experiences. Although it may be possible to identify and detain the terrorist before he commits the planned attack, it might not be possible to explain how he was caught – what guidelines the system used to draw conclusions. The inability to explain why the predictive system produced a particular result might have serious effects on the effectiveness of counterterrorism measures and could result in what Ganor calls a "moral dilemma." In practice, the authorities will not be able to explain based on what criteria and evidence concretely he was identified, implying that the person might be sentenced without knowing why such a decision was made (Ganor 2019, p. 10).

Since terrorism is often considered an existential threat, and Israel is not an exception in how it understands terrorism, having less or insufficient proof as inexplainable predictive program results might be more acceptable to security agencies and government officials than allowing a terrorist attack to happen. In other words, they might prefer to identify someone as a terrorist and restrict his liberty to prevent harm to civilians. As Ganor points out, the violation of human rights is in this case "abstract and embedded in large numbers in vast databases of information," whereas the cost of human lives lost in terrorist attacks is "tangible and concrete." (Ganor 2019, p. 10)

This caution applies to terrorism primarily, as the lives of innocent civilians are at risk. However, the same stands for the opposite case – when a terrorist is not identified, and we are not able to explain why he was overlooked, we will be unable to discover what caused the mistake and correct it.

Regarding the process' transparency as it was defined above – including the awareness of data collection, storage and the knowledge of the functioning of ML techniques used – it is in the case of counterterrorism rather unthinkable as the revelation of such information would be contra-productive, giving the terrorists instructions to what they should avoid and how to proceed to not be detected.

When the identification of terrorists based on such ML techniques as profiling is followed by administrative detentions without trial and arrestation without justification or sufficient evidence, the trust in institutions and justice might be undermined, inciting the grievances among targeted individuals and thus resulting in their radicalisation and decision to commit terrorist attacks (Hasisi et al. 2019, p. 412).

Israel has been criticized for the introduction of strict measures under the pretence of fighting terrorism and for making allegations without having sufficient evidence. In October 2021, Israeli authorities accused six Palestinian non-governmental organisations of supporting terrorism and demanded the international community follow their steps. They sought to prove the NGO's collaboration with the Popular Front for the Liberation of Palestine (PFLP), designated as a terrorist organisation by the international community, including the United States and the European Union. However, according to Michelle Bachelet, the United Nations High Commissioner for Human Rights, Israel did not present sufficient evidence for its accusations and encouraged Israel to withdraw its claims, explaining that the NGO's activities concerned with the human rights of Palestinians are not an act of terrorism. But the Israeli representative insisted they possessed "unequivocal evidence that included video footage, photos, payment receipts that tie the said groups to the backing of terror activity," and thus the evidence was "ironclad." (The Times of Israel 2021) It is questionable to what extent we can deduce from this example that Israeli intelligence services will be satisfied with less transparent results of predictive systems, however, as this case indicates, determining what is sufficient evidence is not easy, and with regard to terrorism, it seems the authorities rather opt for precautionary measures and limit one's activity, although the evidence is not clear.

Intelligence services have always tried to find a modus operandi – a characteristic method of operation typical for terrorists, on basis of which it is possible to reveal their motives and predict the steps they are likely to take. They used to focus on the hierarchical structures of terrorist organizations, which allowed them to proceed from small clues and individual traces to the detection of wider terrorist networks. However, these days, the number of terrorists who act independently without any affiliation to a particular organization is on the rise. Such terrorists do not follow the orders to commit the attack from their superiors and their decision to cause harm is based solely on their affiliation to ideology and their conviction or inspiration. Concrete hierarchical structures and signs that would lead the security agencies to the information that the individual will commit an attack are missing. As stated by the Israeli deputy director of the School for Intelligence of the Israeli Security Agency, it is crucial to make use of the big amount of data collected about the individual to find "a needle in a haystack" – the fragments and clues that would indicate the person is inclined to commit a terrorist attack. Such clues could be the change in behaviour or appearance, change in the volume and type of activity, suspicious transactions and travels, and so forth. The intelligence services needed to adapt to these changes and learn to extract signs and fragments from big data to put them together and find useful correlations. Therefore, the security services do not need to prove that the threat really stems from the suspect – it is sufficient to find correlations rather than dependencies to be able to make use of these relations, which are also easier to prove (M. 2018, p. 63-64).

The biased algorithms and the production of false positives and false negatives might represent a grave issue if the results of predictive programs are used as evidence to accuse or neutralize the person for their terrorist intentions. To illustrate, one of the most well-known methods of the Israeli fight against terrorism used to be targeted killing. Targeted killing is the action undertaken to neutralize the individuals, approved by the government after intelligence and security services prove his guilt. Since the accused person is condemned to death without due process, Israeli officials' role is in this case "reduced to the role of hitmen." Targeted killing was frequently used after 2000 when it has become the "main method" supposed to prevent suicide attacks (Falk 2015, p. 2). Israeli government officials believe that if there is proof that someone is a terrorist offender, the security forces are authorized to act preemptively and neutralize the enemy before the attack claims the lives of innocent citizens. While this method has become a subject of criticism for violating human rights, Israel considers it to be legitimate and necessary as long as it is ethically justifiable – therefore as long as its aim is to prevent civilian casualties (David 2003, p. 111). The efficiency of this method is a controversial topic, however, it is widely agreed that targeted killing might give rise to new waves of hatred, and to the creation of new volunteers who want to become "martyrs" by committing suicide attacks (Falk 2015, p. 2).

Taking targeted killing and insufficient allegations of terrorism as an example, this part aimed to prove that when it comes to terrorism, transparent evidence to be presented in front of the court is not always necessary in Israel, when the neutralisation of terrorists is required. This part also aimed to demonstrate that even the lack of transparency and inability to explain ML systems' results might not be an obstacle for Israeli security agencies to delegitimize or neutralize a suspect. On the other hand, they might perceive the correlations of ML systems results as evidence. Hence, the security forces will not always be able to explain how they found the links between the screened person and his connection to a terrorist organisation or to his plan to commit a suicide attack, because it is impossible with some ML methods. Although it is possible to trace data input and output data, the learning process makes it hard to verify what capabilities have been developed to predict the behaviour. This also makes it impossible to determine who should be responsible for a wrong prediction (Haas and Fischer 2017, p. 288). When a suspected person is killed without providing tangible profs, the sentiment of grievance and injustice may arise among the next would-be terrorists.

5.3. Algorithms' accuracy and false positives/negatives

The features analysed by the police to detect crime are different from those that are examined by the security forces to fight terrorism. Many terrorist groups but also lone-wolf terrorists that do not show typical characteristics of terrorist organisations differ in many aspects as in motives, preparation, and the way the attack is realised. This generates a great amount of particular and very distinct data. Hence, the traces each terrorist and potential terrorist leave differ considerably, making the training process too complex. To make sure the algorithm does not identify wrong patterns due to the diversity of attacks, the analysts may opt to extend training data sets. However, as mentioned above, increasing the dimensions of data sets does not always increase results' accuracy, and might lead to the identification of false positives. Moreover, compared to the abundant quantity of data about regular crime, the data about terrorist attacks might not be sufficient for the extension of data sets, not only because of the particularity of each attack but also because they do not occur on a daily basis (Verhelst, Stannat and Mecacci 2020, p. 2978, Ganor 2019, p. 4).

In the case of terrorism, the quality of the data and thus the resulting accuracy of the algorithms can be affected by the way the information is handled. If the information is gradually shared by different security agencies, it may be misinterpreted at the end of the process. Likewise, insufficient data or wrong data type can produce inaccurate correlations (Moeckli and Thruman 2011, p. 26-28). Although the ML algorithms are taught to operate with slightly imbalanced data sets, the data about a country's citizens are too imbalanced regarding the number of terrorists and non-terrorists (Verhelst, Stannat and Mecacci 2020, p. 2978).

With regard to reducing either false negatives or false positives in order to optimize the algorithm's accuracy, we should be aware that decreasing the occurrence of false positives increases the occurrence of false negatives and vice versa. Increasing false positives might lead to the wrong identification of innocent persons as terrorists (UNICRI and UNOCT 2021b, p. 43).

DeRosa stresses that counterterrorism could be misused by security agencies and governmental institutions as a pretext to use repressive measures such as detention, arrestation, or to impose other kinds of restrictions on liberty (DeRosa 2004, p. 14). The situation becomes alarming when the jurisdiction system allows administrative detention that does not usually require evidence presented to the trial about how the suspected person violated the law or, in our case, that the person was preparing for a terrorist attack. Yet it might be difficult to delineate what should be considered terrorist behaviour. There is no certitude that someone with extremist behaviour on social networks will really commit a terrorist attack in the future – thus, we cannot accuse him of terrorism, although the predictive programs predict that he is inclined to terrorism on the basis of data we possess about him. In an even more serious scenario, the results generated by predictive systems could be used to completely neutralize the person. Security services might assess that a terrorist offender located in a crowd, carrying a bomb and ready to detonate it, must be eliminated before he causes harm to the persons passing by. In case the predictive programs' results are used as evidence to neutralize a person, it could result in a loss of life of a suspect who has been wrongly identified as an attacker as a false positive (Ganor 2019, p. 4, 11-12).

Thwarted "potential lone-wolf attacks"

During his speech in Parliament's Foreign Affairs and Defense Committee in 2017, the former ISA's head Nadav Argaman mentioned that ISA had thwarted 400 terrorist attacks and 1100 "potential lone-wolf attacks" in 2017. Compared to the previous year, the number of realised attacks decreased by half (Knesset 2017, Times of Israel 2018). The former premier Netanyahu also mentioned that such success was achieved thanks to the development of machine learning and artificial intelligence in general (Times of Israel 2018). 1100 potential offenders that were in their phase of planning of the attack were identified by ISA and Military service also thanks to the monitoring of Palestinians on social networks. IDF and ISA allegedly detained more than 400 potential attackers, while some of them were summoned to a trial, and some were sentenced. Those, who were not arrested, received a warning. In case the suspect was young, his parents were notified. This action reportedly discouraged potential offenders and lead to a decrease in the number of attacks – due to the fear of being revealed. Palestinians who committed the attack in this period in most cases acted without any connection to a terrorist organisation (Harel 2017).

The main controversy of this prediction is that we cannot identify a person as a terrorist when the individual is in his first stadiums of planning or radicalisation. What if the person changes his mind and gives up his decision later, therefore does not commit the attack? (Ganor 2019, p. 15) The Israeli daily Haaretz confirms that the intelligence services' greatest obstacle during the monitoring of social networks was to recognize who “only” behaves in an extremist manner and who really intends to commit the attack (Harel 2017).

Ganor also warns that a high number of suspected and arrested persons should not be the criterion to measure the efficiency and success of predictive programs. The high numbers might point out that the filter used to identify offenders is extremely broad, resulting in the inclusion of those, who actually did not plan to commit any attack. Moreover, the concerns increase as the identification of someone as a “potential lone terrorist” is very vague – it is not clear whether and how many of these 1100 potential terrorists would become terrorists at the end, so it is uncertain whether the person should be regarded as a terrorist, and thus what measures should be employed. Accordingly, it is not clear how the result of the predictive system could be used as adequate evidence to convict a person in a court and how the suspect can prove his innocence when the AI system, whose decisions are sometimes inexplainable, determined that the person is likely to commit a terrorist attack. These questions remain unanswered (Ganor 2019, p. 15-16). Indeed, when the system turns out to produce a high number of people who have been wrongly identified as terrorists or non-terrorists, it raises doubts as to whether the investment in these systems is useful, and in the case of false negatives, also the investment in the additional investigation of the suspect. It can also increase the pressure on policymakers, who will have to defend why they are relying on such resources when they are not so efficient (National Research Council 2008, p. 40).

Similarly, the production of false positives might raise concerns when the intelligence services use the results to justify the administrative detention without any previous trial. It usually has neither a defined time limit nor are the reasons for detention made public. Israel has come under criticism for detaining suspected offenders, including terrorists, without providing further reasons for detention and without determining the length of their stay in prison, which can be actually extended without limit. As a result, 482 Palestinians were reportedly held in administrative detention in November 2021. Although the detention order should be approved by the judges and the suspects have formally the right to appeal to the Supreme Court, the judges usually approve the detention without any further examination of the case and without allowing any appeal by the detainee (Btselem 2021).

False negatives – a nightmare of intelligence services

On the other hand, increasing false negatives could result in the loss of lives as the offender would escape the detection systems. If the system incorrectly identifies someone as a non-terrorist, the terrorist attack will not be prevented and lives will be in danger. False negatives thus seem to be “the nightmare of the intelligence analyst” (National Research Council 2008, p. 40). It is reasonable that the production of false negatives in countering terrorism is regarded as more dangerous, as the individual who was wrongly identified as negative could cause significant harm, including a high number of casualties. It is up to analysts to adjust the line between the number of false positives and false negatives, according to the available resources and goals and the strategy of each government (Python et al. 2021).

The US National Research Council emphasizes that although false negatives appear to be more unacceptable than false positives, it is necessary to realise that opting for either false negatives or false positives is not one-to-one trade, but that per one false negative terrorist there will be tens of thousands of false positives. Accordingly, the Council warns that “a society that tolerates too much harm to innocent people based on a large number of false positives is no longer a society that respects civil liberties.” (National Research Council 2008, p. 40-41)

Lone-wolf terrorism

Furthermore, it might be challenging to identify lone wolves, who do not manifest typical signs of terrorism. Such an individual gets into a car and rams into the crowd of people or stabs the persons around him with a knife grabbed in his kitchen. In this case, we do not have any data about him purchasing the weapons or expressing himself in an extremist way. He could have visited websites with extremist content and his friends could have been part of an extremist organization, but as this behaviour alone is not the reason for arresting the individual, the predictive system will probably not assess this person as a terrorist. To predict this kind of lone-wolf terrorist attack, the use of predictive systems is in question, although as demonstrated in the Israeli case, Israeli intelligence services reportedly prevented hundreds of such attacks thanks to big data and analytical technologies. Another aspect undermining the accuracy of such programs is that terrorism is constantly evolving and the categories associated with the models of behaviour that are analysed and valid today, might not be accurate in the future (Ganor 2019, p. 4-5).

In 2021, 54 “significant terrorist attacks” were conducted in Israel, which is 14 more than during the previous year. Overall, Israel has noticed a slight rise in terrorist attacks since 2019. Most of them made part of “popular resistance” attacks that include activities that do not require firearms as car-ramming or stabbing. Moreover, around 1700 rock-throwing and 350 Molotov cocktail attacks were conducted. Most of them were also carried out by lone wolves that were not connected to any terrorist group. Compared to that, fewer terrorist attacks were carried out by Palestinian terrorist organizations thanks to the “high quality of the Israeli security forces’ counterterrorism activities and their collaboration with the PA security forces.” In general, during the last years, the most widespread “popular resistance” terrorism attacks were stabbing attacks. These do not require complex preparation and planning, as only a regular kitchen knife is needed, and this is also the reason why they are difficult to be predicted and prevented as they do not comply with typical profiles of terrorists, so the intelligence might not have any clue that would lead them to identify this person (The Meir Amit Intelligence and Terrorism Information Center 2022, p. 3-15).

5.4. Adversarial ML

As explained in the first part, there are several ways to conduct an attack against the ML systems. An adversary can disrupt the input data before they are uploaded into the system as well as the training process itself to either cause the production of a specific outcome or any other outcome than the right one. All these scenarios can be discussed in the context of terrorism prediction and prevention. The adversarial attack aimed at distorting the results to not identify terrorists could concern many of the domains in this sphere – the image recognition of terrorists, financing of terrorism as well as finding the correlations between suspected persons. For instance, the system that assesses whether a person is trustworthy and thus should get the loan can be through adversarial examples during the training process taught to approve providing loan even to such people that do not meet the criteria of the banking institution (Solomon 2020b). The same could be applied to systems supposed to detect suspicious transactions and financing terrorism.

Therefore, with unceasing advancement, the opportunities come not only for the development and improvement of the systems but also for the enhancement of the methods supposed to find the vulnerabilities, deceive the systems and produce incorrect results. The following paragraphs will discuss the possibilities of an adversarial environment in counterterrorist operations. To illustrate the adversarial examples in a counterterrorism context, image recognition and content analysis will come under scrutiny.

Image recognition

Image recognition systems are nowadays heavily connected to biometric databases and by screening the pictures, they are able to associate the identity to a screened person. However, the vulnerability of ML recognition systems might be attractive for the offenders that could use the same technology to target them.

Sharif et al. (2016) are concerned with the adversarial attacks that are "physically realizable" and "inconspicuous." These attacks can change a person's visage to deceive recognition systems but at the same time look natural and not attract human and recognition system's attention. Such an attack could take the form of using makeup or accessories. With glasses designed by the ML program, the authors scrutinize to which extent the ML recognition systems are robust against such adversarial attacks. According to them, it is harder to manipulate the input data that flows into the systems, as they would have to manipulate the physical scene recorded by cameras that can be also influenced by such factors as the light or the distance from the sensor. It is also necessary for the attackers to not be noticed by a human passer-by when trying to manipulate physical reality. The bystanders could notice and suspect the person who has too much makeup or even a mask on his face. The attacker might aim to confuse the system to look like another person – for instance, a high ranked person as the company's director – to be enabled to access the places and the information limited only to selected persons. Alternatively, his goal could be only to avoid being identified as the correct person, and to be labelled as unknown or as anyone else. This type of attack also seems to be more feasible as it is easier to make the system identify the person as someone else than a concrete person (Sharif et al. 2016, p. 1528-1529).

Except for the use for malicious intentions, this kind of attack can be also employed by people who want to protect their privacy rights and who feel threatened by massive surveillance. The attempts to fool the surveillance recognition systems have been successful and are not new, however, such practices usually consist in covering the face or printing the mask from the internet, which can be easily noticed by other persons. The authors, therefore, designed and printed the glasses that helped to misclassify people in 80% of the cases. In this type of attack, the offender influences only the composition of the inputs based on the knowledge he has about the system – the knowledge he gained through observation of the behaviour of the system and the results it produces, or, he knows the algorithm that is used for the classification (Sharif et al. 2016, p. 1529-1530).

In another study, the Israeli researchers designed an AI system that suggests a makeup with which people are not recognized by facial recognition systems. With such makeup, the persons were recognized only in 1,22% of cases (Guetta et al. 2021). Although we may assume that the terrorist will not use makeup in such a way because the machine told them to do so, these examples demonstrate that AI systems can be deceived by similar technologies. In the case of terrorism, it does not necessarily have to be a matter of applying certain makeup or glasses, but similar and innovative methods of physically adjusting the appearance and face-covering that will not be noticeable and suspicious to bystander humans.

The former chief executive director of the Israel National Cyber Directorate's Technology Unit Hudi Zack claims that such systems that would enable malicious actors reveal vulnerabilities of security agencies' systems are not very real at the moment, but as the technology development is so fast, the National Cyber Directorate needs to be prepared in advance. The Israeli Cyber Directorate is thus working on the methods and criteria for companies and other AI users that will help them to build in their technologies the components that will ensure secure use of the systems. According to Hudi Zack, not all companies are aware of the risks AI systems pose, as their pursuit of economic profit outweighs any other considerations (Solomon 2020b).

On the other hand, malicious actors have already learned to misuse AI technologies for their own purposes. For instance, the UN Report warns about the morphing method through Generative Adversarial Networks that has been already used by terrorists to deceive biometric data systems supposed to recognize forged passports. In practice, several passports might contain the same profile photo and might be used by more than one person, without making the biometric recognition systems notice it (UNICRI and UNOCT 2021a, p. 44-45). Likewise, the ML techniques able to learn and imitate someone's voice have already been employed for purposes such as phishing. The system taught to speak in the same voice as a real person can ask human agents to act in the desired way. During the phishing attacks, offenders recorded the company's CEO's voice and trained the ML system to talk with the same voice, enabling them to give orders regarding sending money to a certain account (Solomon 2020b). Therefore, even though the use of adversarial attacks for terrorist purposes has been rather a matter of academic research at the moment, the software developers and security companies using the advanced technologies should not neglect these threats, because the malicious actors, as we could have seen, have already learnt to misuse the ML technologies to accomplish their goals and it is the matter of the time when they will use the adversarial machine learning to deceive the predictive systems supposed to detect terrorism.

Text analysis

In general, text classification allows categorizing text into appropriate groups according to the words and expressions contained. It is used to conduct online discussion analysis, detect hate comments, inciting to terrorism online, religious extremism, explosives related texts and bomb-making recipes (Li et al. 2018).

IntuView, another Israel-based company, offers various tools for data analysis, including document analysis or social media analysis. The document analysis DOCEX for Homeland Security is able to translate the texts, but also to understand the meaning behind them. Its main goal is to "provide the intelligence and counterterrorism community with an ability to understand radical Islamic and terrorist-related texts without the need to have well versed Arabic literate analysts on hand." It should be capable of considering religious, social and political contexts that might not be always noticed by human translators who often focus only on the language translation without deeply understanding the logic behind the text. The program can analyse around 1 million texts per day, making it not only more precise but also more rapid than human analysts (IntuView 2018a). Similarly, the Social Media Monitor can scan social media to identify the main trends and ideas spreading around the internet and again consider different social and political contexts. Another IntuView product, the IED Recipe Detector analyses the text and scans whether it talks about chemicals related to the explosives. These texts are then matched with already known terrorist procedures and recipes to make the explosives. The program is also able to assess what kind of explosive exactly is a recipe for and how it will be probably used – whether as a car bomb or a bomb on a vest of a suicide bomber. Except for the identification of explosives-related texts, the systems can be integrated into SIGINT for eavesdropping, control of websites and other open-source intelligence (IntuView 2018b, IntuView 2018c).

Although the wider public cannot access the precise information about the model and the algorithms used in such a system as well as to what extent it could be secured, it cannot be ruled out that it could become the target of adversarial examples. Li et al. designed the system TEXTBUGGER, which can generate texts with content that stays comprehensible for humans but is undetectable for machine classifiers. The authors tested the efficiency of existing online Text Understanding applications using Deep Learning, namely sentiment analysis and the detection of aggressive and hateful content. The attacks were in 100% of cases successful when used to disturb IMDB dataset Amazon Web Services and Microsoft Azure under the black-box conditions. Moreover, they proved that the adversarial model is transferable and when used to disturb offline models, they can be transferred to online Deep Learning-based Text

Understanding applications. The system has a marginal influence on human understanding of a given text. It first finds keywords, for which "bugs" were created. These "bugs" are slightly modified or misspelt expressions that replace the original words. The words were changed in several ways – by adding a space between the letters of a word, deleting a letter from a word, mistaking two characters that should be correctly adjacent (which is a typical human error when writing a text on a computer), replacement of letters by characters that resemble the letter, or replacement by another similar letter. The usual real-world dictionary comprises a significantly smaller number of words than is the number of possible combinations of letters that could create such a misspelt word. Therefore, the strategy is very simple – only making a typo is enough to make the classification system identify the word as "undetected" or "unknown." (Li et al. 2018, p. 4) Undoubtedly, the developers of the programs supposed to translate and read the meaning of the text must be somehow prepared for such adversarial disruption. However, we are not able to evaluate to what extent these systems are robust and there will always be the possibility that offenders find a way to deceive them.

As there is a large number of companies concerned with the development of AI and ML products, Israeli authorities should be aware of the extent to which the services are trustworthy and their systems robust. The Israeli National Initiative for Secure Intelligence Systems emphasizes that since AI systems are dependent on computers, they become more vulnerable to cyberattacks (Ben-Israel, Matania and Friedman 2020b, p.10). In the words of one of the Initiative's authors, Ben-Israel, "the more we become dependent on AI – the more successful AI is globally – the more vulnerable we will be to cyber attacks." This Initiative emphasizes the capabilities of intelligent systems and calls for not neglecting their development and for considering both their benefits and risks, as they could be critical for Israeli security, its economy and the well-being of Israeli citizens. Although the plan was finalized in 2019 and since this time it has been waiting for the government's approval, the organizations are allowed to implement the program in advance. Being aware of the potential of the development of AI, the IDF has already implemented the plans and started developing the technologies for their own needs (Hennessey, 2022).

To outline, the law enforcement agencies and intelligence services should not forget that as with almost every technology, the AI and ML systems are vulnerable and might become a target of an adversarial attack, causing the machine to produce incorrect results. Taking into account the wide scope of AI and ML applications in counterterrorism operations, from facial recognition to the detection of financing terrorism, it is safe to say that all of these techniques can be deceived by an adversarial attack. Accordingly, the same technology used to detect and prevent terrorism might be used by terrorists and criminals to commit criminal activity.

It is possible to create a defence against these adversarial attacks or to attenuate their effects, however, they cannot make the systems entirely secure and robust. The existence of adversarial examples demonstrates that these technologies might be circumvented by simple methods, as well as by advanced technology. In the case of facial recognition systems, it is possible to put on the accessories or makeup that will confuse the system, in the case of text analysis, it can be deceived by various typing errors. With regard to more complex adversarial attacks, AI techniques can discover vulnerabilities of data analysis systems and exploit them in order to produce incorrect outcomes. In a real-world scenario, it could lead to the failure of recognition systems and allow the terrorist to enter the country or to gain access to secret and vulnerable information. Regarding the text analysis, it might not allow the system to detect explosives recipes and incitement to terrorism online, and thus not prevent the recruitment of terrorists, planning of the attacks and its share on social media. Another important aspect that this part aimed at demonstrating is that humans, including terrorists, are adaptable and that such systems that are supposed to detect terrorist content or recognize terrorists according to their records, can be attacked by the same technology. Being one of the most developed and investing countries and aspiring to become a world leader in the AI field (Ben-Israel, Matania and Friedman 2020b, p. 8), the Israeli government has to be aware of potential risks and invest in the development and research in the field. As this paper tries to demonstrate, ML products are being used to prevent terrorism although in a large number of cases it is hard to verify to what scale. The former minister Benjamin Netanyahu's award for ML technologies' contribution to terrorism prevention proves that these technologies represent success in the eyes of Israeli authorities, indicating that the government might be interested in their further development and use for these purposes.

5.5. Increased surveillance and the role of private and state sector

Machine learning surveillance in counterterrorism implies the control of individuals through systems with ML technologies such as the recognition cameras or screening of social media to identify and track individuals, who might represent a threat. The unceasing monitoring of social media and citizens' devices enables machine learning technologies to enhance the capabilities to prevent terrorist attacks, including the detection and suppression of radicalization and the spread of extremist ideologies, financing terrorism, profiling of potential victims and offenders, recognising concrete perpetrator and collecting the evidence for his detention as well as it improves the operation of intelligence agencies using drones for information gathering to locate terrorists and to predict where are they likely to be located in future (Meijer and Wessels 2019, p. 1034, Ganor 2019, p. 2, McKendrick 2019, p. 10).

Human rights defenders question the use of ML technologies for such surveillance, as the unceasing monitoring of people violates privacy rights and as the ML applications enhance the capabilities to produce information about every single moment of the life of the individuals, including those who are not necessarily involved in criminal or terrorist activity. The gathered information can be then easily misused by actors with malicious purposes (Ganor 2019, p. 5). According to M. from ISA, nowadays, it is feasible to pursue people through any modern technological device used in everyday life – not only the biometric IDs that the citizens are obligated to establish and carry with them all the time, but also the platforms and products they are not forced to use by their authority, such as social media, smartwatches, and smartphones through which they share the personal and sensitive data voluntarily and consciously with both the state and private companies (M. 2018, p. 66). Therefore, security in terms of protection against terrorism takes its toll on human rights and privacy. This trade-off stems from decision makers' inclination toward surveillance and means that are often compared to measures of authoritative regimes yet that promise early detection and prevention of terrorist attacks (Ganor 2019, p. 9-10, Ganor 2015, p. 36).

Terrorism is usually perceived as the worst felony possible, and the successful securitization of this phenomenon leads to the justification of the deployment of exceptional resources that would not be accepted if they were used for fighting everyday crime – not only the gathering of sensitive data about people and its storage and analysis but also the use of this information to undertake preemptive measures. As indicated above, the notion of terrorism has not any widely agreed meaning and is usually used to portray harmful and uncertain threat that incites feelings of terror and anxiety. It is also for this reason, that the system can be easily misused – labelling an unwanted action or uncomfortable person as terrorist allows the employment of extraordinary measures supposed to prevent such events from occurring (McKendrick 2019, p. 31). Owing to the attempts to prevent terrorism, the distinction between private and public lives becomes unclear as the governmental agencies, commercial sector and citizen's social media and smart devices become interconnected. Likewise, the boundaries between a state's criminality and international security become blurred as the scope of national counterterrorism programs and everyday policing become overlapped (Zedner 2007, p. 263-264, Egbert and Leese 2021, p. 30).

This part aims at showing how surveillance allows gathering a vast amount of data to find the correlations to predict the future. Machine learning tools enable not only recording of the data but give them also some kind of meaning, extract particular information and find connections and similarities. Therefore, this part will pay attention to the use of biometric databases and recognition systems, through which Israeli authority monitors inhabitants of Israeli-controlled territory. It will be demonstrated that machine learning plays an important role in surveillance and terrorism prediction in that way that it improves the capabilities of security agencies to record and store a vast amount of data, to put them together and connect them to a person's identity in identification databases. Human rights advocates criticise these actions due to the violation of privacy rights, especially towards the Palestinians in the occupied territory, which could also represent a significant attempt to suppress any opposition and disagreement. Lastly, it will be demonstrated that the data captured by surveillance tools cannot be completely secured, as with the growing advancement of technologies also grows the capabilities of the actors with malicious interests. Regarding the Israeli case, it will be proven that the tools used for counterterrorism monitoring have been reportedly employed to collect evidence against suspicious people, therefore representing a functional creep when the technology is applied also for another investigation than to the counterterrorist one.

Biometric databases

Biometric identification documents are usually connected to a database of residents to easily associate a citizen's identity with a screened person. In general, biometric IDs are supposed to avoid forging documents and to protect citizens from the theft of their property, but also for more precise identification of victims of natural catastrophes and armed conflicts. One of the reasons for the implementation of biometric components into identification documents in Israel was also the prevention of illegal entry and residence of foreigners and foreign terrorists on Israeli territories (Spektor 2020, p. 113-114).

Since the start of the discussion about the introduction of biometric systems and biometric databases, the Israeli representatives in Knesset often made a connection to terrorism when they referred to a crime. The implementation of the biometric system has been claimed necessary to prevent illegal migrants to enter the country, as they might include terrorists. Similarly, the Israeli Ministry of the Interior launched promo videos arguing in favour of the use of biometric systems, trying to persuade its citizens that the implementation of the system is inevitable for guaranteeing their security (Spektor 2020, p. 113-144).

As mentioned above, through the biometric system, the state can focus the surveillance on a particular group. Weitzberg and Spektor emphasize that the original purpose of using surveillance technologies is the monitoring of minority groups who also serves as samples for experimenting. The authors demonstrate that before being implemented on Israeli citizens, the biometric system was tested on the citizens in the occupied territory (Weitzberg 2021, p. 8, Spektor 2020, p. 109-116).

Furthermore, the use of biometric systems has also become a target of criticism due to the fear of data leaks from the biometric database. According to Knesset's former member and opponent of biometric surveillance Michael Eitan, the leak of the data from these databases would be "worse than the Chernobyl disaster – it is ten times Chernobyl," (Joint Committee, 2009, p. 66 as cited in Spektor 2020, p. 117). The advanced technologies have become not only a tool to improve data analysis but have also given the opportunity to malicious actors to access the system and to steal the data or to publish them.

As the following part demonstrates, these biometric databases connected to recognition systems have become a crucial tool for the identification of Israel's inhabitants and the detection of terrorists.

Surveillance in the Israel-controlled territory

Except for the use of facial recognition systems at checkpoints discussed in the first part, Israel has been reportedly practising unethical surveillance of Palestinians in the occupied territory. AnyVision, later renamed Oosto, is another Israel-based company that has provided recognition systems to Israel in the West Bank checkpoints and the territory of the Palestinian Authority (Solomon 2020a). The surveillance project was nicknamed Google Ayosh, referring to the capabilities of Google that can find anyone and anything, while Ayosh is the expression for occupied Palestinian territories. Oosto's program deployed in the West Bank territory is able to keep an eye on one particular person on various cameras at the same moment, recognize suspicious objects and individuals on the records, conduct the analysis and store the data about all the persons that have appeared on the record (Solon, 2019). However, Oosto was also accused of providing the data collected by its surveillance systems to various actors, including IDF (Solomon 2020a). The company's software has also been employed by Israeli police in East Jerusalem, where the majority of its inhabitants are Palestinians (Solon, 2019). Except for its local business, the company has also provided the products worldwide – to US companies, municipalities or private security agencies (Weitzberg 2021, p. 22).

The allegations that Oosto does not observe ethical rules for facial recognition work are believed to be the reason why Microsoft withdrew its investment from the company (Solon, 2019). These rules set by Microsoft include 1) equal treatment of all groups monitored – however, at the same time Microsoft admits that "it is not possible to guarantee that any technology is completely bias-free," 2) transparency regarding technology's capacity and limits, 3) guarantee of accountability, 4) prohibition of unlawful discrimination, 5) the consent of the individuals before they are recorded, and 6) surveillance that respects peoples' democratic freedoms (Microsoft 2018).

It is impossible to get known the exact functioning and the extent of surveillance on the Israeli-controlled territory, however, when Microsoft stopped investing in Oosto, Microsoft's representatives claimed that "the available evidence" proved that Oosto's products had not "previously and does not currently power a mass surveillance program in the West Bank that has been alleged in media reports." (Oosto, 2020). As Weitzberg (2021) notes, the statement does not deny the practice of surveillance, as both expressions "the available evidence" and "mass surveillance" include the possibility of surveillance, although without sufficient proof and on a smaller scale than what could be considered as "mass surveillance" (Solomon 2020, Weitzberg 2021, p. 17). The companies agreed that ending the investment would be beneficial for both sides. Losing the investment for Oosto reportedly did not mean a significant loss for the company and it still has a lot of investors as the demand for recognition camera systems rises worldwide (Solomon, 2020).

At the time of the revelation of the concerns about Oosto's deployment, the representative of the company Adam Devine argued that the Israeli government should be vigilant and not rely on uncredible and unsafe software that is being developed in the AI market. According to him, the government should trust only such systems that can guarantee their safety – prevent the leak of data, privacy rights issues, production of biases as well as to ensure that the systems will not fall into the hands of organizations that would not use them properly. He also argued that a company's algorithms were trained in such a way that they could "equally and accurately recognize every face, every gender and every orientation, without bias." (Solomon 2020)

The public has also been concerned about the facial recognition technology's ability to identify participants of the manifestation and their subsequent detention. Nevertheless, Oosto claims that such a process does not meet its standards and privacy policy rules. Reportedly, the software identifies only the persons that have already committed a crime and thus are on a list of offenders. In order to add a new person to the list, a special authority's approval is required. The program therefore cannot simply upload a picture, identify and track the person if there is not any record about the person yet. Adam Devine also denied that Oosto's algorithms would be biased because they had been "trained on the most diverse datasets in the most diverse, real-world environments." He also claimed that he wished all law enforcement agencies in the world used their software, for the above-mentioned reasons (Solomon 2020). Therefore, despite allegations that the company violates human rights by its massive use of surveillance tools, its representatives believed it was one of the most secure technologies that could recognize and identify offenders and thus prevent crime and terrorism. Although it is not possible to examine in detail the operation of these systems, the demand for the company's services not only from Israeli authorities but also from customers abroad indicates its popularity and hopes their clients put in its products.

Furthermore, the Israeli privacy law does not apply to Palestinians living in the West Bank, as they do not have Israeli citizenship (Brown 2019). The founder of Israeli company Suspect Detection Systems Shabtai Shoval confirmed that "privacy ... is not really an issue in the West Bank," and that "everybody is monitoring everybody, because everybody's afraid of everybody." (Estrin, 2019)

Jerusalem has also long attired the attention of companies from abroad, such as Cisco, or Mer Group that established cooperation with the Israeli government for the development and innovations in visual surveillance technologies. In the Old City of Jerusalem, where the majority of inhabitants are Palestinian, around 400 CCTV cameras Mabat 2000 has been deployed in 2000. During this period already, the defenders of human rights pointed out that the law enforcement agencies have not developed any code of use that would protect individual rights and provide lawful use (Weitzberg 2021, p. 14).

In 2014, 43 so-called "refuseniks" – reservists from the SIGINT agency's Unit 8200 signed and published a letter expressing their disagreement with their missions conducted on the territory of the West Bank due to the massive surveillance of nothing-knowing Palestinian citizens that are not involved in the military conflict. The information gathered through this monitoring was reportedly later used to blackmail Palestinians and to force them into collaboration. The Unit's workers were told to focus on sensitive information such as sexual orientation or infidelity to frighten the individuals (G. Waters 2019, p. 575).

They pointed out that what the Israeli Defense Forces do in the occupied territory is not always defence as the title of the organization indicates, but rather taking control over the inhabitants in order to protect Israeli society. The reservists also asserted that in terms of Israel's intelligence, Palestinian citizens did not have the same rights as the Israeli ones and that the regime that took control over people over years had become desperate for "infiltrating every aspect of their life." To access the data about the Israeli people, the intelligence service usually must ask the court for permission, which does not apply to the Palestinians. They also confirmed that the gathered information served as proof to put the Palestinians under administrative detention (The Guardian, 2014).

The reservists also warned that in such a case, "there is no one who is responsible." Neither the soldier who shot down the pursued person nor the intelligence analyst who provided information about his location. By publishing the letter, the authors wanted to show what is considered a normal daily situation accompanying the occupation in Israel (The Guardian, 2014). This case demonstrates that the problems of determining accountability have been present in intelligence services even before the introduction of advanced machine learning systems and considering how these systems work and that they have been used more and more for data analysis and crime prediction, we might assume that this accountability gap will become even deeper.

Pegasus software

Lastly, although the Israeli regulations do not allow gathering data on Israeli people without the consent of the authority, concerns about the use of surveillance technology have been raised also among Israeli citizens. Therefore, not only inhabitants of the occupied territory but also the Israeli people can become the target of the state's monitoring.

Pegasus is spyware developed by an Israeli-based NSO group, that was accused of helping to pursue activists in non-democratic states and of spying on important political leaders. The Pegasus program by NSO is considered to be one of the most effective and at the same time the most dangerous surveillance tools. It enables to activate camera or microphone as well as file downloading, without being noticed by the targeted person. Although the Israeli software developers have argued that the technology was not intended to repress activists, the governments usually have their own interpretations of offensive action – what is considered human rights activism in one country could be portrayed as terrorism in another (Ari Gross, 2021).

According to NSO, the company produces technology that helps governments save human lives against crime and terrorism (NSO, 2022). However, Pegasus was also blamed for being used to spy on Palestinian human rights defenders. Three of the pursued activists were also members of organizations that were claimed terrorists by the Israeli state. The NSO group reported the company could not confirm the identities of targeted persons, as the company only provided a license yet did not have details about targeted individuals. Two of the pursued persons were then put in administrative detention – without trial and sufficient evidence. According to Human Rights Watch, such surveillance not only threatens the rights to privacy of the watched persons but also could discourage persons interested in taking part in political activism (HRW 2021).

In response to the outrage over the misuse of NSA products to arrest human rights activists and opponents of governments in other countries, in December 2021 the Israeli government announced plans to strengthen regulation of the export of surveillance technologies. The Defense Ministry's Defense Export Control Agency decided to introduce a certificate, which must be completed by Israeli technology companies. The document defines the concepts of terrorism and serious crime more clearly – it states that "an act of expressing an opinion or criticism… shall not, in and of itself, constitute a Terrorist Act" or a "Serious Crime." In this way, Israel claims it seeks to prevent Israeli products from being used to "inflict harm on an individual or a group of individuals, merely due to their religion, sex or gender, race, ethnic group, sexual orientation, nationality, country of origin, opinion, political affiliation, age or personal status." In case of a breach of these regulations, Israel has the right to withdraw from the export licence. However, this measure has been also criticized because the form will be filled in by the Israeli company selling the product, not by the end-user who buys and uses it (Ari Gross, 2021).

At the beginning of 2022, the Israeli business newspaper Calcalist revealed that Israeli citizens have also become the target of Pegasus' surveillance. These persons were ex-premier Benjamin Netanyahu, government officers and members of NGOs. The surveillance was conducted without any approval of the court or any other legal authority. To illustrate, Pegasus has been installed on the device of a person close to the public official to check whether he is involved in criminal activity and if so, to collect the evidence. The person was therefore spied on in advance, making the surveillance rather preventive than reactive (Ganon, 2022a).

Israeli police did not confirm the information published by Calcalist and claimed it to be "untrue," arguing the police obey the law during its operations, and, that their action is always under the control of the responsible authority. They also added that they would continue to work in this way with all the material they had access to, both in the “physical and online spaces,” in order to “fight crime in general, and organized crime in particular.” (Ganon, 2022a)

There is no supervision over how the data has been collected, used, and stored by law enforcement agencies and how they have been shared with various actors. When the SIGINT gather the data, they pass the information to the police who provide it to other investigative institutions, who do not necessarily know how the Police acquired the data. Indeed, the information collected can be later easily used for blackmailing or forcing the surveilled ones into the activity they would otherwise refuse to take part in (Ganon, 2022a).

Moreover, the Calcalist informed that in some cases, external hackers were hired to perform Police’s investigative work. Therefore, access to the surveillance cameras was given to persons that did not dispose of any kind of security clearance. In the past, SIGINT hired at least three hackers to collect data that the Police did not arrive to gather via Pegasus software. They hacked the private Wi-Fi and downloaded the data from smart devices and recognition cameras. There is no guarantee that such persons will not misuse the information they gained during the monitoring and that they will not transfer it to the third party, as they did not pass any security check to verify that they were trustworthy (Ganon, 2022b).

The case of Pegasus confirms that data gathered by surveillance spyware can be used as proof against human rights defenders or other uncomfortable persons and as a justification to put them into administrative detention. Such practice is expected to continue with the employment of ML counterterrorist surveillance tools, that will provide the predictions with a veil of truth and precision. Thanks to these technical capabilities and a big amount of data, surveillance is omnipresent and focused on everyone, including the persons about whom we do not know for sure that they have committed the act of crime. This is valid also for counterterrorism measures – even those who are not likely to commit terrorist acts in the present are under control in case they will do so in the future. Finally, this part also showed that access to data in Israel is not always limited to authorised persons and that the information is often shared between various actors – which can in the case of the use of ML predictive programs affect the accuracy of their results and endanger their proper handling.

6. Findings

The thesis aimed at presenting the challenges of using AI/ML technologies for predicting and preventing terrorism and at demonstrating them in practice on the Israeli case.

The prediction and the subsequent prevention of terrorism are expected to be achieved through the monitoring of citizens and the analysis of their behaviour. The analysis is provided by machine learning technologies that enable the systems to act with some degree of autonomy and to make predictions based on gained experience. While the thesis focused mostly on the attempts to identify terrorists to prevent them from causing damage and to detect extremist and terrorist content, it should not be forgotten that similar issues will be encountered when trying to predict areas vulnerable to terrorist attacks, terrorism financing and so forth. To answer the question to which extent can we rely on AI and ML systems to predict terrorism, the thesis focused on five key challenges the predictive approach and the use of ML might encounter and that were obtained from the examination of predictive policing literature and the functioning of ML technologies: algorithm bias, algorithms' accuracy, lack of transparency and inability to explain the outcomes, the vulnerability of the systems and adversarial examples, and the increased surveillance.

First, the thesis discussed the consequences of biased algorithms in case the ML predictive systems are employed to identify terrorists. We can expect that this bias will be incorporated in the results of predictive systems, which will not be able to assess the situation objectively, as the data inserted are the reflection of past practices and prejudices rooted in the society. The Palestinian population and the situation in the Palestinian Authority in general, are extensively monitored. The surveillance focused on this minority may lead to the concentration of resources including the predictive programs to the control of Palestinians, resulting in the production of a high number of Palestinians wrongly suspected of terrorism. At the same time, if this targeted surveillance is apparent and those who are controlled are aware of that, it might lead to the dissatisfaction of Palestinians and their inclination towards extremist ideologies. Very dangerous could be such systems for data analysis that proceed only from such criteria as nationality, appearance, or accent. Although Israeli representatives deny any kind of discrimination against Palestinians, the fact that their privacy rights cannot be guaranteed in practice and the discourse that often labels another ethnicity as a threat does not support their assertion. There is a significant number of companies concerned with data analysis and although we do not have concrete information about which of them and to what extent the Israeli security agencies use them, the limits discussed in this section could apply similarly to all ML predictive systems.

Second, the paper reviewed the possible effects of the opacity of ML predictive systems reasoning and results. Not being able to explain that a suspected person is taking part in terrorism subverts the notion of evidence as it has been long understood. The evidence is not something that we can clearly show up but is based on some criteria generated by a machine and learned from the analysis of past data. And this is valid not only for the case that an innocent person is evaluated as a terrorist but also for the situation when a person is incorrectly assessed as a credible person, possibly leading to a terrorist attack. In other words, it will be impossible to prove why the system failed and based on what criteria it labelled a person as a non-risk person. The examples showed that the detention of a suspect of terrorism in Israel does not necessarily need to be preceded by demonstrating sufficient evidence or a trial. When the machine learning program, for which it is not always possible to demonstrate a transparent procedure of data gathering as well as the machine's inducting, the result and practice will be very similar. The inability to explain why an individual is believed to represent a danger may not be considered by Israeli authorities as a reason to not detain him.

Predictive programs cannot have absolute accuracy and when applied to populations in Israeli-controlled territory, even the systems with high accuracy could produce a high number of false positives or false negatives. It seems reasonable that the Israeli or any other authorities are willing to tolerate rather the production of false positives and to detain innocent persons than produce many false negatives and endanger the lives of citizens. The question is then, how many false positives and how many false negatives the Israeli government is willing to accept? Another puzzle is how the detainees can then prove they are false negatives – how will they demonstrate that they are the small percentage of the innocent ones when their claims are in contradiction with the machine's results?

Furthermore, the terrorists can use ML technologies to discover the vulnerabilities of predictive programs or to design tools that make the system produce incorrect results. The perpetrator may but not need to know the entire functioning and parameters of the programs to be able to deceive them and might alter the data before they are inserted into the system or after that. The capability to attack the system and ensure it will produce incorrect outputs extends the scope of terrorists' opportunities in various areas. In the case of Israel, it could be the recognition systems and content analysis systems, as the Israeli security agencies use them. To illustrate, a terrorist could enter the country illegally because his ID or passport was wrongly recognized by the recognition system that had been attacked. The recognition of illegal migrants that could include terrorists was one of the reasons for the introduction of biometric systems in Israel and the installation of checkpoints for non-Israeli citizens passing across the territories. Although these controls were introduced a long time ago, nowadays, human controllers are replaced by intelligent systems that are expected to be faster and likely to not be deceived. Likewise, the adversarial attacks may be realized against the text analysis or voice recognition systems. We have seen cases when the malicious actors managed to circumvent these measures by different methods such as covering the face to confuse recognition cameras or by using the technologies to forge passports and deceive the controlling systems. We can expect that with the increasing use of intelligent systems for terrorism detection, terrorists will learn to use the technologies and exploit their vulnerabilities for their own goals.

Finally, the paper examined the effects of machine learning surveillance conducted to prevent terrorism. The increased targeted surveillance in the Israeli-controlled territories is often questioned in terms of the privacy rights of non-Israeli inhabitants. However, we have also seen the case of an Israeli company that provided spyware to track high ranked Israeli officials and was also sold out abroad. Thus, the up-to-now practice of surveillance in Israel includes the control of both inhabitants of Israel and occupied territories with an emphasis on the latter one. The Israeli security agencies have also been alleged for misusing the information gathered by surveillance technologies for persuading and blackmailing surveyed targets. The introduction of ML surveillance systems that can connect information from various sources and analyse the correlations, but at the same time is not perfect and might produce inaccurate results, represents the reinforcement of the problem if surveilled yet innocent people are pursued, blackmailed or forced to collaborate. The availability of a big amount of data therefore not only provides a great amount of information that can be used for developing the ML systems yet handling these data can be also misused for other purposes. The state and private sectors are interconnected as the technologies are often provided by private companies to Israeli security agencies. This interconnection magnifies the scope of surveillance, which has been also subject to criticism due to improper handling of sensitive data.

Finally, through examining the predictive policing approach and the challenges the law enforcement agencies and intelligence services encounter when trying to prevent crime and terrorism, the thesis aims to emphasize the differences and similarities between predictive policing and terrorism prediction.

Scarce data about terrorism

Compared to data about crime used to train predictive policing models, there is in general less data about terrorism. In most countries, terrorist attacks are rather rare and as mentioned above, they differ in so many aspects and thus can be translated into so many different variables that it is hard to build general models for predictions (Munsk 2017, Verhelst, Stannat and Mecacci 2020). Although there might be a high number of terrorist attacks in Israel, it is important to realize that for training such systems, a vast amount of data is needed. However, due to the complexity of terrorism, the data might contain too many variables and might not be perfectly accurate.

Another challenge is the increasing popularity of lone-wolf attacks. This type of attack is both low-cost and difficult to be predicted, as it might be carried out by a person without warning by predictive systems that do not find any clues indicating the person is likely to commit a terrorist attack. Many of these attackers were not caught up until they pulled out the knife in public when they were noticed by security forces and were subsequently detained or shot. In the case of a car-ramming attack, it might be even harder to stop the attacker in a vehicle before he carries out the attack.

The adaptation of malicious actors

The criminals and terrorists are likely to discover the assumptions the security agencies work with to predict criminal and terrorist activities, which will make them capable of adapting to the system of predictions. Terrorism is evolving and adapting to the epoch and possible capabilities. Plane hijacking pushed Israeli authorities to increase the security controls at the airports, which successfully discouraged terrorists from plane hijacking but made them invent a new strategy – the attacks against embassies. After increasing the degree of security at the embassies, the terrorists started to attack public spaces by stabbing individuals and car-ramming crowds (Aravind, 2016). Therefore, it is no exaggeration to say that terrorists have always found a way to conduct their activities and cause harm. Like anyone else, they have learnt to use modern technologies as intelligent devices, including drones and automated weapons (UNICRI and UNOCT 2021a, p. 25-26).

The inferences about criminal activity and terrorism are not based on real criminal activity yet on past patterns of policing that inform us about what training data we were able to gather. Therefore, the continuity of police practice rather than the prediction of future development is expected. The continuity will be reflected in predictions, making the criminals able to adapt to the system and circumvent it. This debate is applicable to terrorism as well. What is evolving is not only the tools and a way of their use but also the methods the terrorists recruit and attract new sympathisers. Thus, counterterrorism strategies need to be adapted to these changes as well. According to their own words, it is what the Israeli security agencies do

(Aravind 2016). However, it is necessary to realise that the systems depend on past data that provide a starting point for the training process. To make the predictive programs more accurate, they would have to be able to adapt to changes and historic, social and political contexts.

The significant difference between predictive policing and terrorism prediction is, that if we assume that the environment will adapt to the predictive measures, the adaptability of terrorists will be lower as the prediction does not make part of public discussion to such extent as we could have seen in the case of the use of predictive policing by law enforcement agencies – the access to the system of predictions will be more limited and it will be significantly harder for the terrorists or any other malicious actors to exploit it or to bypass it through committing other forms of terrorist actions than those that can be predicted by the predictive systems.

Understanding the context

Although some companies using ML technologies claim that their systems can read between the lines and take into account religious and political background and historical development, we should not forget that they are still the machines that, put simply, calculate the results according to what they have learnt from past data. They miss any kind of understanding of the social world and in contemplating terrorism, even human researchers do not reach a consensus and every theoretical framework approaches every level of the phenomenon differently, answers different questions and reveals another perspective on the issue. These are the reasons why the systems cannot deeply understand what happens in the world of human reasoning, including the deep understanding of terrorism that would contribute to its prevention.

It follows that ML technologies might help the intelligence sector to find correlations between the terrorists, reveal what are their capabilities and objectives, and track and discover the structures of terrorist organizations as well as their functioning. However, although ML systems sometimes give the appearance that they are all-knowing as they use a huge amount of data to draw conclusions, there are various gaps consisting in what ML systems will not tell us about terrorism, namely the understanding of terrorism as a social phenomenon, understand the ideology spread by terrorists, and their motivation and considerations while deciding about committing an attack (Ganor 2019, p. 8).

Another question will be to what extent the systems of the Israeli companies that were trained on data gathered in Israeli-controlled territory will be efficient if they are sold out to the governments or security agencies abroad and thus applied to other populations.

Misuse of the technologies

Most technologies are undoubtedly a double-edged sword. Machine learning technologies can help detect terrorists as well as they might allow terrorists to deceive predictive programs and other computer systems by revealing their vulnerabilities. Likewise, state authorities can use yet also misuse the technology. Many authors, including for example Ganor 2019, Brundage et al. 2018, DeRosa, McKendrick 2019 warn that the technologies could be employed by governments or private actors against the will of the citizens and in such a way that harms the individuals. As an example, we can refer to the accusations against the Israeli Defense Forces for searching for such sensitive information, which can be employed to blackmail or threaten individuals.

Conclusion

The thesis aimed at answering two research questions: To which extent can we rely on ML technologies supposed to predict and prevent terrorism? And what are the implications of their use for security in Israel?

By using the approach of predictive policing and the explanations of the operating of the ML systems, the paper discussed key challenges the law enforcement agencies could encounter when they seek to predict and prevent crimes. The empirical part then discussed these aspects in the context of terrorism in general and in the context of counterterrorism in Israel.

The first part of the thesis examined the risks that might arise during the attempts to predict the future, namely criminal activity, which is challenging even without the use of advanced technologies. It is very difficult if not impossible to harmonize the principles of the precautionary approach and pre-crime with the principles of democracy, such as privacy rights, the presumption of innocence and the right to a fair trial. Technologies in this case amplify these challenges in the way they automatize the processes and at the same time provide the veil of scientism and objectivity. Therefore, the first part pointed out that although the ML systems promise precise and fast analysis of vast amount of data, they bear significant risks that could negatively affect the accuracy of their results. The limits of predictive ML technologies encountered by law enforcement agencies are expected to be encountered by intelligence when trying to predict terrorism as well, although the dynamic might be different, as the terrorism and crime are two distinct phenomena. Thus, security agencies should not place all hopes in technology but rather stay vigilant and take into consideration these risks, especially when they aim to use their results as evidence or justification for restricting a suspect's freedom.

To answer the second question, the thesis examined the practices of Israel in countering terrorism and found out that the presence of a minority group in the Israeli-occupied territory and the approach of Israel toward Palestinians might be a source of inaccurate and biased results.

Biased algorithms cannot be removed from the data as long as society has prejudices and diverse opinions. This bias might have serious consequences on those who are in any way disadvantaged by society or seen as different and designated as a threat. In the case of Israeli-controlled territories, the paper attempted at demonstrating that non-Israeli citizens and particularly the Palestinians could become the victims of the systems due to biased algorithms. This bias, together with other phenomena such as the curse of dimensionality, imbalanced number of positive and negative cases, or spurious correlations, might lower the accuracy of the systems and could produce a high number of incorrectly classified results. To emphasize this type of challenge, the thesis demonstrated that Israel's part of terrorist profiling has used to be the affiliation to ethnonational groups that are still treated differently compared to Israeli citizens.

The paper also emphasized the lack of transparency and inexplicability that accompanies advanced machine learning systems. It was shown that in the case of terrorism in general and in the case of Israel in particular, the authorities opt for pre-emptive measures against suspected persons even though their culpability has not been proved and confirmed by legal authority. Such measures can be mass surveillance or administrative detentions. Therefore, the difficulty to explain how the machine identified potential terrorists might not represent an obstacle to Israeli security in detaining such suspects.

The vulnerability of ML systems was examined and domains that could be affected by adversarial examples were discussed. The malicious actors might be able to change the form of the text so it is unrecognisable for document analysis systems to detect terrorist content while keeping the meaning understandable for humans. It is also possible to disturb the process of ML reasoning and make the program produce inaccurate results even without changing the appearance of the words. The recognition cameras might be deceived in a similar way. To do so, the malicious actor does not need to have full knowledge about the functioning of the system, but can also operate in the black box scenario and make the system produce wrong results.

To train the models, a big amount of data is needed. This leads to the increase of surveillance, which can be enhanced by ML systems, as in the CCTV cameras that are able to develop their own criteria to detect suspicious behaviour. In Israel, this increased surveillance is often questioned because of being focused on minorities in the Palestinian Authority and because of allowing to handle the data about everyone's personal and sometimes sensitive information. At the same time, we encounter the problem regarding transparency of this process, as it is not always clear how the data was handled and by whom and thus also who might use or misuse them.

Finally, the thesis highlighted the differences and similarities between crime and terrorism prediction that could signify further challenges for the intelligence agencies. Compared to crime, the data about terrorism are rather scarce and include many different features, making it harder to build typical profiles of terrorists, undermining the accuracy of such predictions. Indeed, there is a different risk in the form of lone-wolf attacks that are even harder to be predicted on time. This issue is also related to the possible adaptation of malicious actors to the prediction systems, because as we could have seen, the methods of terrorist attacks have been changing throughout the years and these days, the most popular ways to conduct an attack in Israel is car-ramming and stabbing that do not require any special preparation and tools, and, are thus difficult to be predicted. The question is then, to what extent the systems will be able to adapt to these changes and how fast they will learn to identify these lone-wolf terrorists that do not manifest any typical signs of affiliation to a terrorist organization.

Likewise, the paper highlighted that compared to humans, the ML systems cannot fully understand and take into account the contexts, making them not suitable for further explanations. Therefore, the predictive tools should be used in combination with other methods and should not rely exclusively on predictive techniques. At the same time, they should also take into account the benefices and disadvantages of their up-to-date measures – although they seem to be efficient considering the high numbers of thwarted attacks, one should not forget that the high number of thwarted "potential attacks" and high numbers of detainees might not necessarily mean the success in terms of terrorism prevention.

References

AHMAD, Shakeel et al. 2019. Detection and classification of social media-based extremist affiliations using sentiment analysis techniques. *Human Centric Computing and Information Sciences.* Vol. 9, no. 24, p. 1-23. DOI: 10.1186/s13673-019-0185-6

Adalah, 2017. The Discriminatory Laws Database. *Adalah - The Legal Center for Arab Minority Rights in Israel* [online]. September 25, 2017 [cit. 2021-12-15]. Available at: https://www.adalah.org/en/content/view/7771

Amnesty International, 2021. Israel and Occupied Palestinian Territories 2020. *Amnesty International* [online]. [cit. 2021-12-15]. Available at: https://www.amnesty.org/en/location/middle-east-and-north-africa/israel-and-occupied-palestinian-territories/report-israel-and-occupied-palestinian-territories/

AMOORE, Louise, 2013. *The politics of possibility: risk and security beyond probability*. Duke University Press, November 2013, 232 p. ISBN: 978-0-8223-5560-1.

ANGWIN, Julia et al., 2016. Machine bias: There's software used across the country to predict future criminals. And it's biased against blacks. *ProPublica* [online]. May 23, 2016 [cit. 2021-12-15]. Available at: https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

ANTEBI, Liran, 2021. Artificial Intelligence and National Security in Israel. Institute for National Security Studies. Memorandum No. 207. *The Institute For National Security Studies* [online]. February 2021 [cit. 2021-12-15], 137p. Available at: https://www.inss.org.il/wp-content/uploads/2021/02/Memo207_AntebyENG_9.pdf

ARAVIND, Adarsh, 2016. A critical analysis of Israel's counter-terrorism strategy. *Foreign Policy News* [online]. August 8, 2016 [cit. 2021-12-09]. Available at: https://foreignpolicynews.org/2016/08/08/a-critical-analysis-of-israels-counter-terorrism-strategy/

ARGAMAN, Nadav, 2018. Preface. IN: KUPERWASSER, Yossi and David SIMAN-TOV. *Intelligence in Theory and Practice*. Israel Intelligence Community Commemoration and Heritage Center Institute for the research of the methodology of intelligence, Effi Melzer Ltd., No. 3, October 2018, 172p.

ARI GROSS, Judah, 2021. Amid fallout from NSO scandal, Israel imposes new restrictions on cyber exports. *Times of Israel* [online]. December 6, 2021 [cit. 2021-12-09]. Available at: https://www.timesofisrael.com/amid-fallout-from-nso-scandal-israel-imposes-new-restrictions-on-cyber-exports/

BEN-ISRAEL, Isaac, Eviatar MATANIA and Leehe FRIEDMAN, 2020a. IN: BEN-ISRAEL, Isaac et al. *Towards regulation of AI systems.* Council of Europe, December 2020, 197 p.

BEN-ISRAEL, Isaac, Eviatar MATANIA, and Leehe FRIEDMAN, 2020b (Eds.). *The National Initiative for Secured Intelligent Systems to Empower the National Security and Techno-Scientific Resilience: A National Strategy for Israel. Special Report to the Prime Minister.* CRC – Blavatnik Interdisciplinary Cyber Research Center, Tel Aviv University, September 2020. 64 p.

BENNETT MOSES, Lyria and Janet CHAN, 2016. Is Big Data challenging criminology? *Theoretical criminology*. London, England: SAGE Publications, Vol. 20, no. 1, 21-39. DOI:10.1177/1362480615586614.

BENNETT MOSES, Lyria and Janet CHAN, 2018. Algorithmic prediction in policing: Assumptions, evaluation, and accountability. *Policing and Society*. Vol. 28, no. 7, p. 806-822. DOI: 10.1080/10439463.2016.1253695.

Btselem, 2021. Israel holds even Palestinian minors in administrative detention. *Btselem* [online]. November 18, 2021 [cit. 2021-12-09]. Available at: https://www.btselem.org/administrative_detention/20211128_israel_holds_even_palestinian_minors_in_administrative_detention

BRAYNE, Sarah, 2017. Big Data Surveillance: The Case of Policing. American *Sociological Review.* Vol. 82, no. 5, p. 977-1008. DOI: 10.1177/0003122417725865

BROWN, Hannah, 2019. An artificial intelligence company backed by Microsoft is helping Israel surveil Palestinians. *Vox* [online]. October 31, 2019 [cit. 2022-03-05]. Available at: https://www.vox.com/2019/10/31/20937638/israel-surveillance-network-covers-palestinian-territories

BROWNING, Matthew and Bruce ARRIGO, 2020. Stop and Risk: Policing, Data, and the Digital Age of Discrimination. *American Journal of Criminal Justice.* Vol. 46, p. 298–316. DOI: 10.1007/s12103-020-09557-x

CARTELLA, Francesco et al., 2021. Adversarial Attacks for Tabular Data: Application to Fraud Detection and Imbalanced Data. a*rXiv preprint*. arxiv: 2101.08030

CHAMIEH, Joelle et al., 2018. Biometric of Intent: A New Approach Identifying Potential Threat in Highly Secured Facilities. *IEEE Xplore*. 6th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW), p. 193-197, 2017. DOI: 10.1109/W-FiCloud.2018.00037

DASTIN, Jeffrey, 2018. Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters* [online]. October 11, 2018 [cit. 2021-09-12]. Available at:

https://www.reuters.com/article/us-amazon-com-jobs-automation-%20insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-%20idUSKCN1MK08G

DAVID, Steven R., 2003. Israel's Policy of Targeted Killing. *Ethics & international affairs*. New York, USA: Cambridge University Press, 2003, Vol. 17 no. 1, p. 111-126. DOI:10.1111/j.1747-7093.2003.tb00422.x

DELGADO, Yaneisy et al., 2021. Forensic intelligence: Data analytics as the bridge between forensic science and investigation. *Forensic Science International: Synergy*. Vol. 3, 7p. DOI: 10.1016/j.fsisyn.2021.100162

DEROSA, Mary, 2004. *Data mining and data analysis for counterterrorism*. Center for Strategic and International Studies. N.W., Washington, D.C., 2004, 24 p. ISBN 0-89206-443-9.

DING, Fangyu et al., 2017. Understanding the dynamics of terrorism events with multiple-discipline datasets and machine learning approach. *PLoS ONE.* Vol. 12, no. 6, 11p. DOI: 10.1371/journal.pone.0179057.

DOMASHOVA, Jeny and Natalia MIKHAILINA, 2021. Usage of machine learning methods for early detection of money laundering schemes. *Procedia Computer Science volume*. p. 184–192. DOI:10.1016/j.procs.2021.06.033.

EGBERT, Simon and Matthias LEESE, 2021. *Criminal Futures: Predictive Policing and Everyday Police Work.* Routledge, 2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN, 231 p. ISBN: 978-0-429-32873-2.

ESTRIN, Daniel, 2019. Face Recognition Lets Palestinians Cross Israeli Checkposts Fast, But Raises Concerns. *Npr* [online]. August 22, 2019 [cit. 2021-12-09]. Available at: https://www.npr.org/2019/08/22/752765606/face-recognition-lets-palestinians-cross-israeli-checkposts-fast-but-raises-conc

Faception, 2021. Public safety & save city. *Faception* [online]. [cit. 2021-12-09]. Available at: https://www.faception.com/hls-and-public-safety

FALK, Ophir, 2015. Measuring the Effectiveness of Israel’s “Targeted Killing” Campaign. *Perspectives on Terrorism.* Vol. 9, no.1, p. 2–26. ISSN 2334-3745.

FERRARA, Emilio et al., 2016. Predicting online extremism, content adopters, and interaction reciprocity. *arXiv preprint.* arXiv:1605.00659.

GANON, Tomer, 2022a. Israel police uses NSO’s Pegasus to spy on citizens. *Calcalistech* [online]. January 18, 2022 [cit. 2022-03-05]. Available at: https://www.calcalistech.com/ctech/articles/0,7340,L-3927410,00.html

GANON, Tomer, 2022b. Move over NSO: Israeli police is paying private hackers to spy on citizens. *Calcalistech* [online]. January 19, 2022 [cit. 2022-03-05]. Available at: https://www.calcalistech.com/ctech/articles/0,7340,L-3927495,00.html

GANOR, Boaz, 2015. *Global Alert: The Rationality of Modern Islamist Terrorism and the Challenge to the Liberal Democratic World.* New York: Columbia University Press, 218 p. ISBN 0-231-53891-X.

GANOR, Boaz, 2019. Artificial or Human: A New Era of Counterterrorism Intelligence? *Studies in Conflict & Terrorism.* Vol. 44, no.7, 20 p. DOI: 10.1080/1057610X.2019.1568815.

GARCIA-BEDOYA, Olmer, Oscar GRANADOS and José CARDOZO BURGOS, 2020. AI against money laundering networks: the Colombian case. *Journal of money laundering control*. Vol. 24, no. 1, p. 49-62. DOI: 10.1108/JMLC-04-2020-0033.

GOHAR, Faryal, Wasi Haider BUTT, and Usman QAMAR, 2014. *Terrorist group prediction using data classification*. International Conference on Artificial Intelligence and Pattern Recognition. Kuala Lumpur, Malaysia, p. 199–208. ISBN: 978-1-941968-02-4.

GRASSIANI, Erella, 2017. Commercialised occupation skills: Israeli security experience as an international brand. *Security/Mobility: Politics of Movement*. Manchester University Press, 2017. DOI:10.7228/manchester/9781526107459.003.0004

GREEN, Ben and Lily HU, 2018. *The Myth in the Methodology: Towards a Recontextualization of Fairness in Machine Learning* [online]. Presented at the Machine Learning: The Debates workshop at the 35th International Conference on Machine Learning, Stockholm, Sweden. 5p, [cit. 2022-03-05]. Available at: https://scholar.harvard.edu/files/bgreen/files/18-icmldebates.pdf

GOODFELLOW, Ian, Patrick MCDANIEL, and Nicolas PAPERNOT, 2018. Making Machine Learning Robust Against Adversarial Inputs. *Communications of the ACM* . Vol. 61, no. 7, p. 56-66. DOI: 10.1145/3134599

GUETTA, Nitzan et al., 2021. Dodging Attack Using Carefully Crafted Natural Makeup. a*rXiv preprint*. arxiv: 2109.06467.

GUO, Yunhui et al., 2019. SpotTune: Transfer Learning through Adaptive Fine-tuning. a*rXiv preprint*. arxiv: 1811.08737.

G. WATERS, Benjamin, 2019. An international right to privacy: Israeli intelligence collection in the Occupied Palestinian Territories. *Georgetown Journal of International Law*. Vol. 50, no. 2, p. 573-597.

Haaretz, 2016. Israeli Startup Claims to Spot Terrorists With Facial Recognition. *Haaretz* [online]. May 26, 2016 [cit. 2021-12-09]. Available at: https://www.haaretz.com/israel-news/business/israeli-startup-claims-to-spot-terrorists-1.5387498

HAAS, Michael Carl and Sophie-Charlotte FISCHER, 2017. The evolution of targeted killing practices: Autonomous weapons, future conflict, and the international order. *Contemporary Security Policy.* Vol. 38, no. 2, p. 281-306. DOI: 10.1080/13523260.2017.1336407

HASISI, Badi, Simon PERRY and WOLFOWITZ Michael, 2019. Counter-Terrorism Effectiveness and Human Rights in Israel. p. 410-429. IN: SHOR Eran, Stephen HOADLEY (eds.) *International Human Rights and Counter-Terrorism.* Springer Nature Singapore Pte Ltd., 2019, 522 p. DOI: 10.1007/978-981-10-4181-5_22

HASISI, Badi, Yoram MARGALIOTH and Liav ORGAD, 2012. Ethnic Profiling In Airport Screening: Lessons From Israel, 1968–2010. *American Law and Economics Review*. Vol. 14, No. 2, Fall 2012, p. 517–560. DOI:10.1093/aler/ahs009

HAREL, Amos, 2017. Israel Arrested 400 Palestinians Suspected of Planning Attacks After Monitoring Social Networks. *Haaretz* [online]. April 18, 2017 [cit. 2021-12-09]. Available at: https://www.haaretz.com/israel-news/how-israel-uses-big-data-to-fight-palestinian-terror-1.5461381

HENDRIX, Steve, 2020. Israel's Netanyahu turns to anti-terrorism tools in battle against coronavirus. *The Washington Post* [online]. March 15, 2020 [cit. 2021-12-09]. Available at: https://www.washingtonpost.com/world/middle_east/israel-turns-to-anti-terrorism-tools-in-battle-against-coronavirus/2020/03/15/3670bd94-66b9-11ea-b199-3a9799c54512_story.html

HENNESSEY, Zachy, 2022. Israel's critical role in the future of AI. *The Jerusalem Post* [online]. February 8, 2022 [cit. 2022-02-13]. Available at: https://www.jpost.com/business-and-innovation/article-695865

HUSZTI-ORBÁN, Krisztina and Fionnuala NÍ AOLÁÍN, 2020. Use of Biometric Data to Identify Terrorists: Best Practice or Risky Business? *United Nations Human Rights Office* [online]. Regents of the University of Minnesota, Human Rights Center, 45 p, 2020 [cit. 2021-12-09]. Available at: https://www.ohchr.org/Documents/Issues/Terrorism/Use-Biometric-Data-Report.pdf

HRW, 2021. Spyware Used to Hack Palestinian Rights Defenders. *Human Rights Watch* [online]. November 8th, 2021 [cit. 2022-05-03]. Available at: https://www.hrw.org/news/2021/11/08/spyware-used-hack-palestinian-rights-defenders

IBM Corporation, 2021. Explainable AI. *IBM Corporation* [online]. [cit. 2021-10-10]. Available at: https://www.ibm.com/watson/explainable-ai

IntuView, 2018a. DOCEX for Homeland Security. *IntuView* [online]. [cit. 2021-12-09]. Available at: https://www.intuview.com/homelandsecurity

IntuView, 2018b. Social Media Monitor. *IntuView* [online]. [cit. 2021-12-09]. Available at: https://www.intuview.com/socialmediamonitor

IntuView, 2018c. IED Recipe Detector. *IntuView* [online]. [cit. 2021-12-09]. Available at: https://www.intuview.com/ied

ISA, 2020. Information Technology. *The Israeli Security agency* [online]. [cit. 2021-12-08]. Available at: https://www.shabak.gov.il/english/cybertechnology/Pages/technology.aspx

JEREMY BOB, Yonah, 2020. The future of AI in warfare and counterterrorism. *The Jerusalem post* [online]. January 25, 2020 [cit. 2021-12-09]. Available at: https://www.jpost.com/jpost-tech/the-future-of-ai-in-warfare-and-counterterrorism-615112

JOHNSON, Justin, M. and KHOSHGOFTAAR Taghi M., 2019. Survey on deep learning with class imbalance. *Journal of Big Data*. Vol. 6, no. 27, 54 p. DOI: 10.1186/s40537-019-0192-5.

KALAIARASI, Sanskar, et al., 2019. Using global terrorism database (GTD) and machine learning algorithms to predict terrorism and threat. *International Journal of Engineering and Advanced Technology.* Vol. 9, no. 1, p. 5995–6000. DOI: 10.35940/ijeat.A1768.109119.

KFIR, Isaac, 2019. Israel's approach to counter-terrorism. In: JONES, David Martin, Paul SCHULTE, Carl UNGERER and Michael Lawrence ROWAN SMITH. *Handbook of terrorism and counter terrorism post 9/11*. Northampton: Edward Elgar Publishing, 2019, 447 p. ISBN 9781786438027.

KAUR Armaan, Jaspal KAUR SAINI, Divya BANSAL, 2019. Detecting Radical Text over Online Media using Deep Learning. *arXiv preprint.* arXiv:1907.12368.

KEILAF, Omer, 2020. An Oasis Of Mobility Innovation: The Origins Of Israel's Silicon Wadi. *Forbes* [online]. July 3, 2020 [cit. 2021-12-09]. Available at: https://www.forbes.com/sites/forbestechcouncil/2020/07/03/an-oasis-of-mobility-innovation-the-origins-of-israels-silicon-wadi/?sh=c2d17c23a0fa

Knesset, 2017. Shin Bet chief to Foreign Affairs and Defense Committee: ”Relative calm is misleading; Hamas working relentlessly to carry out terror attacks”. *The Knesset* [online]. December 24, 2017 [cit. 2021-12-09]. Available at: https://m.knesset.gov.il/en/News/PressReleases/Pages/Pr13696_pg.aspx

LENNON, Genevieve, 2015. Precautionary tales: Suspicionless counterterrorism stop and search. *Criminology & Criminal Justice.* Vol. 15, no. 1, p. 44–62. DOI: 10.1177/1748895813509637

LEPRI, Bruno, 2018. Fair, Transparent, and Accountable Algorithmic Decision-making Processes. *Philos. Technol*. Vol. 31, p. 611–627. DOI: 10.1007/s13347-017-0279-x

LI, Jinfeng et al., 2018. TEXTBUGGER: Generating Adversarial Text Against Real-world Applications. 15 p. A*rXiv preprint*. arxiv : 1812.05271v1.

LYON, David, 2007. National ID Cards: Crime-Control, Citizenship and Social Sorting. *Policing*. Vol. 1, no. 1, p. 111–118. DOI:10.1093/policing/pam015

M., 2018. Angels in the Skies of Berlin—New Intelligence Questions in a Data-Saturated World. IN: KUPERWASSER, Yossi and David SIMAN-TOV. *Intelligence in Theory and Practice*. Israel Intelligence Community Commemoration and Heritage Center Institute for the research of the methodology of intelligence., Effi Melzer Ltd., No. 3, October 2018, 172p.

MCCULLOCH, Jade and Dean WILSON, 2016. *Pre-crime: Pre-emption, Precaution and the Future.* London: Routledge. ISBN: 978-1-138-78169-6.

Microsoft, 2018. Six principles to guide Microsoft's facial recognition work. *Microsoft* [online]. December 17, 2018 [cit. 2022-02-20]. Available at: https://blogs.microsoft.com/wp-content/uploads/prod/sites/5/2018/12/MSFT-Principles-on-Facial-Recognition.pdf

Middle East Eye, 2021. Meet Blue Wolf, the app Israel uses to spy on Palestinians in the occupied West Bank. *Middle East Eye* [online]. November 9, 2021 [cit. 2021-12-09]. Available at: https://www.middleeasteye.net/news/israel-whats-blue-wolf-app-soldiers-use-photograph-palestinians

MILLER, Kevin, 2014. Total Surveillance, Big Data, and Predictive Crime Technology: Privacy's Perfect Storm. *Journal of Technology Law & Policy*. Vol. 19, no. 1, June 2014, p. 105-146.

MILLER, Tim, 2018. Explanation in artificial intelligence: Insights from the social sciences. p. 1-38. *arXiv preprint arXiv. arxiv:1706.07269.*

MOECKLI, Daniel, James THURMAN, 2011. An evaluation of the use of data mining in counter-terrorism. *Zurich Open Repository and Archive* [online]. University of Birmingham. United Kingdom, 2021, 39 p. [cit. 2021-12-09]. DOI: 10.5167/uzh-51611.

MOHLER George et al., 2015. Randomized Controlled Field Trials of Predictive Policing. *Journal of the American Statistical Association*. Vol. 110, no. 512, p.1399-1411. DOI: 10.1080/01621459.2015.1077710

MUNK, Timme Bisgaard, 2017. 100,000 false positives for every real terrorist: Why anti-terror algorithms don't work. *First Monday.* Vol. 22, no. 9. DOI: 10.5210/fm.v22i9.7126

National Research Council, 2008. *Protecting Individual Privacy in the Struggle Against Terrorists: A Framework for Program Assessment*. Washington, DC: The National Academies Press, 376 p. DOI: 10.17226/12452.

NSO, 2022. NSO GROUP - Cyber intelligence for global security and stability. *NSO* [online]. [cit. 2022-03-21]. Available at: https://www.nsogroup.com/

OECD, 2021. The OECD Artificial Intelligence (AI) Principles. *OECD* [online]. OECD, 2021 [cit. 2021-12-15]. Available at: https://oecd.ai/en/ai-principles

O'DONNELL, Renata, 2019. Challenging racist predictive policing algorithms under the equal protection clause. *New York University law review (1950).* New York: UNIV SCHOOL LAW, Vol. 94 No. 3, p. 544-580. ISSN 0028-7881.

OLESKSER, Ronnie, 2014. Law-making and the securitization of the Jewish identity in Israel. *Ethnopolitics*. Vol. 13, no. 2, p. 105–121. DOI: 10.1080/17449057.2013.773156

Oosto, 2020. Ethics Update. *Oosto* [online]. March 20, 2020 [cit. 2022-03-05]. Avaiable at: https://oosto.com/press/ethics-update/

PAPERNOT, Nicolas et al. 2015. The Limitations of Deep Learning in Adversarial Settings. p.1-16. *arXiv preprint arXiv. arxiv:* 1511.07528v1

PEARSALL, Beth, 2010. Predictive policing: The future of law enforcement. *National Institute of Justice Journal*. Vol. 266, no.1, p. 16–19.

PERRY, Walter L. et al., 2013. *Predictive Policing: The Role of Crime Forecasting in Law Enforcement Operations*. Santa Monica, CA: RAND Corporation, 2013, 186p. DOI: 10.7249/RR233

PYTHON, Andre et al., 2021. Predicting non-state terrorism worldwide. *Science Advances.* Vol. 7, no. 31, 13p. DOI: 10.1126/sciadv.abg4778

RAPAPORT, Amir, 2015. Quite a few Terrorists lost their lives owing to Big Data. *Israel Defense* [online]. January 3, 2015 [2021-12-09]. Available at: https://www.israeldefense.co.il/en/content/quite-few-terrorists-lost-their-lives-owing-big-data

RING, Tim, 2016. Privacy in peril: is facial recognition going too far too fast? *Biometric Technology Today*. Vol. 2016, No. 7–8, p. 7-11. DOI: 10.1016/S0969-4765(16)30123-0

ROCHA-SALAZAR, José-de-Jesús, SEGOVIA-VARGAS, María-Jesús, CAMACHO-MINANO María-del-Mar, 2021. Money laundering and terrorism financing detection using neural networks and an abnormality indicator. *Expert systems with applications*. Vol. 169, no.4, 15 p. DOI: 10.1016/j.eswa.2020.114470

REDDY MEKALA Rohan, Sai YERRAMREDDY, Adam PORTER, 2021. Metamorphic Adversarial Detection Pipeline for Face Recognition Systems. AAAI-22 AdvML Workshop LongPaper. *Openreview* [online]. November 22, 2021 [cit. 2022-02-12], 8p. Available at: https://openreview.net/forum?id=defs2E_Mrf0

ROTH, Emma, 2021. The Israeli army is using facial recognition to track Palestinians, former soldiers reveal. *The Verge* [online]. November 8, 2021 [cit. 2021-12-15]. Available at: https://www.theverge.com/2021/11/8/22769933/israeli-army-facial-recognition-palestinians-track

RUMSFELD, Donald, 2010. Known and Unknown: Author's Note. *Rumsfeld* [online]. December 2010 [cit. 2021-12-15]. Available at: https://www.rumsfeld.com/about/default.asp?name=authors-note

SHARIF, Mahmood et al., 2016. Accessorize to a Crime: Real and Stealthy Attacks on State-of-the-Art Face Recognition. CCS '16: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, October 24-28, 2016, Vienna, Austria, p. 1528-1540. DOI: 10.1145/2976749.2978392.

SHEEHEY, Bonnie, 2018. Algorithmic paranoia: the temporal governmentality of predictive policing. *Ethics and Information Technology*. Vol. 1, no. 1, p. 1-10. DOI: 10.1007/s10676-018-9489-x.

SOLOMON, Shosanna, 2020a. Bias claims spur debate on need for curbs, oversight of facial recognition tech. *The Times of Israel* [online]. July 2, 2020 [cit. 2021-12-15]. Available at: https://www.timesofisrael.com/bias-claims-spur-debate-on-need-for-curbs-oversight-of-facial-recognition-tech/

SOLOMON, Shosanna, 2020b. AI a new and ‘frightening’ battlefield in cyber war, experts warn. *The Times of Israel* [online]. February 16, 2020 [cit. 2021-12-15]. Available at: https://www.timesofisrael.com/ai-a-new-and-frightening-battlefield-in-cyber-war-experts-warn/

SOLON, Olivia, 2019. Why did Microsoft fund an Israeli firm that surveils West Bank Palestinians? *NBC News* [online]. October 28, 2019 [cit. 2021-12-15]. Available at: https://www.nbcnews.com/news/all/why-did-microsoft-fund-israeli-firm-surveils-west-bank-palestinians-n1072116

SPEKTOR, Michelle, 2020. Imagining the Biometric Future: Debates Over National Biometric Identification in Israel. *Science as culture*. Vol. 29, no. 1, p. 100-126. DOI: 10.1080/09505431.2019.1667969.

STRIKWERDA, Litska, 2021. Predictive policing: The risks associated with risk assessment. *The Police Journal*. Vol. 94, no. 3, p. 422–436. DOI: 10.1177/0032258X20947749.

TAWIL-SOURI, Helga, 2012. Uneven Borders, Coloured (Im)mobilities: ID Cards in Palestine/Israel. *Geopolitics*. Vol. 17, no. 1, p. 153-176. DOI: 10.1080/14650045.2011.562944.

The Guardian, 2014. Israel's Unit 8200 refuseniks: 'you can't run from responsibility'. *The Guardian* [online]. September 12, 2014 [cit. 2021-12-09]. Available at: https://www.theguardian.com/world/2014/sep/12/israel-unit-8200-refuseniks-transcript-interview

The Meir Amit Intelligence and Terrorism Information Center (2022). Palestinian Terrorism, 2021: Summary, Types and Trends. *The Meir Amit Intelligence and Terrorism Information Center* [online] January 2022 [cit. 2022-03-29], 82p. Available at: https://www.terrorism-info.org.il/app/uploads/2022/02/E_186_21.pdf

The White House, 2002. President Bush Delivers Graduation Speech at West Point. *The White House* [online]. June 1, 2002 [cit. 2021-12-09] Available at: https://georgewbush-whitehouse.archives.gov/news/releases/2002/06/20020601-3.html

Times of Israel, 2018. Shin Bet thwarted 500 terror attacks in 2018, Netanyahu says. *The Times of Israel* [online]. December 5, 2018 [cit. 2021-12-15]. Available at: https://www.timesofisrael.com/shin-bet-thwarted-500-terror-attacks-in-2018-pm/

Times of Israel, 2021. UN rights chief slams 'arbitrary' Israel decision to outlaw Palestinian groups. *The Times of Israel* [online]. October 26, 2021 [cit. 2021-11-13]. Available at: https://www.timesofisrael.com/france-calls-for-clarifications-after-israels-terror-listing-of-rights-groups/

KRAFFT, Peter M. et al., 2020. Defining AI In Policy Versus Practice. a*rXiv preprint*. arXiv:1912.11095v1.

UNICRI and UNOCT 2021a. Algorithms and terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes. *United Nations* [online]. A Joint Report by UNICRI and UNCCT, [cit. 2021-11-13]. Available at: https://www.un.org/counterterrorism/sites/www.un.org.counterterrorism/files/malicious-use-of-ai-uncct-unicri-report-hd.pdf

UNICRI and UNOCT 2021b. Countering Terrorism Online with Artificial Intelligence. An Overview for Law Enforcement and Counter-Terrorism Agencies in South Asia and South-East Asia. *United Nations* [online]. A Joint Report by UNICRI and UNCCT, [cit. 2021-11-13]. Available at: http://unicri.it/sites/default/files/2021-06/Countering%20Terrorism%20Online%20with%20AI%20-%20UNCCT-UNICRI%20Report.pdf

United Nations, 2018. Secretary-general's Strategy on New Technologies. *United Nations* [online]. September 2018 [cit. 2021-11-13]. Available at:

https://www.un.org/en/newtechnologies/images/pdf/SGs-Strategy-on-New-Technologies.pdf

VERHELST, Hugo, Alexander STANNAT and Giulio MECACCI, 2020. Machine Learning Against Terrorism: How Big Data Collection and Analysis Influences the Privacy-Security Dilemma. *Science and Engineering Ethics*. Vol. 26, p. 2975–2984. DOI: 10.1007/s11948-020-00254-w

VERMA, Chaman, Sarika MALHOTRA, and Vineeta VERMA, 2018. Predictive modeling of terrorist attacks using machine learning. *International Journal of Pure and Applied Mathematics.* Vol. 119, no. 15, p. 49-61. ISSN: 1314-3395.

VOGIATZOGLOU, Plixavra, 2019. Mass surveillance, predictive policing and the implementation of the CJEU and ECtHR requirement of objectivity. *European Journal of Law and Technology*. Vol. 10, no.1, 20 p.

WEITZBERG, Keren, 2021. Biometrics and counter-terrorism. Case study of Israel/Palestine. *Privacynational* [online]. [cit. cit. 2021-11-13]. Available at: https://privacyinternational.org/sites/default/files/2021-06/PI%20Counterterrorism%20and%20Biometrics%20Report%20Israel_Palestine%20v7.pdf

ZHANG, Xun et al., 2018. On the risk assessment of terrorist attacks coupled with multi-source factors. *ISPRS International Journal of Geo-Information*. Vol. 7, no. 9. DOI: 10.3390/ijgi7090354.

ZHENG, Yu-jun et al., 2017. Airline Passenger Profiling Based on Fuzzy Deep Machine Learning. *IEEE transaction on neural networks and learning systems*. PISCATAWAY: IEEE, 2017, Vol. 28, no. 12, p. 2911-2923. DOI: 10.1109/TNNLS.2016.2609437

ZEDNER, Lucia, 2007. Pre-Crime and Post-Criminology? *Theoretical Criminology*. Vol. 11, no. 2, p. 261–281. DOI: 10.1177/1362480607075851.

ZUSSMAN, Asaf, 2013. Ethnic Discrimination: Lessons from the Israeli Online Market for Used Cars. *The Economic journal (London)*. HOBOKEN: Blackwell Publishing, 2013, Vol. 123, no. 572, F433-F468. DOI: 10.1111/ecoj.12059.