Artificial Intelligence in Cities: Securing Our Future

Report 2025

CTPN
COUNTER TERRORISM
PREPAREDNESS NETWORK

CTPN
COUNTER TERRORISM PREPAREDNESS NETWORK

www.london.gov.uk/ctpn

Contributors and Reviewers

Balques Al Radwan
Cybersecurity and New Technologies Unit
United Nations Office of Counter Terrorism

Professor Saskia Petra Bayerl
Professor of Digital Communication and Security
Deputy Director of CENTRIC
(Centre of Excellence for Research into Terrorism, Resilience, Intelligence and Organised Crime)
Sheffield Hallam University

Kjell Brataas
Crisis Management Expert
Civil Protection Group Member
North Atlantic Treaty Organisation (NATO)

Blagoj Delipetrev
Artificial Intelligence Expert
Science for Peace and Security
Innovation, Hybrid and Cyber Division
North Atlantic Treaty Organisation (NATO)

Donald Dudenhoeffer
Cyber Security Research Engineer
Centre for Digital Safety and Security
Austrian Institute of Technology

Lauren Fricke (Researcher and Author)
Programme Coordinator, Counter Terrorism Preparedness Network
District of Columbia Homeland Security and Emergency Management Agency

Dr Paul Martin CBE
Distinguished Fellow
Royal United Services Institute
Former Head of the UK Centre for the Protection of National Infrastructure

Dawn Morris
Independent Protective Security Expert
Mayor's Office for Policing and Crime, London

Judy Pal
Independent Communications Expert
Former Assistant Commissioner, New York Police Department

Professor John Parkinson OBE
Professor of Security and Counter Terrorism at the Global Secure Societies Institute
Executive Managing Director of the Leadership in Counter Terrorism Alumni Association

Dr Valeria Spizzichino
Senior Researcher
Italian National Agency for New Technologies, Energy and Economic Development

Alex Townsend-Drake (Author and Editor)
Head of the Counter Terrorism Preparedness Network
Greater London Authority

Mark Wittfoth
Head of Observatory, Innovation Laboratory
European Union Agency for Law Enforcement Cooperation (EUROPOL)

Senior Officer (Name Undisclosed)
National Digital Exploitation Service
UK Counter Terrorism Policing

1 Introduction

“As technology continues to evolve, its impact grows deeper, presenting both opportunities and challenges that shape our future.”

Technology is an integral part of modern life, influencing the way people live, work, and interact. From the convenience of smartphones to the power of Artificial Intelligence (AI), these innovations affect our daily routines and societal progress. As technology continues to evolve, its impact grows deeper, presenting both opportunities and challenges that shape our future. AI is at the forefront of this, accelerating across sectors faster than ever before.

In 2021, the United Nations Office of Counter Terrorism (UNOCT), in collaboration with the United Nations Interregional Crime and Justice Research Institute (UNICRI), published *Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes.* This report examined the threats posed by AI and similar technologies if used by individuals intending to commit acts of terrorism. AI-related threats were identified and divided into three branches of security priorities: malicious cyber activity, the enabling of physical attacks, and political interference.[1] This was endorsed by The Alan Turing Institute in 2023, which also noted that there are three broad categories of threat actors. The first is the state actor which may use AI as part of a wider armoury; the second is the non-state actor, such as an organised crime group that seeks to exploit the public and undermine the rule of law; and the third is lone-actors that may use AI to inflict harm.[2]

This annual report by the Counter Terrorism Preparedness Network (CTPN) – *Artificial Intelligence in Cities: Securing Our Future* – builds on the categories outlined by UNOCT. However, as cyber threats were recognised in a previous CTPN report on *City Preparedness for Cyber-Enabled Terrorism*, this report will specifically focus on the physical and political threats AI-enabled technologies pose.[3] This recognises that there is a degree of overlap or interchange since physical threats may serve as political threats (or ideological/religious ones) and vice versa. This report otherwise directly complements additional CTPN reports on *Mis- and Disinformation: Extremism in a Digital Age,* and *Preparing for Hostile Drones in Urban Environments*, which respectively highlight AI as an enabler for disinformation and targeted attacks by drones.[4,5]

This report seeks to further understand AI and the potential threats it poses in the context of terrorism. It also considers the benefits AI offers to counter terrorism and build both security and preparedness. Of course, there are many other uses of AI in different sectors, which are not discussed in this document. It is also accepted that the speed of change and development in the field of AI, and technology more broadly, is fast. What is written today may not apply in the same way tomorrow. Therefore, whilst this report refers to current research and examples, it avoids overly technical details and offers a high-level perspective for strategic planning in cities.

By applying a city lens and narrowing the focus to a smart city perspective, this report seeks to drive awareness by highlighting both the threats and benefits of AI, and how terrorists could exploit them. It evidences the need to continually address vulnerabilities and enhance security and preparedness against terrorist threats powered by AI, before concluding with overarching implications and a set of high-level recommendations aimed at city authorities.

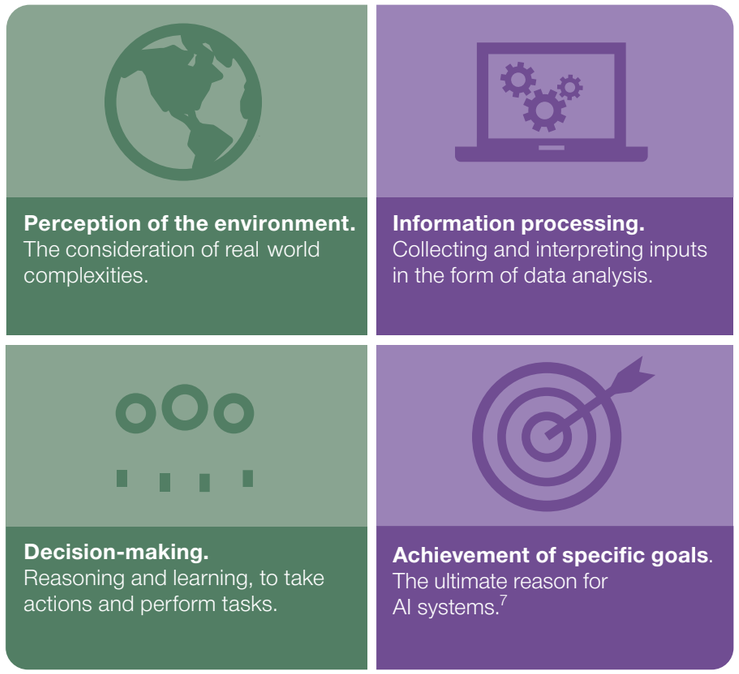
> **AI is not a single entity; rather, it is a field of diverse technologies that can be used individually or collectively.**

2 Understanding Artificial Intelligence

According to the European Commission's report *AI Watch: Defining Artificial Intelligence*, an AI system refers to "software that is developed…for a given set of human-defined objectives, [to] generate outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with."[6]

It is important to note that AI is not a single entity; rather, it is a field of diverse technologies that can be used individually or collectively. The field has evolved since 1956, when universities originally established laboratories to research AI. Understanding of the technology grew, and AI broke into the realms of speech translation and recognition.

Common Features of AI



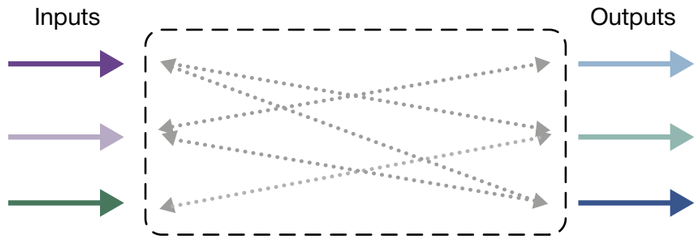
Moreover, the UK National Cyber Security Centre considers AI systems to be "computer systems which can perform tasks usually requiring human intelligence. This could include visual perception, speech recognition or translation between languages."[8] This usually involves training machines to simulate human thought processes and decision-making.[9,10]

At the turn of the 21st century, the internet and its abundance of data became mobile, allowing AI to be more widely integrated.[11,12] Today, automation and deep learning represent the latest revolutions in AI.

Large Language Models (LLMs) and multi-modal models are examples of deep learning. They are pre-trained on vast quantities of data.[13]

This is one category of generative AI, a form of machine learning that can be broken into two types: supervised – where the correct answer is provided – and unsupervised – where the computer interprets patterns in the data to come to a result. Supervised learning has a greater ability to compute complex data because its inputs are known.

The Black Box Effect



This also increases the accuracy and reliability of its results. Unsupervised learning otherwise relies on patterns and clusters of data to "self-supervise" decision-making.[14] Some types of LLMs are capable of unsupervised training by simultaneously processing entire sequences of data, often with hundreds of billions of parameters and sources.[15] However, one of the main challenges is human comprehension.

LLMs are based on neural network frameworks, resulting in extremely large and complex models once built. Their size can be difficult to comprehend, and in some cases, the user will be entirely blind to the LLM's decision process, aware only of the inputs and outputs. This opaqueness is known as the "black box effect".[16] This stresses the need for clear parameters and criteria that enable more control over what goes into the box.

It also demonstrates how, today, the capabilities of AI have expanded significantly. AI is used extensively on most internet services, including language translation to image, voice, and facial recognition. Consider, for instance, how after analysing thousands of images of dogs, a generative AI model can create a new, synthetic image of a dog that, while entirely fictional, looks indistinguishably real. Generative AI and platforms like ChatGPT and Google Gemini are transforming how we do things, from co-piloting aircrafts to developing computer code.

In fact, the 2024 Nobel Prize in Chemistry was awarded for the development of an AI algorithm that solved the fifty-year protein structure prediction challenge, and the 2024 Nobel Prize in Physics was awarded for developing machine learning technology using artificial neural networks.[17]

Such advancements are remarkable. In recent months, the AI company "OpenAI" shared early test results from a new model called o3. "These results indicated significantly stronger performance than any previous model on several of the field's most challenging tests" from programming to abstract and scientific reasoning, outperforming many human experts.[18]

For this reason, AI has been argued to reduce labour, enable space for humans to explore more creative pursuits, and drive societal progress. However, it is also considered to risk increasing unemployment, damaging industry, and encouraging an over-reliance on technology whilst posing threats to security whether cyber-based, physical, or political.

How an AI system is regulated, developed, implemented, used, and managed is key. This hinges on humans' ability to maintain control over the technology and, thus, its impact. AI's impact – in the field of counter terrorism or otherwise – will depend on how well it is understood and embedded into the fabric of society.[19] However, its effectiveness also depends on factors such as governance, regulation, ethics, and the adaptability of security frameworks. Likewise, its impact – when used by malicious actors (e.g., hostile states, terrorists, organised crime groups, or other criminals) – will depend on the capabilities they have access to, how they decide to use them, how technologically skilled they are, and how robust counter measures are.

> **AI's impact – in the field of counter terrorism or otherwise – will depend on how well it is understood and embedded into the fabric of society.**

> **The integration of AI as an enabler for terrorism represents a growing threat, allowing for more sophisticated operations, broader reach, and increased impact.**

3 Threats Posed by Artificial Intelligence

AI development collaborations between nation-states reduce the unit cost of mutually beneficial technologies, which is driven down further by commercial market competition. Consequentially, as AI technologies become more affordable, they are more easily acquired or exported, including by non-state actors or through illegal markets. This rapid diffusion of technology presents significant security challenges.[20]

The *2024 European Union Terrorism Situation and Trend Report* noted terrorists "have the capability to strategically integrate the most recent developments in digital technology." This includes AI "to spread propaganda, recruit, plan attacks, and evade detection by law enforcement."[21] It further noted how LLMs and deepfakes are exploited to "create false identities, spread disinformation and bolster propaganda campaigns."[22]

Trends indicate that AI is primarily being used by criminals to enhance current cybercrime activities such as deepfake scams, phishing assistance, vulnerability research, as well as target reconnaissance.[23] Such AI-enhanced activities could be adapted to support terrorist attack planning. Herein lies an important distinction. Terrorists may be AI-enabled (by maliciously using generative AI, for example), and their attacks may be AI-enabled (such as with drones). The rise in new or novel attacks using AI, however, has been slow so far but will increase as knowledge expands and associated accessibility grows for malicious actors.

Adversarial AI, for example, could be used by a terrorist group to obtain sensitive data, recreate sensitive models, or adversely bias model results. Thus, the use of AI itself offers a potential attack vector for an adversary.

AI can be used by malicious actors in a myriad of ways. Consider political interference, disinformation (fabricated or deliberately manipulated content), or deepfakes that leverage AI-generated propaganda and chatbot-driven radicalisation; the use of AI-enabled 3D-printing to assist in producing untraceable assault weapons; synthetic identity fraud, using generative AI to create fake personas for illicit financing, infiltration, or recruitment; AI use in malware generation or automated cyber-attacks; and the development of "how to" guides, or the application of technologies to plan, facilitate, or deliver attacks.

For example, the *International AI Safety Report 2025* found that "systems have displayed some ability to provide instructions and troubleshooting guidance for reproducing known biological and chemical weapons and to facilitate the design of novel toxic compounds."[24] The CTPN report on *Bioterrorism: Applying the Lens of COVID-19* also recognised this possibility.[25] Of course, this requires a degree of knowledge, investment, access, and resources, but AI – notably LLMs – could offer significant assistance.

Terrorists are already testing ways AI can be used. A 2023 EUROPOL report highlighted the ability of LLMs to answer contextual questions, making it "significantly easier for malicious actors to better understand and subsequently carry out various types of crime."[26] These systems can be tricked into providing guides tailored to the goals and resources the terrorist desires.[27]

The recent 'cyber-truck' explosion in Las Vegas raises further concerns about the accessibility of generative AI tools and their potential misuse to help build explosive devices. Kevin McMahill, Sheriff of the Las Vegas Metropolitan Police Department said, "This is the first incident that I'm aware of on US soil where ChatGPT [was] utilised to help an individual build a particular device. It's a concerning moment."[28] The integration of AI as an enabler for terrorism (whether non-state or state-sponsored) represents a growing threat, allowing for more sophisticated operations, broader reach, and increased impact.

Rebecca Weiner, the New York Police Department's Deputy Commissioner for Counter Terrorism, said that "AI takes existing problems and magnifies them."[29] Indeed, terrorists have a track record of using everyday technology in unconventional and exploitative ways. Whilst views vary on the capability of terrorists to effectively employ AI – and low-tech physical approaches such as blunt and bladed weapons, firearms, and explosives are still considered to be more likely – the rapid evolution of technology coupled with low barriers to access suggest the threat of AI should be a concern. Its synergy with terrorist intent, especially as an enabler, is evident.

Whilst there is no internationally agreed definition of terrorism, it can be understood as the intentional use of force – either physical, emotional, or psychological – against a certain group to advance a political, ideological, or religious agenda.[30,31] This definition has become increasingly blurred with hostile state activity and the resurgence of state-sponsored terrorism, serving as a dangerous interplay in the context of AI. The following sections will look at how this translates to physical and political threats in more detail.

Potential Threats Posed by AI Technologies

1 New attack patterns and tactics that use AI in yet unknown ways.	4 Hostile reconnaissance, intelligence collection, and operational planning/instructions.
2 Semi or fully automated attacks using autonomous and remote systems.	5 Enhanced and higher volumes of propaganda and disruption through disinformation.
3 Radicalisation and recruitment of individuals by targeting vulnerable groups.	6 Higher tempo of cyberattacks, lowering the entry point for those that are more sophisticated.

4 Physical Threats

> **City infrastructure itself can not only become terrorist targets but also a tool for terrorists or other malicious actors.**

Physical threats encompass AI-related changes to the threat landscape as it relates to the physical world. This goes beyond the use of AI to push disinformation, which is explored in the next section, instead focusing on its potential to directly enable system failures or attacks. Examples include the use of AI for attack planning or impacts stemming from AI-driven activity in cyberspace, materialising through hard technologies or physical world implications. For a threat originating in cyberspace (such as AI-generated ransomware), the focus would be on its physical consequences (such as the shutting down of an energy grid).

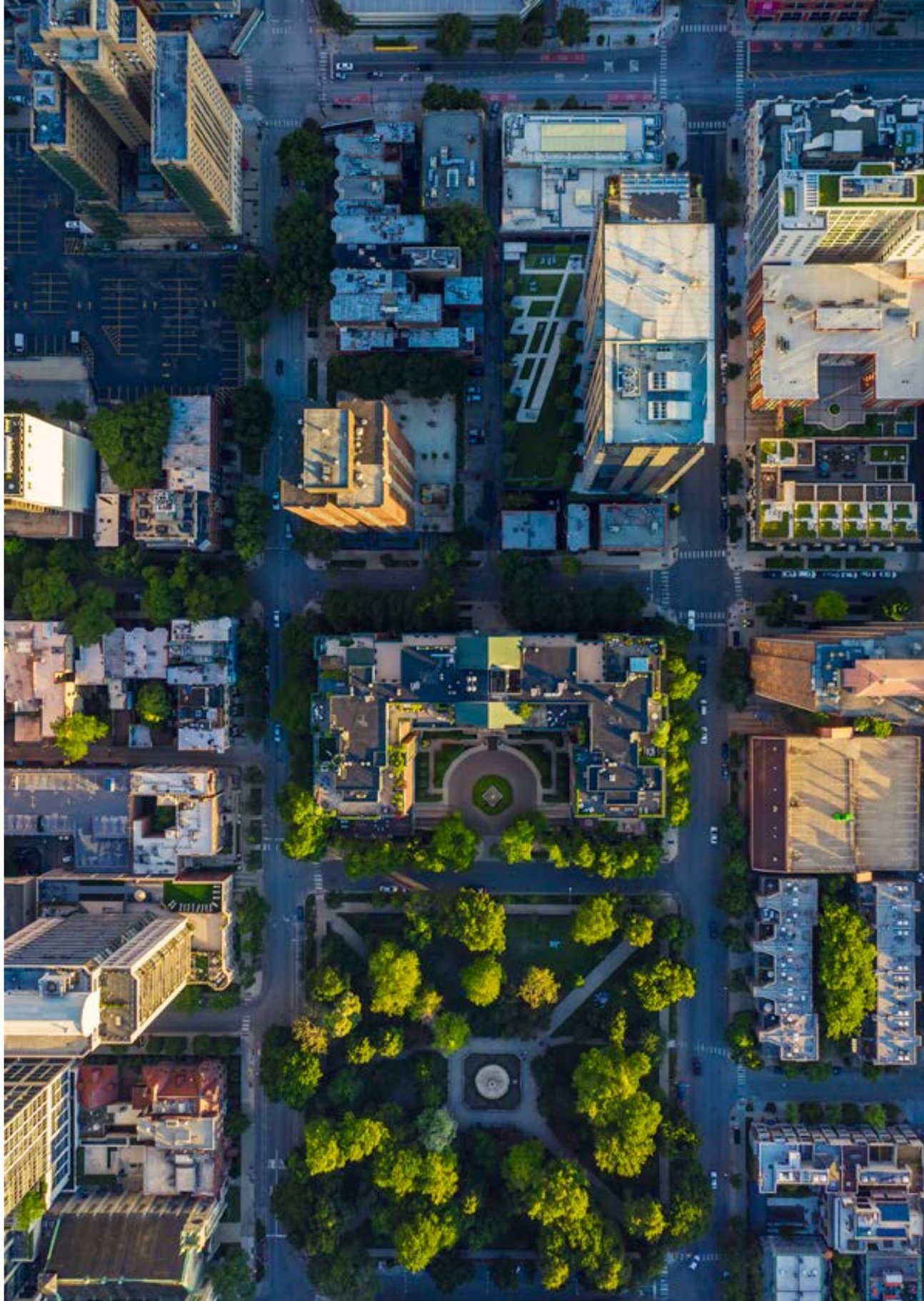
It follows that, in smart cities, AI-integration into the built environment can develop into an attack vector. Structures may be "designed, constructed, and maintained making use of advanced, integrated materials, sensors, electronics, and networks which are interfaced with computerised systems comprised of databases, tracking, and decision-making algorithms."[32] Technologies, and especially AI, are an integral part of creating and operating smart infrastructures, in everything from energy and water systems to buildings, streets, and lighting.

For instance, packed with 28,000 sensors, 'The Edge' in Amsterdam has been described as the "smartest building in the world." A smartphone app enables constant connectivity with this office space. It "checks your schedule, and the building recognises your car when you arrive and directs you to a parking spot…"Then the app finds you a desk. Wherever you go, the app knows your preferences for light" and temperature, and it tweaks the environment accordingly, in a bid to drive efficiencies.[34] Future smart cities will likely increase this type of AI-integration.

The subsequent extent of vulnerability to attacks may be determined by the level of integration, adaptivity, and automation of AI components.[34] In this context, city infrastructure itself can not only become terrorist targets but also a tool for terrorists or other malicious actors. Today's architecture deploys AI systems, for instance, for predictive cooling/smart heating, ventilation, and air conditioning – systems that can be hacked and repurposed across a building or whole city block.

> **Data poisoning is the intentional manipulation of training data to introduce biases or control the outcome of a system.**

The potential physical manifestations of data poisoning offer another example. Data poisoning is the intentional manipulation of training data to introduce biases or control the outcome of a system. It is particularly relevant for machine learning, specifically deep learning, because of the heavy reliance on training data. Authorities also have difficulty identifying what is contained in the training data and may not be able to forensically detect any changes made. By subtly altering the data used to train models, attackers can exploit this blindness to degrade performance, cause incorrect predictions, or even embed hidden backdoors which could be accessed later.[35]



Indeed, it may be possible to disrupt or derail decision-making by forcing machine learning systems to misclassify data, thus influencing, the machine's learning and actions, and teaching it to create harm or hold unacceptable viewpoints.[36] Such could result in improper flags for innocent individuals or the lack of identification of known terrorists.[37] Theoretically, blind spots could also be created for detectors, such as concealing weapons from x-rays or misidentifying a license plate.[38] Whilst this would be a sophisticated attack, it demonstrates how open-source AI platforms, which rely on open-source data, could be manipulated with major consequences. This underscores the importance of ensuring credible and reliable data to train AI systems. This relates to major concerns regarding data privacy and human rights, as explored later.

Dr Paul Martin, a Distinguished Fellow at the Royal United Services Institute, has also highlighted how organisations now face a growing security risk from artificial insiders. Dr Martin referred to agentic artificial insiders that can achieve outcomes independently, albeit mostly in and from the virtual world. Like the human insider threat, these are digital insiders or imposters that could commit "fraud; blackmail; theft of data, money or intellectual property; covert influencing; physical or cyber sabotage; violence; leaking [of confidential information]; terrorism; espionage; and so on," which can manifest with real-world consequences that are physical and emotive in nature.[39]

He further stated, "it seems almost inevitable that protective security practitioners will increasingly be required to defend organisations against AI insiders. The problem, however, is that protective security practitioners are mostly not thinking about AI, and AI experts are mostly not thinking about insider risk."[40] Herein is a need to merge disciplines.

Multiple cases have also evidenced the threat of AI as a tool to create false personas that support malicious actors infiltrating workplaces, such as the hiring of a North Korean fake IT worker in 2024.[41]

On the other hand, AI weaponization and programming of physical technologies deserves attention. Drones, for example, have already been used by terrorists to deliver explosives, serve as flying projectiles, and gather intelligence on future targets.[42,43] AI continues to significantly enhance the capabilities of drones, transforming them from remote-controlled devices into intelligent systems capable of completing data analysis and making autonomous decisions. The more precise and efficient AI makes drones, the greater the risk malicious actors will use them to target individuals, ethnic groups, or specific locations. This could include critical infrastructure and densely populated areas, as seen taking place in Russia and Ukraine.

> **Some AI programmes already use data where police uniforms and vehicles could be set as automatic target identifiers.**

It has been reported that, under certain conditions, a software change could allow a drone to engage a target without a human having to make the decision.[44] In conflict zones, there have already been cases suggesting that drones have selected a target autonomously.[45] This refers to a drone being programmed to attack targets without requiring data connectivity between the operator and the munition.[46] The potential threats posed by drones are explored further in the CTPN report on *Preparing for Hostile Drones in Urban Environments.*[47]

A German arms manufacturer has also revealed a vehicle that can locate and destroy targets on its own.[48] Such advances raise concerns about how these types of technologies could transfer to domestic settings. For example, there are several AI software packages on the market with object detection. These offer commercial benefits but carry the risk that a programme could be taught to identify specific targets or to maximise hostile reconnaissance. Some AI programmes already use data where police uniforms and vehicles could be set as automatic target identifiers. This could be amended to anything, including aeroplanes

As technology advances, the emergence of driverless and autonomous vehicles also introduces new challenges domestically. Experience shows that vehicles have been used in terrorist attacks due to their accessibility and potential for widespread impacts. Cars, trucks, and vans are easy to obtain – they are typically less regulated than firearms or explosive materials – and simple to operate. Vehicular attacks usually involve deliberately driving vehicles into crowds and infrastructure or using them to deliver explosives. They have been used by self-initiated terrorists – individuals operating without support from terrorist networks or organisations – as well as groups which do not have, or do not wish to spend the time and resources on, other methods.[49] In the future, driverless vehicles may appeal to terrorists because of their potential to streamline and enhance attack effectiveness without the perpetrator being physically present.

In theory, automation systems that enable vehicles to be driverless could also be susceptible to manipulation. Driverless vehicles use a combination of deep learning AI and sensor technology to detect objects, interpret road signals and traffic lights, and make decisions.

Attackers could target specific components of a vehicle, such as its sensors or communication systems, or they could compromise the control system as a whole.[50] This would depend on a set of exploitable circumstances occurring, but the following examples apply:

1. 2019 Tesla autopilot hack: Researchers tricked the cars AI by placing small stickers on the road, causing it to misinterpret lane markings and change lanes unexpectedly. This experiment showcased how the external manipulation of sensors could pose risks.[51,52]

2. 2015 Chrysler hack: Researchers gained access to the cars system including the vehicle's steering. Chrysler issued a recall for 1.4 million vehicles to address this security flaw.[53,54]

Vulnerabilities could allow attackers to alter a vehicle's functions remotely, causing accidents or enabling them to use the vehicle as a weapon. Furthermore, the possibility of coordinating and executing a series of attacks from a single location or through remote operations should be considered.[55] However, attacks targeting vehicle automation systems have a much higher barrier to access due to the advanced technical knowledge and resources required, making them much less of an immediate threat.[56]

More pressing is the potential impact of AI-generated swatting.[57,58] This describes the action of making hoax phone calls, which can be cross-border generated, to report serious crimes to emergency services. The aim is to activate emergency responses of different scales to unsuspecting victims or areas.[59] Whilst swatting itself is not new, the threat accelerates when it is AI-generated. Consider auto-generated calls en masse pulling resources to one geographic area whilst a real attack occurs elsewhere. The consequences could be significant.

A swatting case at Boulder Valley School District in February 2023 used AI voice technology enhanced with AI-generated gunshots in the background.[60] Any such call would demand a response. This could have serious implications if emergency resources were improperly diverted and risks creating vulnerabilities elsewhere. It also poses the risk of miscalculation by first responders – such as assessing threat and using force based on false or deliberately misleading information. To this end, the Federal Bureau of Investigation (FBI) has formed a national database to track and try to prevent swatting.[61]

The use of swatting as a form of disinformation relates to the political threats discussed in the next section where, there is again, a complicated nexus between hostile states, organised crime, and terrorism. Indeed, the landscape is increasingly complex with reducing barriers to access and expanding criminal operations, which are becoming more scalable and harder to detect.[62] These issues are compounded by the dark web and the ability to purchase crime-as-a-service.[63]

Artificial Intelligence in Cities: Securing Our Future – Report 2025

“[It is] estimated that by 2026, 90% of social media content will be synthetically generated, launching what some experts call an “information apocalypse”.”

5 Political Threats

Social media and other internet-based platforms are now intertwined with political life. These media platforms offer avenues for political engagement and democratic participation, yet “they are increasingly perceived as conducive to the creation of ideological echo chambers” whilst being “used in attempts to covertly influence the political choices of citizens, thus sapping their democratic credentials.”[64] Political threats in this context largely relate to the spread of disinformation and propaganda using AI-generated narratives, chatbots, and deepfakes.

Deepfakes are a form of synthetic media which use AI deep learning to alter the original message or create fictional content.[65,66] They prey upon an individuals’ implicit trust in what they see and hear.[67] Deepfakes can be created for images, video, and audio, and are often so convincing that they cannot be readily identified.[68] Even low-technology alternatives, known as “cheapfakes”, can appear convincing to the casual viewer.[69] Consider the picture of a fake explosion at the Pentagon, which went viral in 2023.[70]

Despite the term’s negative connotation, deepfakes are not inherently malicious. The technology used to create deepfakes today was originally called “video rewrite” and intended for syncing lip movements with spoken words in dubbed movies.[71] Positive modern uses include completing a television or movie character’s story after the actor has died or “humanising” robotic voices.[72] An AI-generated video available on YouTube titled *This is not Morgan Freeman* offers a short example of how realistic this technology can be.[73] In recent years, it has become “next to impossible” to differentiate between sophisticated AI deepfakes and real footage.[74] This was demonstrated in a video created by Emmy-award winning journalist Johnny Harris.[75]

Individuals are especially easy to target using deepfakes. Deepfakes have been used to carry-out financial scams and, in one case, were used to simulate a kidnapping by spoofing the voice of the victims’ daughter.[76,77] Such highly convincing false narratives pose a significant threat. There are now countless examples of how deepfakes have been used to impersonate political leaders or public figures.[78,79] Alternatively, Ukraine has introduced an AI-generated spokesperson to provide information on consular affairs, as interviewed by Radio Free Europe.[80]

Because deepfakes can produce realistic yet entirely fabricated content, their existence can derail the implicit trust in media and communications that societies rely on.

Because deepfakes can produce realistic yet entirely fabricated content, their existence can derail the implicit trust in media and communications that societies rely on. This presents unique challenges for public services in safeguarding public trust and legitimacy. The scale of the issue was pinpointed by EUROPOL, which estimated that by 2026, 90% of social media content will be synthetically generated, launching what some experts call an “information apocalypse.”[81]

In such an environment, public trust in evidence – including images, video, and audio – may erode, leading to widespread apathy where people struggle to determine what is real and what is not. This erosion of trust is particularly concerning for law enforcement, as credibility and transparency are foundational to its role in maintaining public safety and justice.

Bots and automated accounts further amplify this challenge by spreading false narratives at an unprecedented scale and pace. AI-generated content paired with bot networks can create and propagate misleading stories about police actions or incidents, making it difficult to correct the record once false information goes viral. This amplification can shape public perception in harmful ways, fostering distrust in law enforcement, inflaming tensions in already divided communities, and undermining legitimate efforts to hold malicious actors accountable. If the public perceives the police as an unreliable source of truth – or as complicit in spreading or succumbing to disinformation – efforts to build bridges with the community could falter.

AI’s role in crafting hyper-realistic but fake material exacerbates this trust deficit. In the past, evidence like body-worn camera footage or surveillance video was widely seen as indisputable proof of events. However, as synthetic media becomes harder to detect, juries and the public may begin to question even verified evidence. This doubt can hinder the justice system, where the ability to prove facts is paramount, and may lead to scepticism of both law enforcement and the legal process itself. When people no longer believe what they see, they may stop believing altogether – a dangerous development in a society that relies on shared truths.[82]

When people no longer believe what they see, they may stop believing altogether – a dangerous development in a society that relies on shared truths.

The more convincing false media and deepfakes prove to be, the less open a population will be to believing any media, and the more influential propaganda could become.[83] Researchers on the Stanford University Deepfake Research Team recommended turning the same deep learning techniques used to create deepfakes against them. An automated inspection tool created using this premise was found to be 97% accurate when inspecting material.[84] However, this detection percentage is, unfortunately, considered to be much lower when the generating model and data source are unknown, such as for a random deepfake encountered online.

Another method to combat false media is known as blockchain. Blockchain is a tool which stores digital signatures and can be used to track changes and identify the original, raw data.[85] Vidprov – a blockchain programme developed by a Stanford University team – can track the version editing of videos.[86] By working backward, Vidprov can identify the parent video and the source which produced it. Videos from reputable sources are then given proof of authenticity. Another is DebunkBot, which has proven to successfully use AI to debunk conspiracy theories.[87] Despite these examples, technology to readily identify and counter deepfakes is not yet on par with the ability to create them. Until this gap is filled, deepfakes and other forms of disinformation will continue to spread and pose threats.[88]

Former UK Justice Secretary Sir Robert Buckland urged the government to do more to tackle what he sees as a “clear and present danger” to democracy, a view supported by former US Deputy Attorney General Lisa Monaco.[89] Ms Monaco emphasised how AI could supercharge disinformation in elections. She described AI as the “ultimate double-edged sword” because it can deliver profound benefits to society or be used to sow chaos and incite violence.[90]

This was evident in relation to the riots at the United States Capitol in 2021 and, more recently, the 29th of July 2024 knife attack in Southport, England, after which public disorder spread nationally as the far-right convened to protest the presence of migrants, particularly those from predominantly Islamic countries.[91] Notably, less than three hours after the stabbing attack, an AI-generated image was posted on social media platform ‘X’ by an account called Europe Invasion. It depicted bearded men in traditional Muslim dress outside the Houses of Parliament,

one waving a knife, behind a crying child in a Union Jack t-shirt. The tweet, which has since been viewed over 900,000 times, was captioned: "We must protect our children!"[92]

Far-right groups have promoted the malicious use of AI on social media platforms including Gab and 4chan. A Gab Telegram channel instructed white nationalist communities to "imagine the possibilities" of using AI deepfakes to create convincing, instantaneous media at no cost.[93] The message board site 4chan targeted the conflict in Israel and Gaza by utilising Bing's AI image generator to create Nazi media and propaganda.[94]

Islamist terrorist groups are also capitalising on the offerings of deepfakes. Accessed via the AI Incident Database, a 2024 Washington Post article titled *These ISIS news anchors are AI fakes: Their propaganda is real*, evidenced this shift. It referred to a video referencing the ISIS attack on a Russian concert venue in March 2024, using an ISIS newscast channel that now releases regular AI-generated dispatches. This use of AI enables the promotion of ISIS operations globally, creating professional propaganda videos that can be pushed at a higher volume and lower cost than ever before. Indeed, Pro-ISIS media outlets are seeking people to work with AI. "O mujahideen of media," a recruitment advert said; "the media is waiting for your attack."[95]

An al-Qaeda affiliated group also announced it would start hosting online AI workshops. The next day, it partnered with another al-Qaeda affiliate organisation to release a fifty-page guide titled *Incredible Ways to Use Artificial Intelligence Chat Bots*.[96] A big concern here is also the interaction between recipient and propagandist, as it is now possible to have LLM-powered deepfakes and virtual propagandists interact with users autonomously. Tasked with pushing a certain narrative or individualised propaganda, the potential impact on terrorism recruitment could be significant.

The use of bots in generating disinformation and driving malicious narratives is well documented.[97] Recent studies note that advances in AI, natural-language processing, and machine learning mean some bots can closely mimic human behaviour, generating original language and content.[98] These bots can then be exploited by malicious actors to spread disinformation in a more knowledgeable, persuasive, and dangerous manner, and to manipulate groups at scale. A foiled assassination plot against the British Queen Elizabeth in 2021, for example, was found to have been discussed and supported by a virtual chatbot.[99] Other chatbots have pushed antisemitic ideas and harmful, inappropriate content.[100]

This frames how groups of people can be misinformed, manipulated, and mobilised through AI-enabled technology. This is analysed further in the CTPN report on *Mis- and Disinformation: Extremism in the Digital Age*.[101] Indeed, in the contexts of radicalisation and extremism, society is already witnessing how AI can pave the way for acts of violence and terrorist attacks.

Generative AI has become centred around easy-to-use consumer products which require little expertise, thus lowering the barrier to entry. Andrew Rogoyski, a Director at the Institute for People-Centred AI, said these advances – with image, audio, and video generating tools now widely available online – mean "anyone can create anything."[102] The next step for deepfakes is the creation of entirely AI generated content that produces entirely synthetic realities from scratch with a single prompt. This is a significant danger from a terrorism point of view as highly realistic and contextual propaganda can be created easier than ever before. In addition, it is possible to generate deepfake videos in most languages, giving them a far greater reach.

This places an onus upon the developers of AI models to add restrictions. It also highlights the importance of the human eye for vigilance and oversight. It is the technical experts and specialist operators of AI-based tools that will give public authorities an advantage over malicious actors. But this demands national and city-level strategies and action plans; investment in highly trained teams; and the careful procurement of technology.

This correlates with the previous section on the physical threats posed by AI. Ultimately, the potential threats hinge on public-private sector accountability; how sophisticated software restrictions and technological guardrails are; and the capabilities to identify, contain and counter malicious actors, which involves harnessing the positive applications of AI.

Rita Katz of SITE Intelligence Group concluded her article *Extremist Movements are Thriving as AI Tech Proliferates* by writing, "I've tracked terrorists for 25 years now and I've seen the game-changing advantages they harnessed from Internet forums in the 2000s and then social media and smart phones in the 2010s. But what I'm seeing today with AI worries me far more. Those past shifts may pale in comparison to how AI can help extremists inspire attacks, sow hate, harass, and aggravate social divisions. We are sprinting into uncharted territory."[103]

On the other hand, the Chief Executive Officer of Google, Sundar Pichai, said: "Every technology shift is an opportunity to advance scientific discovery, accelerate human progress, and improve lives," noting that "AI will be the most profound in our lifetimes."[104]

6 Implications for Security, Preparedness, and City Operations

> **AI [is the] potential "ace up our sleeve" to ensure the safety and security of the global community.**

This report has explored some of the potential physical and political threats posed by AI. These threats have significant international, national, and local implications, but the focus is now on what they could mean for authorities in terms of security and preparedness at the city level. This includes both protection against AI-enabled threats and the protection of AI systems.

Ironically, this will likely require further investment in AI-based technologies but with clear regulation, governance, guardrails, and procedures for human oversight and control across all categories: machine learning, natural language processing, speech, vision, expert systems, and robotics. Smart cities are already using data, collected in large quantities and analysed by AI, to identify trends and improve safety and services.[105] To this end, the UK National Cyber Security Centre published *Guidelines for secure AI system development*, designed to help providers build safe and functional AI systems.[106]

A joint report by INTERPOL and UNICRI, titled *Towards Responsible AI Innovation*, frames AI as the potential "ace up our sleeve" to ensure the safety and security of the global community. The report contends AI to have the potential to "alter the very nature of policing and enhance efficiency and effectiveness to, for instance, identify persons of interest in crowded spaces; forecast and predict violence; automatically sort, tag and classify large police operational data such as evidence or harmful materials; and monitor drivers of radicalisation."[107]

The Potential of AI

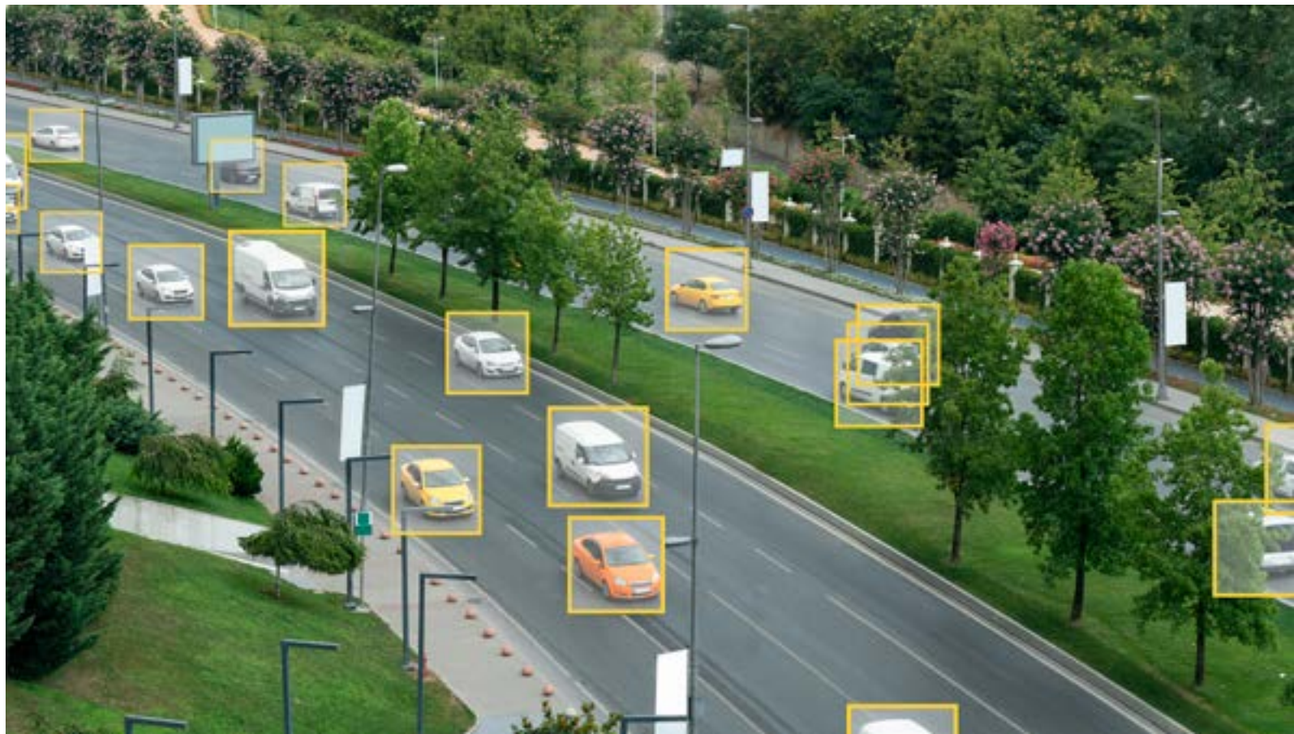
The Australian Federal Police worked with a deep learning model that could recognise, tag, and cluster harmful online material. This automated recognition was coupled with a data airlock system which helped reduce the exposure of officers to these materials.[108] It is easy to see the benefits of AI in detecting and removing extremist content from online platforms, or helping to identify at-risk individuals through online behaviour, and enhancing investigations.

To varying degrees, AI is already used in security tools and infrastructure (closed circuit television and control and operations rooms, for example) that play pivotal roles in the protection of cities. AI has revolutionised the areas of image and video processing to human and object recognition or movement tracking (consider facial or license plate recognition). AI-enabled visual processing systems can also be used to help identify abnormal behaviour and to facilitate or limit entry into specific buildings or closed events, such as concerts and festivals.[109]

The Government of Japan initiated cooperation with several technology companies in a bid to maximise major event security through AI analytics. Police in Tokyo were also building AI pilot projects focused on identifying areas of high crime risks to determine optimal patrol routes or crime prevention techniques, known as hot-spot mapping.[110] There are several examples of AI being used to help automatically populate maps for situational awareness. AI-enabled drones can also provide real-time aerial monitoring of large areas. Potential threats can then be transferred to police or security on the ground using interactive command platforms.[111]

This type of data transfer will become quicker as technology continues to advance, enabling real-time secure communications at high-speed to reduce costs, increase efficiency, and improve safety, as well as decision-making. To this end, 6G is expected to become one of the first AI-native networks that will integrate AI directly into the networking infrastructure.[112]

These types of operational efficiencies can apply to counter terrorism and optimisation for emergency services and city operations more broadly. They extend to route planning and resource dispatch.[113]



However, to understand and predict optimal decisions, "a system needs to know what 'optimal' is and how to calculate it. Furthermore, it needs to know when, where and how incidents occur; how resources can be deployed; how well deployed resources will perform; how long cases take; how other variables, such as traffic, day or time of the week and weather affect incident patterns and responses."[114] Therefore, truly optimal systems require a high-volume of accurate and verified data to succeed.

Predictive analytics could also be used. By analysing patterns in historical data, social media activity, and other intelligence sources, theoretically, AI could help predict potential terrorist attacks and inform interventions. AI can process intelligence and integrate large datasets from sources like surveillance footage, digital platforms, and financial transactions against potential threats, but this still demands 'human-in-the-loop' – an approach based on human input.[115]

AI-based predictive technology can also support planners to anticipate various scenarios and assess their impacts. One resource is the Global Terrorism Database (GTD). The GTD is an open-source database that compiles historical data on terrorist events across the globe starting in 1970 including the date, location, target, and perpetrators.[116] New models have been developed which use the abilities of machine learning to identify complex patterns in security data and predict casualties to support the planning and response of security operations.[117,118]

These types of techniques are increasingly employed in counter terrorism; for example, by classifying and correlating data to anticipate terrorist attacks.[119] However, it is important to note that despite the ability of AI platforms to detect patterns in complex datasets, they have limited ability to predict specific future events because of their inherently unpredictable nature.

Thus, whilst machine learning techniques could help to identify attack probabilities by target location and method, this approach demands caution. In practice, it is more about understanding patterns and trends rather than predicting. However, this alone may support agencies to improve their planning processes and better mitigate risks associated with terrorist attacks by further informing decision-making and strategic planning.

The opportunities for AI in security and preparedness more broadly are vast. IMPETUS (Intelligent Management of Processes, Ethics and Technology for Urban Safety) was a European project that sought to provide authorities with new means to improve the security of public spaces in smart cities. The project results included AI-enabled prototypes relating to firearms and bacteria detection, cyber threat intelligence and response, and emergency management. There was an evacuation optimiser that could provide advice to emergency staff on how to effectively manage an evacuation, based on simulations of different evacuation scenarios. Others included a social media detection tool for scanning online threats and an urban anomaly detector to continuously monitor and gather data from multiple city sensors to detect cases deviating from the norm that might indicate cause for concern.[120]

This also applies to aviation and maritime operations, which have benefitted significantly from AI technologies. In aviation, AI already "enhances safety and efficiency through predictive maintenance, aiding air traffic management, and refining pilot training with advanced insights and simulations."[121] In this setting, AI continues to be embraced as an opportunity to transform flight operations, airspace management, airport infrastructure, and border security.

In the ISOLA project, funded by the European Union, AI was used to support decision-making for security incidents on-board passenger ships, fusing multiple input data streams to detect threats and identify appropriate actions for the crew, security staff, and passengers with a view to improving safety for all.[122] Transport for London has otherwise said it is prioritising work on new AI technology designed to keep passengers safe on station platforms.[123] The adoption of AI in video surveillance and security in these environments is increasing.

The NATO DEXTER (detection of explosives and firearms to counter terrorism) project is a flagship initiative of the NATO Science for Peace and Security Programme. Using AI-enabled technologies, it works to identify firearms and explosives among moving pedestrians, remotely and in real-time. A large-scale trial of the technology took place in a subway station in Rome, where the system also proved its ability to anonymise identities, instead focusing on detecting anomalies.[124] Intelligent video surveillance, tested in Australia, could also blur faces.[125] These types of technologies offer a step-change in transport hub security where large and dense crowds can be vulnerable as they converge around waiting areas or pinch-points like ticket barriers.

The identification of anomalies transfers to broader city operations, crowd movement dynamics, and behaviour. In the context of securing busy urban areas or major events, these types of tools could prove highly beneficial. Added security layers, such as biometrics (systems that use the biological characteristics of an individual to verify their identity), scanning equipment, and real-time alerts for weapon identification, offer investment options for authorities.

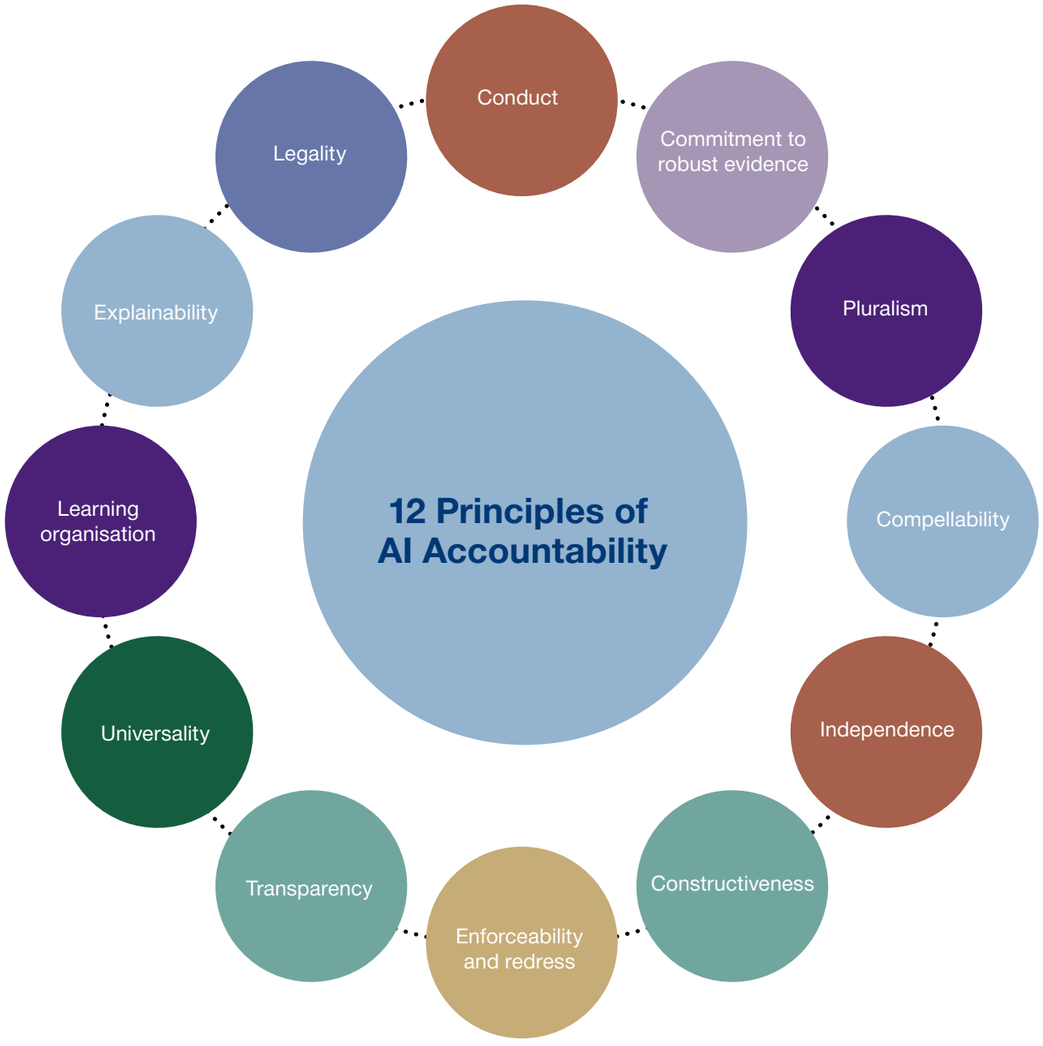
However, they all demand robust regulation, processes for testing andprocurement, and watertight approaches toward governance, ethics, and implementation.

In addition, developing technologies could help with victim identification following a terrorist attack by matching physical descriptors collected from witnesses at an attack site with descriptions of missing people provided by loved ones. In the future, this type of technology could expand to hospitals and waiting rooms to improve service delivery for victims' family and friends.

Another example of the positive use of AI was the development of an "aidbot" that sought to narrow the gap between the demand and supply of aid for displaced people in Lebanon. This localised initiative was designed to communicate with its users online via WhatsApp. It was programmed to ask simple questions about the types of aid people required along with their names and locations. This information was then recorded onto a Google spreadsheet to assist in the distribution of food, blankets, medicine, and clothes.[126]

These types of approaches can help optimise real-time identification, information flows, and decision-making.

More broadly, smart traffic lights and motorways demonstrate how AI innovations are helping to address traffic congestion.[127,128] Wireless sensor networks can detect the number of vehicles per lane, compare them, and adjust traffic lights or alter speed restrictions to reduce backup. Emergency vehicles fitted with wireless communication tags can ride "green waves", ensuring they receive green lights.[129] In addition, rail networks use AI in maintenance and inspection. This includes defect and fault detection, failure prediction, and maintenance planning.[130] There have also been large gains in healthcare, as well as ample reported benefits spanning education, industry, services, and lifestyle convenience.[131,132,133,134] As cities continue to become 'smarter', their reliance on technology and AI will also continue to grow.



> **As cities continue to become 'smarter', their reliance on technology and AI will also continue to grow.**

There are multiple known and unknown implications for security, preparedness, and city operations. The application of AI technologies in this regard must be underpinned by legal frameworks and ethical considerations, as well as common approaches towards procurement policies and practices for AI-based tools reinforced by standards and duties upon public authorities. In October 2023, the International Citizen Consultation on AI Accountability in Policing published *Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain*. This project sought to assign responsibility for AI and define the auditing of its systems. AP4AI proposed twelve principles for AI accountability: conduct, legality, explainability, commitment to robust evidence, learning organisation, universality, transparency, enforceability and redress, constructiveness, independence, compellability, and pluralism.[135]

Public services must apply such principles and take additional steps to safeguard public trust. These include adopting innovative technologies to authenticate evidence and collaborating with independent experts to verify the integrity of critical footage. Transparency is also key; openly explaining how evidence is gathered, verified, and protected against tampering can help reassure the public. Educating government officials and communities about the dangers of synthetic media and partnering with technology companies to detect and counteract AI-generated content are other vital strategies. Perhaps most importantly, law enforcement must demonstrate its commitment to truth and accountability every day, especially when illusion is increasingly hard to distinguish from reality. This means earning legitimacy and trust with the public and maintaining positive relationships with traditional media and community leaders. There is, of course, the need for the mainstream media and public authorities to expose and prosecute disinformation that contributes towards extremism and violence in any form.

> **Authorities should consider AI within risk and threat assessments, grouped with other Emerging Disruptive Technologies (EDTs) like cyber, drones, and 3D-printing.**

Likewise, authorities should consider AI within risk and threat assessments, grouped with other Emerging Disruptive Technologies (EDTs) like cyber, drones, and 3D-printing. Establishing expert working groups that convene multi-agency partners accountable for the ongoing review, planning, and mitigation for potential EDT threats should also be considered.

There is a further need to link threat intelligence and cyber security with multi-agency communication structures in the context of disinformation, including deepfakes and chatbots. The capability to identify, understand, and address these in a credible, timely, and authoritative manner will become increasingly important in minimising negative impacts, reputation management, as well as for public awareness and reassurance.

Within the smart city context, further consideration should be given to the role of built environments and how they can be used for – and protected from – concerted attacks by malicious actors (including by hostile states or through cyber warfare). Infrastructures, and their interactions with the people using them – living within or around them – need to be recognised as vulnerabilities. Collaboration with experts on infrastructure design, engineering, and architecture will be increasingly important to understand, predict and counter these vulnerabilities.

Finally, community engagement in relation to the threats posed by AI is necessary, especially in the online environment. This could include generic campaigns, targeted training for vulnerable groups, and sessions integrated into school curriculums. These should embody awareness of data privacy and human rights to help protect people online.

7 Data Privacy and Human Rights

> **“Without... safeguards, the risk of biased systems exacerbating existing inequalities or causing secondary societal consequences remains a serious threat to the future of AI and our security.”**

Maintaining the privacy of personal data and ensuring human rights are core concerns and challenges when it comes to the application of AI. Whilst this report is not intended to address these issues, they must be recognised as they are deserving of dedicated urgent attention.

Data privacy and human rights are absolute priorities with significant implications for the developers of AI-software and the public authorities that may use it. Indeed, there is a fundamental need to continually balance threats to human rights against technological opportunity.

The way people use tools like Google and Facebook show that most are prepared to sacrifice privacy for the benefits they provide. The case is similar for the big LLMs (ChatGPT, Claude, Gemini, etc.), with others including DeepSeek also having remarkable popularity.

This leans into the importance of robust data protection measures, including encryption, access controls, and strict governance policies; compliance with legal frameworks and international data protection regulations; and ethical approaches towards safety, responsibility, and transparency. To this end, the Oslo Police District of Norway explored non-intrusive surveillance, testing a system in which pattern recognition (for example, trends in human behaviours) was combined with the means of automated anonymisation.[136] It is possible that this “fusion of AI and biometrics can enhance criminal identification accuracy while protecting the privacy of nonrelevant individuals.”[137] This is a progressive approach that seeks to combine the benefits of AI with ethical application, but there is a long way to go.

There is also a need to mitigate bias in AI decision-making, which is necessary to build trust and ensure the responsible deployment of AI technologies. The inability to account for the complexities of human relationships is a single point of failure for AI. People come with their own biases and prejudices which must be taken into account.[138] The introduction of bias – be it consciously or unconsciously – into AI is a significant risk that can undermine the fairness, accuracy, and reliability of machine learning.

Bias in AI can be divided into three main types: human-induced or world bias, data driven bias, and algorithmic or machine self-learning bias. World bias is an accidental bias introduced while the system is being trained, resulting from societal inequalities and prejudices. Data bias is created by skewed or incomplete training data which then causes the system to come to a misguided conclusion. Algorithmic bias is self-produced. As the system changes and checks itself, missed incorrect connections can be reinforced, eventually snowballing to create algorithmic bias in the data.[139] These biases will intersect and build on one another, eventually leading to an AI system which makes unfair and/or unethical decisions.

For example, if a set of training data over-represents a certain group of people when identifying terrorist threats, the model trained using that data could then overstate the threat that group poses. Recommendations based on the model, especially if monitoring or targeting individuals, could result in racial or ethnic profiling, or prejudice against specific groups. Such occurred in Australia where a tool designed to predict future crime in terrorist offenders considered them at greater risk of offending if they were autistic despite having no empirical basis to do so. A report, titled *Testing the Reliability, Validity and Equity of Terrorism Risk Assessment Instruments,* found the tool to have “potentially serious implications for [its] validity and reliability” and found it was “extremely poor” at predicting risk.[140]

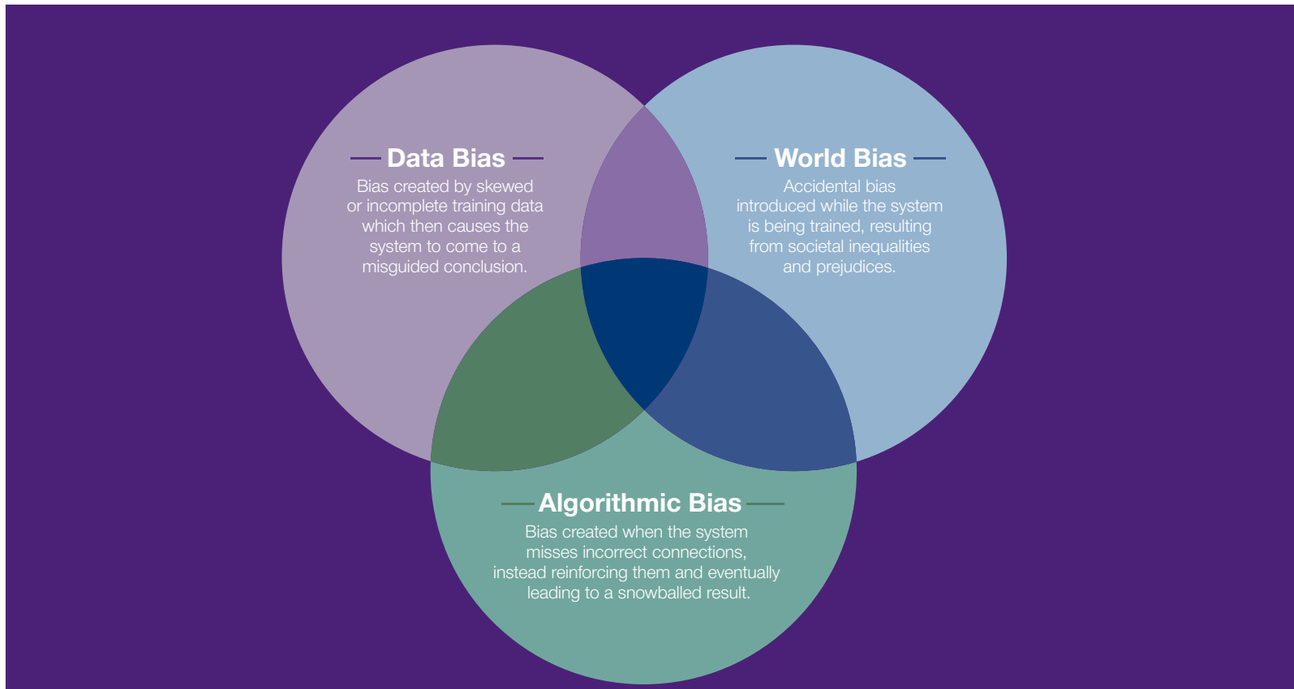
It is for this reason that the black box effect, as mentioned earlier, is so concerning. If the decision-making process cannot be seen, then there is no way for humans to trace and correct mistakes.[141] These issues lean into serious moral, ethical, and human-rights considerations whilst raising questions regarding the appropriateness of AI in certain contexts.[142]

Take, for example, surveillance technologies, including speech and facial recognition systems. They are crucial to counter terrorism. These systems rely on machine learning algorithms to accurately identify individuals and track suspicious activities. However, it is important to emphasise that the accuracy rate of current capabilities is variable, and there are false-positive cases due to data/algorithmic bias. This can pose major human rights violations when it comes to the use of unreliable or inaccurate AI systems, and raises further concerns about data collection, storage, and potential misuse. This makes the use of AI, particularly in security and threat profiling, extremely sensitive and political.

That said, human behaviour can also be regarded as biased. A person’s thoughts and actions cannot always be explained in terms of how their brain is wired; yet humans clearly can learn, be trained, and change. Similarly, AI outputs can be refined using RLHF (Reinforcement Learning from Human Feedback) without needing to understand precisely how the AI arrived at its outputs. This line of thought implies that truly bias-neutral AI systems may be hard to achieve in practice, suggesting robust governance and accountability are the key.

Therefore, responsible public authorities and governments must ensure the meticulous development

The Three Types of Bias in AI



and procurement of any AI-based tools and always employ the human-oversight rule. This requires close cooperation with the private sector and suppliers; extensive testing, verification, and auditing; comprehensive training and awareness for personnel; as well as robust checks and balances for verification before application.

This must be underpinned by legislation and regulation and framed by clear policies and procedures. Without such safeguards, the risk of biased systems exacerbating existing inequalities or causing secondary societal consequences remains a serious threat to the future of AI and our security.

Only by purposefully ensuring datasets are high-quality, diverse, and representative; anonymising and securing personal data both before an AI system is used and after; building transparency into algorithm decision-making; and ensuring the continuous monitoring of applications, can AI systems begin to operate more fairly, ethically, and equitably.

The United Nations Educational, Scientific and Cultural Organisation (UNESCO) report on the *Ethics of Artificial Intelligence* highlighted that AI must maintain human rights, fundamental freedoms, and human dignity to ensure it does not reinforce discrimination or bias.[143] These principles apply to national and city levels, as well as organisations using AI-technology. They should help form the foundation of the inevitable growth in AI, driven and upheld only by true accountability. One project, AI Accountability for Policing and Security (AIPAS), aims to design the practical mechanisms and software tools needed to assess and implement AI Accountability for AI applications.[144] It is intended to support AI accountability during deployments and the design and procurement stages. These initiatives warrant further attention.

8

Conclusion and Recommendations

“The potential benefits of creating intelligence are huge. We cannot predict what we might achieve, when our own minds are amplified by AI.”

This report has explored AI, largely focusing on its potential uses in terrorism versus opportunities for application in security, preparedness, and city operations. The ongoing debate about whether the risks of AI technologies surpass its rewards, or vice versa, underscores the critical need for societal discourse and informed decision-making. The undeniable fact, however, is that AI technologies are rapidly being integrated into a wide range of digital technologies and applications. To follow New York Police Department Deputy Commissioner for Counter Terrorism, Rebecca Weiner, society is “at the curve.”[145] The way society leans today, will determine tomorrow.

In the words of Stephen Hawking, “The potential benefits of creating intelligence are huge. We cannot predict what we might achieve, when our own minds are amplified by AI.”[146] Unfortunately, this works both ways. The transformative capability of AI offers waves of creativity and innovation, and AI has been found to have the most impact potential on technological improvements relating to the United Nations’ Sustainable Development Goals.[147,148] But its potential to progress society for the better is matched by its capacity to accelerate violent extremism and terrorism, which will be hard to trace, contain, and counter.

In recent years, extremist propaganda has become more interactive. “Extremist video games, social media content, and music have found their way onto a variety of internet platforms” with games developed by neo-Nazi groups encouraging “players to engage in violent behaviour towards minorities from a first-person shooter perspective.”[149] AI takes this to a new level in terms of AI-generated media and propaganda that could further contribute to the growth of extremism and serve as a direct enabler for physical attacks.

There are three potential relationships between AI and terrorism: (1) AI is restricted through countermeasures, policy interventions, or technological failures, (2) AI is leveraged to prevent and counter terrorism, and (3) AI is exploited to facilitate or execute terrorist activities.[150] Of course, these are not mutually exclusive, and all could happen concurrently.

How AI-enabled threats evolve – and how its benefits are harnessed – remain to be seen. Regardless, public authorities must stay ahead of the curve by assessing these threats in terms of planning, preparedness, and security.

The necessity to counter AI-enabled threats with AI-based solutions highlights the complexities of the AI security dilemma.

From real-time situation mapping to surveillance and behavioural analysis; from simulation training to predictive policing; from data analysis, integration, and prioritisation to risk profiling and report writing, AI offers many potential benefits. It absolutely has the potential to actively contribute to more prepared, secure, resilient, and adaptive societies and can be harnessed to enhance services, bolster safety and security, and facilitate city operations.

What is clear is that the effectiveness and safety of AI is dependent upon governance and accountability; the sophistication of its own technology; the depth and accuracy of its datasets; and how it is programmed, applied, and operated, and by whom. This context is key, given that society is already accelerating along a digital trajectory from which there is little or no return.

This reality has resulted in several high-level documents. In 2017, UNESCO published a report on robot ethics, followed by an Ethical AI Framework in 2021. A 2019 framework by the European Commission sought to define "Trustworthy AI", and the European Parliament released three resolutions to push against giving AI systems personalities in 2020. [151] In October 2023, the International Citizen Consultation on AI Accountability in Policing published *Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain* as noted earlier, and in March 2024, the European Parliament approved the *Artificial Intelligence Act*.[152,153]

The AI Act, the first of its kind, is a full regulatory framework that establishes boundaries for AI, based on varying levels of risk. AI systems with unacceptable risk are detailed as those which threaten individual rights or take advantage of individual vulnerabilities. They are banned with the only exception being for law enforcement purposes. The AI Act further outlines transparency and testing requirements.[154]

Also, in March 2024, the UN General Assembly adopted its first resolution on AI, titled *Seizing the opportunities of safe, secure, and trustworthy artificial intelligence systems for sustainable development* which accompanied the UN Secretary Generals Pact of the Future.[155] This Global Digital Compact "includes the first truly universal agreement on the international governance of artificial intelligence."[156]

At a national level, governments have established AI Safety Institutes and declared AI Action Plans.[157,158] This shows demonstrable progress but there remain many gaps, threats, and challenges. The stampede to enjoy the benefits of AI is likely to run roughshod over efforts to regulate and control it. Indeed, the current effectiveness of legislation and regulation is questionable given that it is outpaced by technological advancements and the fact that malicious actors will often find ways to work around these.

The position on AI also varies significantly from nation-to-nation, and, by its very nature, AI transcends borders and is thus very hard to control by individual nation-states. Therefore, it is even more important that any regulation over AI ensures meaningful and localised human control. In practice, this means that brakes need to be built into AI technologies to force human oversight and evaluation.[159] Efforts should also focus on measures that impede the adverse consequences of AI enabled technology, such as software restrictions.[160]

A top-down regulatory approach must meet bottom-up responsibilities as these will fall on the shoulders of a culmination of stakeholders – including developers, suppliers, data custodians, and regulatory bodies – to decide AI's boundaries and acceptable limits of accountability.

By extension, there is an onus upon public authorities to ensure they are operating with the public interest at heart. That is to identify, contain, and counter the threats posed by AI and harness the tools offered in proportionate, legal, ethical, transparent, and appropriate ways. For city administrations, services, and operations, this could – over time – result in unprecedented change. Whether that is for better or worse depends on decisions made today.

Recommendations

Note: This is an international report designed for an international audience at a city level. It is recognised that arrangements and resources will differ from city-to-city. It is therefore anticipated and accepted that different recommendations will apply in different contexts. The list is non-exhaustive, and stakeholders are advised to consult the relevant lead agencies and specialist authorities whilst aligning with local, national, and international policy.

1	Prioritise ethical AI development rooted in governance, accountability, human rights, and data privacy. Advocate for transparency in all matters, including operating procedures.
2	Work toward collectively strengthening the legislation, regulation, and standards surrounding AI. This should include public-private cooperation to ensure the necessary safeguards or guardrails are built into software, online platforms, and technologies.
3	Cultivate professionals in the science, technology, engineering, architecture, urban design, and math fields to build a pool of technical experts that can provide human oversight.
4	Develop situational awareness and key risk indicators to identify threat actor use of AI.
5	Establish a multi-agency expert working group that is accountable for monitoring, reviewing, and working to mitigate, as well as prepare for, threats posed by AI and other Emerging Disruptive Technologies. This would be beneficial at national and city levels.
6	Plan for specific threats like AI-generated swatting and disinformation campaigns.
7	Review organisational policies and procurement procedures for AI-based technologies.
8	Invest in AI security and preparedness solutions by implementing advanced technologies designed to identify and mitigate threat actors or support response and city operations.
9	Enhance internal business continuity arrangements to include managing the impacts of disinformation or chatbots targeting an organisation or individuals within it.
10	Bolster security vetting and awareness training for staff to minimise insider threats.
11	Understand the potential for artificial insiders and build infrastructure to counter these.
12	Connect threat intelligence and multi-agency communication teams in the context of identifying and responding to disinformation in a timely manner.
13	Collaborate with trusted media partners in understanding the impacts of widespread disinformation and how this can be managed from a communications perspective.
14	Ensure communication teams can act quickly to dispel disinformation generated by AI.
15	Educate employees, elected officials, and community leaders about the risks of AI, and conduct community training and awareness campaigns to inform the public.
16	Horizon scan for developments in AI, identify new and emerging threat vectors, and work with cross sector stakeholders to address these.
17	Train and exercise against scenarios that incorporate AI on a regular basis.
18	Participate in related research and innovation projects to capture and share best practices.

Reference List

1. United Nations Office of Counter-Terrorism, United Nations Interregional Crime and Justice Research Institute. (2021). 'Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes', *United Nations*.
2. Janjeva, A., Harris, A., Mercer, S., Kasprzyk, A., and Gausen, A. (2023). 'The Rapid Rise of Generative AI: Assessing risks to safety and security.' *The Alan Turing Institute: Centre for Emerging Technology and Security*. pp.38.
3. CTPN. (2022). 'City Preparedness for Cyber-Enabled Terrorism', *Counter Terrorism Preparedness Network*.
4. CTPN. (2023). 'Mis- and Disinformation: Extremism in the Digital Age', *Counter Terrorism Preparedness Network*.
5. CTPN. (2024). 'Preparing for Hostile Drones in Urban Environments', *Counter Terrorism Preparedness Network*.
6. Samoili, S., López Cobo, M., Delipetrev, B., Martínez-Plumed, F., Gómez, E., and De Prato, G. (2021). 'AI Watch: Defining Artificial Intelligence 2.0. Towards an operational definition and taxonomy for the AI landscape', *Publications Office of the European Union*, Luxembourg, JRC126426, pp. 10.
7. Samoili, S., López Cobo, M., Delipetrev, B., Martínez-Plumed, F., Gómez, E., and De Prato, G. (2021). 'AI Watch: Defining Artificial Intelligence 2.0. Towards an operational definition and taxonomy for the AI landscape', *Publications Office of the European Union*, Luxembourg, JRC126426, pp.10.
8. UK Government. (n.d.). 'AI and cyber security: what you need to know', *National Cyber Security Centre*.
9. Hearth, M. and Mittal, M. (2022). 'Adoption of Artificial Intelligence in Smart Cities: A Comprehensive Review', *International Journal of Information Management Data Insights,* 2.
10. Caldwell, M., Andrews, J. T. A., Tanay, T., and Griffin, L. D. (2020). 'AI-Enabled Future Crime', *Crime Science*, 9:14.
11. Zhang, C. and Yang, L. (2021). 'Study on Artificial Intelligence: The State of the Art and Future Prospects', *Journal of Industrial Information Integration*, 23.
12. SAP. (n.d.). 'The Complete History of AI', *LeanIX*.
13. Amazon Web Services. (n.d.). 'What is LLM (Large Language Model)?'
14. Caldwell, M., Andrews, J. T. A., Tanay, T., and Griffin, L. D. (2020). 'AI-Enabled Future Crime', *Crime Science*, 9:14.
15. Amazon Web Services. (n.d.). What is LLM (Large Language Model)?
16. Rai, A. (2020). 'Explainable AI: From Black Box to Glass Box', *Journal of the Academy of Marketing Science*, 48, pp. 137-141.
17. Li, B., and Gilbert, S. (2024) 'Artificial Intelligence awarded two Nobel Prizes for innovations that will shape the future of medicine', *npj Digital Medicine*, 7:336.
18. UK Government. (2025). 'International AI Safety Report', *Department for Science, Innovation and Technology and AI Safety Institute*, pp. 11.
19. Hearth, M. and Mittal, M. (2022). 'Adoption of Artificial Intelligence in Smart Cities: A Comprehensive Review', *International Journal of Information Management Data Insights*, 2.
20. CTPN. (2022). 'City Preparedness for Cyber-Enabled Terrorism', *Counter Terrorism Preparedness Network*.
21. European Union. (2024). 'Terrorism Situation and Trend Report (TE-SAT)', pp. 7.
22. European Union. (2024). 'Terrorism Situation and Trend Report (TE-SAT)', pp. 7.
23. ENISA. (2024). 'ENISA Threat Landscape 2024', *European Union Agency for Cybersecurity (ENISA).*
24. UK Government. (2025). 'International AI Safety Report', *Department for Science, Innovation and Technology and AI Safety Institute,* pp.18.
25. CTPN. (2021). 'Bioterrorism: Applying the Lens of COVID-19', *Counter Terrorism Preparedness Network*, pp. 28.
26. EUROPOL. (2023). 'ChatGPT - the impact of Large Language Models on Law Enforcement', pp. 7.
27. Nelu, C. (2024). 'Exploitation of Generative AI by Terrorist Groups', *The International Centre for Counter-Terrorism (ICCT).*
28. Venegas, N. (2025). 'Matthew Livelsberger Used Generative AI in Planning Truck Blast: Police', *NewsWeek*.
29. NBC News. (2024). 'How the NYPD is fighting AI-powered terrorism threats and criminal activity. Tom Winter interviews NYPD Deputy Commissioner Rebecca Weiner.'
30. Khan, F. A., Li, G., Khan, A. N., Khan, Q. W., Hadjouni, M., and Elmannai, H. (2023). 'AI-Driven Counter-Terrorism: Enhancing Global Security Through Advanced Predictive Analytics', *Digital Object Identifier*, v. 11.
31. David, A. (2021). 'Artificial Intelligence and the Fight Against International Terrorism', *National Military Intelligence Foundation*, 38:2, pp. 63-73.
32. R. E. Hall, J. Braverman, J. Taylor, and H. Todosow, "The Vision of A Smart City.," in 2nd International Life Extension Technology Workshop (Paris, France, Sep 28), 2000, pp. 1–6.
33. Randall, T. (2015). 'The Smartest Building in the World: Inside the connected future of architecture', *Bloomberg*.
34. P. S. Bayerl and V. Butot, "Smart City Configurations: A Conceptual Approach to Assess Smart City Practices and Outcomes," 2021 6th International Conference on Smart and Sustainable Technologies (SpliTech), Bol and Split, Croatia, 2021, pp. 01-07.
35. Schwarzschild, A., Goldblum, M., Gupta, A., Dickerson, J. and Goldstein, T. (2021). 'Just How Toxic is Data Poisoning? A Unified Benchmark for Backdoor and Data Poisoning Attacks,' *Proceedings on Machine Learning Research*, 139.
36. CTPN. (2022). 'City Preparedness for Cyber-Enabled Terrorism', *Counter Terrorism Preparedness Network*.
37. Caldwell, M., Andrews, J. T. A., Tanay, T., and Griffin, L. D. (2020). 'AI-Enabled Future Crime', *Crime Science*, 9:14.

38. Caldwell, M., Andrews, J. T. A., Tanay, T., and Griffin, L. D. (2020). 'AI-Enabled Future Crime', *Crime Science*, 9:14.
39. Martin, P., and Mercer, S. (2025). 'We Need to Talk About the Insider Risk from AI', *RUSI*.
40. Martin, P., and Mercer, S. (2025). 'We Need to Talk About the Insider Risk from AI', *RUSI*.
41. Sjouwerman, S. (2024). 'How a North Korean Fake IT Worker Tried to Infiltrate Us', *KnowBe4*.
42. Pledger, T. (2021). 'The Role of Drones in Future Terrorist Attacks', *The Association of the United States Army*, 137.
43. Kallenborn, Z., Ackerman, G., and Bleek, P. (2023). 'A Plague of Locusts? A Preliminary Assessment of the Threat of Multi-Drone Terrorism', *Terrorism and Political Violence*, 35:7, pp. 1556-1585.
44. Ro, C. (2023). 'On the warpath: AI's role in the defence industry', *BBC News*.
45. Cramer, M. (2021). 'A.I. Drone May Have Acted on Its Own in Attacking Fighters, U.N. Says', *The New York Times*.
46. United Nations. (2021). 'Libya arms embargo 'totally ineffective': UN expert panel'.
47. CTPN. (2024). 'Preparing for Hostile Drones in Urban Environments', *Counter Terrorism Preparedness Network*.
48. Ro, C. (2023). 'On the warpath: AI's role in the defence industry', *BBC News*.
49. Ganor, B. (2021). 'Understanding the Motivations of "Lone Wolf" Terrorists', *Perspectives on Terrorism*, 15:2, pp. 23-32.
50. Bathla, G., Bhadane, K., Singh, R., Kumar, R., Aluvalu, R., Krishnamurthi, R., Kumar, A., Thakur, R., and Basheer, S. (2022). 'Autonomous Vehicles and Intelligent Automation: Applications, Challenges, and Opportunities', *Mobile Information Systems.*
51. Tencent Keen Security Lab. (2019). 'Experimental Security Research of Tesla Autopilot.'
52. Hao, K. (2019). 'Hackers trick a Tesla into veering into the wrong lane', *MIT Technology Review.*
53. Villasenor, J. (2015). 'Jeep Cherokee hack offers important lessons on the "Security of Things"', *Forbes*.
54. Greenberg, A. (2015). 'Hackers Remotely Kill a Jeep on the Highway—With Me in It', *Wired*.
55. Caldwell, M., Andrews, J. T. A., Tanay, T., and Griffin, L. D. (2020). 'AI-Enabled Future Crime', *Crime Science,* 9:14.
56. Bathla, G., Bhadane, K., Singh, R., Kumar, R., Aluvalu, R., Krishnamurthi, R., Kumar, A., Thakur, R., and Basheer, S. (2022). 'Autonomous Vehicles and Intelligent Automation: Applications, Challenges, and Opportunities', *Mobile Information Systems.*
57. Slater, J. (2023). 'Coordinated 'swatting' effort may be behind hundreds of school shooting hoaxes', *The Washington Post.*
58. Andone, D. (2019). 'Swatting is a dangerous prank with potentially deadly consequences. Here's what you need to know', *CNN News*.
59. Prasad, M. (2024). 'Swatting: A Fictitious Threat Generating Real-World Hazards', *Global Network on Extremism and Technology.*
60. Jaquish, C. (n.d.). 'AI and Swatting: Navigating Technology's Impact', *Future Policing Institute.*
61. Ward, J., and Kolodny, L. (2023). 'The FBI has formed a national database to track and prevent 'swatting'', *NBC News.*
62. Aleksandrowicz, M. and Williams, A. (2025). 'Europol warns of AI-driven crime threats', *Reuters*.
63. United Nations Office of Counter-Terrorism, United Nations Interregional Crime and Justice Research Institute. (2024). 'Beneath the Surface: Terrorist and Violent Extremist use of the Dark Web and Cybercrime-as-a-Service for Cyber-Attacks', *United Nations*.
64. Staniforth, A. (2023). 'Hybrid threat vectors: Artificial intelligence to counter digital disinformation', *Crisis Response Journal*, 18:2.
65. Appel, M. and Preitzel, F. (2022). 'The Detection of Political Deepfakes', *Journal of Computer-Mediated Communication*, 27:4.
66. United Nations Office of Counter-Terrorism, United Nations Interregional Crime and Justice Research Institute. (2021). 'Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes', *United Nations*.
67. Appel, M. and Preitzel, F. (2022). 'The Detection of Political Deepfakes', *Journal of Computer-Mediated Communication, 27:4.*
68. Bazarkina, D. Y. and Pashentsev, E. N. (2020). 'Malicious Use of Artificial Intelligence', *New Psychological Security Risks in BRICS Countries*, 18:4, pp. 154-177.
69. Diakopoulos, N. and Johnson, D. (2021). 'Anticipating and Addressing the Ethical Implications of Deepfakes in the Context of Elections', *SAGE*, 23:7.
70. Collins, B. (2023). 'Fake picture of explosion at Pentagon spooks Twitter', *NBC News*.
71. Bregler, C., Covell, M., and Slaney, M. (1997). 'Video Rewrite: Driving Visual Speech with Audio', *Association for Computing Machinery Special Interest Group on Computer Graphics and Interactive Techniques*.
72. Chang, C. J. and Chien, W. C. (2024). 'Towards a Positive Thinking About Deepfakes: Evaluating the Experience of Deepfake Voices in Emotional and Rational Scenarios', Kurosu, M. and Hashizume, A. Human-Computer Interaction. Lecture Notes in Computer Science, pp. 311-325.
73. Diep Nep. (2021). 'This is not Morgan Freeman - A Deepfake Singularity', *YouTube*.
74. Wakefield, J. (2024). 'Tackling deepfakes 'has turned into an arms race'', *BBC News*.
75. Harris, J. (n.d.). 'I Deep Faked Myself, Here's Why It Matters.' *YouTube*.
76. Goodwin, L. (2024). 'Romance scammer duped £17k from me with deepfakes', *BBC News*.

77. Williams, A. (2023). 'AI clones child's voice in kidnapping scam', *The Standard*.

78. Bayer, J., Holznagel, B., Lubianiec, K., Pintea, A., Schmitt, J., Szakacs, J., and Uszkiewicz, E. (2021). 'Disinformation and Propaganda: Impact on the Functioning of the Rule of Law and Democratic Processes in the EU and its Member States', *European Parliament*, PE 653.633.

79. Bazarkina, D. Y. and Pashentsev, E. N. (2020). 'Malicious Use of Artificial Intelligence', *New Psychological Security Risks in BRICS Countries*, 18:4, pp. 154-177.

80. YouTube (no date). 'We Interviewed Ukraine's New AI-Generated Spokesperson', *Radio Free Europe*.

81. EUROPOL. (2022). 'Facing reality? Law enforcement and the challenge of deepfakes, an observatory report from the Europol Innovation Lab', *Publications Office of the European Union,* Luxembourg.

82. Twomey, J., Ching, D., Aylett, M., Quayle, M., Linehan, C., and Murphy, G. (2023). 'Do Deepfake Videos Undermine our Epistemic Trust? A Thematic Analysis of Tweets that Discuss Deepfakes in the Russian Invasion of Ukraine', *PLoS One*, 18:10.

83. Appel, M. and Preitzel, F. (2022). 'The Detection of Political Deepfakes', *Journal of Computer-Mediated Communication,* 27:4.

84. Yu, S. and Carrol, F. (2021). 'Implications of AI in National Security: Understanding the Security Issues and Ethical Challenges', Montasari, R. and Jahankhani, H., Artificial Intelligence in Cyber Security: Impact and Implications. *Springer*, pp. 157-172.

85. Krichen, M., Meryem, A., Mihoub, A., and Almutiqu, M. (2022). 'Blockchain for Modern Applications: A Survey', *Sensors*, 22:5274.

86. Yu, S. and Carrol, F. (2021). 'Implications of AI in National Security: Understanding the Security Issues and Ethical Challenges', Montasari, R. and Jahankhani, H., Artificial Intelligence in Cyber Security: Impact and Implications. *Springer*, pp. 157-172.

87. Costello, T., Pennycook, G., and Rand, D. (2024). 'Durably reducing conspiracy beliefs through dialogues with AI', *Science*, v. 385, 6714.

88. Bazarkina, D. Y. and Pashentsev, E. N. (2020). 'Malicious Use of Artificial Intelligence', *New Psychological Security Risks in BRICS Countries,* 18:4, pp. 154-177.

89. Corera, G. and Wheeler, B. (2023). 'Fears UK not ready for deepfake general election', *BBC News*.

90. Corera, G. and Wheeler, B. (2024). 'AI could 'supercharge' election disinformation, US tells the BBC', *BBC News*.

91. Krusch, M. (2024). 'Fomenting Civil War?: Disinformation Narratives after the Southport Attack', *Global Network on Extremism and Technology.*

92. Quinn, B. and Milmo, D. (2024). 'How TikTok bots and AI have powered a resurgence in UK far-right violence,' *The Guardian.*

93. Katz, R. (2024). 'SITE Special Report: Extremist Movements are Thriving as AI Tech Proliferates', *SITE Intelligence Group Enterprise.*

94. Kaplan, A. (2023). '4chan users are generating images with Nazi imagery and other "propaganda" via Microsoft Bing's AI tool', *MediaMatters.org*.

95. Washington Post. (2024). 'These ISIS news anchors are AI fakes. Their propaganda is real.', *The AI Incident Database*.

96. Washington Post. (2024). 'These ISIS news anchors are AI fakes. Their propaganda is real.', *The AI Incident Database*.

97. Tech Against Terrorism. (2024). 'How TikTok bots and AI have powered a resurgence in UK far-right violence.'

98. CTPN. (2022). 'City Preparedness for Cyber-Enabled Terrorism', *Counter Terrorism Preparedness Network*.

99. Weaver, M. (2023). 'AI chatbot 'encouraged' man who planned to kill queen, court told', *The Guardian*.

100. Wells, D. (2024). 'The Next Paradigm-Shifting Threat? Right-Sizing the Potential Impacts of Generative AI on Terrorism', *The Middle East Institute*.

101. CTPN. (2023). 'Mis- and Disinformation: Extremism in the Digital Age', *Counter Terrorism Preparedness Network*.

102. Quinn, B. and Milmo, D. (2024). 'How TikTok bots and AI have powered a resurgence in UK far-right violence,' *The Guardian.*

103. Katz, R. (2024). 'SITE Special Report: Extremist Movements are Thriving as AI Tech Proliferates', *SITE Intelligence Group Enterprise.*

104. Pichai, S. and Hassabis, D. (2023). 'Introducing Gemini: our largest and most capable AI model', *Google*.

105. Hearth, M. and Mittal, M. (2022). 'Adoption of Artificial Intelligence in Smart Cities: A Comprehensive Review', *International Journal of Information Management Data Insights,* pp.2.

106. UK Government. (n.d.). 'Guidelines for secure AI system development', *National Cyber Security Centre*.

107. INTERPOL and United National Interregional Crime and Justice Research Institute (UNICRI). (2020). 'Towards Responsible AI Innovation: Second INTERPOL-UNICRI Report on Artificial Intelligence for Law Enforcement', pp. 3.

108. INTERPOL and United National Interregional Crime and Justice Research Institute (UNICRI). (2020). 'Towards Responsible AI Innovation: Second INTERPOL-UNICRI Report on Artificial Intelligence for Law Enforcement', pp. 26.

109. INTERPOL and United National Interregional Crime and Justice Research Institute (UNICRI). (2020). 'Towards Responsible AI Innovation: Second INTERPOL-UNICRI Report on Artificial Intelligence for Law Enforcement', pp. 16-17.

110. INTERPOL and United National Interregional Crime and Justice Research Institute (UNICRI). (2020). 'Towards Responsible AI Innovation: Second INTERPOL-UNICRI Report on Artificial Intelligence for Law Enforcement', pp. 30.

9 **Reference List**
continued

111. United Nations Office of Counter-Terrorism, United Nations Interregional Crime and Justice Research Institute. (2021). 'Algorithms and Terrorism: The Malicious Use of Artificial Intelligence for Terrorist Purposes', *United Nations*.

112. EUROPOL. (2023). 'The Second Quantum Revolution – The impact of quantum computing and quantum technologies on law enforcement, Europol Innovation Lab Observatory Report', *Publications Office of the European Union*, Luxembourg, pp. 49.

113. INTERPOL and United National Interregional Crime and Justice Research Institute (UNICRI). (2020). 'Towards Responsible AI Innovation: Second INTERPOL-UNICRI Report on Artificial Intelligence for Law Enforcement', pp. 18.

114. INTERPOL and United National Interregional Crime and Justice Research Institute (UNICRI). (2020). 'Towards Responsible AI Innovation: Second INTERPOL-UNICRI Report on Artificial Intelligence for Law Enforcement', pp. 19.

115. Nelu, C. (2024). 'Exploitation of Generative AI by Terrorist Groups', *The International Centre for Counter-Terrorism (ICCT).*

116. National Consortium for the Study of Terrorism and Responses to Terrorism. (2022). 'Global Terrorism Database'.

117. Khan, F. A., Li, G., Khan, A. N., Khan, Q. W., Hadjouni, M., and Elmannai, H. (2023). 'AI-Driven Counter-Terrorism: Enhancing Global Security Through Advanced Predictive Analytics', *Digital Object Identifier,* v. 11.

118. Huamani, E. L., Alicia, A. M., and Roman-Gonzalez, A. (2020). 'Machine Learning Techniques to Visualize and Predict Terrorist Attacks Worldwide using the Global Terrorism Database', *International Journal of Advanced Computer Science and Applications*, 11:4.

119. Khan, F. A., Li, G., Khan, A. N., Khan, Q. W., Hadjouni, M., and Elmannai, H. (2023). 'AI-Driven Counter-Terrorism: Enhancing Global Security Through Advanced Predictive Analytics', *Digital Object Identifier,* v. 11.

120. IMPETUS. (2023). 'Intelligent Management of Processes, Ethics and Technology for Urban Safety, Project Results Booklet', *European Union's Horizon 2020 Research and Innovation Programme.*

121. UK Civil Aviation Authority. (2024). 'AI in Aviation: A Technology Outlook', *CAA Horizon Scanning and Insight,* pp. 2-6.

122. CENTRIC. (2024). 'Research to Reality Report. The Centre of Excellence in Terrorism, Resilience, Intelligence and Organised Crime Research (CENTRIC)', *Sheffield Hallam University*, pp.13.

123. Vickers, N. (2025). 'TfL aims to use more AI to improve platform safety', *BBC News*.

124. NATO. (n.d.). 'DEXTER: Detection of Explosives and firearms to counter TERrorism', *NATO Science for Peace and Security (SPS) Programme.*

125. Barthelemy, J., Iqbal, U., Qian, Y, Amirghasemi, M., and Perez, P. (2024). 'Safety After Dark: A Privacy Compliant and Real-Time Edge Computing Intelligent Video Analytics for Safer Public Transportation', *Sensors*.

126. Kansara, R. (2025). 'The woman who built an 'aidbot' for displaced people in Lebanon', *BBC News*.

127. Ahmed, O., and Ekanoye, F. (2023). "Intelligent Traffic Lights Based on Artificial Intelligence", *International Journal of Advancement in Education, Management, Science and Technology,* 6(2).

128. Hearth, M. and Mittal, M. (2022). 'Adoption of Artificial Intelligence in Smart Cities: A Comprehensive Review', *International Journal of Information Management Data Insights*, 2.

129. Yusuf, A., Arifin, A., and Zulkifli, F. (2020). 'Recent development of smart traffic lights', *IAES International Journal of Artificial Intelligence*, 10 (1), pp. 224-233.

130. Tang, R., Donato, L, Besinovic, N., Falmmini, F., Goverde, R., Lin Z., Liu, R., Tang, T., Vittorini, V., and Wang, Z. (2022). 'A literature review of Artificial Intelligence applications in railway systems', *Transportation Research Part C*.

131. Chen, L., Chen, P., and Lin, Z. (2020). 'Artificial Intelligence in Education: A Review', *Institute of Electrical and Electronics Engineers,* 6.

132. Racine, E., Boehlen, W., and Sample, M. (2019). 'Healthcare Uses of Artificial Intelligence: Challenges and Opportunities for Growth", *Healthcare Management Forum*, 32:5, pp. 272-275.

133. Nguyen, T., Larrivée, N., Bilaniuk, O., and Durand, R. (2021). 'Use of Artificial Intelligence in Dentistry: Current Clinical Trends and Research Advances', *Journal of the Canadian Dental Association*, 87:17.

134. UK Government. (2023). '£21 million to roll out artificial intelligence across the NHS', *Online Press Release*.

135. Bayerl, P., Obst, M., and Akhgar, B. (2023). 'Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain', *International Citizen Consultation on AI Accountability in Policing*, pp.3.

136. INTERPOL and United National Interregional Crime and Justice Research Institute (UNICRI). (2020). 'Towards Responsible AI Innovation: Second INTERPOL-UNICRI Report on Artificial Intelligence for Law Enforcement', pp. 24.

137. EUROPOL. (2023). 'The Second Quantum Revolution – The impact of quantum computing and quantum technologies on law enforcement, Europol Innovation Lab Observatory Report', *Publications Office of the European Union*, Luxembourg, pp. 10.

138. Yigitcanlar, T., Mehmood, R., and Corchado, J. (2021). 'Green Artificial Intelligence: Towards an Efficient, Sustainable and Equitable Technology for Smart Cities and Futures', *Sustainability*, 13, 8952.

139. Yu, S. and Carrol, F. (2021). 'Implications of AI in National Security: Understanding the Security Issues and Ethical Challenges', Montasari, R. and Jahankhani, H., Artificial Intelligence in Cyber Security: Impact and Implications. *Springer*, pp. 157-172.

140. Bucci, N., and Knaus, C. (2023). 'Australian terrorism prediction tool considered autism a sign of criminality despite lack of evidence', *The Guardian*.

141. Yu, S. and Carrol, F. (2021). ‘Implications of AI in National Security: Understanding the Security Issues and Ethical Challenges’, Montasari, R. and Jahankhani, H., Artificial Intelligence in Cyber Security: Impact and Implications. *Springer*, pp. 157-172.

142. Marcelline, M. (2022). ‘UK Police Use of Facial Recognition Fails to Meet ‘Legal And Ethical Standards’’, *PC Mag*.

143. United Nations Educational, Scientific and Cultural Organization. (2022). ‘Recommendation on the Ethics of Artificial Intelligence’. *United Nations*.

144. CENTRIC. (2024). ‘Research to Reality Report. The Centre of Excellence in Terrorism, Resilience, Intelligence and Organised Crime Research (CENTRIC)’, *Sheffield Hallam University,* pp.15.

145. NBC News. (2024). ‘How the NYPD is fighting AI-powered terrorism threats and criminal activity. Tom Winter interviews NYPD Deputy Commissioner Rebecca Weiner.’

146. UK Civil Aviation Authority. (2024). ‘AI in Aviation: A Technology Outlook’, *CAA Horizon Scanning and Insight*, pp. 3.

147. United Nations Department of Economic and Social Affairs. (2024). ‘The 17 Goals’.

148. Bolón-Canedo, V., Morán-Fernández, L., Cancela, B. and Alonso-Betanzos, A. (2024). ‘A Review of Green Artificial Intelligence: Towards a More Sustainable Future’, *Neurocomputing*, 599: 128096.

149. Siegel, D., and Doty, M. (2023). ‘Weapons of Mass Disruption: Artificial Intelligence and the Production of Extremist Propaganda’, *Global Network on Extremism and Technology.*

150. Caldwell, M., Andrews, J. T. A., Tanay, T., and Griffin, L. D. (2020). ‘AI-Enabled Future Crime’, *Crime Science*, 9:14.

151. Yu, S. and Carrol, F. (2021). ‘Implications of AI in National Security: Understanding the Security Issues and Ethical Challenges’, Montasari, R. and Jahankhani, H., Artificial Intelligence in Cyber Security: Impact and Implications. *Springer*, pp. 157-172.

152. Bayerl, P., Obst, M., and Akhgar, B. (2023). ‘Accountability Principles for Artificial Intelligence (AP4AI) in the Internal Security Domain’, *International Citizen Consultation on AI Accountability in Policing*.

153. Yakimova, Y. and Ojamo, J. (2024). ‘Artificial Intelligence Act: MEPs Adopt Landmark Law’, *European Parliament News*.

154. Yakimova, Y. and Ojamo, J. (2024). ‘Artificial Intelligence Act: MEPs Adopt Landmark Law’, *European Parliament News*.

155. United Nations. (2024). ‘Seizing the opportunities of safe, secure, and trustworthy artificial intelligence systems for sustainable development’, *UN General Assembly*, Seventy-eighth session, Agenda item 13.

156. Lederer, E. (2024). ‘UN nations endorse a ‘Pact for the Future,’ and the body’s leader says it must be more than talk’, *Associated Press World News*.

157. UK Government. (n.d). ‘AI Safety Institute’, *Gov.uk*.

158. McMahon, L., Kleinman, Z., and Edwards, C. (2025). ‘PM plans to ‘unleash AI’ across UK to boost growth’, *BBC News*.

159. Ro, C. (2023). ‘On the warpath: AI’s role in the defence industry’, *BBC News*.

160. Kreps, S. (2021). ‘Democratizing Harm: Artificial Intelligence in the Hands of Nonstate Actors’, *Foreign Policy at the Brookings Institute.*

CTPN
COUNTER TERRORISM
PREPAREDNESS NETWORK