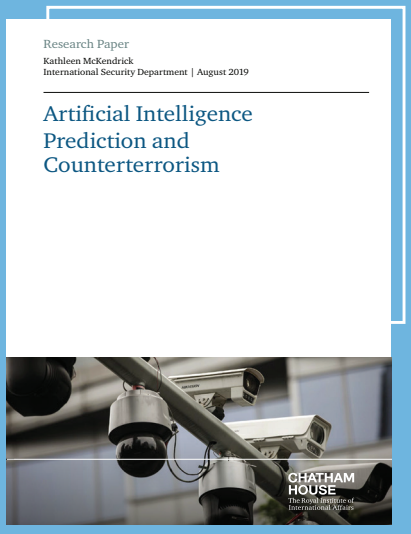
التحالف الإسلامي العسكري لمحاربة الإرهاب
ISLAMIC MILITARY COUNTER TERRORISM COALITION
General Directorate of Planning and Coordination

Spotlight On

9
Special Reports
Jan. 2020

Artificial Intelligence and Counterterrorism

Limitations, Opportunities and Risks




التحالف الإسلامي العسكري لمحاربة الإرهاب
ISLAMIC MILITARY COUNTER TERRORISM COALITION

9

Special Reports

Jan. 2020

Spotlight On

Artificial Intelligence and Counterterrorism

Limitations, Opportunities and Risks

Artificial Intelligence (AI) refers to the ability of digital machines and computers to carry out specific tasks similar to what smart organisms do, such as thinking and planning, learning and creating, adapting and interacting, improving procedures, extracting knowledge and forecasting large and varied digital data along with other operations that require precise mental processes.

Counterterrorism experts explain that there are two methods to prevent terrorist attacks: the first method is to protect infrastructure and individuals and implement security controls; the other method is to deprive terrorists of the ability to launch attacks, by arresting them before carrying out their plans, and combating extremism and terrorist recruitment.

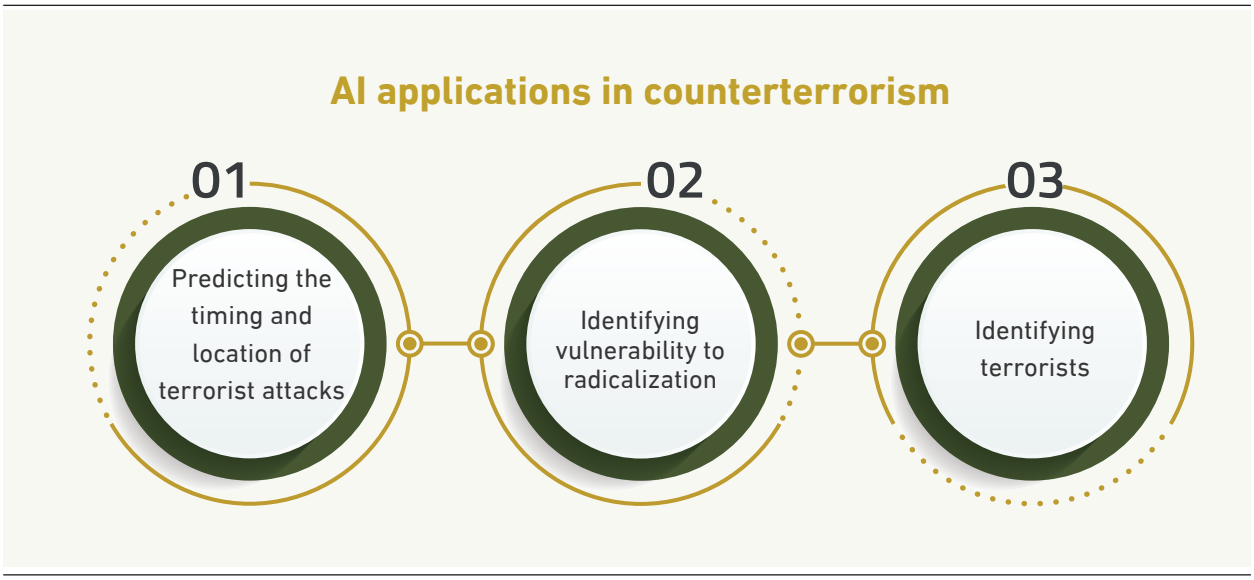
The report "Artificial Intelligence and Counterterrorism", which was produced by Kathleen McKendrick and published by the Chatham House Institute in August 2019, discusses the multidimensional relationship between terrorism and AI, highlighting the possible opportunities and potential risks.

When re-examining and waking up to underlying reality, a valid question may spring to mind about the relationship between AI and the fight against terrorism. The answer lies in one word: "prediction", which is one of the most important uses of AI. Simply put, prediction triggered by AI contributes to preventing terrorist attacks by providing physical protection to the infrastructure, and improving the allocation and provision of resources to potential sites targeted by terrorists. AI-assisted prediction also helps prevent terrorists from launching attacks, by arresting them before such potential attacks are carried out. More importantly, it also helps to combat terrorist recruitment, and helps to identify the individuals vulnerable to extremism or recruitment by terrorist organizations.

Uses of AI in counterterrorism centre on generating accurate predictions that help direct resources for countering terrorism more effectively. Predictive AI might also minimize unnecessary intrusion on the majority of the population and mitigate human bias in decision-making.

In this regard, an effective counterterrorism aims to provide security for the majority of citizens, with minimal infringement of rights and freedoms. Prediction allows discretion in the application of preventive measures, minimizing the effect on the population. Effective prediction might, for example, have the effect that only violent terrorists are met with coercive force or restrictions, while conciliatory measures are directed towards individuals vulnerable to radicalization. Thus, the prediction provided by AI is a means to improve the allocation of resources in the fight against terrorism most likely to be targets in this area.

More importantly, the uses of AI in counterterrorism may generate accurate predictions that help direct resources for countering terrorism more effectively and may also minimize unnecessary intrusion on the majority of the population and thus mitigate human bias in decision-making, by carefully drawing attention to the areas or individuals most vulnerable to threats, and reducing the number of citizens who are subject to more surveillance. On the flip side of AI, the lack of adequate safeguards on the use of AI, and on the huge reservoirs of data on which it depends, could lead not only to its misuse by autocratic governments to impose control over their citizens, but also to excessive infringement on rights such as of privacy and freedom of expression. The approach to using accurate prediction techniques of AI in counterterrorism can be subject to conflicting goals; however, this does not preclude looking at the possibilities and costs of such an approach, and how to organize and regulate this nascent field.



Counterterrorism Applications

The predictive capabilities of AI have become widely recognized in the domain of counterterrorism, but its application is still on a small scale. Security and intelligence services use automatic data analysis to assess the risks of air travelers and to reveal the links between terrorist organizations and their members. The police also use security and intelligence services to analyze criminal gang networks, while some technology companies use advanced predictive measures to monitor and disable terrorist activity on social media platforms, and AI is used in the financial services sector to report any suspicious activity. Here are some examples of AI predicative ability to fight terrorism:

the lack of adequate safeguards on the use of AI, and on the huge reservoirs of data on which it depends, could lead not only to its misuse by autocratic governments to impose control over their citizens, but also to excessive infringement on rights such as of privacy and freedom of expression.

1-Timing and Location of Terrorist Attacks

Artificial intelligence can be used to predict terrorist operations based on communication data, financial transaction information, travel patterns and internet surfing activity. Models have been developed that predict the location and timing of terrorist attacks. In 2015, for instance, a technology startup claimed that its predictive model was able to predict suicide attacks with an accuracy of 72%. Some other models have also relied on open source data for individuals who use social media and applications on their mobile phones, including a preemptive event recognition system that incorporates the results of different separate predictive models to predict specific events. It is critically important to understand that more data does not necessarily mean that the quality of the prediction is improved; rather, it must be validated.

2-Vulnerability to Radicalization

Some technology companies have developed tools to assess vulnerability to violent extremist ideologies. One of the technology companies launched a project (Redirect Method), which targets users of video-sharing sites who may be susceptible to propaganda from terrorist groups and redirects them to videos espousing a credible counter-narrative.

3-Identifying Terrorists

Some leaked information of the US National Security Agency (SKYNET) program shows that an AI-based algorithm was used to analyze the metadata of about 55 million mobile phone users in Pakistan in 2007; the result was that approximately 15,000 individuals could become terrorists out of a population of 200 million at the time. Although the model used was not effective per se, it revealed the predictive value of the data when identifying close links with terrorism. These examples of using predictive AI capabilities

in counterterrorism are still valid possibilities, and AI is not expected to provide immediate, comprehensive and accurate answers to complex questions. However, the ability to develop AI tools for this purpose remains with those who have the ability to access or protect data. With improved AI performance, there will be more room to extract accurate predictions of terrorism in the future and increase its associated uses in counterterrorism.

Challenges

There are two challenges related to the use of AI in counterterrorism: the first one is related to human rights problems, and the second one is related to the practical implications.

First: Human Rights Problems

1. Lack of well-established norms for the use of AI technology

There is no agreed international position on the limits of the use of AI, and this puts the rights and freedoms of citizens at stake, so the need for adequate safeguards for the use of AI by governments and security services increases, and a review of the measures established to protect the privacy and fundamental freedoms of citizens.

In December 2014, the United Nations General Assembly passed Resolution 68/167 on the right to privacy in the digital age that states shall have the responsibility to ensure that their activities comply with international law. In December 2016, the European Court issued a ruling against Britain for its failure to fulfill the basic rights guaranteed by the European Union in data retention practices.

2. Randomness and Disproportionality of Data Collection

Many practices relevant to the use of AI predictive technology in counterterrorism depend on its application to the population as a whole, making it random and disproportionate, as well as violating the privacy of the general public. Again, legislative frameworks and governance principles related to the use of data tend to take care of data access powers, and often ignore the regulation of how that data is used or protected from misuse.

3. An expanding but weakly regulated private sector role

The responsibility for AI predictive uses to counterterrorism is spread among a wide range of actors, including technology companies, some of which have turned public information into a commodity that they can sell to whomever pays! Law enforcement agencies may be able to obtain such data legally, which means that corporate obligations to protect customer privacy may be in doubt.

4. Lack of Transparency

AI methods are used to combat and counter terrorism, and it is difficult to obtain legal safeguards of transparency in the fight against terrorism.

Second: Practical Concerns

AI methods are used to combat and counter terrorism, and it is difficult to obtain legal safeguards of transparency in the fight against terrorism.

1.The inability to achieve adequate prediction

Terrorism has many different paths and manifestations, and this simply means that it is impossible to prepare a final list of definitive indicators of involvement in or exclusion from terrorism. The small number of terrorists from the general population makes access to wide characteristics of terrorists based on profiling no predictive value. The low incidence of terrorism, and the tendency for terrorist means to develop rapidly, makes it difficult to build good predictive models.

However, research has shown that it may be possible to use AI to analyze communication and distinctive characteristics, such as the degree of extremism, or aggressive intent, and this means that the ability to predict terrorist involvement is no longer impossible at this time. We can improve the accuracy of predictive models based on one source of data by combining results with other summaries, but restricting access to data may limit the efficiency of AI in achieving these tasks.

The challenges of using AI in counterterrorism

2. Lack of common validation standards

Important data access per se alone does not guarantee success in building accurate predicative models. Again, validation and testing are necessary to measure predictive accuracy and assess the proportionality of their use. Given the wide range of actors involved, common standards are of great importance. The uptake of techniques used by intelligence agencies in law enforcement and the use of AI by the private sector simply mean that actors are increasingly involved in the development and use of predictive AI for counterterrorism. A group of recent research has focused on the extent of AI susceptibility, and confirmed that predictive results of AI may be more scientifically objective than human assessments that are affected by cultural bias.

Opportunities and Risks

The challenges mentioned above stem from the unregulated development of AI as a means of predicting terrorism. Is it possible then to use the AI predictive power in a legitimate way as a tool used by the state to identify and detect terrorism?

This simply means providing government agencies with broader access to public data, with a clearer regulation of how such data is used. This approach requires the collection and analysis of citizen data, which is unacceptable in many countries.

First: Opportunities

1. Reducing infringement and increasing efficiency

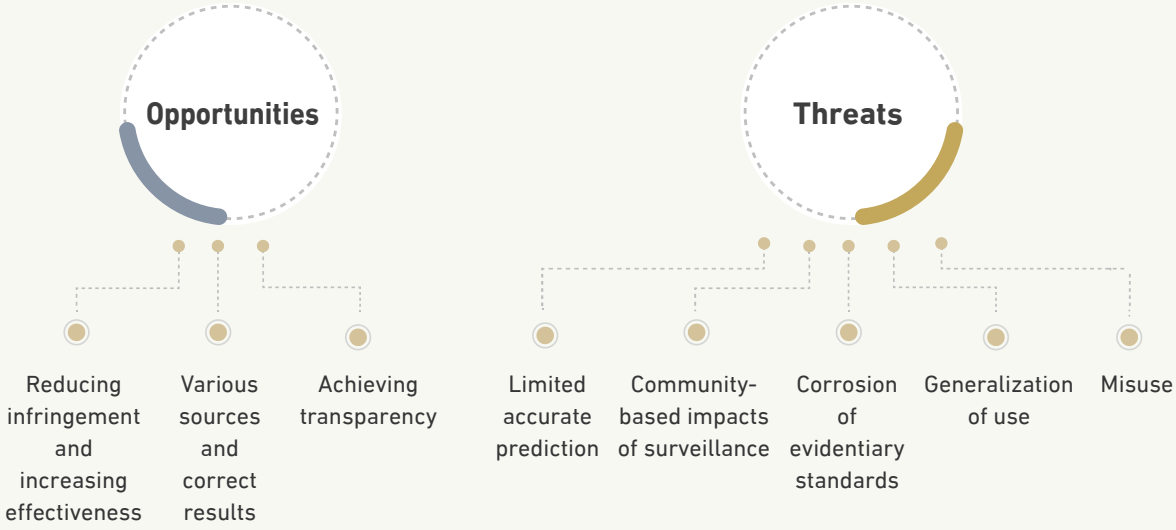
The process of automated data analysis is marred by quite a few infringements, but it nonetheless contributes to reducing the infringements of privacy of citizens compared to human analysis. When using a predictive model to automate population data analysis, it may be necessary to require validation, i.e. providing a causal relationship with a specific objective, with no alternative options, while proving that the benefits outweigh the costs associated with the infringements of rights.

Telecommunications service providers may be more willing to meet requests for access to their data if they realize that the standards for these requests are high and that they relate to a legitimate aim and achieve good results.

2. Numerous sources and correct results

AI can develop separate sources for gathering information and validating the results, rather than using a centralized system that includes all individuals, whether they are terrorists or vulnerable to extremism. Although validation is necessary, there is still room for improvement, and the use of automated methods provides an

The opportunities and threats for using al in counterterrorism



opportunity to assess model performance and measure its feasibility.

There is no doubt that the treatment of individuals vulnerable to extremism and real terrorists within the system itself causes disruption that prevents maintaining real discrimination between them except through predictive systems based on the analysis of data digitally created and preserved, which are neutral, reliable and effective data that can be used to direct early non-coercive interventions, such as interventions that prevent youth from being drawn to and tempted by violent extremist ideology.

An increase in the availability of data may lead to corrosion of existing evidentiary standards. The capabilities offered by the ability to collect and use data for prediction will result in greater interest in preventive and proactive operations and punish the accused before the crime occurs!

3. Achieving transparency

Quantitative standards provide a measure of technical transparency, and the exchange of quantitative standards can increase confidence and create conditions for a better exchange of information among actors and international agencies. It is also important to enhance the transparency of the legal framework surrounding predictive analysis to better improve model performance, and to have a clear, consistent and sound understanding of what the analysis is and what the cause for it is.

Centralization of the responsibility for overseeing this analysis and working with government agencies improves the ability to achieve a clear understanding of the analysis; making this information available improves oversight and addresses grievances related to discrimination. For example: airport arrest procedures are developed according to a model that combines suspicious travel patterns, unusual drivers, and physiological signs of discomfort. This has reduced the total number of people subject to searches by 50%. When travelers are aware of this fact, their fears will be reduced that these measures may be based on racial or ethnic bias. The shared quantitative standards can create the best conditions for information exchange between actors and international agencies. For example, telecommunications service providers may be more willing to meet requests for access to their data if they realize that the standards for

these requests are high and that they relate to a legitimate aim and achieve good results.

Second: Risks

1. Limitation of accurate prediction

Predicting terrorist attacks and being vulnerable to extremism using mathematical methods alone may prove to be inaccurate. Therefore, methods must be proposed, tried and assessed first, and dispensed with if they do not achieve sufficient predictive value. AI is successfully used as showcased in industry by improving prediction capabilities using machine learning and dynamic adaptation, including adapting to changes in data within the system, imposing pre-use validation standards, and progressive updating. These measures provide the precondition for the coordinated use of AI in countering terrorism, while bearing in mind that the results of any predictive system provide only a possibility, not evidence.

2. Community impact of surveillance

Automated data analysis often reduces infringement into privacy on an individual level, but increases the likelihood that everyone will feel monitored at all times. A number of academics have studied the way in which digital surveillance can cause widespread fear of dealing with political issues or activities, expressing a dissenting opinion, criticizing ideas, or disagreeing with prevailing standards.

This issue was a source of concern for civil society groups, especially after the disclosure of the details of the spy program (PRISM) in 2013 by Edward Joseph Snowden, a programmer for the CIA, who unveiled and divulged the size of national surveillance programs in America. The historical experience of citizen behavior under authoritarian regimes provides compelling evidence of spine-chilling findings of mass surveillance. On the other hand, there are those who believe that the continued political opposition, freedom of expression and the increase in the voluntary exchange of information via social media, provide evidence of exaggerating the negative effects on such surveillance.

3. Corrosion of the evidentiary standards

There are warnings that an increase in the availability of data may lead to corrosion of existing evidentiary standards. The capabilities offered by the ability to collect and use data for prediction will result in greater interest in preventive and proactive operations, than in the past. These new predictive techniques will enable the authorities to punish the accused before the crime occurs! Here comes the pre-crime ghost, which can only be avoided by recognizing that any predictive model is a possibility and not a guide; again, intervention according to such models should ultimately be based on a human decision made by an expert on the limits and capabilities of such models.

The use of AI capabilities for predictive purposes in counterterrorism is neither inherently good nor bad; rather, it critically depends on the way in which these AI predicative capabilities are used.

4. Generalizing the use of AI in crimes

Precedents in which AI systems used to counter terrorism can be observed and identified in fighting other crimes, which are what the New York City Police did when they combined 3000 closed-circuit television cameras with other sensors. As a matter of fact, the availability of a large body of public data is an invaluable treasure for those involved in investigating criminals or interested in social interventions.

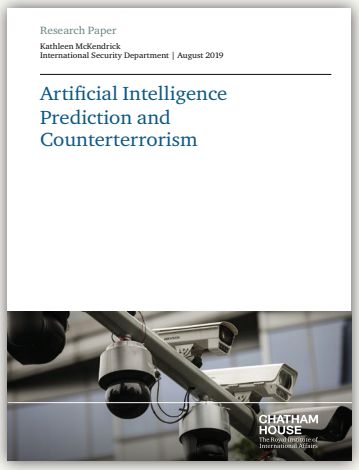
5. Abuse of AI

Even with regulating the use of AI to combat terrorism and imposing restrictions on such use, the interpretation of its outputs remains of a personal nature, which raises fears that the term "terrorist" may be used to suit political purposes. Allowing the collection and analysis of public data carries a potential risk of abuse. The advantage of limitations and controls imposed on the use of AI in data analysis are diminished by the lack of a general understanding of the amount of data created and the method of use thereof.

Author

Major Kathleen McKendrick, a British army officer, has served in Iraq and Afghanistan, and has provided education, awareness and counter-terrorism training courses at the Center for Excellence in Defense Against Terrorism (COE-DAT) of the North Atlantic Treaty Organization (NATO) in Ankara, Turkey. In 2017 and 2018, McKendrick worked as a research fellow in the Department of International Security at the Royal Institute of International Affairs (Chatham House).

McKendrick obtained a Bachelor Degree in Air Systems Engineering from Cranfield University, and a Master Degree in International Relations from the London School of Economics and Political Science. Her research areas of interest include defense and security and AI application to military operations.




التحالف الإسلامي العسكري لمحاربة الإرهاب
ISLAMIC MILITARY COUNTER TERRORISM COALITION


التحالف الإسلامي العسكري لمحاربة الإرهاب
ISLAMIC MILITARY COUNTER TERRORISM COALITION
General Directorate of Planning and Coordination