CHAPTER 6

USING ARTIFICIAL INTELLIGENCE TO DISRUPT TERRORIST OPERATIONS

Colonel Christopher J. Lowrance and Dr. C. Anthony Pfaff

Editor's Note

As the introduction to this volume points out, denying terrorists operational success, whether by preventing and disrupting attacks or denying terrorists their effect, is an important part of a counterterrorist deterrent strategy. This chapter examines how artificial intelligence can play a role in such a strategy by disrupting the stages of the terrorist attack cycle. The chapter begins by outlining the typical phases terrorists follow, from ideation and motivation through to execution and post-attack exploitation. Understanding these stages is crucial for effectively applying AI to detect and mitigate threats. By leveraging AI's capabilities, security forces can detect subtle signs of terrorist activities that were previously difficult for humans to identify. But success may come at a cost. As described in the introduction, the logic of terrorism suggests terrorist leaders are less risk-averse when things are going badly than when things are going well. So, the more successful counterterrorists are, the more dangerous terrorists can become. Mitigating this dynamic may lead to a *revolution in counterterror affairs* analogous to the "revolution in terrorist affairs" described in chapter 4. As this chapter recommends, such a revolution might consider revisions to AI architecture to combine multiple AI models or algorithms to collaborate on a specific task and improve decision-making or predictive accuracy.

Introduction

Previous chapters have explored various methods by which terrorists might use AI to facilitate their attacks, with the preceding chapter changing directions and specifically highlighting AI's role in recent counterinsurgency operations. This chapter adopts a broader approach to combating terrorism, focusing on harnessing AI to counteract and disrupt terrorist

[*] **The information and views expressed in this publication are solely those of the author and do not necessarily represent the views and policies of NATO, COE-DAT, NATO member states or institutions with which the author is affiliated.**

operations more generally. We begin by reviewing the systematic stages most terrorist attacks follow; this foundational understanding frames our discussion and allows us to pinpoint how AI can be strategically applied at each stage of the attack cycle to thwart terrorist activities. Moreover, thoroughly comprehending AI capabilities—including those of generative AI—is crucial for devising optimal counterterrorism deployment strategies.

This chapter assumes the reader has a general understanding of AI's capabilities, as described in the introduction. It focuses on how AI's capacities may be applied to counterterrorism. It is worth emphasizing the importance of integrating the proposed approach in this chapter to enhance the capabilities of military operations, law enforcement, and intelligence analysts. Such integration aims to foster a synergistic human-machine collaboration that performs more effectively in combating terrorism. The ensuing sections detail specific ways in which AI can augment human efforts to counter terrorist threats effectively. We acknowledge the data and information proposed for AI's use in disrupting terrorist operations may raise ethical and privacy concerns, which will be addressed in the subsequent chapter.

The Terrorist Attack Cycle

Terrorist operations often adhere to a cyclical pattern from conception to execution.[261] Recognizing and understanding the stages of terrorist operations is crucial for effective intervention. Within each stage, terrorists typically engage in actions that may inadvertently leave indicators of their operations.[262] Historically, these subtle traces were challenging for humans to detect. However, detecting terrorists' actions through intelligent machines has become more feasible in the age of AI. Artificial intelligence excels in analyzing data to detect patterns within information streams.[263] As illustrated in Figure 6-1, AI can analyze and exploit specific data to detect and disrupt attack planning at each point in the terrorist attack cycle. Those cycles are described below:

- **Ideation and Motivation:** This initial phase involves the formation of the ideological motivations and intentions behind terrorist activities. A 2015 study states radicalization occurs in four phases, each with associated behaviors. However, no individual exhibited all associated behaviors at every stage.[264] Thus, discerning whether an individual is becoming radicalized can be difficult even when behaviors are detected. For example, even though neighbors in Manchester, United Kingdom, alerted police when a local youth made statements that supported terrorism and suicide bombing, the neighbors failed to connect the youth to the Da'esh-sponsored radicalization efforts.

261 "JCAT Counterterrorism Guide for Public Safety Personnel," Office of the Director of National Intelligence (website), n.d., accessed on August 6, 2024, https://www.dni.gov/nctc/jcat/index.html; and Marc Sageman, *Understanding Terror Networks* (University of Pennsylvania Press, 2004).

262 Rick Samples, "A Police Officer's Guide to the Terrorist Attack Cycle," Lexipol (website), January 8, 2021, https://www.lexipol.com/resources/blog/a-police-officers-guide-to-the-terrorist-attack-cycle/.

263 Ron Schmelzer and Kathleen Walch, "The Seven Patterns of AI," Project Management Institute (website), n.d., accessed on August 6, 2024, https://www.cognilytica.com/the-seven-patterns-of-ai/.

264 Jytte Klausen, *A Behavioral Study of the Radicalization Trajectories of American "Homegrown" Al Qaeda-Inspired Terrorist Offenders* (US Department of Justice, 2016), i.

The youth would go on to detonate a bomb at a local arena, killing 22 and injuring 116.[265] Therefore, detecting radicalization early may require identifying specific behavior patterns in individuals to understand better which combinations of actions are linked to the radicalization process.

- **Target Selection:** During this stage, potential targets are identified based on their symbolic, economic, or political significance. As described in the previous chapter, AI can analyze repeating factors such as weather patterns; symbolic events; increased activity around cultural, historical, and religious centers; and locations of past activity to predict future attacks.
- **Intelligence Gathering:** During this stage, terrorist operatives collect information pertinent to the selected targets, including security details and operational timelines. Terrorists often conduct surveillance and gather intelligence to prepare for attacks, and these activities frequently take place near the terrorists' homes, which are typically close to their selected targets.[266] As a result, AI could monitor data surrounding potential targets, scanning for suspicious or anomalous behavior. AI could be enough to thwart an attack when combined with other anomalies in other stages. As another form of risk mitigation, AI could assess one's vulnerabilities, especially regarding critical infrastructure, symbolic events, and other high-value targets.
- **Planning and Preparation:** This stage involves the strategic planning and logistical arrangements needed for the attack.[267] Intelligence analysis is a function of assigning probabilities to possible courses of action. Finding a way to improve the accuracy of probability assessments should help avoid similar intelligence failures in the future.
- **Logistics and Deployment:** Resources and personnel are mobilized to execute the planned attack. During this phase, operatives will leave safe areas, collect and assemble weapons and explosives, form teams, and move to the attack site.[268] At this point, terrorists are at their most vulnerable. To the extent AI can enhance counterterrorists' ability to detect, analyze, and intervene during this phase of the cycle, more terrorist attacks can be disrupted.
- **Attack Execution:** The actual implementation of the terrorist act. At this point, what counterterror measures are possible depends on the nature of the attack. Once the shooting starts or the bombs go off, counterterrorists can do little about it. But not all attacks begin with a bang. On September 11, 2001, ten of the 19 terrorists were

265 "Manchester Attack: What We Know So Far," BBC (website), June 12, 2017, https://www.bbc.com/news/uk-england-manchester-40008389.

266 Brent Smith, "A Look at Terrorist Behavior: How They Prepare, Where They Strike," *National Institute of Justice Journal*, no. 260 (July 2008).

267 Oren Liebermann and Tamar Michaelis, "Calls for Accountability Grow over October 7 Failures, but Israel's Leadership Is Unlikely to Act," *CNN* (website), June 20, 2024, https://www.cnn.com/2024/06/20/middleeast/israel-gaza-intelligence-accountability-analysis-intl-cmd/index.html.

268 "Defining the Terrorist Attack Cycle," RANE Worldview (website), February 23, 2012, https://worldview.stratfor.com/article/defining-terrorist-attack-cycle.

selected by the Computer-Assisted Passenger Prescreening System for additional screening. But this selection only meant the terrorists' luggage would not be loaded until they boarded the plane. Some of the terrorists' carry-on luggage also set off alarms as they went through X-ray machines. Still, airport security personnel did not find weapons or anything suspicious.[269] This example suggests the ability to aggregate anomalies better should improve capabilities to intervene once an attack has been initiated.

- **Escape and Exploitation:** At this stage, terrorists will execute post-attack strategies, including escape routes (if applicable) and exploitation. Assuming the terrorist operation is not a suicide attack, terrorist operatives will need a way out of the attack zone—where they would have to confront military, law enforcement, or other first responders—and will need to go back to a safe area. Additionally, terrorist groups will want to take advantage of the attention the attack draws to express and advance their goals.

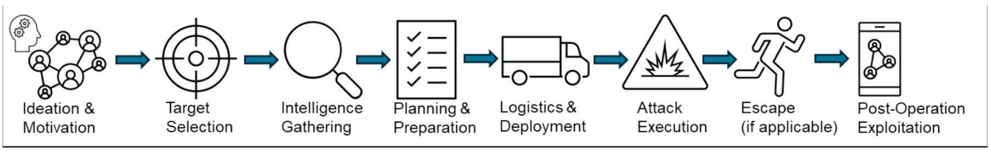

Figure 6-1. General stages of a terrorist attack cycle

This discussion illustrates how counterterrorism strategies can be significantly enhanced by identifying where AI can be interjected effectively within the stages of the terrorist attack cycle.

Artificial Intelligence Applications in Counterterrorism

Using AI to prevent or disrupt terror attacks is not new. As discussed in the previous chapter, in Iraq and Afghanistan, the United States used Project Maven to analyze massive amounts of imagery to enhance human analysts' ability to identify people, vehicles, and installations associated with terrorist activity.[270] In the United States, AI applications have been used to conduct facial recognition, analyze video, monitor social media, map criminal networks better, and conduct predictive policing both to solve crime and to prevent terrorist activity.[271] But what AI applications do not do is aggregate across the range of the terrorist attack cycle. When one analyzes terrorist-related activity across the cycle's range, new

269 National Commission on Terrorist Attacks upon the United States, *The 9/11 Commission Report* (National Commission on Terrorist Attacks upon the United States, 2004), 1–4, 451.

270 Marcus Weisgerber, "The Pentagon's New Artificial Intelligence Is Already Hunting Terrorists," Defense One (website), December 21, 2017, https://www.defenseone.com/technology/2017/12/pentagons-new-artificial-intelligence-already-hunting-terrorists/144742/.

271 Ángel Díaz, "New York City Police Department Surveillance Technology," Brennan Center for Justice (website), October 4, 2019, https://www.brennancenter.org/our-work/research-reports/new-york-city-police-department-surveillance-technology; and Issie Lapowsky, "How the LAPD Uses Data to Predict Crime," *WIRED* (website), May 22, 2018, https://www.wired.com/story/los-angeles-police-department-predictive-policing/.

patterns emerge, potentially improving counterterrorists' ability to disrupt, if not prevent, attack planning.

Detecting, Attributing, and Countering the Terrorists' Narrative

The initial stage outlined in Figure 6-1, "Ideation & Motivation," involves forming ideological motivations and intentions that drive terrorist activities. During this phase, a terrorist or terrorist group seeks to persuade individuals to join their cause or adopt their ideology.[272] For example, Da'esh used a combination of traditional networks, social media platforms like Twitter and YouTube, and messaging applications like Telegram to identify potential recruits and highlight operational successes to legitimize its cause and attract new followers. According to a study by the International Center for the Study of Violent Extremism, Da'esh leaders have used in-person meetings reinforced by social media messaging to reach recruits. Those most vulnerable to joining tend to fit a particular pattern, such as experiencing poverty, unemployment, criminal activity and substance abuse, among other factors. Overall, the study identified 22 factors that, in various combinations, can make individuals more vulnerable to terrorist messaging.[273]

Similarly, during the final stage, 'Post-Operation Exploitation', terrorists pursue analogous goals, often by claiming their actions were successful and disseminating propaganda across various mediums, including social media, to exploit the attention their attacks bring. After the Paris attacks of 2015, Da'esh used social media to claim responsibility and propagate videos and messages that glorified the acts, positioning the attackers as martyrs and amplifying their ideological messages.[274] Essentially, the terrorist attack cycle is bookended by similar processes aimed at influencing others through strategic information campaigns. These campaigns typically focus on recruiting members and garnering support for terrorists' political or ideological causes.

The similarities between the first and last stages of a terrorist operation suggest similar AI tools will effectively detect and disrupt terrorist activities. Terrorists frequently use Internet websites and social media platforms to garner support and disseminate propaganda. Artificial intelligence can automatically scrape data from these sites and leverage generative AI capabilities to interpret text-based data, identifying where terrorists attempt to communicate and disseminate information. Once terrorist sites or social media accounts are identified, they can be flagged for tracking and further analysis. Over time, and with additional data, AI could progress to identifying and attributing the website or social media account to specific individuals or groups. With that information, counterterrorist organizations could collect

[272] Alex P. Schmid, ed., *Handbook of Terrorism Prevention and Preparedness* (International Centre for Counter-Terrorism, July 2021).

[273] Anne Speckhard and Molly D. Ellenberg, "ISIS in Their Own Words: Recruitment History, Motivations for Joining, Travel, Experiences in ISIS, and Disillusionment over Time—Analysis of 220 In-Depth Interviews of ISIS Returnees, Defectors and Prisoners," *Journal of Strategic Security* 13, no. 1 (2020): 93–96.

[274] J. M. Berger and Jonathon Morgan, "The ISIS Twitter Census: Defining and Describing the Population of ISIS Supporters on Twitter," Brookings Institution (website), March 5, 2015, https://www.brookings.edu/articles/the-isis-twitter-census-defining-and-describing-the-population-of-isis-supporters-on-twitter/.

more data, find more patterns, and better map the terrorist network. Artificial intelligence could also find correlates of attack planning, which could be useful in predicting future attacks. Moreover, those correlates may not always be direct. For example, the data company Fraym used vaccination rates as a proxy for state effectiveness to determine areas of Burkina Faso that were vulnerable to terrorist attacks. With sufficient data, Fraym could predict, with a high degree of accuracy, which areas would fall under terrorist control.[275]

To achieve this level of prediction, AI could develop models that recognize the signature behaviors or trademarks of a group based on its language, grammatical style, message content, ideological stance, and potentially lower-level digital information such as the physical location of the web server and the geographic region associated with its Internet Protocol address. Importantly, obtaining some digital footprint data, such as Internet Protocol addresses, might require a legal warrant and cooperation from the platform provider. Moreover, this approach could be compromised if terrorists use virtual private networks to appear to be in regions other than their actual location. Nonetheless, detectable clues often remain that can help identify and possibly attribute terrorist information campaigns.

If AI is to disrupt the "Ideation & Motivation" and "Post-Operation Exploitation" stages effectively, a significant part of its role will involve countering terrorist narratives. Generative AI, with its ability to interpret text in virtually any language, can play a pivotal role.[276] It can scrape social media feeds and websites to identify accounts or platforms that appear to be recruiting members for terrorist plots or advocating terrorist ideologies. Additionally, generative AI can interpret spoken language, translate it into text, and summarize conversations. This functionality enables AI to be deployed in targeted online forums or social media spaces where live, verbal conversations occur. In such scenarios, AI could monitor forums suspected of harboring sympathies toward terrorists by eavesdropping to gather intelligence that assesses the group's sentiment or provides more targeted evidence of potential targets or threats.[277]

In addition to monitoring social media feeds and listening to online public forums, generative AI can adopt a more proactive role in disrupting terrorists' operations by directly countering their narratives. If AI systems detect an account or social media post that aligns with terrorist ideologies, they could automatically generate responses that challenge the logic or ideology behind terrorism. For instance, the AI could craft and post replies that undermine the messages propagated by terrorist agents or advocates. Similarly, if the AI tools participate in live conversations within social media spaces, they could employ an artificial yet authentic, human-sounding voice to articulate arguments against terrorism and

---

275 Judd Devermont, "How Fraym Explains Extremist Violence in Burkina Faso," *Fraym* (blog), n.d., accessed November 19, 2024, https://fraym.io/blog/violence-in-burkina-faso/.

276 Priyank Mathur et al., "The Radicalization (and Counter-Radicalization) Potential of Artificial Intelligence," International Centre for Counter-Terrorism (website), May 1, 2024, https://www.icct.nl/publication/radicalization-and-counter-radicalization-potential-artificial-intelligence.

277 Rocco Bellanova and Marieke de Goede, "Co-Producing Security: Platform Content Moderation and European Security Integration," *Journal of Common Market Studies* 60, no. 5 (2022).

its underlying philosophies in the native tongue of the participants.

Achieving this level of automation is not without its challenges. As scholars Rocco Bellanova and Marieke de Goede observe, using AI to assess and remove extremist content from Internet websites requires technological and legal mechanisms that require human-machine teaming. In their study of the EU's Internet Referral Unit, which operates under the regulation for "preventing the online dissemination of terrorist content," referred to as the Terrorist Content Online Regulation, adopted in April 2021, Bellanova and de Goede point out, "there is no 'magic algorithm' for identifying terrorist content on the internet," which forces online platforms and government agencies to make judgments about what content should be removed. Critical to the process of identifying online terrorist content, in addition to the algorithm, are the right flagging templates that effectively prioritize suspected content.[278]

The ability to quickly process and flag alarming content is a major advantage of using AI to search for terrorist-related content. One of the significant challenges today is the sheer volume of information generated daily on the Internet and social media platforms. Law enforcement or intelligence analysts cannot realistically track every open conversation on the Internet that is suspected to involve terrorists or those sympathetic to their causes. Additionally, some online discussions occur in languages unfamiliar to the officers, further complicating intelligence-gathering efforts. These challenges can be mitigated through AI, which operates at machine speed and can understand text and spoken languages.

Consequently, AI significantly lowers the threshold for automating searching and monitoring terroristic communications in public online forums. Artificial intelligence can scan vast online communications to detect radicalization patterns and identify emerging threats. Artificial intelligence tools can flag content that suggests extremist indoctrination or recruitment activities by analyzing data across social networks, forums, and websites.

Detecting Terrorist Planning and Staging Activities

As terrorists progress through target selection, intelligence gathering, planning and preparation, logistics and deployment, and attack execution, they inevitably leave traces or signatures that could theoretically be detected before an attack occurs. These steps usually involve some form of coordination and communication, as well as acquiring and transporting goods and people.[279] The primary challenge is detecting these subtle signals of communication and logistics, particularly as terrorists strive to conceal or disguise their activities to evade detection. Terrorists often use encrypted communication methods or camouflage their

---

[278] Bellanova and de Goede, "Co-Producing Security."

[279] Don Reisinger, "The Many Ways Terrorists Communicate Online," *Fortune* (website), May 3, 2016, https://fortune.com/2016/05/03/terrorists-email-social-media/; and Department of Homeland Security, *Department of Homeland Security Strategic Framework for Countering Terrorism and Targeted Violence* (Department of Homeland Security, September 2019).

preattack activities within the mundane routines of civilian life.[280] But historical analyses of terrorist events have shown terrorists inadvertently produce detectable signs that can forewarn about potential attacks.

Experts have identified eight common forms of the detectable signs of terrorism: surveillance, inquiries, security tests, fundraising, acquiring supplies, suspicious behavior, dry runs, and deployment of assets.[281] Additionally, research has demonstrated terrorist attacks are often preceded by criminal acts such as illegal arms possession, robbery, wire fraud, money laundering, espionage, trespassing, and irregular immigration.[282] This research suggests recognizing the temporal and spatial correlation between such criminal activities can help trigger timely and accurate alerts about potential terrorist plots.

Given the eight common warning signs previously mentioned and the fact terrorists commonly commit crimes before they strike, data sources exist through which AI can filter to search for red flags. Some of these potential sources include international or federal crime databases, local or state crime databases, local site security or event logs, flagged financial-transaction logs, digital scanners of goods at controlled entry points, immigration and customs databases, international or federal travel databases, local city enforcement street cameras, and local site surveillance video. A breakdown of where AI could fit into each of the remaining stages in figure 6-1 is listed below.

- **Target Selection and Intelligence Gathering:** Given AI algorithms can process and analyze large datasets to uncover unusual patterns of behavior, they could be deployed to operate using data within travel and customs databases, financial-transaction logs, and other data sources to uncover reconnaissance activities or anomalous purchases that may indicate preparations for a terrorist attack. For example, increases in meetings, phone calls, supply purchases, and fundraising, which could include both criminal and noncriminal means, are typically associated with preparation for a terrorist operation.[283] Artificial intelligence can collect, aggregate, and flag relevant patterns, allowing counterterrorists to disrupt attacks before planning progresses too far.
- **Planning and Preparation:** Artificial intelligence systems could be deployed to monitor communications and dark web forums where planning might occur. For example, a 2008 study on international and environmental terrorists found groups using uncoordinated violence—characterized by lone actions, strong commitment, and minimal contact between leaders and operators—began planning about 15 days

[280] Robert Graham, "How Terrorists Use Encryption," *CTC Sentinel* 9, no. 6 (June 2016): 20–25.

[281] "Eight Signs of Terrorism," Colorado Division of Homeland Security and Emergency Management (website), n.d., accessed on August 3, 2024, https://dhsem.colorado.gov/ciac/eight-signs-of-terrorism.

[282] Brent L. Smith et al., "The Terrorists' Planning Cycle," in *The Handbook of the Criminology of Terrorism*, ed. Gary LaFree and Joshua D. Freilich (Wiley-Blackwell, 2016), 62–76; and Smith, "Terrorist Behavior."

[283] Brent L. Smith et al., *Geospatial Analysis of Terrorist Activities: The Identification of Spatial and Temporal Patterns of Preparatory Behavior of International and Environmental Terrorists* (University of Arkansas, 2008), 53.

before an attack, whereas groups with more deliberate strategies started planning around 92 days in advance.[284] With those indicators in mind, machine learning models could identify coded language and other suspicious correlations human analysts might overlook.

- **Logistics and Deployment:** Surveillance camera systems and logistics scanning equipment enhanced with AI could detect known suspects, suspicious activity such as loitering and unusual photography, and smuggled explosives. Similarly, facial recognition and anomaly-detection algorithms can alert authorities about unusual gatherings or movements at critical locations or suspected terrorist targets.
- **Attack Execution:** The real-time AI-driven analytics and AI-enabled surveillance equipment mentioned above can assist in immediately detecting threats and coordinating response efforts. These systems can analyze surveillance footage and sensor data to identify attacks as they unfold, enabling rapid deployment of emergency services. In particular, facial recognition may increase in importance as terrorist organizations rely more on lone wolf attacks, which may not have as large a signature as more complex attacks.[285]

To illustrate the utility of AI, scholars from the University of Sciences and Humanities in Lima, Peru, used the University of Maryland's Global Terrorism Database to visualize and accurately predict numbers and types of terrorists by region. They used classifier models known as "decision trees" and "random forests" to analyze terrorist attacks using 100 classifying characteristics. A decision-tree model consists of branches that describe the attributes of attacks and leaves that classify each attribute so the algorithm can analyze it. The random-forest model uses multiple decision trees with random attributes to overcome a single decision tree's limitations. The result of the decision-tree model was a prediction accuracy of 75.45 percent for the attack region and 79.24 percent for the type of attack, whereas the random-forest method reached 89.54 percent for the attack region and 90.41 percent for the type of attack.[286]

A Layered Artificial Intelligence Architecture for a Holistic View of Potential Risks

The previous sections highlighted various data sources AI could leverage to detect hallmark signs of terrorist activity, guided by the distinctive stages of terrorist operations and insights from research on such indicators. Building on this foundation, we propose a nuanced deviation from the traditional AI architecture known as an ensemble: a concept familiar

---

284 Smith et al., *Geospatial Analysis*, 51.

285 Tony Porter, "Facial Recognition Technology in the Fight Against Terror," *Counter Terror Business* (website), February 10, 2023, https://counterterrorbusiness.com/features/facial-recognition-technology-fight-against-terror.

286 Enrique Lee Huamaní et al., "Machine Learning Techniques to Visualize and Predict Terrorist Attacks Worldwide Using the Global Terrorism Database," *International Journal of Advanced Computer Science and Applications* 11, no. 4 (2020): 562–70.

to AI researchers and practitioners.[287] Ensemble learning is a technique that aggregates several models to improve accuracy.[288] Our adapted ensemble approach is designed to ingest and synthesize information from diverse, siloed data sources, providing a comprehensive and coherent analysis of potential terrorist threats. By integrating these varied inputs, the architecture aims to offer a more detailed and accurate portrayal of the dynamics underpinning terrorist activities.

An AI ensemble traditionally combines multiple AI models or algorithms to improve decision making or predictive accuracy. These models usually operate using the same dataset, pooling their insights to produce an accurate and robust output by capitalizing on each model's strengths and mitigating the models' weaknesses.[289] Ensembles use aggregation functions, such as majority voting or averaging, to combine the outputs of the individual models. For example, in a classification task, an ensemble with a majority voting mechanism would select the class upon which most independent classifiers agree.

But the approach proposed in this paper modifies the traditional ensemble concept. In our approach, we were inspired by the concept of an ensemble, where multiple AI models collaborate on a collective task. But we diverge from the traditional ensemble because our system follows a multitiered structure. At the lower tier, individual AI models operate using narrowly defined, specific datasets, with each dataset potentially having its own traditional ensemble designed to detect patterns or anomalies unique to its data. After processing, AI models (or independent ensembles) pass their flagged detections to a higher-level metamodel, distinguishing our approach from a conventional ensemble. This metalevel AI agent synthesizes flagged events from various sources to create a clearer, more comprehensive picture of potential terrorist activities. By analyzing aggregated signals across multiple data streams, the metalevel model aims to reveal patterns that may be overlooked in isolation, thereby enhancing the system's predictive accuracy and providing deeper insights into the structure and plans of terrorist organizations.

287 "A Comprehensive Guide to Ensemble Learning," GeeksforGeeks (website), updated December 26, 2023, https://www.geeksforgeeks.org/a-comprehensive-guide-to-ensemble-learning/.

288 Jacob Murel and Eda Kavlakoglu, "What Is Ensemble Learning?," IBM (website), March 18, 2024, https://www.ibm.com/topics/ensemble-learning#:~:text=Ensemble%20learning%20is%20a%20machine,than%20a%20single%20model%20alone.

289 Murel and Kavlakoglu.

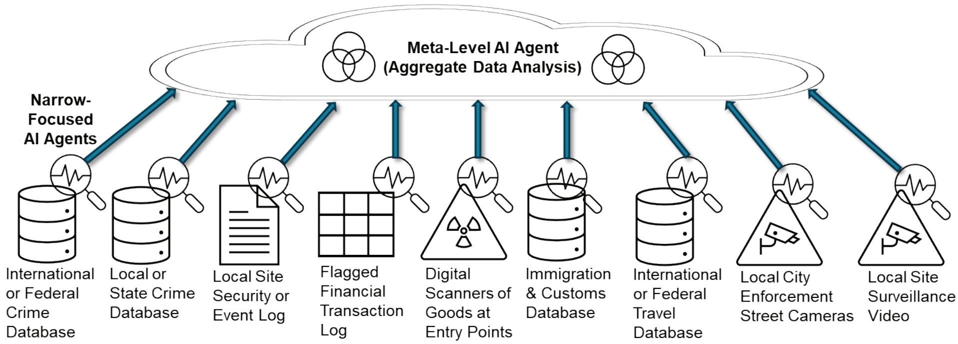

Figure 6-2. Architecture of a two-level AI ensemble.

The diagram in Figure 6-2 demonstrates the integration of diverse machine learning models at the first level, each independently processing input data. The second-level AI agent analyzes outputs to identify potential patterns suggestive of a terrorist plot. This ensemble approach reduces the complexity of the problem space compared to a flattened architecture, in which a single AI agent would need to analyze all raw data across various sources simultaneously.

Figure 6-2 also illustrates the adapted ensemble approach, specifically designed for AI to detect and disrupt terrorist operations. This architecture comprises two levels of AI agents for optimal effectiveness: one at the lowest level, narrowly trained to recognize patterns within specific datasets, and another at a higher level, tasked with operating on the output—or red flags—provided by each AI agent at the lower level. For instance, one narrow AI agent might analyze surveillance camera feeds to identify suspicious behaviors, such as by recognizing the license plates of vehicles frequently spotted in high-security zones or individuals loitering in sensitive areas. Another agent could sift through financial transaction logs to flag unusual patterns indicating funding for terrorist activities, such as sudden, large withdrawals or international transfers linked to regions known to be terror-prone.

Whenever the narrowly focused AI agents detect an anomaly or suspicious activity, they relay the flagged information to the metalevel for comprehensive analysis. This metalevel AI agent stitches together clues from multiple sources, enhancing the capacity to discern more complex patterns that may not be evident from a single data stream. For example, combining insights from local crime databases, immigration records, and travel logs could reveal the movements of suspected individuals and their networks, offering a more detailed profile of potential threats.

Each narrowly trained AI agent plays a critical role in this system by providing specific clues contributing to a holistic security overview. Although a single tip from one AI source could potentially prevent an attack, integrating multiple data points can offer a clearer picture of the perpetrators, their methods, and the timing and location of potential attacks. Thus, the meta-level AI agent is crucial for synthesizing these disparate pieces of information to form a comprehensive intelligence picture. Our proposed approach underscores the significant strength of AI in ingesting large, multidimensional data sources and discerning patterns across them, providing a strategic advantage in counterterrorism efforts.

In counterterrorism, the upper echelons of our two-level AI ensemble architecture are where generative AI truly shines. A generative AI agent operating at this metalevel, fine-tuned with comprehensive counterterrorism research and doctrine, possesses an extensive knowledge repository. Such an advanced AI capability allows for an in-depth analysis of the nuanced signals the initial model layer detects. By synthesizing vast amounts of literature and previous studies on terrorist activities, generative AI can effectively identify and piece together subtle signs that might individually seem benign but, when viewed holistically, indicate a potential terrorist plot. The system's integration of generative AI both enhances its detection capabilities and makes implementing sophisticated AI ensembles more feasible than ever before, promising a significant leap forward in proactive counterterrorism measures.

Mitigating the Challenges

One significant challenge in deploying AI for counterterrorism is training each AI agent effectively. In supervised learning, where algorithms are trained on labeled datasets, training typically requires a substantial set of labeled training examples to teach AI to recognize specific events.[290] But labeling data is often expensive and labor-intensive. Moreover, terrorist activity data within various repositories might be scarce, complicating the development of a robust, generalized model. One approach to addressing these challenges is to leverage unsupervised learning techniques, where algorithms learn from unlabeled datasets.[291] Unsupervised learning methods enable AI agents to detect anomalies within data streams without labeled examples. By analyzing patterns and deviations from norms, unsupervised learning can highlight potential threats autonomously.

Alternatively, another viable solution is synthetic data, which is artificial data generated to mimic the attributes of real data, allowing one to increase the size of a database for more effective output or to mask real-world data to prevent unauthorized access.[292] Through simulation or generative AI, the synthetic-data method can augment sparse real data, providing AI agents with ample diverse scenarios to learn from, thus enhancing their ability to generalize across different situations. The human-machine teaming aspect of our proposed system is crucial for continuous improvement. Regular updates are essential for optimal

290 "Supervised Machine Learning," GeeksforGeeks (website), updated October 23, 2024, https://www.geeksforgeeks.org/supervised-machine-learning/.

291 "Supervised and Unsupervised Learning," GeeksforGeeks (website), updated September 23, 2024, https://www.geeksforgeeks.org/supervised-unsupervised-learning/.

292 "What Is Synthetic Data Generation?," K2view (website), updated May 7, 2024, https://www.k2view.com/what-is-synthetic-data-generation/.

performance in any dynamic machine-learning application.[293] The nature of data signatures, or their characteristics, can evolve over time. In terrorism contexts, tactics and techniques may shift due to attempts to evade detection, or as new technologies emerge. Thus, the AI system must adapt to changes by learning from new data, with humans playing a pivotal role in guiding and correcting the AI, particularly in instances of false positives or missed detections.

Humans also play a role in continuously improving the system by publishing new research on the tactics, techniques, and procedures of terrorist operations. As the new research becomes available, generative AI can ingest it to integrate new insights into terrorist strategies, continually enhancing the system's effectiveness. Another challenge is gaining access to data indicative of terrorist plots. The most revealing data types are often shielded by privacy laws, encrypted communications, or deliberate obfuscation by terrorists. The ethical, moral, and legal implications of using such data for AI-based surveillance will be further explored in the subsequent chapter.

Conclusion

Particularly with advancements in generative AI, AI introduces unprecedented opportunities to disrupt terrorist operations. Terrorist attacks often follow a discernible sequence of events involving detectable actions such as communications, transactions, and movements. Terrorists' efforts to conceal their activities make distinguishing these actions amid a vast sea of legitimate data formidable, even with AI assistance. Yet, the capabilities of AI to sift through and interpret complex data patterns at an unparalleled scale make detecting terrorists' actions more feasible than ever before. This chapter explored the staged process of terrorist attacks, identifying eight detectable signs of impending threats. It discussed the potential of AI, especially generative AI, to monitor and analyze data from various stages of a terrorist attack cycle, from ideation and influence—where AI could monitor social media and online forums—to the execution stages, where specific suspicious behaviors might be detected.

We propose a multitiered AI system in which each data source is overseen by a specialized AI agent trained to recognize signs of terrorism within the data stream. Alerts from specialized AI agents could be escalated to a higher-order, metalevel AI agent that integrates signals across multiple data modalities. This generative AI agent, leveraging a vast repository of counterterrorism knowledge, could predict details about terrorists' plans, enabling preemptive actions. Humans remain crucial to the proposed system's success. They are responsible for making final decisions on actions such as surveillance, arrests, or interventions and are vital in refining AI's performance. Human oversight is essential to correcting false positives and omissions, labeling new data, and integrating the latest research into AI models.

---

293 "AI Model Maintenance: A Guide to Managing a Model Post-Production," Appen, February 26, 2021, https://www.appen.com/blog/ai-model-maintenance-guide-to-managing-model.

Finally, using data sources, as described, may raise significant privacy concerns, highlighting the perennial security-privacy trade-off. Although this chapter does not cover privacy issues in depth, it acknowledges the potential for misuse of such surveillance technologies and the importance of maintaining ethical standards in deploying AI-driven counterterrorism tools. This proposed AI-enhanced approach to counterterrorism reflects a sophisticated integration of technology and human expertise, aiming to detect and to disrupt terrorist plots before they come to fruition. Maintaining a balance between innovative security measures and ethical considerations will be paramount as we advance.