UNIDIR

Securing Cyberspace for Peace

Insights into Cyberthreats and International Security in 2025

SECURITY AND TECHNOLOGY PROGRAMME

Acknowledgements

Support from UNIDIR's core funders provides the foundation for all of the Institute's activities. Work of the Security and Technology Programme on international cybersecurity is funded by the Governments of Canada, Czechia, France, Germany, Italy, the Netherlands, Norway, Switzerland and Microsoft.

About UNIDIR

The United Nations Institute for Disarmament Research (UNIDIR) is a voluntarily funded, autonomous institute within the United Nations. One of the few policy institutes worldwide focusing on disarmament, UNIDIR generates knowledge and promotes dialogue and action on disarmament and security. Based in Geneva, UNIDIR assists the international community to develop the practical, innovative ideas needed to find solutions to critical security problems.

Notes

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the Secretariat of the United Nations concerning the legal status of any country, territory, city or area, or of its authorities, or concerning the delimitation of its frontiers or boundaries. The views expressed in the publication are the sole responsibility of the individual authors. They do not necessarily reflect the views or opinions of the United Nations, UNIDIR, its staff members or sponsors.

About the authors

This report was produced by the **UNIDIR Security and Technology Programme**. Pavel Mraz drafted the report; Giacomo Persi Paoli, Andrea Gronke, Dominique Steinbrecher, Federico Mantellassi, Shimona Mohan and Sarah Grand-Clément contributed to the report.

Citation

UNIDIR Security and Technology Programme, "Securing Cyberspace for Peace: Insights into Cyberthreats and International Security in 2025", UNIDIR: Geneva, 2026.

Cover Image: Digital map of the world representing advanced cybersecurity technology (generated with AI). Credit: Adobe Stock / SKIMP Art.

Foreword by the UNIDIR Director

In 2025, cyberspace continued to evolve both as a catalyst for human advancement and as a domain fraught with escalating threats to international peace and security. The rapid integration of artificial intelligence (AI) and the advent of quantum computing offer transformative potential, enabling new solutions to global challenges. Yet, these same technologies can also be weaponized to exploit vulnerabilities in critical infrastructure, destabilize economies and fuel geopolitical tensions. Navigating this complex technological landscape necessitates collective action, grounded in collaboration, objective analysis and shared responsibility.

UNIDIR is committed to supporting these efforts, serving as a bridge-builder between the technical, diplomatic and policy communities. Our role is to connect diverse perspectives, advance research on emerging technologies, and foster dialogue on maintaining peace and security in the digital age. At the *2024 UNIDIR Cyber Stability Conference*, we convened global experts to analyse the implications of the rapidly changing cyberthreat landscape and to explore possible pathways for greater resilience and cooperation. These discussions reaffirmed that, while the challenges are formidable, the international community has the tools, expertise and frameworks to mitigate these risks – if we work together.

This report builds on the discussions at the conference and reflects UNIDIR's ongoing commitment to supporting the international community through independent, neutral and evidence-based research. It aims to provide diplomats, policymakers and practitioners with timely analysis of key developments, emerging risks and evolving trends in cyberspace that may affect international peace and security.

From cyberattacks targeting critical infrastructure via ransomware to disinformation campaigns that erode trust in governance, this report highlights the most pressing cyberthreats that shaped the international security landscape in recent years. It also examines how malicious activities are increasingly blurring the lines between criminal and state-sponsored activities and how emerging technologies – including AI and quantum computing – are transforming the field of cybersecurity.

The stakes could not be higher. The international community should continue to deepen its understanding of evolving cyberthreats and reinforce its collective commitment to cooperation. As states move toward the operationalization of the newly established *United Nations Global Mechanism on developments in the field of ICTs in the context of international security and advancing responsible State behaviour in the use of ICTs*, we hope this report will contribute to awareness-raising and capacity-building around evolving cyberthreats and support collaborative international responses. Together, we can ensure that technology serves as a force for peace and shared progress in an open, peaceful, secure, accessible and stable cyberspace.

Dr. Robin Geiss
UNIDIR Director

Acronyms & Abbreviations

AGI	Artificial general intelligence
AI	Artificial intelligence
APT	Advanced persistent threat
BEC	Business email compromise
CAAS	Cybercrime-as-a-service
COVID-19	Coronavirus disease 2019
DDOS	Distributed denial-of-service
DEX	Decentralized exchange (cryptocurrency trading platform)
EUROQCI	European Quantum Communication Infrastructure
GDP	Gross domestic product
GGE	Group of Governmental Experts
GPS	Global Positioning System
HNDL	Harver now, decrypt later
ICRC	International Committee of the Red Cross
ICTS	Information and communications technologies
IHL	International Humanitarian Law
IT	Information technology
MAAS	Malwave-as-a-service
MSP	Managed service
OEWG	Open-ended Working Group
OSINT	Open-source intelligence
OT	Operational Technology
PQC	Post-quantum cryptography
QKD	Quantum key distribution
RAAS	Ransomware-as-service
SCADA	Supervisory Control and Data Acquisition
UN	United Nations
UNICC	United Nations International Computing Centre
UNICRI	United Nations Interregional Crime and Justice Research Institute
UNIDIR	United Nations Institute for Disarmament Research
UNOCT	United Nations Office for Counter-Terrorism

Glossary of Terms

Advanced persistent threat (APT)
A highly resourced, often state-linked group conducting prolonged, targeted cyber operations for strategic gain

Adversarial AI attack
A technique that manipulates AI systems into making errors, often used to evade cyber-security tools

AI-driven
Describes cyber tools or processes powered by artificial intelligence to automate, enhance or scale actions

Attacker/malicious cyber actor
Any entity – state or non-state – conducting harmful cyber activity; the two terms are used interchangeably in this report

Botnet
A network of compromised devices controlled remotely to conduct coordinated cyberattacks

Commitments versus obligations
For the purposes of this report, obligations refer to legally binding duties under international law, while commitments refer to voluntary, consensus-based agreements – such as the norms of responsible state behaviour in cyberspace – which, while not legally binding, are politically significant and universally endorsed by United Nations Member States

Critical infrastructure and essential services
Relevant United Nations processes have highlighted each state's prerogative to determine which infrastructures it designates as critical; for the purposes of this report, the term refers to infrastructure types highlighted in reports of United Nations Open-Ended Working Groups and Groups of Governmental Experts as examples essential to societal functioning (e.g., medical facilities, communications, financial services, energy, water, transportation and sanitation).

Cyber-enabled influence operation
An activity that uses digital technologies to shape, manipulate or disrupt public opinion, political processes or social cohesion; tactics include disinformation, deepfakes and hack-and-leak campaigns

Cyber operation
An activity in or through cyberspace conducted by a state or a state-controlled proxy

Cyberattack
A deliberate act – conducted by state or non-state actors – targeting the confidentiality, integrity or availability of ICT systems or data

Cybercrime-as-a-service (CaaS)
A model where malicious tools and services (e.g. ransomware kits) are sold or rented to other threat actors

Cybercriminals
Non-state actors engage in cyber activity primarily for financial gain (e.g. ransomware)

Cybersecurity service providers
Private entities offering defensive and, in some cases, offensive cybersecurity services

Disinformation
False information spread intentionally to mislead or manipulate

Distributed denial of service (DDoS) attack
A cyberattack that floods a system with traffic to make it unavailable

Double extortion
A ransomware tactic combining data encryption and threats to leak stolen data

Ethical hackers
Cybersecurity professionals, penetration testers and good faith hackers who test systems to find and report ICT vulnerabilities

False flag cyber operation
A cyberattack staged to appear as if conducted by another actor or state

Hack-for-hire services/cyber mercenaries
Commercial actors offering custom cyberattack tools for clients for a fee

Hacktivists/patriotic hackers
Actors motivated by ideological or nationalist causes who conduct cyberattacks to advance them

Harvest now, decrypt later (HNDL)
The practice of stealing encrypted data for future decryption with advanced tools (e.g., quantum computing)

Managed service provider (MPS)
A company that remotely manages an organization's IT or cybersecurity infrastructure and services – such as networks, servers, endpoints, or security monitoring – typically under a subscription or service contract.

Misinformation
Inaccurate information spread without intent to deceive

Obfuscation techniques

Tactics used to hide malware from detection (e.g., encryption, packing)

Phishing/spear phishing

Deceptive emails or messages used to trick users into revealing sensitive data; spear phishing denotes a highly targeted and personalized form of phishing

Polymorphic malware

Malware that changes its code without human intervention to avoid detection

Post-quantum cryptography (PQC)

Encryption methods designed to remain secure against future quantum-enabled decryption

Pre-positioning in cyberspace

The covert placement of malware in systems that signals potential future disruption or coercion

Quantum key distribution (QKD)

A secure communication method using quantum physics, where any interception attempt is detectable

Ransomware-as-a-service (RaaS)

A model where ransomware tools are licensed to affiliates in exchange for a share of ransom payments

Script kiddies

Unskilled individuals using pre-built hacking tools to launch attacks

State actors

Governments or government-integrated entities conducting cyber operations, including military and intelligence services

Supply chain attack

A form of cyberattack targeting a third-party vendor or provider to access downstream systems

United Nations Framework of Responsible State Behaviour in Cyberspace

A set of universally endorsed commitments aimed at promoting responsible conduct by states in the use of information and communication technologies (ICTs); developed through consensus in United Nations processes and endorsed by all Member States; includes norms, confidence-building measures and capacity-building efforts, and affirms that international law applies in cyberspace

Zero-day exploit

A vulnerability unknown to the software vendor and exploited before a patch is available

Table of Contents

EXECUTIVE SUMMARY: IMPACT OF CYBERATTACKS ON INTERNATIONAL PEACE AND SECURITY IN 2025 9

INTRODUCTION 13

1. EVOLVING CYBERTHREATS: FROM TARGETED ATTACKS TO SYSTEMIC DISRUPTIONS 15

1.1. Critical infrastructure and essential services: A persistent target 15

1.1.1. The United Nations system and humanitarian organizations 18

1.2. Cybercrime as-a-service: The underground industry fuelling ransomware and fraud 19

1.3. Supply chain insecurity: From targeted attacks to global disruptions 24

1.4. Cyber-enabled influence operations: From hack-and-leak to AI-generated manipulation 29

2. CYBERTHREAT ACTORS: BLURRED LINES AND GROWING ACTOR COMPLEXITY 33

2.1. Resources and capabilities 34

2.1.1. Highly capable and resourceful actors 34

2.1.2. Moderately capable actors 34

2.1.3. Low-resource actors 35

2.2. Spectrum of state involvement 36

2.3. Objectives and motivations 42

3. EMERGING TECHNOLOGIES RESHAPING CYBERSPACE: THE DOUBLE-EDGED SWORD 45

3.1. Cybersecurity and artificial intelligence 45

3.1.1. AI-powered cyberattacks 46

3.1.2. Adversarial attacks on AI-systems 48

3.1.3. AI-powered cyberdefenses 48

3.1.4. AI–cyber nexus 49

3.2. Cybersecurity and quantum computing 51

CONCLUSIONS 56

ENDNOTES 57

Image of a cyberattack occurring on a network (generated with AI). Credit: Adobe Stock / YM Creative Studio.

Executive summary: Impact of cyberattacks on international peace and security in 2025

This report examines key developments in the cyberthreat landscape in 2025, with a focus on their implications for international peace and security. Drawing on public reporting, expert insights and the outcomes of the *2024 UNIDIR Cyber Stability Conference*, it explores three interrelated dimensions:

- **The evolving nature of cyberthreats**
- **The changing landscape of threat actors**
- **The role of emerging technologies in cybersecurity**

By highlighting the increasing scale, complexity and impacts of cyberattacks, the report aims to support diplomats, policymakers and practitioners in better understanding the dynamics of cyberspace and in responding to the challenges posed by malicious cyber activities.

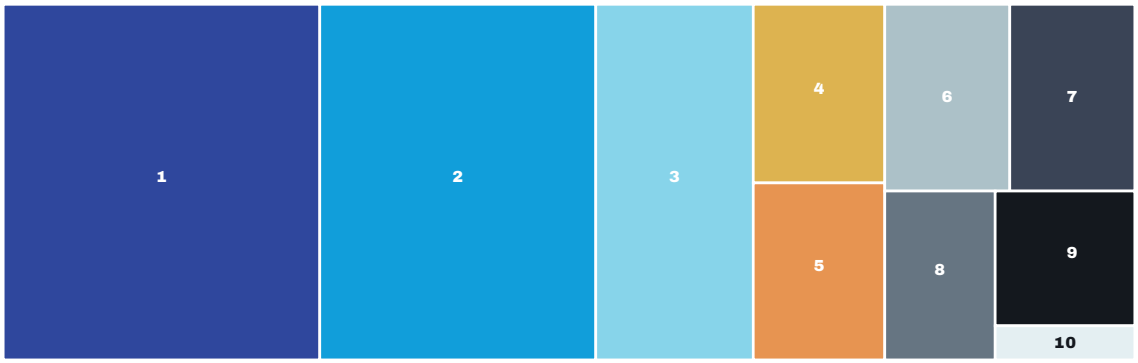
This report details how the rising prevalence and sophistication of cyberattacks in 2025 underscored their growing role as a disruptive force in the international system. Cyberattacks, which were once limited in scope and impact, have **evolved into a multidimensional threat that transcends borders and targets critical systems** essential to the functioning of societies and economies. The global reliance on interconnected networks and digital infrastructure has

amplified the potential for national, regional and global disruptions, making cybersecurity an increasingly important cornerstone of international stability.

One of the most significant impacts of cyberattacks has been the **destabilization of critical infrastructure**. Critical infrastructure worldwide is reported to have faced over 420 million cyberattacks last year, 30 per cent increase from the previous year. Cyberattacks in 2025 that targeted energy grids, healthcare systems, water distribution networks and transportation hubs highlighted vulnerabilities in digital systems that underpin delivery of essential public services (see Figure 1). The cyberattacks targeting power distribution and telecommunications networks also exemplified the cascading effects of such disruptions that can lead to widespread blackouts, economic losses and heightened public anxiety. These incidents revealed the fragility of interconnected systems and underscored the potential for cyberattacks to escalate into humanitarian crises – particularly when essential public services and deliveries of humanitarian assistance are compromised.

FIGURE 1.

Sectors most targeted by cyber operations in 2023/2024



1	IT	24%	6	CONSUMER RETAIL	5%
2	EDUCATION AND RESEARCH	21%	7	FINANCE	5%
3	GOVERNMENT	12%	8	MANUFACTURING	4%
4	THINK TANKS AND NGOS	5%	9	COMMUNICATIONS	4%
5	TRANSPORTATION	5%	10	ALL OTHERS	1%

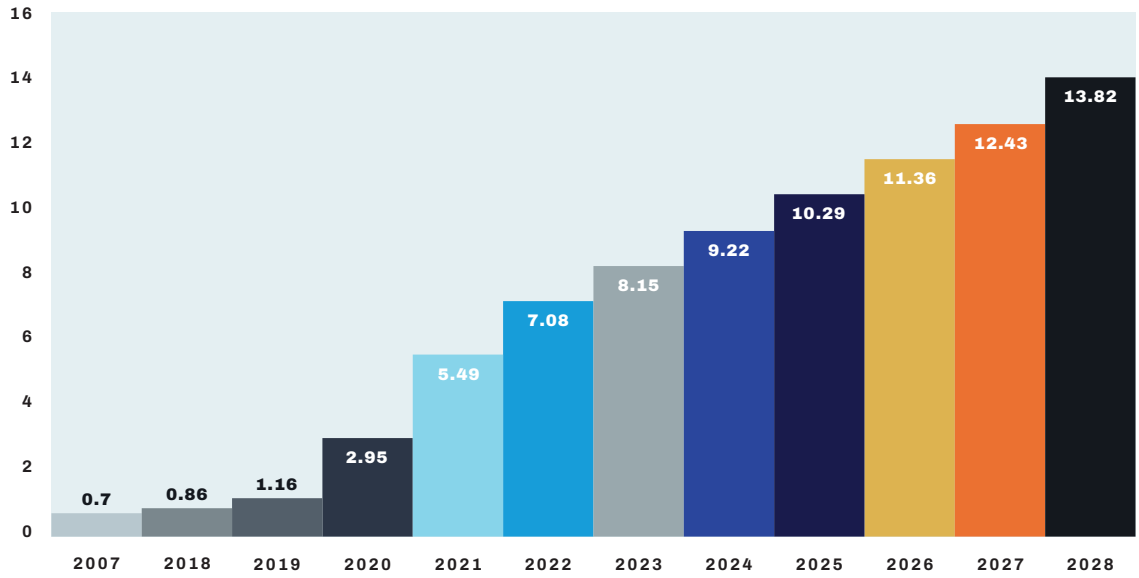
Source: Microsoft Digital Defense Report, 2024.

The financial sector also faced heightened threats, with **cyberattacks targeting global financial systems and causing disruptions with increasing monetary costs**. Ransomware attacks – which increased approximately threefold between 2023 and 2024 – and cryptocurrency theft have eroded trust in digital transactions and undermined economic stability. The global economy lost close to US$10 trillion in economic value due to cybercrime disruptions in 2025 (see Figure 2). Such financial **disruptions can ripple through global markets, exacerbating economic inequalities**, draining already stretched government budgets and straining diplomatic relations between states. Moreover, **cyberattacks have become a tool for geopolitical manipulation and conflict**. Some states have reportedly used information and communications technologies (ICTs) alongside kinetic weapons in the context of armed conflict. Additionally, both state and non-state actors are reported to have deployed cyberattacks to achieve various strategic objectives, ranging from espionage and disruption of state political processes to circumvention of United Nations Security Council sanctions.

In 2025, state-affiliated actors were also reported to use criminal tools and tactics – and even criminals themselves – to advance their objectives. This **blurring of the lines between state-backed cyber operations and cybercriminal activity** can complicate attribution and efforts to hold malicious actors accountable. This growing actor ambiguity may also delay responses and foster mistrust among states, increasing the risk of miscalculations, retaliatory actions and escalating tensions. The **spread of disinformation and misinformation through cyber-enabled campaigns** has further undermined international stability. Cyber influence operations targeting elections and public discourse threatened to exacerbate political polarization and erode trust in democratic institutions just as half the world's population, in over 60 countries, went to the polls in 2024.

FIGURE 2.

Real and projected global cost of cybercrime disruptions, 2017–2028



Source: Cybersecurity Ventures.

These cyber-enabled campaigns, some of which now use artificial intelligence (AI) to craft ever-more convincing narratives, may have a destabilizing effect on governance, weakening the social fabric of affected states. This **amplified ability of malicious actors to manipulate public opinion and exploit societal divisions through technology** may pose a challenge for policymakers seeking to preserve social cohesion and maintain legitimacy.

Cyberattacks have also posed challenges to international development and humanitarian efforts, increasingly targeting international organizations, the United Nations system and implementing partners. Humanitarian organizations have reported a significant increase in cyber incidents targeting their operations, with attackers seeking to disrupt aid delivery or steal sensitive data on people in vulnerable positions. Such attacks can have profound implications for such people, as they can delay or derail critical support in conflict zones and disaster-affected areas. Furthermore, the **growing misuse of cyber tools by terrorist groups** to coordinate attacks or fund their activities has intensified security challenges, complicating global counterterrorism efforts.

Last year, the **rise of cybercrime-as-a-service (CaaS) also expanded the availability of advanced malicious cyber tools** and techniques, which enabled a broader range of actors to launch more persistent and impactful attacks. This **commoditization of cybercrime not only increased the frequency of cyberattacks but also their severity**, as unskilled actors gained access to more advanced capabilities. In the future, the proliferation of malicious cyber tools may create a security dilemma for states, as the lowering of barriers to entry for malicious activities continues, challenging existing defensive and law enforcement measures.

Despite ongoing efforts within the United Nations and other multilateral forums to promote responsible state behaviour in cyberspace, the **rapid evolution of cyberthreats continues to outpace regulatory and enforcement measures**. If left unchecked, evolving cyberthreats may not only destabilize the digital ecosystem but may also erode the very foundations of trust that sustain diplomatic, economic and security cooperation between states. Without a concerted and unified global effort to strengthen cyber resilience, build capacities and foster cross-border collaboration to hold malicious cyber actors accountable, the proliferation of malicious cyber activities may undermine international peace and stability.

Introduction

In an increasingly interconnected world, **cyberspace plays a pivotal role in global peace and security.** Cyberattacks may threaten critical infrastructure, destabilize economies, undermine trust in governance and fuel geopolitical tensions. As technological advancements accelerate, cyberthreats are becoming more sophisticated, challenging traditional frameworks for the maintenance of international peace and security. This report aims to provide an in-depth overview of contemporary cybersecurity trends, offering insights into their implications for international stability and potential pathways for maintaining peace and security in the digital age.

The significance of this topic extends beyond technical boundaries: **cyberthreats are inherently transnational, and so require coordinated technical, diplomatic and legal responses**. By fostering global collaboration and awareness, diplomats and policymakers can better navigate the complexities of the evolving cyberthreat landscape and work towards an open, secure, stable, accessible and peaceful cyberspace.

In 2025, the rapidly evolving cyberthreat landscape presented new critical challenges for the international community's efforts to maintain international peace and security. As digital technologies integrated further into critical infrastructure, economies and governance, they simultaneously became targets and tools for malicious actors. The UNIDIR's flagship **2024 Cyber Stability Conference** convened experts, diplomats and policymakers from the multi-stakeholder community to analyse these challenges – across types of cyberthreat, actor and technology used – and to explore strategies for cyber resilience. These discussions underscored the importance of cooperative international frameworks, with the United Nations Framework of Responsible State Behaviour in Cyberspace at the centre, and the urgent need to continue to assess emerging technological risks.

This report provides in-depth analyses across three key dimensions explored during the 2024 Cyber Stability Conference. As well as a detailed technical overview of evolving **cyberthreats**, it offers a breakdown of **cyberthreat actors** and their motivations and an **exploration of the emerging technologies** that are driving both offence and defence in cyberspace. The analysis is informed by latest insights from leading cybersecurity vendors, public reporting, outcomes of United Nations processes, and discussions and conclusions drawn from the **2024 Cyber Stability Conference**.

Section 1, **"Evolving Cyberthreats: From Targeted Attacks to Systemic Disruptions"**, examines the most pressing contemporary cyberthreats in detail, ranging from attacks on critical infrastructure and supply chains to the rise of ransomware and cyber-enabled influence operations. While these threats are not new, their scale, sophistication and impact expanded significantly over the year, amplifying risks to global stability. Section 1 also provides an in-depth exploration of how these threats operate at the technical level as well as examples of how each threat can undermine international peace and security.

Section 2, **"Cyberthreat Actors: Blurred Lines and Growing Actor Complexity"**, shifts the focus to those behind evolving cyberthreats. It unpacks their possible motivations, their tactics and the growing entanglement of state and non-state actors in cyberspace. The section

highlights how states, organized criminal groups and independent actors are leveraging advanced cyber capabilities to pursue various objectives in cyberspace – often in overlapping ways – making attribution and response efforts more difficult. Section 2 also examines the use of proxies, the rise of hacktivists, "hack-for-hire" services and the evolving role of private sector entities in securing digital infrastructure amid rising geopolitical tensions.

Finally, Section 3, **"Emerging Technologies Reshaping Cyberspace: The Double-Edged Sword"**, examines the impact of artificial intelligence (AI) and quantum computing on cybersecurity. It provides concrete examples of how AI is transforming both offensive and defensive cyber capabilities, from AI-powered malware that evades detection to autonomous cybersecurity systems that can pre-emptively neutralize threats. Section 3 also examines the future role of quantum computing in the field of cybersecurity, exploring how quantum advancements could both strengthen cybersecurity for some while simultaneously threatening it for many by rendering widely used encryption methods obsolete.

Taken together, the three sections represent a **conceptual framework developed by UNIDIR for understanding the evolving cybersecurity landscape**.[1] They explore the threats that are shaping digital security, the actors behind them, and the technologies enabling new strategies for attack and defence. By structuring the report in this way, it is hoped that policymakers, diplomats and cybersecurity practitioners can gain a greater understanding of the forces shaping cyberspace today and their broader implications for international peace and security. Throughout the report, explanatory text boxes demystify how various types of cyberattack work, outline notable case studies, and highlight emerging technical and policy-relevant trends. These boxes are designed to support readers in understanding the operational mechanics behind the evolving cyberthreats, threat actors and technological developments as well as their implications for national, regional and global security.

Geneva Cyber Week Conference flags on the Pont du Mont-Blanc, May 2025, Geneva. Credit: UNIDIR / A.Tardy.

1. Evolving cyberthreats: From targeted attacks to systemic disruptions

This section provides a detailed examination of specific threats that defined the cybersecurity landscape in 2025. Contemporary **cyberthreats grew more systemic, targeted and consequential**, with increasingly severe implications for international peace and security. No longer limited to isolated technical events, cyberattacks are now capable of disrupting essential services, paralysing public institutions and undermining international stability.

This **section examines the most significant types of cyberthreat observed up to date** emphasizing how they exploit technological, institutional and geopolitical vulnerabilities in ways that can create cascading effects across borders and domains. It highlights the growing use of cyber operations as a tool of coercion and disruption, in both peacetime and conflict, and underscores the urgent need for sustained international cooperation and continued efforts to advance norms, resilience and accountability.

The section is organized around several threat categories that have taken on growing importance in international cybersecurity discussions:

- **Cyberattacks on critical infrastructure and essential services,** including healthcare, energy and water systems, space assets, cloud computing and undersea cables
- **The industrialization of cybercrime,** particularly through ransomware, cybercrime-as-a-service (CaaS) models, hack-for-hire services and financially motivated attacks that have disrupted public services and national institutions
- **Supply chain compromises,** which exploit trusted digital and physical vendor relationships to enable large-scale, systemic breaches
- **Cyber-enabled influence operations,** including disinformation, deepfakes and electoral interference campaigns that exploit AI and digital platforms to distort public discourse and destabilize societies.

Together, the content of this section provides a foundational understanding of contemporary cyberthreat landscape and a basis for the subsequent analysis of the types of threat actor, technological advancements and governance challenges.

1.1. Critical infrastructure and essential services: A persistent target

Critical infrastructure remained a central focus of cyberattacks in 2025, with nearly **40 per cent of state cyber operations targeting critical sectors**[2] such as energy, water, healthcare, cloud services, information technology (IT), education, government, transportation, finance and telecommunications.[3] These attacks are reported to have grown in sophistication, leveraging a combination of ransomware, data exfiltration, supply chain compromises and stealthy pre-positioning techniques.[4] While cybercrime now poses greater risks to critical infrastructure, some

state-sponsored operations are reported to have shifted towards long-term strategic pre-positioning within the critical infrastructure of other states.[5] This trend raised concerns around potential disruptive and destructive cyberattacks in the event of geopolitical escalation.[6] It also raised pressing questions about the classification of pre-positioning activities under international law.[7]

BOX 1.

How critical infrastructure attacks work

Critical infrastructure attacks exploit vulnerabilities in systems that manage essential services such as energy grids, water supplies and transportation networks. Threat actors often target operational technologies (OT) such as Supervisory Control and Data Acquisition (SCADA) systems, which are designed to manage industrial control processes. By injecting malicious code or exploiting outdated software, attackers can cause disruption or physical damage. For instance, a cyberattack might disable a power grid's ability to distribute electricity, leading to widespread blackouts.

Several cyberattacks targeting essential services to the public,[8] particularly **healthcare and water systems**,[9] raised concerns around the vulnerability of critical infrastructure that should remain operational at all times.[10] Healthcare organizations, already under strain from resource constraints and outdated IT systems, have become particularly attractive ransomware targets due to their **inability to afford prolonged downtime and a higher likelihood of paying ransom demands to quickly restore patient care services**.[11] Malicious actors have exploited these weaknesses, with notable attacks targeting hospital networks, emergency response systems and pharmaceutical supply chains.[12] Similarly, reported **cyber intrusions into water treatment facilities and waste water management systems** have raised concerns about potential public health risks,[13] including the potential contamination of drinking water supplies or disruption of irrigation systems essential for food production.[14]

Cloud computing infrastructure[15] and space infrastructure also became more frequently targeted. Cloud platforms host vast amounts of sensitive data and are integral to the operations of industries, governments and critical services worldwide.[16] Some cyberattacks on major cloud service providers resulted in cascading intrusions across multiple sectors,[17] demonstrating the **systemic risks posed by malicious cyber activities that target the transnational critical infrastructure delivering services across borders**.[18]

Space infrastructure, including satellites used for communications, navigation and intelligence gathering, also faced heightened cyberthreats.[19] This exacerbated concerns about the militarization of cyberspace and fears of disruptions of space-reliant services, including in the context of an armed conflict.[20] Looking ahead, the potential for cyberattacks to **disrupt satellite operations, interfere with GPS signals or compromise remote sensing capabilities may pose serious risks to global security**, emergency response efforts and international conflict management.[21]

BOX 2.

Spill-over effects in space: The ViaSat cyberattack

In February 2022, coinciding with the onset of the armed conflict in Ukraine, a cyberattack targeted Viasat's KA-SAT satellite network, which was providing Internet coverage during concurrent physical disruptions to the country's communications infrastructure.[22] The attack appeared to be aimed at disrupting services in Ukraine, but its effects spread far beyond the likely intended target. **Notably, it disrupted remote monitoring of approximately 5,800 wind turbines in Germany and caused Internet outages for nearly 9,000 subscribers in France.** The incident underscored the escalating cyberthreats against space-based assets and highlighted how cyberattacks on transnational infrastructure can have unintended cross-border consequences, affecting civilian and critical services far beyond the initial target.[23]

Meanwhile, **the topic of strategic pre-positioning of malware within adversarial critical infrastructure** has emerged as a notable concern related to state-backed cyber operations.[24] Unlike traditional espionage, where the goal is intelligence gathering, pre-positioning attacks may **suggest an intent to hold infrastructure at risk for future coercion or in potential conflict scenarios**.[25] Cybersecurity firms that track these activities reported an increase in state-affiliated actors embedding stealthy malware within the power grids,[26] telecommunications networks,[27] financial institutions, and transportation systems of geopolitical rivals.[28] As geopolitical tensions escalate, pre-positioning operations may become more frequent. This **blurring of the line between cyber espionage and coercive cyber operations raises new policy challenges**, including how such actions could be classified under international law.[29]

Addressing the growing cyber risks to critical infrastructure and essential services, states participating in the General Assembly's **Open-ended Working Group (OEWG) on ICT security reaffirmed that cyberattacks against critical sectors such as healthcare, financial services, energy and transport can have cascading effects at national, regional and global levels**.[30] States have also highlighted the threat posed by malicious cyber activities that target undersea cables and orbit-based communications networks, warning that disruptions to such infrastructure delivering services across borders could have a severe impact on telecommunications and the integrity and availability of the Internet.[31] In response to these developments, **states and international organizations have called for enhanced cooperation to secure transnational critical infrastructure and reinforce international norms** prohibiting attacks on essential services.[32]

While discussions in the United Nations have reinforced the importance of securing critical infrastructure and essential services, **challenges remain in translating these commitments into concrete protections against cyberthreats**. The difficulty of implementing international norms is further complicated by the challenges around attribution and the related uneven technical and policy capacity across different states. As geopolitical tensions rise, securing critical infrastructure may require greater diplomatic engagement, stronger public–private partnerships and increased investment in cybersecurity resilience measures.

1.1.1. The United Nations system and humanitarian organizations

Cyberattacks also **posed a growing threat to international organizations (including the United Nations system) and humanitarian organizations**. Concretely, the United Nations International Computing Centre (UNICC) has reported a significant rise in cyberthreats targeting United Nations agencies, funds and programmes. Latest analysis shows that targeted phishing emails remain the most common attack vector, accounting for 57 percent of observed malicious ICT activity, followed by the exploitation of software vulnerabilities at 11 percent. In terms of motivation, financial gain represents approximately 51 percent of identified malicious ICT activity, while information gathering activities, often associated with advanced persistent threat actors, account for around 29 percent, with hacktivism comprising a further 13 percent.[33] For example, in March 2024, **a cyberattack on a United Nations Development Programme server exposed personal information of past and current personnel**.[34]

Humanitarian organizations have also reported a marked increase in cyber incidents,[35] which have interfered with aid delivery, disrupted critical operations[36] and put people in vulnerable positions at risk.[37] While no state currently designates humanitarian organizations as a distinct category of critical infrastructure, their essential role – particularly in conflict and crisis settings – has been widely recognized. Additionally, **some humanitarian functions (e.g., distribution of food or medicine) may also fall under existing critical infrastructure categories under national frameworks**, such as healthcare, emergency services or public safety.[38] Furthermore, during emergencies such as the Covid-19 pandemic,[39] several states granted humanitarian workers operational status comparable to that of providers of essential services.[40]

BOX 3.

Cyberattack on the International Committee of the Red Cross

In January 2022, the International Committee of the Red Cross (ICRC) experienced a sophisticated cyberattack that compromised personal data of over 515,000 vulnerable people across at least 60 national Red Cross and Red Crescent societies. The attackers, by exploiting an unpatched critical vulnerability, gained unauthorized access to sensitive information, including details of missing persons, detainees, and families separated by conflict or disaster. This breach forced the ICRC to temporarily shut down its Restoring Family Links programme, which is vital for reuniting families separated by crisis. The incident underscored how escalating cyberthreats facing humanitarian organizations can have a severe impact on global humanitarian efforts and aid delivery.[41]

Recognizing these growing threats, states have begun to articulate shared concerns and commitments across several international forums. **Within the OEWG, states have expressed concern over malicious cyber activities that target international and humanitarian organizations**, warning that such incidents could disrupt their mandates, compromise the safety and independence of their operations and undermine trust in their work.[42] Similar concerns have been echoed by the **United Nations Security Council in resolution 2730 (2024)**, which expresses concern over rising cyber incidents – including data breaches and information

operations – that target humanitarian organizations and disrupt their relief efforts, threaten the security of humanitarian personnel and assets, and erode trust in the neutrality of the United Nations and its partners.[43]

Concurrently, the **34th International Conference of the ICRC** adopted a consensus **resolution calling on states and parties to armed conflicts to allow and facilitate impartial humanitarian activities during armed conflict, including those that rely on ICTs, in accordance with international legal obligations**.[44] These recent developments reflect a growing international recognition that malicious cyber activities targeting humanitarian and multilateral actors pose not only operational risks to delivery of essential aid and assistance but also legal and normative challenges to the safety, trust and integrity of international and humanitarian action.

1.2. Cybercrime as-a-service: The underground industry fuelling ransomware and fraud

Cybercrime has evolved into a **highly organized, industrial-scale enterprise**, with ransomware continuing to dominate the cyberthreat landscape.[45] **The commoditization of cybercrime through cybercrime-as-a-service models – a key trend of the contemporary cyberthreat landscape** – has lowered barriers to entry for malicious actors.[46] Criminal networks now offer **ransomware-as-a-service (RaaS)**, phishing kits, credential theft tools and illicit hacking services on Dark Web marketplaces, making sophisticated attack methods **more accessible and scalable than ever**.[47]

Due to this ever-more-sophisticated division of labour, would-be **malicious actors can now purchase cybercrime subscription services on the Dark Web**,[48] granting them access to a menu of attack methods – including ransomware, phishing and distributed denial of service (DDoS) techniques – along with supporting services such as a 24/7 troubleshooting hotline.[49] These **subscription-based cybercrime platforms (which can cost as little as US$500 per year)**[50] franchise the global cybercrime economy, thereby allowing even unskilled attackers to launch professional-grade cyberattacks with minimal effort. The rapid growth of **black-market demand for ICT vulnerabilities** further accelerates this trend,[51] as cybercriminals capitalize on undisclosed software flaws and ICT vulnerabilities to develop, commercialize and expand their attack capabilities.[52]

As a result, cybercriminal activity surged last year, causing concerns among states that cybercrime is no longer just a crime and law enforcement issue but a serious threat to economic prosperity and national security. In concrete numbers, estimated **cybercrime-related financial losses – including ransom payments, operational disruptions and forensic investigation costs – exceeded US$10 trillion in 2025**, 30 per cent more than in 2022.[53] To put the global economic losses from cybercrime into perspective, if cybercrime were a country, it would **have the world's third-largest economy, trailing only the United States and China in terms of total GDP**. [54] If left unabated, this level of financial drain may divert critical resources away from economic development, innovation and essential public services and widen the existing digital divides.

In terms of specific cybercrime threats, **ransomware remained one of the most pervasive and financially damaging cybercrime threats, with an annual increase of 275 per cent in 2024** according to the 2024 Microsoft Digital Defense Report.[55] Last year, approximately 59 per cent of surveyed organizations suffered ransomware attacks,[56] and 70 per cent of the attacks resulted in data encryption – leaving affected organizations with few options beyond costly recovery efforts or paying ransoms.[57] Double extortion tactics, where attackers both encrypt and threaten to leak stolen data, have also been on the rise and proved to be especially effective, coercing many organizations into **compliance with ransom demands to avoid reputational, regulatory and operational fallout**.[58]

BOX 4,

How ransomware works

Cybercriminals use ransomware to encrypt victims' data and then demand payment in exchange for restoring access. Many ransomware operations now employ double extortion, where attackers also threaten to leak sensitive information unless ransom demands are met. These attacks are typically executed through phishing emails, malicious attachments or the exploitation of software vulnerabilities. Ransomware groups generally require payment in cryptocurrencies, which offer varying degrees of anonymity and make it more difficult for law enforcement to trace illicit funds. To further obscure the flow of money, attackers often use several techniques:

- **Mixing services (or "tumblers")** blend the attacker's cryptocurrency with funds from many other users, making it harder to identify the original source of the ransom payment.
- **Multihop transactions** involve moving funds through multiple cryptocurrency wallets in rapid succession, sometimes across chains, to complicate tracing efforts and break investigative trails.
- **Decentralized exchanges (DEXs)** allow users to swap cryptocurrencies directly without relying on a centralized intermediary that might implement identity checks or monitoring, giving attackers additional opportunities to launder or convert their proceeds with reduced oversight.

Last year, ransomware attacks continued to have an impact on a wide range of sectors, with certain industries experiencing higher frequencies of incidents. **The business services sector was the most targeted**, accounting for 24.1 per cent of reported ransomware cases, followed by the retail sector (with 15.2 per cent of cases) and manufacturing (10.5 per cent).[59] Traditionally, healthcare has been a prime target for ransomware due to its critical nature and often outdated security infrastructure.[60] **Cryptocurrencies remained central to ransomware operations, facilitating swift and anonymous transactions**.[61] Despite global efforts to combat illicit financial flows, cybercriminal networks continue to **exploit cryptocurrency laundering techniques,** using mixing services, multi-hop transactions and decentralized exchanges to obscure the origins of illicit gains.[62]

BOX 5.

Case study: The Costa Rica ransomware crisis

In 2022, Costa Rica became the first country to declare a national state of emergency in response to a ransomware attack. A major cybercriminal group paralysed critical government services overnight, disrupting customs operations, healthcare, education and social security systems. The attack not only caused financial and administrative chaos but also highlighted the strategic vulnerability of public institutions to ransomware threats. Faced with a prolonged crisis, Costa Rica turned to international cooperation and received cybersecurity assistance from other states and the private sector, among others. Since the attack, Costa Rica has strengthened its cybersecurity posture and is now emerging as a regional leader in cyberdefence, using its experience to advocate for stronger global responses to ransomware. The case serves as a warning of how cybercrime can destabilize governments and why cross-border cooperation is essential to strengthening cybersecurity worldwide.[63]

Beyond ransomware, cybercriminals have **expanded their reach into large-scale online scams, financial fraud[64] and business email compromise (BEC) attacks**. BEC scams alone resulted in billions of dollars of annual financial losses, with fraudsters impersonating executives or trusted contacts to **deceive victims into making unauthorized transactions**.[65] Similarly, investment scams, fraudulent cryptocurrency schemes and Ponzi operations have surged, with cybercriminals leveraging deepfake technology and fake endorsements from public figures to deceive investors.[66] As described in greater detail in Section 3, **AI-enabled social engineering tactics** are enhancing the effectiveness of such scams, making them harder to detect and mitigate.[67]

At the same time, **ransomware operators, online fraud syndicates and large-scale scam factories** are reported to have taken root in certain jurisdictions that either tolerate their operations or lack the technical, legal or financial capacity to disrupt well-entrenched cybercriminal enterprises. This **industrialization of online scams** – ranging from fraudulent job offers to financial fraud and cryptocurrency Ponzi schemes – has become a multibillion-dollar global industry, with entire call centres dedicated to delivering ransomware payloads and deceiving victims on an international scale.[68]

These operations often intersect with human trafficking, as criminal organizations coerce or traffic individuals into working in exploitative conditions in cybercrime factories, particularly in regions where weak governance enables such practices to persist.[69] The **widespread availability of stolen personal data on Dark Web marketplaces** further fuels these criminal operations, allowing cybercriminals to craft highly personalized scams that exploit victims' specific preferences, social networks or vulnerabilities.[70] As cybercrime continues to expand beyond purely digital domains into organized transnational criminal networks, the challenge of disrupting these operations is expected to grow, underscoring the **need for stronger cross-border cooperation, legal harmonization and capacity-building efforts** to hold accountable both the perpetrators and the systemic enablers of cybercrime.

BOX 6.

Thailand dismantles online scam centre in Myanmar

In a significant law enforcement operation, Thai authorities rescued hundreds of individuals representing over 30 nationalities from a scam call centre compound in Myanmar. The operation, which took place in early 2025, highlighted the growing crisis of cybercrime syndicates operating large factories. Many of the rescued were victims of human trafficking, coerced into executing large-scale online scams targeting international victims.[71] Thai law enforcements officials estimate that tens of thousands could be held in illegal scam compounds in Southeast Asia.[72]

Beyond traditional cybercriminals, **terrorist groups have also turned to cyber tools to fund and facilitate their activities**. Reports by the United Nations Office of Counter-Terrorism (UNOCT) and the United Nations Interregional Crime and Justice Research Institute (UNICRI) highlight how **violent extremist organizations are exploiting the Dark Web and CaaS models** to execute cyberattacks, raise and transfer funds, and recruit new members.[73] This trend not only complicates global counterterrorism efforts but raises the prospect of **malicious cyber tools developed for financial extortion being repurposed for destructive operations** with national, regional and global security implications.[74]

Recognizing the escalating global threat of cybercrime, **United Nations Member States have reaffirmed their commitment to tackling cyber-enabled criminal activities, particularly ransomware**. The OEWG on ICT security has emphasized concerns over the frequency, scale and severity of ransomware attacks, the proliferation of commercially available intrusion tools, and the use of cryptocurrencies to finance malicious activity. States have also emphasized the

Raided gang-run internet ‘scam farm’ in South East Asia, 2024. Credit: UNODC / Laura Gil.

need for comprehensive responses, including **disrupting the enabling infrastructure of ransomware, addressing illicit financial flows, and countering the misuse of vulnerabilities and CaaS models** that fuel broader threats to international peace and security.[75]

BOX 7.

Possible ransomware response and mitigation measures

As ransomware attacks grow in scale and complexity, governments and organizations of all sizes should prepare for the possibility of a successful attack. While prevention is essential, a robust response strategy is also critical to minimize damage and ensure continuous delivery of services. By implementing the following measures, organizations can strengthen their resilience, minimize disruptions and protect critical assets from ransomware threats.

- **Ransomware Response Plan:** Establish and regularly update an incident-response strategy covering ransomware-specific scenarios like data encryption and exfiltration
- **Secure Backups:** Keep encrypted, offline or immutable backups of critical data and test them frequently
- **Strengthened Network Security:** Apply least privilege access, patch systems regularly and segment internal networks to hinder lateral movement once a system is penetrated
- **Employee Training:** Conduct regular cybersecurity-awareness sessions, drills and mock phishing attacks to mitigate social engineering risks
- **Clear Reporting Protocols:** Establish clear channels for coordination and incident reporting to law enforcement and cybersecurity agencies

A major milestone was the **formal adoption of the United Nations Convention Against Cybercrime**, which was opened for signature in Hanoi in October 2025.[76] The Convention seeks to establish global legal standards for investigating and prosecuting cybercriminal activities, while facilitating information-sharing and law enforcement collaboration between states.[77] **Operationalizing the Convention and ensuring capacity-building support for developing countries will be key** to its success.

While these **consensus outcomes of United Nations processes mark progress, significant enforcement challenges remain**. The complexity of attribution, jurisdictional limitations and diverging national approaches to cybercrime prosecution may continue to hinder international responses. Moreover, **the persistence of safe havens for cybercriminals** – where some states may either tolerate cybercriminal operations or lack the capacity to disrupt them – highlights the complexity of addressing cybercrime globally. Sustained political will, private sector engagement and investment in cybersecurity resilience will also be essential as ransomware, financial fraud and cyber-enabled terrorism continue to evolve, threatening global stability.

1.3. Supply chain insecurity: From targeted attacks to global disruptions

Last year, **supply chain threats evolved from an emerging concern among a few major powers to a central cybersecurity battleground**, exacerbated by the increasing complexity and interconnectedness of global supply networks (see Figure 3). Threat actors target supply chains to exploit vulnerabilities in third-party vendors, subcontractors and software providers in the hope of **compromising many downstream customers in government**, industrial production and defence. Such attacks often bypass traditional cybersecurity defences, leveraging trusted relationships between vendors and customers to gain unauthorized access to critical systems.

BOX 8.

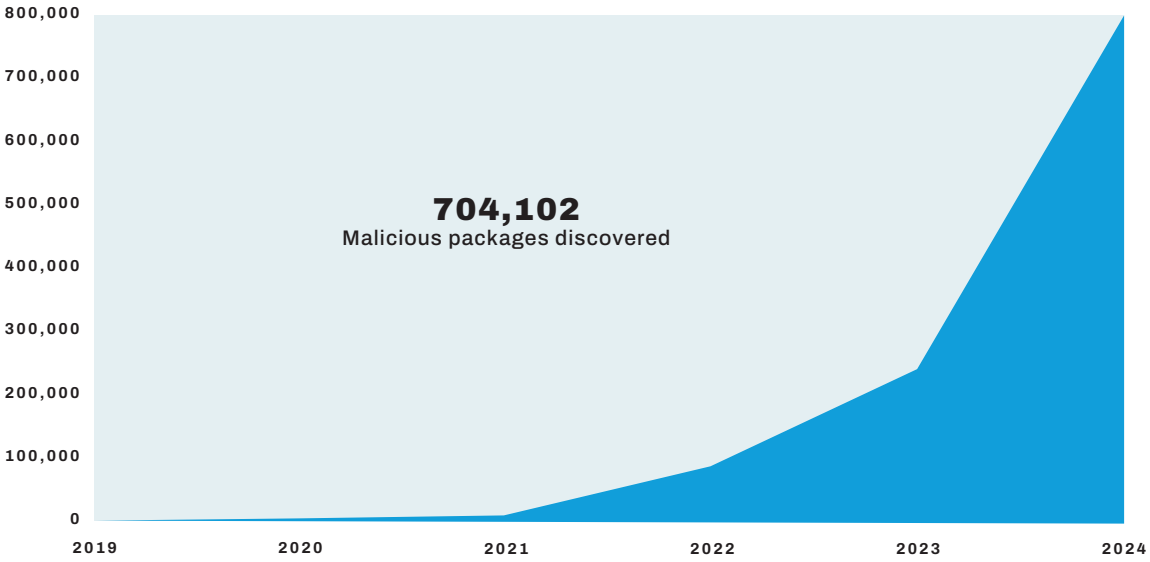
How supply chain attacks work

A supply chain attack compromises a third-party vendor or a software update in order to infiltrate a larger set of organizations or specific downstream customers. The attacker injects malicious code into trusted software or exploits vendor credentials to gain unauthorized access. Such attacks exploit the interconnected nature of modern supply chains, allowing a single breach to cascade across multiple systems and organizations on the basis of the "*attack one, compromise many*" principle.

Supply chain attacks not only increased in frequency but also grew in sophistication.[78] Attackers moved beyond traditional software exploits to hardware manipulation,[79] cloud-based attacks[80] and the insertion of back doors into widely used enterprise IT solutions. Attackers also started to focus on **embedding malicious code in widely used software platforms** before they reach end users, with the aim of compromising thousands of organizations through a single breach. Concerningly, attackers also **exploited legitimate software updates and patches as a vector of attack**,[81] raising alarms in the cybersecurity community over both the indiscriminate nature of such incidents and the undermining of trust in the very process – patching – that is meant to keep the digital ecosystem secure.[82]

FIGURE 3.

Next-generation software supply chain attacks, 2019–2024



Malicious actors have also started to distribute online pirated versions of legitimate software that contain back doors with greater frequency. By exploiting users who seek free or unauthorized software copies, these actors can gain covert access to multiple systems.[83] The back-doored versions not only provide direct access to compromised systems but are also used to silently enlist machines into botnets,[84] which are later leveraged for large-scale ransomware campaigns, DDoS attacks and further supply chain compromises. **Hardware supply chain compromises – where adversaries tamper with physical components prior to distribution – also rose in strategic relevance**, highlighted by recent incidents involving widely circulated consumer devices.[85] This has raised concerns not only over long-term espionage, hardware compromise and pre-positioning risks,[86] but also over the indiscriminate and widespread targeting of such intrusion tactics.[87]

BOX 9.

Supply chain attack vectors

Supply chain attacks exploit the interconnected nature of modern digital ecosystems to infiltrate multiple organizations through a single point of compromise. This infiltration can follow one of the following vectors:

- **Injecting malicious code into software updates** or firmware patches, infecting downstream users when the compromised product is deployed
- **Targeting third-party vendors**, suppliers or service providers, exploiting trusted relationships to bypass security controls
- **Tampering with hardware components or manufacturing processes**, embedding vulnerabilities before devices reach end users
- **Leveraging stolen credentials or remote-access privileges** from suppliers to gain entry into primary targets

These methods allowed attackers to infiltrate multiple organizations simultaneously, often going undetected for months before discovery.[88] Notably, **the number of software supply chain attacks doubled last year**,[89] underscoring the vulnerabilities of interconnected systems and the growing reliance on third-party software components in modern digital infrastructure. Key recent examples include the exploitation of widely used software platforms and cloud services,[90] where a single compromise can indiscriminately affect a multitude of downstream entities.

As **supply chain attacks have become more successful and scalable**,[91] both state actors and cybercriminal organizations are reported to have intensified their focus on exploiting third-party vendors.[92] **State-affiliated groups** have been accused of compromising technology and service providers to pre-position in adversary networks for long-term intelligence collection or **future disruptive operations**.[93] These actors may rely on supply chain attacks to infiltrate **critical infrastructure sectors** (e.g., telecommunications, energy and defence industries), enabling them to maintain persistent access within high-value targets.[94]

Meanwhile, as sophisticated cyberattacks became more widely available through CaaS models, cybercriminals **have also adopted supply chain attack techniques** as part of ransomware and financial fraud operations.[95] By targeting managed service providers (MSPs) and cloud platforms, attackers can **encrypt or steal data across entire client networks simultaneously, maximizing the scale of extortion attempts**.[96] This trend has raised alarms within the private sector and government agencies, as a single breach in a supply chain can now escalate into a systemic crisis with cascading national, regional or global consequences.[97]

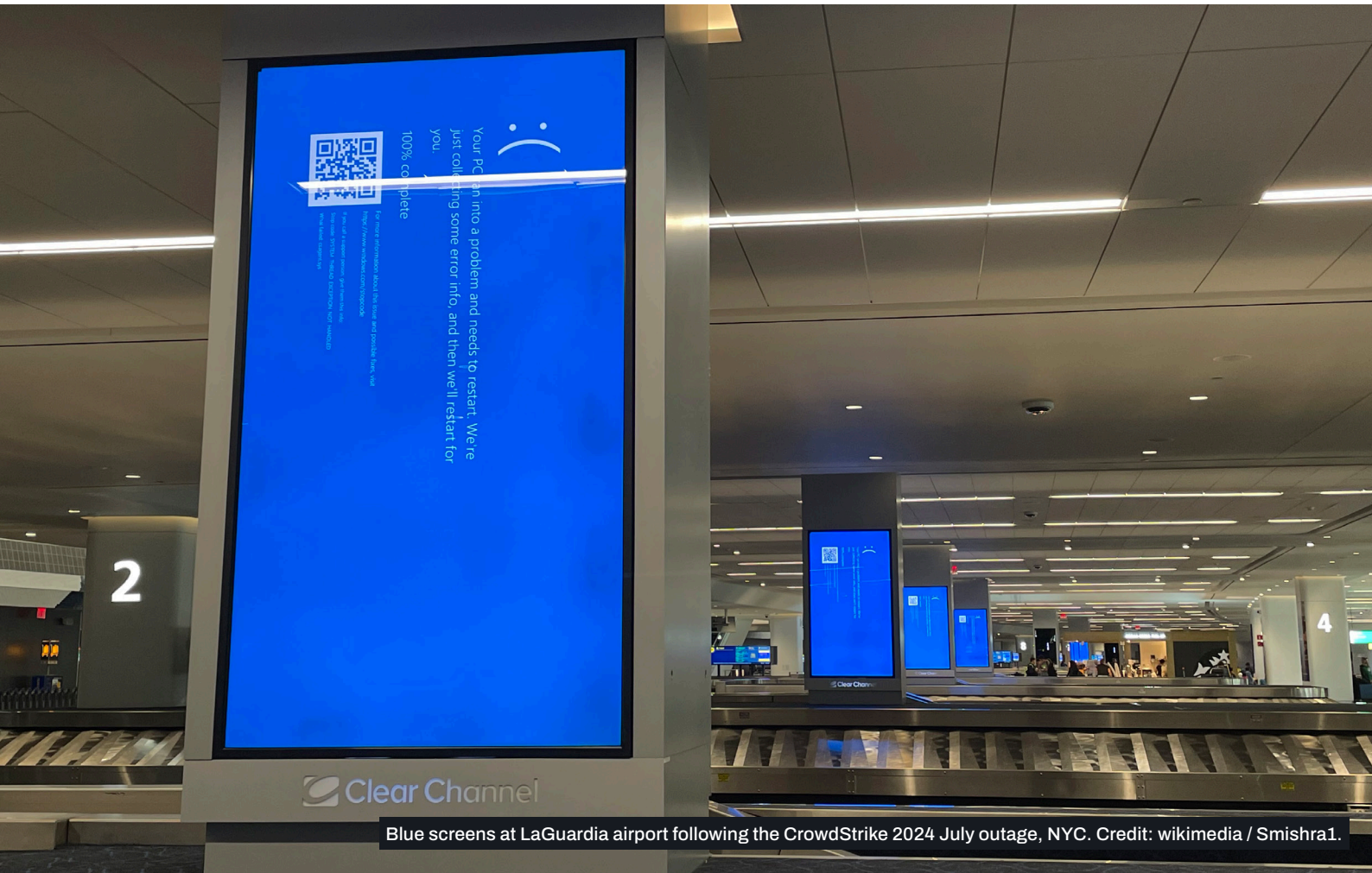
Several high-profile supply chain attacks and disruptions last year illustrated the severity of this threat vector. In one instance, attackers **compromised widely used IT-management software** and inserted malicious updates that infiltrated thousands of companies and government agencies worldwide. **Cloud service providers** faced several breaches during the year that exposed sensitive corporate and consumer data stored across multiple downstream customers. The financial sector was targeted through software supply chain compromises of **payment processing networks, fintech platforms and digital banking services**,[98] which can potentially threaten the security of billions of financial transactions.[99] As supply chains become more digitalized and interconnected, attackers **are likely to exploit this transformation to move beyond traditional attack vectors**, and to compromise software vendors instead of directly targeting high-security financial institutions.[100]

BOX 10.

A single faulty software update, global chaos: The 2024 CrowdStrike incident

In July 2024, a faulty software update from cybersecurity firm CrowdStrike led to one of the most extensive IT outages in history, affecting approximately 8.5 million systems worldwide. This incident, although not a cyberattack, caused widespread disruptions across multiple sectors. Airlines around the world experienced widespread flight cancellations, financial institutions faced operational disruptions and healthcare services were interrupted. This event demonstrated how a single point of failure within a supply chain – stemming from an altered code – can precipitate global operational paralysis, leading to significant economic losses and jeopardizing public safety.[101]

The consequences of **supply chain breaches can extend far beyond individual companies or financial losses to affect critical infrastructure and national security.** Malicious actors have reportedly targeted international supply networks serving governments, public services and military organizations,[102] creating security and escalatory risks that transcend national borders.[103] In particular, sectors reliant on international vendors – such as energy, telecommunications, healthcare and defence – are facing heightened risks of supply chain compromises.[104] **Developing states may remain particularly vulnerable to supply chain risks due to limited resources, regulatory gaps and a lack of the cybersecurity expertise needed to effectively monitor and secure supply chains**.



Blue screens at LaGuardia airport following the CrowdStrike 2024 July outage, NYC. Credit: wikimedia / Smishra1.

Finally, **cyberattacks on supply chains pose growing risks to global trade**: cyber disruptions to key sectors such as semiconductor manufacturing and cloud infrastructure could trigger widespread financial, manufacturing and operational consequences. **As demonstrated by past semiconductor shortages**,[105] **cyber-induced delays and the exploitation of supply chain vulnerabilities can present serious challenges** for both economic resilience and security policy.

In short, supply chain threats **are no longer an emerging risk affecting few states or a tool available to only a few highly skilled state-affiliated actors – they are now a central battleground in cybersecurity**. As attackers refine their tactics to exploit the complexity of modern supply chains, ensuring the security of these intertwined digital ecosystems will require **a proactive, multilayered approach**. Governments, businesses and the technical community should work together to strengthen supply chain defences, minimize systemic risks, and uphold the security of global commerce and critical infrastructure.

Recognizing the threat of cyber intrusions in supply chains, **United Nations Member States have expressed concern over the exploitation of ICT product vulnerabilities**, the use of harmful hidden functions, and the potential impact of compromised supply chains on international peace and security.[106] Discussions in the OEWG have emphasized that attacks on supply chains not only threaten individual organizations but also have cascading consequences across critical infrastructure, financial systems and global markets. In this context, states in the OEWG has **underscored the risks posed by third-party dependencies, lack of transparency in ICT manufacturing and the proliferation of malicious cyber tools**, all of which contribute to the increasing complexity of supply chain security.[107]

In response, **states have called for stronger cooperation and common security standards to ensure supply chain integrity**. Agreed recommendations include **developing risk-management frameworks**, encouraging ICT vendors to implement s**ecure-by-design principles** and **avoiding the introduction of harmful hidden functions** that could compromise system security. The OEWG discussions have also highlighted the need for independent certification mechanisms to validate ICT product security, as well as enhanced public–private partnerships to foster transparency and facilitate the exchange of best practices.[108] Strengthening international information-sharing mechanisms on supply chain security and good practices were also central elements of these discussions.[109]

Despite these commitments, **challenges remain in translating agreements into concrete enforcement mechanisms**. Diverging national policies, resource disparities, and concerns over economic and technological dependencies complicate the implementation of supply chain security measures. Some states also emphasized the need to ensure that supply chain security efforts do not unintentionally disadvantage developing economies or reinforce existing digital divides.[110] Moving forward, **sustained diplomatic engagement, regulatory harmonization and investment in capacity-building** will be necessary to secure global supply chains and to mitigate emerging cyber risks.

1.4. Cyber-enabled influence operations: From hack-and-leak to AI-generated manipulation

Cyber-enabled influence operations grew more sophisticated in recent years, becoming a potent tool for undermining public trust in institutions, destabilizing societies and shaping geopolitical landscapes.[111] The **fusion of AI-generated content with advanced social engineering tactics** has enabled threat actors – including state-backed entities, terrorist networks and hacktivist groups – to amplify their influence campaigns with new precision, speed and reach.[112] This subsection examines the main cyber-enabled influence operations and provides an overview of the impact of these tactics.

BOX 11.

How cyber-enabled influence operations work

Cyber-enabled influence operations may exploit digital platforms, AI-generated content and cyber intrusions to manipulate public perception and achieve strategic objectives. The key methods include:

- **Hack-and-Leak Tactics:** Cyber intrusions can be used to steal sensitive documents, which are then selectively leaked, – altered or manipulated before release – to serve political or strategic aims.
- **Deepfake Propaganda:** AI-generated fake videos, audio clips and images impersonating public figures may spread misleading narratives.
- **Automated Disinformation Campaigns:** Bot networks and coordinated troll operations can flood online spaces with content to sway public opinion.
- **Electoral Interference:** Cyberattacks on voter registration systems, election commission websites or digital voting infrastructure can undermine confidence in the integrity of elections, even without directly altering results.
- **Algorithmic Amplification:** Exploiting or altering social media recommendation algorithms can amplify divisive narratives and reinforce societal biases.

The year **2024 was notable for its extensive electoral activity**, with more than 60 countries – home to over half the world's population – holding elections. This concentration of electoral events heightened concerns about the potential misuse of AI to disrupt democratic processes [113] through disinformation campaigns that could erode confidence in election results by spreading false claims about electoral fraud,[114] manipulate public opinion polls and launch smear campaigns.[115] Despite initial fears, the direct impact of AI-generated content on the 2024 election cycle was reported to be less significant than anticipated.[116] Experts believe that a combination of regulatory measures, industry self-regulation and public scepticism towards AI content may have mitigated its effectiveness for now.[117] However, AI-generated dis- and misinformation was detected in elections across various regions, indicating a **potential trend towards more sophisticated and harder-to-detect cyber influence operations** in the future.[118]

BOX 12.

Cyber-enabled dis- and misinformation

Disinformation campaigns involve the deliberate spread of false information to manipulate public opinion, while misinformation refers to the unintentional spread of inaccurate information. Cyber-enabled influence campaigns may spread and amplify certain narratives using a combination of cyberattacks, content placement on social media platforms, bots and AI-generated content (e.g. deepfakes). These campaigns may exploit existing algorithmic biases to target specific demographics in order to influence elections and policies and affect social cohesion.[119]

Notably, the use of hack-and-leak operations in cyber-enabled influence campaigns was reported in the 2024 election cycle.[120] These involved cyber intrusions into entities such as government agencies, political parties, media organizations or public figures, followed by the selective leaking of stolen data to shape public narratives.[121] The operations sought to exploit the perceived authenticity of leaked documents, even when altered or taken out of context, in order to interfere in elections, discredit political figures or influence public policy.

The **growing accessibility of generative AI** has enabled threat actors to enhance the plausibility of disinformation campaigns by creating hyper-realistic deepfake videos, synthetic voice recordings, altered images and automated fake news articles.[122] Last year, AI-generated content impersonating political figures,[123] fabricating statements[124] and spreading false narratives was recorded across different regions.[125] Technology has also amplified the **ability of foreign actors to generate convincing narratives in local languages**, allowing disinformation campaigns to be more persuasive and tailored to specific cultural and political contexts.[126] While there is little evidence to suggest that these efforts decisively altered recent electoral outcomes, **as AI tools become more refined and accessible, their role in disinformation is expected to expand**.[127] This will make detection and mitigation ever more challenging.[128]

Additionally, reported cyber intrusion **attempts targeting voting infrastructure have risen**,[129] with attackers focusing on voter-registration databases, digital voting systems and election commission websites.[130] These incidents may not only threaten electoral integrity but may also provide a pretext for discrediting democratic processes and undermine public trust in elections.[131]

BOX 13.

Case study: Cyberattack on election infrastructure

In April 2024, the computer infrastructure of a county in Georgia, United States, was compromised in a cyberattack, which prompted state-level officials to sever the county's access to Georgia-wide election systems. The incident involved unauthorized cyber activity within the county's IT infrastructure, leading to concerns about the security of election-related systems. While there was no evidence of data exfiltration, the attack underscored vulnerabilities in local election infrastructures and highlighted the potential risks posed by cyberthreats to electoral integrity.[132]

Beyond elections, **social media platforms remained among the primary vectors for many cyber-enabled influence campaigns**.[133] Threat actors sought to exploit platform algorithms to amplify divisive narratives, using bot networks and coordinated disinformation campaigns to manipulate online discourse.[134] Algorithms that are designed to maximize engagement may inadvertently prioritize sensational and polarizing content,[135] and so create fertile ground for disinformation and manipulation. The proliferation of bots and fake accounts has added to the difficulty for users to differentiate between legitimate discourse and orchestrated influence operations.

By leveraging existing algorithmic biases, malicious actors can also maximize the probability of misleading content reaching the most susceptible target audiences and can maximize the political and social impact of deliberately polarizing content. **Biases reported to be embedded in some algorithmic systems may exacerbate these challenges**, particularly in contexts involving gender and race.

However, **societal impacts of cyber-enabled influence campaigns** can affect not just electoral politics, but also public health, social stability and even conflict dynamics.[136] Health misinformation, for example, has hindered public health initiatives, exacerbating crises such as the Covid-19 pandemic by spreading false narratives about treatments, vaccines and government responses.[137] Similarly, **disinformation campaigns that target marginalized communities** have deepened social inequalities, fuelling hate speech, violence and societal polarization.[138] In the context of armed conflict, c**yber-enabled influence operations may be used to distort reporting on humanitarian crises**, manipulate narratives around military actions and undermine trust in international institutions.[139] The ability to fabricate or selectively alter digital content at scale has made it easier for malicious actors to sow confusion, justify aggression and erode public confidence in verified information, further complicating crisis response and diplomatic efforts.

BOX 14.

AI-generated war imagery

In 2023, AI-generated images depicting fictionalized scenes from the Israel–Hamas conflict – air strikes, destroyed buildings and conflict zones – were mistakenly used by multiple news outlets after being uploaded to a stock image platform.[140] These synthetic images, which were created by generative AI tools, raised concerns about the integrity of war journalism. While not necessarily intended as disinformation, their unintended use undermined public trust in factual war reporting, allowing various actors to question the credibility of war coverage and downplay the scale of devastation. This incident highlighted how AI-generated content, even when not maliciously deployed, can fuel influence operations by distorting public perception, eroding confidence in verified information, and complicating efforts to maintain accurate crisis reporting.

Recognizing the escalating risks posed by cyber-enabled influence operations, **United Nations Member States have underscored the urgent need for collective action to protect the integrity of democratic processes** and information ecosystems. The OEWG on ICT security has emphasized that malicious ICT activities, particularly those undermining trust in electoral processes, present a concern for international peace and security.[141] Many s**tates and international organizations**[142] **have also highlighted the convergence between cyberthreats and influence operations**, warning that these campaigns may not only disrupt domestic stability but also exacerbate geopolitical tensions and erode public trust in governance.[143]

In response to these challenges, several **measures aimed at mitigating threats posed by cyber-enabled influence operations** can be considered. Strengthening election security has been a primary focus, with an emphasis on improving the cybersecurity defences of digital election infrastructure,[144] **enhancing resilience against cyberattacks** and c**ountering hack-and-leak disinformation campaigns**.[145] **Promoting digital literacy and public awareness** has also been identified as a key priority,[146] equipping individuals with the skills needed to recognize and resist manipulative content and algorithmically amplified disinformation. Furthermore, addressing algorithmic biases requires improving representative data sets for training algorithms, auditioning systems to identify and fix biased code,[147] providing gender-sensitization to policymakers,[148] and encouraging multi-stakeholder collaboration to distil good practices to guide the development, deployment and use of social media algorithms.[149] Finally, **fostering greater transparency around AI-generated content** – through regulatory frameworks,[150] public–private partnerships, and enhanced content scanning – has been recognized as essential to curbing the spread of deceptive narratives.[151]

However, significant challenges remain in implementing effective mitigation measures. The rapid advancement of generative AI tools, combined with diverging national policies on information regulation, complicates regulatory and enforcement efforts. As highlighted by recent trends, threat actors continue to refine their tactics while detection and response mechanisms struggle to keep pace. Additionally, the decentralized nature of digital platforms and the global reach of cyber-enabled disinformation campaigns make it difficult to hold actors accountable.

The 2024 "AI election year" may not have seen the widespread effects of disinformation that were feared, but the trends **emerging suggest that the next wave of influence operations may be even more sophisticated**, deceptive and difficult to counter. As adversaries refine their methods, **cyber-enabled influence operations may no longer be just a challenge for media regulation or election oversight**; they also represent an important concern for cybersecurity professionals. These recent developments also underscore the importance of governments, technology companies and civil society working together to build resilience against cyber-enabled influence campaigns while preserving the integrity of democratic processes, public trust and international stability.

2. Cyberthreat actors: Blurred lines and growing actor complexity

The contemporary cyberthreat landscape is shaped by an **increasingly diverse and interconnected web of actors**, spanning state-led operations, cybercriminal syndicates, ideological hacktivists, commercial cyber mercenaries and even civilian volunteers. These actors operate with varying levels of resources, technical sophistication and links to states. This variety challenges traditional frameworks for attribution, responsibility and accountability in cyberspace. The result is a **more crowded and fragmented threat actor environment** in which the lines between state and non-state activity, between legitimate and illegitimate conduct, and between civilian and combatant roles are becoming ever more difficult to define.

This **blurring of boundaries between state and non-state cyber actors is being driven by several overlapping trends**. States continue to develop advanced cyber capabilities, but some may also rely on proxies, private contractors or may tolerate non-state actors carrying out certain cyber operations, often to maintain plausible deniability. Meanwhile, the proliferation of cybercrime-as-a-service models and offensive tools for hire highlighted in Section 1 has allowed even low-resourced actors to execute disruptive attacks with global consequences. **Civilian actors – including individuals and private companies – have also been reported to play more prominent roles**, particularly during armed conflict, where they may support cyber defence, sustain critical infrastructure or, in some cases, engage directly in cyberattacks against perceived adversaries. These developments raise not only operational and legal questions but also broader concerns about future cyberspace stability and the protection of civilians engaged in cyber activities in the context of armed conflict.

To help navigate this increasingly complex landscape, this section examines cyberthreat actors through three key dimensions:

- **Resources and Capabilities:** Exploring the range of technical sophistication, access to tools and operational reach among actors, from highly resourced state-linked groups to small-scale actors empowered by commoditized malware services
- **Spectrum of State Involvement:** Analysing how states engage with, tolerate or enable different cyber actors, including through proxies and partnerships with private and civilian entities
- **Motivations and Objectives:** Mapping the varied drivers behind malicious cyber activities, including financial gain, strategic competition, political disruption and hybrid uses of cybercrime for state ends

Taken together, these trends suggest that understanding the cyberthreat landscape today requires not only identifying who is behind an attack, but also assessing how actors operate, whom they serve and what they aim to achieve – and whether existing international frameworks are equipped to respond.

2.1. Resources and capabilities

The cyberthreat actor landscape is marked by a diverse array of entities with varying resources and capabilities. These actors can be broadly categorized into three groups:

1. **Highly Capable and Resourceful Actors:** This group includes a few state actors with substantial resources and sophisticated capabilities. They may be engaged in long-term, strategic cyber operations targeting critical infrastructure and sensitive government data. Some major private sector actors playing supportive roles can also fall into this category.
2. **Moderately Capable Actors:** Organized criminal groups, most hack-for-hire services and some state actors with nascent cyber capabilities may fall within this category. These entities utilize moderately advanced tools for financial gain or strategic advantage. They may engage in ransomware attacks, supply chain compromises or espionage.
3. **Low-Resource Actors:** This group encompasses individual hackers, hacktivists and small-scale criminal networks engaging in cyberattacks for profit or to advance specific political and social causes. AI and CaaS models have particularly empowered these types of actors by lowering the technical expertise required to conduct effective cyberattacks, which tend to target either individuals or small and medium-sized organizations.

2.1.1. Highly capable and resourceful actors

The first category, **highly capable and resourceful actors**, consists of state actors and some non-state entities with access to significant resources and advanced technological capabilities. These actors are often engaged in long-term, strategic operations that target or pre-position in critical infrastructure and supply chains in order to collect sensitive data. Highly capable actors may still leverage proxies, such as hacktivist groups, for certain cyber operations to maintain plausible deniability. At the same time, some highly capable and well-resourced **private cybersecurity firms and cloud and satellite service providers have been thrust into conflict dynamics**.[152] Some actively **defend critical national infrastructure**[153] or provide **support to states on one or both sides of armed conflict**.[154]

2.1.2. Moderately capable actors

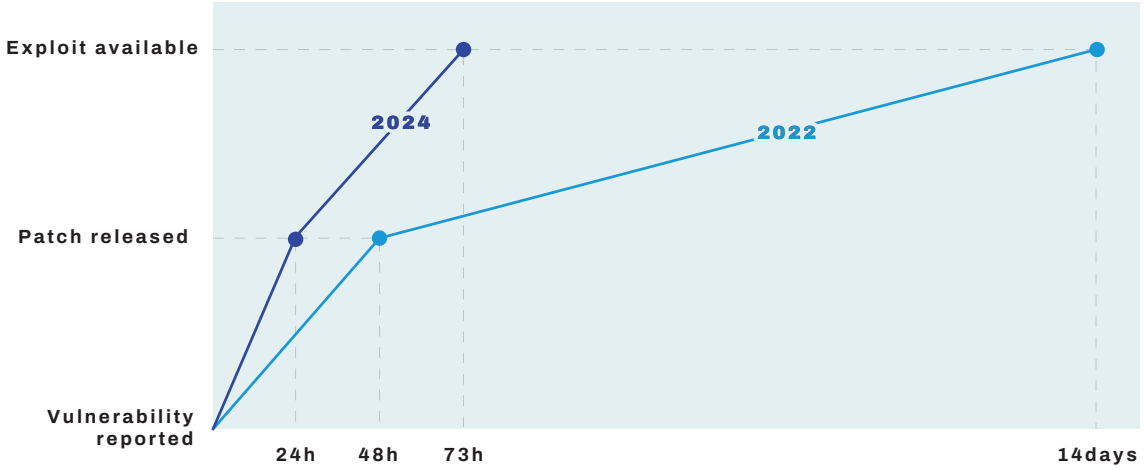
The category of **moderately capable actors** includes organized criminal groups, hack-for-hire services and some states with nascent cyber capabilities. These actors are primarily motivated by financial gain or the gaining of strategic advantage. They often engage in activities such as ransomware and distributed denial-of-service attacks, or corporate espionage and surveillance. As the cybercrime ecosystem becomes more sophisticated, services such as malware-as-a-service (MaaS) are increasingly enabling actors in this category to launch more effective attacks (e.g., supply chain compromise) that were previously reserved to few highly capable and resourceful actors. Moderately capable actors have presented an increasing threat due to proliferation of the AI tools that enable them to enhance operational impact across their activities.

As highlighted in Section 1, one of the most significant developments in recent years was the **evolution of cybercrime into fully developed CaaS and RaaS business models**. As a result, many moderately capable actors now benefit from sophisticated division of labour, with one

group focusing exclusively on hunting for vulnerabilities or developing malware payloads, another specializing in crafting convincing phishing messages or executing attacks, and yet another managing financial transactions and ransom negotiations. This creates an increasingly efficient, specialized and interdependent ecosystem.[155]

FIGURE 4.

Comparison of speed of ICT vulnerability exploitation, 2022–2024



The **rapid proliferation of ICT vulnerabilities** (see Figure 4) illustrates how the increasingly sophisticated division of labour among moderately resourced cybercriminal groups has significantly enhanced their ability to exploit weaknesses at scale, increasing the effectiveness of their attacks. **Some cybercriminal groups now specialize solely in vulnerability research**, hunting for weaknesses from public disclosures and released patches and selling them on the black, grey and even white markets.[156] As a result, the time window for patching critical vulnerabilities shrank from 14–30 days in 2022 to just 24–72 hours in 2024, as cybercriminals are now able to more rapidly identify, sell and weaponize ICT vulnerabilities.[157] This **escalating arms race between cybersecurity vendors and cybercriminal groups** has placed immense pressure on organizations to patch vulnerabilities immediately, as delays of even a few days can leave systems exposed to immediate exploitation. However, the **challenge of patching systems quickly may be compounded by cyberattacks targeting software updates** – if supply chain attacks continue to compromise update mechanisms, testing patches may take longer or users may be hesitant to apply them, which could further amplify cybersecurity risks.

2.1.3. Low-resource actors

The final category, **low-resource actors**, encompasses **individual hackers, hacktivists and small-scale criminal networks**. While these actors typically lack the resources and sophistication of the other two categories, advancements in technology and now widely available CaaS services have lowered the barriers to entry for this category of actors. Low-resource actors can now execute more impactful attacks with minimal technical expertise.

Low-resource cyber actors traditionally rely on a range of low-cost yet disruptive cyber tactics to maximize their impact despite lacking the advanced capabilities of highly capable actors. **DDoS attacks remain a preferred method of hacktivists**, as they can overwhelm government,

corporate and media websites, temporarily disabling services and drawing public attention to various political or societal causes.[158] **Website defacements**, where hackers replace official content with political messages, propaganda or nationalist symbols, are also widely used to generate visibility.[159] Additionally, **malware attacks**, including **AI-enhanced phishing campaigns**, have become more prevalent among this type of actor, enabling infiltration of networks and exfiltration of sensitive data for ransom, public release or both.[160] Recent trends suggest that some **low-resource actors are also moving beyond purely symbolic disruptions and are increasingly targeting critical infrastructure**[161] (e.g., energy grids, transportation networks, water infrastructure[162] and financial institutions) through coordinated waves of DDoS attacks that can escalate tensions and introduce wider security risks.[163]

2.2. Spectrum of state involvement

The relationship between cyberattacks and those behind them is often complex, defying a clear-cut division between state and non-state actors. In practice, **cyber actors can operate along a broad spectrum of relationships with states**, ranging from fully independent non-state actors acting without any government awareness, to entities that are directly integrated into state structures such as security, military or intelligence services. Between these ends lies a grey zone in which non-state actors may receive tacit approval, indirect support or explicit coordination from state actors, forming a varied **nexus of relationships** (see Figure 5). While cyber activities are theoretically expected either to fall under formal state authority or to be criminalized

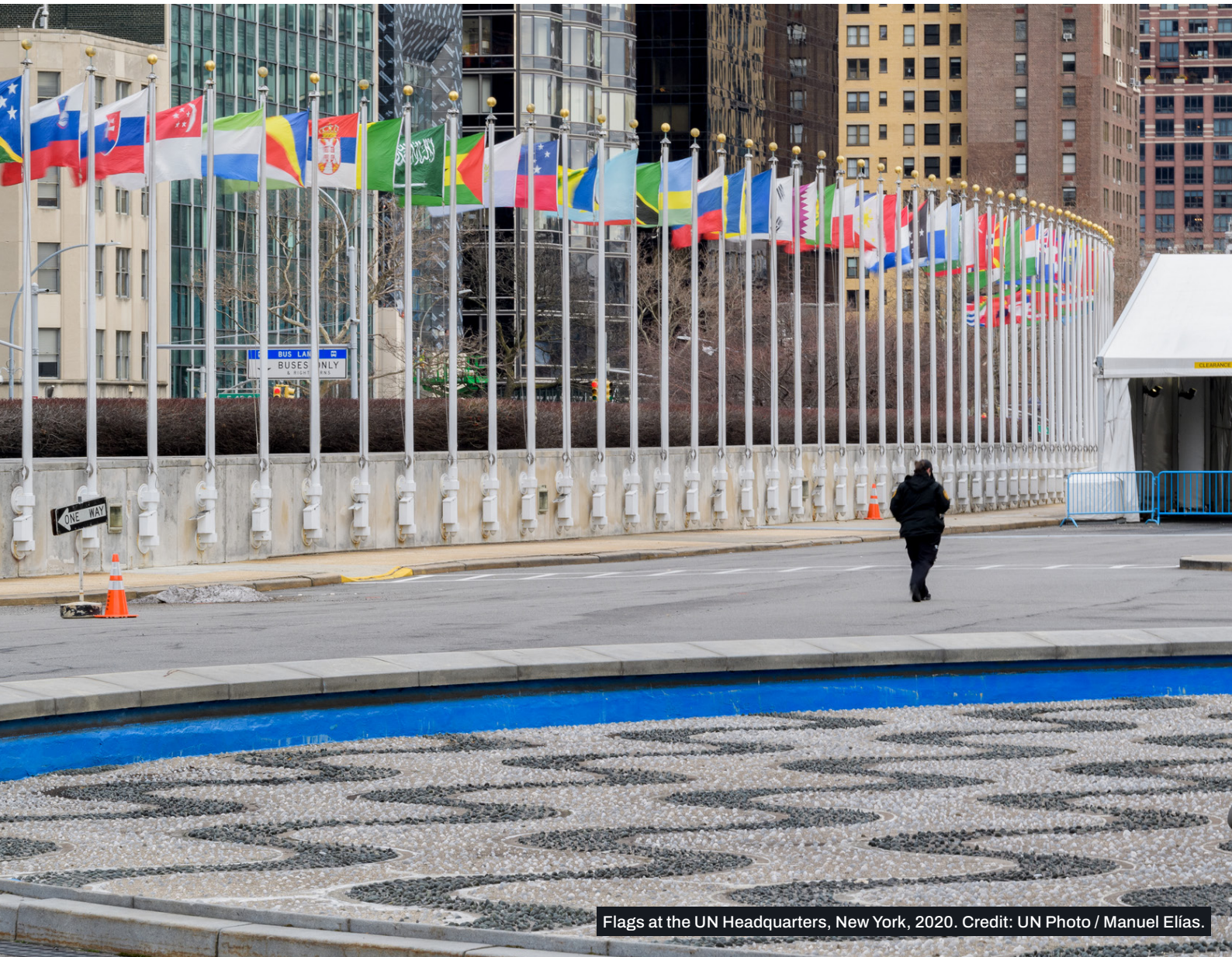
FIGURE 5.

Spectrum of state involvement in cyber activities

Note: This taxonomy is a conceptual tool to provide granularity beyond the binary 'state' vs. 'non-state' distinction.

when undertaken by non-state actors, the reality includes a continuum of possible arrangements where states may turn a blind eye to malicious activities originating from their territory, provide selective encouragement, or even integrate non-state groups into state-directed operations.

These varied relationships have significant implications for understanding motivations and assessing responsibility. States may selectively suppress malicious cyber activities when they conflict with national interests or legal obligations, yet in other circumstances may ignore, tolerate or support such activities when they offer strategic advantage or political utility. This **layered landscape complicates attribution by obscuring the degree of state involvement**, challenges effective law enforcement where jurisdiction and intent are unclear, and affects diplomatic responses when states deny direct control yet benefit from the outcomes. Recognizing the spectrum of possible state–actor relationships is therefore essential for evaluating the role of proxies in cyberspace and for upholding the *United Nations Framework for Responsible State Behaviour in Cyberspace*.



Flags at the UN Headquarters, New York, 2020. Credit: UN Photo / Manuel Elías.

BOX 15.

The spectrum of state involvement in cyber activities

While malicious cyber activities are often framed in binary terms – either "state" or "non-state" driven – the reality is more complex. In practice, a wide range of relationships may exist between cyber actors and states, from full state control to complete independence. The following conceptual taxonomy is provided to add granularity to that oversimplified dichotomy. It is intended as an analytical tool to illustrate the spectrum of possible state involvement.

- **State-Prohibited:** The government opposes cyberattacks emanating from within its territory and actively works to prevent, disrupt or prosecute perpetrators, including through legislation and other measures.
- **State-Prohibited but no Enforcement:** The government either lacks awareness of cyber activities within its territory or lacks the capacity or resources to effectively prevent or stop attacks.
- **State-Tolerated:** The government is aware of third-party cyber activities within its territory but takes no meaningful action to stop them, allowing cybercriminals or hacktivist groups to operate freely.
- **State-Encouraged:** The government does not directly control cyber actors but tacitly supports or publicly praises their actions as being aligning with national interests.
- **State-Assisted:** The government provides material assistance, intelligence or operational resources to cyber actors without direct control over their operations.
- **State-Coordinated:** The government influences and coordinates cyber activities by third parties, including setting operational guidelines, or providing strategic guidance on types of targets.
- **State-Directed:** Cyber actors operate under explicit state orders, executing attacks based on government directives while maintaining nominal independence for deniability.
- **State-Executed:** The government carries out cyber operations directly through official state-controlled cyber units, military forces or integrated allied capabilities.

These categories are not intended to correlate with legal definitions for the purposes of attribution under the international law of state responsibility.

Importantly, **these relationships may not be static and do not always correlate directly with the scale or impact of cyberattacks**. A state that initially turns a blind eye or tacitly supports certain cyber activities by non-state actors may later distance itself and take law enforcement action due to evolving diplomatic, legal or strategic concerns. Conversely, **cyber actors operating independently may, over time, develop closer ties to government entities**, particularly if their actions align with national security objectives. This fluidity is evident in some reported cases where non-state groups evolved from loosely affiliated actors into more structured operations with implicit or explicit state backing.[164] In some cases, **a state may avoid taking direct action against cyber activities emanating from within its territory by signalling tacit acceptance** through inaction or ambivalence.[165] In other cases, a cyber group that demonstrates operational effectiveness may attract growing levels of state interest, oversight and coordination over time.[166]

Apart from the blurring of lines between state and non-state activities in cyberspace, another notable trend is the **growing involvement of civilian cyber actors in armed conflicts**. These civilian actors may include hacktivists, "patriotic" hackers, cybersecurity professionals and private companies, among others. **On the defensive side, some private companies have assumed an active role in armed conflict**,[167] either voluntarily or due to existing legal obligations and contractual arrangements. Their roles can include safeguarding national critical infrastructure from cyberthreats, restoring essential services following DDoS attacks or sharing cyberthreat intelligence with government authorities.[168] For example, cloud service providers have migrated critical data outside conflict zones to protect key government assets,[169] while satellite operators have maintained Internet connectivity and provided satellite imagery to mitigate the impact of damaged physical infrastructure and to support intelligence efforts.[170] Cybersecurity firms, meanwhile, have played a crucial role in both countering cyberattacks and restoring the functionality of critical government services and operations following successful cyber intrusions.[171]

While efforts by private sector actors can help ensure the delivery of essential services, the actors' involvement in the context of armed conflict may also interfere with or impede an adversary's military cyber operations, raising complex legal, ethical and security implications.[172] For instance, **such interference may result in such an actor losing protected civilian status under international humanitarian law (IHL)**,[173] and so potentially increases the risk that cybersecurity professionals and infrastructure operated by the private sector are exposed to potential retaliation.[174] This evolving dynamic underscores the **need for greater clarity around the roles and responsibilities of private sector actors in conflict dynamics** and the need to balance the necessity of cyberdefence with the principles of distinction between civilians and combatants under IHL.

On the offensive side, civilian involvement is often informal but can also have significant impact, particularly when individuals coordinate through digital platforms to conduct attacks on behalf of their country or an ideological cause. So far, **civilian hackers seem to have favoured disruptive tactics over strategic objectives**,[175] prioritizing high-visibility targets such as DDoS attacks on government websites.[176] Some states are reported to have **actively encouraged participation of civilian hackers** and to have distributed tools or guidance to volunteers, creating further ambiguity in distinguishing between civilians and combatants during armed conflict.[177] However, the **use of civilian hackers presents risks to states themselves**

– they may act unpredictably, escalate conflicts or execute attacks that do not align with state objectives. Apart from raising strategic, law enforcement[178] and ethical questions,[179] the involvement of civilian hackers – including those operating from far beyond conflict zones – also raises the possibility of their losing protected civilian status under IHL.[180] This may put hackers and those around them at risk of retaliation.[181]

This trend also raises pressing **questions about the long-term trajectory of civilian hackers once conflicts subside**. Individuals who acquire offensive cyber capabilities during wartime may later shift into cybercriminal activity, offer mercenary hacking services, or continue conducting disruptive operations that undermine post-conflict stability. Without structured efforts to demobilize and reintegrate these hacktivists into civilian life, post-conflict periods could therefore see a sustained increase in cybercrime and other disruptive cyber activities.

These risks suggest that governments and international organizations may need to consider mechanisms that address several interrelated challenges: **de-escalation**, to ensure that civilian cyber mobilisation does not persist after hostilities end; **accountability**, to clarify responsibility for harmful activities conducted during conflict; and **reintegration**, to help individuals with newly acquired cyber skills transition away from harmful or illegal behaviour. Developing approaches in these areas may be important for **preventing the normalization of civilian involvement in cyber operations** and for reducing the likelihood that wartime hacktivism contributes to long-term instability in cyberspace, both within and beyond the immediate theatre of conflict.

Despite the risks outlined above, state actors may still encourage the involvement of non-state actors in cyber activities for several strategic reasons. **Plausible deniability** is a key factor for some offensive cyber activities, as civilian hackers can provide governments with a degree of separation from cyberattacks. By making technical, legal and political attribution more challenging, this reduces the risk of direct political or military repercussions.[182] Additionally, **civilian involvement may offer a strategic advantage**, allowing states to offset gaps in their existing cyber defensive and offensive capabilities by drawing on non-state actors who possess relevant expertise or resources.[183] Finally, **civilians can serve as short-term force multipliers:**[184] mobilizing them – whether through informal networks, online communities, direct state encouragement or formal partnerships – may often be faster and more cost-effective than establishing or expanding formal state cyber units. This can enable states to rapidly scale up their cyber capabilities in times of heightened geopolitical tensions or active conflict while maintaining a degree of separation.

Building on these dynamics, it is increasingly important to understand the broad spectrum of non-state actors that may be drawn into cyber operations, either independently or through varying degrees of state involvement. The **wide spectrum of non-state cyber actors** ranges from ideologically motivated hacktivists to private cybersecurity providers and organized criminal networks (see Figure 6 for a taxonomy). As cyberspace becomes more entangled with state interests and strategic objectives, understanding who these actors are – and how they relate to state structures – is essential for assessing risks, ensuring accountability and upholding the United Nations Framework for Responsible State Behaviour in Cyberspace.

BOX 16.

Taxonomy of non-state cyber actors

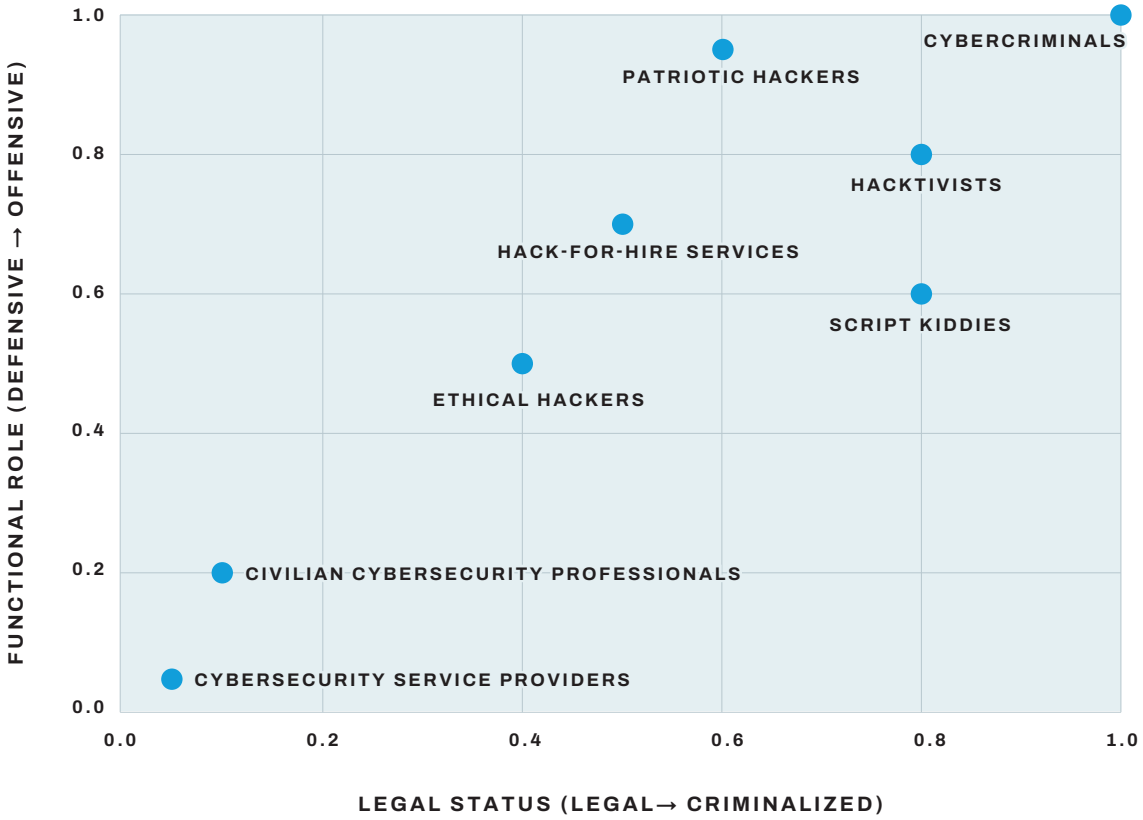
The cyberthreat landscape includes a diverse range of non-state actors, each with distinct motivations and tactics. These actors may operate legally, illegally or in a grey zone and may focus on either offensive or defensive cyber activities or a combination of these.

- **Cybercriminals:** Individuals and organized crime groups engaging in cyberattacks predominantly for financial gain
- **Hacktivists:** Ideologically driven hackers who conduct cyberattacks to promote political or social causes
- **Patriotic Hackers:** Nationalist actors who carry out cyberattacks in support of their country
- **Cybercriminals:**xIndividuals and organized crime groups engaging in cyberattacks predominantly for financial gain
- **Script Kiddies:** Inexperienced hackers using pre-existing tools to develop their technical skill
- **Ethical Hackers:** Security professionals and penetration testers who may enter systems with or without authorization to search for and report ICT vulnerabilities
- **Hack-for-Hire Services / Cyber Mercenaries:** Organizations or individuals offering offensive cyber capabilities and commercial spyware to both government and non-government clients for a fee
- **Cybersecurity Service Providers:** Private companies offering defensive cybersecurity services to customers
- **Civilian Cybersecurity Professionals:** Individuals defending national networks, sometimes voluntarily or as contractors or in direct employ of states

Global network cybersecurity (generated with AI). Credit: Adobe Stock / Farros.

FIGURE 6.

Taxonomy of non-state cyber actors



2.3. Objectives and motivations

The **objectives driving cyberthreat actors may be diverse and sometimes interrelated.** While financial gain remains a primary driver for criminal groups, state actors may be predominantly motivated by strategic objectives such as gaining competitive advantage, achieving disruption or increasing geopolitical influence. Hacktivists and ideological groups may seek to advance specific political and social causes, while individual actors may be driven by a combination of curiosity, notoriety or personal gain.

However, the **motivations driving cyberthreat actors are becoming increasingly complex** as the boundaries between state-sponsored, financially motivated and politically driven cyber operations continue to blur. **Some cybercriminal groups have been accused of collaborating with state actors to conduct strategic disruptions**, leveraging off-the-shelf ransomware tools to conduct cyber intrusions to serve geopolitical objectives.[185] At the same time, some hack-for-hire services, which traditionally operated for financial gain, have been accused of enabling states to conduct cyber operations that target political dissidents, journalists and human rights activists across multiple jurisdictions under the guise of law enforcement.[186] In 2024, a United Nations panel of experts concluded that a state actor may have used ransomware not only as a means of economic disruption but as a tool for financial gain, sanctions evasion and covert funding of illicit activities.[187]

BOX 17.

Possible motivations of state and non-state actors

Non-state actors

- **Financial Gain:** Many cybercriminal groups operate purely for monetary profit, leveraging ransomware, data theft and financial fraud.
- **Political Goals (Hacktivism):** Ideologically motivated groups launch cyberattacks to further political causes, protest against injustices or challenge governments and corporations.
- **Societal Disruption:** Some actors seek to maximize chaos, targeting essential services or critical infrastructure.
- **Intellectual Property Theft:** Malicious actors may steal sensitive military and corporate secrets in order to sell them or to publicly disclose them to harm the targeted organisations.
- **False-Flag Operations:** Some non-state actors can conduct cyberattacks designed to be misattributed to states, escalating geopolitical tensions.

State actors

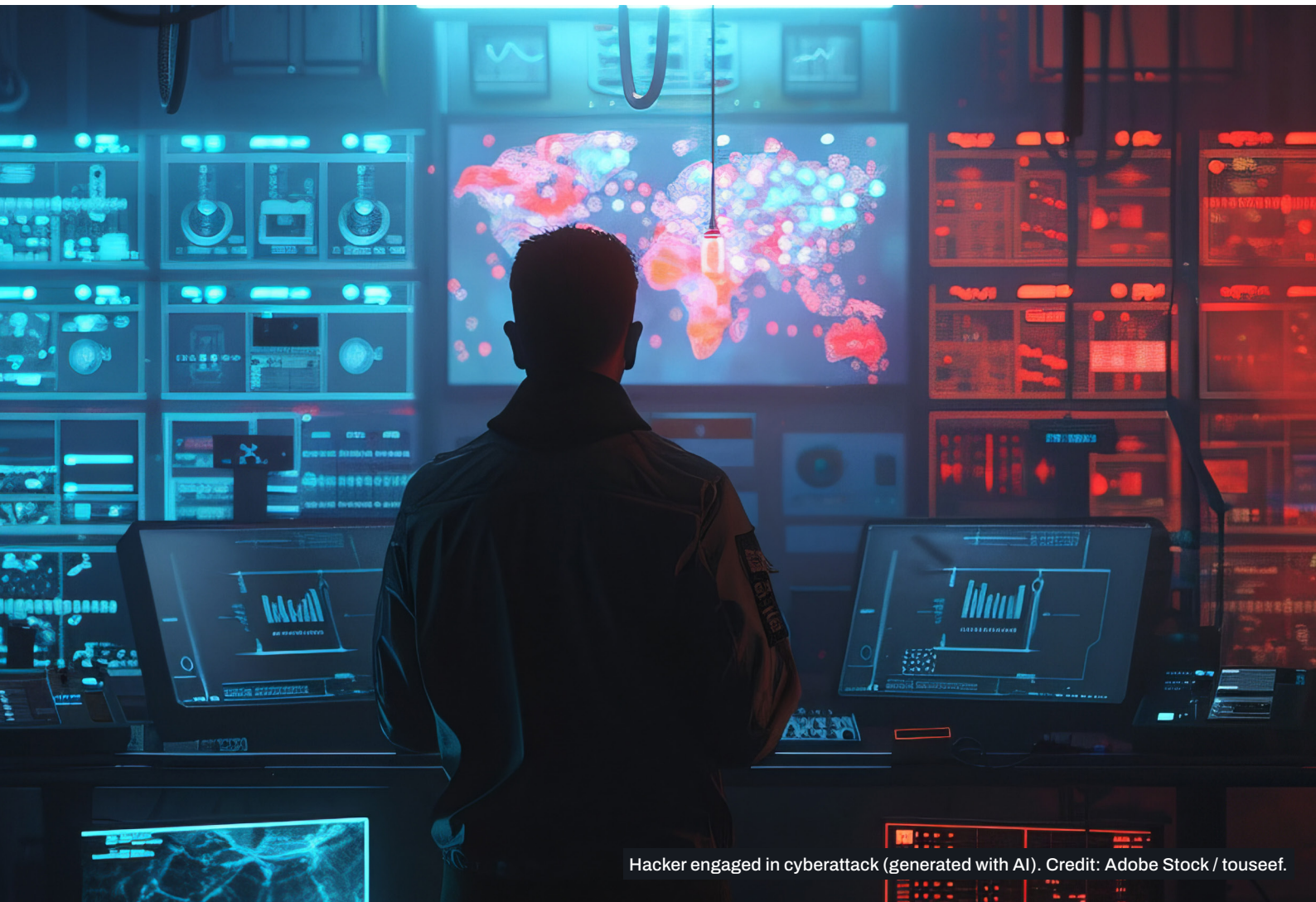
- **Strategic Signalling:** Cyber operations can serve as tacit warnings or tit-for-tat responses in times of geopolitical tensions.
- **Plausible Deniability:** States can use cybercriminal proxies to conduct cyberattacks while maintaining plausible deniability.
- **Financial Gain:** Some states, particularly those subject to sanctions, can use cyber operations for financial gain.
- **Cyber Espionage:** Cyber operations can be used for intelligence collection.
- **Future Advantage:** Cyber operations can be used for pre-positioning in critical infrastructure of adversaries to gain possible future strategic advantage, such as in the event of military escalation.
- **Force-Multiplier:** Cyberattacks may be used alongside conventional military tactics to weaken adversaries in the context of armed conflict.

Overall, the cyberthreat landscape has **fundamentally changed**, with the **traditional distinction between state and non-state actors fading.** Cybercriminals can **act as geopolitical tools**, civilians may act as force-multipliers for cyber operations during armed conflict and private corporations can **play an active role in cyberdefence**. This **blurred landscape of cyber actors, their capabilities and relationships with states presents significant challenges for attribution, governance and preserving stability in cyberspace**. These challenges are particularly acute in the context of armed conflict. While civilian actors may not be formally affiliated with any military, their actions can have real strategic impacts, potentially

putting them, their surroundings, and the digital infrastructure they use at risk of losing protection under international humanitarian law.

These trends highlight the **need for international action that addresses the role of non-state actors** and enforces the existing state obligations on non-use of proxies reaffirmed in the United Nations Framework of Responsible State Behaviour in Cyberspace.[188] Recognizing these challenges, the United Nations Secretary-General has explicitly warned against the rise of cyber mercenaries and hack-for-hire services, emphasizing their potential to exacerbate conflicts, destabilize states and operate beyond legal oversight.[189] **A number of states have also established rules around the use of hack-for-hire services**, such as through the Pall Mall Process Declaration.[190] Similarly, the private sector has taken steps to curb the growing trend of cyber mercenaries.[191] At the same time, the ICRC has published a legal and ethical framework with eight principles for civilian hackers involved in conflicts and outlined four key obligations for states to prevent civilian-led cyber activities from putting civilians and infrastructure at risk.[192]

These initiatives reflect a growing consensus on the urgent **need to address the evolving role of non-state cyber actors by strengthening international cooperation** and clarifying guardrails around the use of cyber capabilities that could put civilians at risks, lead to escalation, and undermine international peace and security. The rapid proliferation of advanced cyber capabilities beyond a few highly resourced actors further underscores the importance of capacity-building initiatives, particularly in developing countries, to enhance understanding of the quickly evolving threat actor landscape in cyberspace.



Hacker engaged in cyberattack (generated with AI). Credit: Adobe Stock / touseef.

3. Emerging technologies reshaping cyberspace: The double-edged sword

The year 2025 saw rapid advances in technologies that are fundamentally reshaping the cybersecurity landscape. In particular, **artificial intelligence and quantum computing gained further significant traction**, influencing both the thinking behind and execution of cyberthreats and defence strategies. These technological developments are not only changing the field of cybersecurity but may also have far-reaching implications for national security and international peace and stability. As digital infrastructure becomes intertwined with global stability, policymakers should stay informed about the ways in which these technologies are transforming the cyberthreat landscape.

AI is now a central element in cutting-edge cybersecurity defence strategies. It can enhance the speed and accuracy of threat detection, can automate responses and can strengthen resilience of digital infrastructure.[193] However, it is also being leveraged by attackers to carry out more sophisticated and scalable cyberattacks.[194] Meanwhile, **quantum computing, while still emerging, received renewed investment and achieved research breakthroughs in 2025**.[195] This raised pressing questions about its potential to break widely used encryption methods and disrupt secure communications and critical systems.

This **section examines how technological advances in AI and quantum are shaping the field of cybersecurity** and outlines their broader implications for national and international security. It explores the ways in which AI is transforming both cyberoffence and cyberdefence, from AI-enhanced phishing, malware development and adversarial attacks on AI systems, to AI-driven anomaly detection and security automation in cybersecurity applications. It also **assesses the potential impact of quantum computing**, particularly in the context of encryption, data protection and the long-term security of digital communications. These issues have also recently gained prominence in ongoing diplomatic discussions on international ICT security[196] as states seek to manage cyber risks and ensure that technological progress does not undermine international peace and stability. The following subsections provide an overview of these evolving trends and their significance for international cybersecurity dialogues.

3.1. Cybersecurity and artificial intelligence

In the last few years, **AI fundamentally transformed both the nature of cyberthreats and the defensive tools** available to counter them. Three key trends stand out:

- **AI-powered cyberattacks,** in which malicious actors leverage AI to enhance the scale and sophistication of their operations
- **Adversarial attacks on AI systems,** which exploit vulnerabilities in machine learning models to manipulate their outputs
- **AI-powered cyberdefences,** which use AI-driven analytics and automation to detect and mitigate cyberthreats more effectively

This subsection explores each of these trends in detail, examining their implications for cybersecurity and international stability.

BOX 18.

How AI-powered cyberthreats work

Artificial intelligence enhances the technical sophistication of cyberattacks by automating tasks that were once manually intensive. For example, AI can generate highly convincing phishing emails by analysing data patterns and tailoring messages to specific individuals. AI now also plays a key role in executing many password-spraying attacks. By leveraging AI, attackers can analyse vast data sets of leaked credentials to predict password patterns, evade detection mechanisms and optimize attack timing to avoid triggering security alerts. AI-driven automation further allows malicious actors to develop new malware variants and to conduct cyberattacks on a much larger scale, increasing the chances of success while minimizing the risks of detection.[197]

3.1.1. AI-powered cyberattacks

First, there was a notable increase in **AI-powered cyberattacks** targeting governments, critical infrastructure operators, industries and individuals.[198] Specifically, attackers integrated AI into their operational practices to automate large-scale phishing campaigns, accelerate **password-spraying attacks** and enhance malware capabilities. For example, Microsoft reported that password spraying attacks – one of the most potent strategies for gaining unauthorized access to systems and delivering a ransomware payload – increased from around 3 billion attempts per month in 2023 to 30 billion per month in 2024, a trend driven largely by AI-enabled automation.[199]

BOX 19.

What is a password-spraying attack?

A password-spraying attack uses a trial-and-error approach in which an attacker attempts to access multiple accounts by trying a few commonly used passwords against many usernames, rather than trying many passwords against a single account. This method avoids account lockouts that occur after multiple incorrect login attempts, making it an effective way for attackers to exploit weak password policies.

There was also a widespread deployment of **AI-driven social engineering tactics last year,** which significantly amplified the effectiveness of cyberattacks.[200] AI-generated phishing emails in particular became more advanced, making them harder to detect and more convincing to victims.[201] While phishing is not a new method, until recently sophisticated attempts required time, knowledge of local context and language, and social skills. Recently, **malicious actors started relying on AI to generate highly tailored and personalized email phishing**

messages. A successful phishing attack can allow a malicious actor to infiltrate one trusted email account within an organization and use that account to send emails with harmful links and attachments to other users within the same organization. Since many phishing emails are now sent internally, multiple users can – and often do – fall victim to phishing attacks by clicking on links that come from legitimate and trusted sources. This leads to faster and more widespread proliferation of successful cyberattacks across the digital landscape.

BOX 20.

What is a phishing attack?

A phishing attack is a cyberattack in which a malicious actor sends fraudulent emails, messages or other communications to a large number of individuals or organizations in an attempt to dupe the recipient into reveailing sensitive information, such as login credentials or financial details. These campaigns often use deceptive social engineering tactics, including impersonating trusted entities, embedding malicious links in emails or prompting users to download harmful attachments.

Malicious actors also **used AI to develop advanced malware variants that can evade detection and adapt to security measures in real time**.[202] AI enables attackers to automate the creation of polymorphic malware, which continuously changes its code to bypass traditional antivirus defences.[203] Additionally, AI-driven obfuscation techniques, such as code encryption and packing, make malicious code harder to analyse and remove.

BOX 21.

How does obfuscation work

Obfuscation refers to common techniques used by attackers to bypass antivirus software and prevent or delay detection of malicious code by analysts. The most frequently employed obfuscation techniques include:

- **Code encryption,** which scrambles the malware's code to hide its true function from security tools, preventing easy detection
- **Packing,** which compresses malware into a different format so that it appears harmless until it is unpacked and activated on the target system

Recent research has demonstrated that AI-powered tools can generate thousands of malware variants – enough to overwhelm traditional security systems. **Some AI-generated malware can even learn from defensive responses and then modify its behaviour to remain undetected**.[204] As AI-assisted development of malware continues to evolve, it may present a new challenge for cybersecurity professionals, such as a need to adopt AI-driven defences to counter sophisticated threats.

3.1.2. Adversarial attacks on AI-systems

Second, **adversarial AI attacks**, where threat actors manipulate AI systems to generate errors, also pose a growing concern as AI systems continued to be deployed at ever-larger scale.[205] These attacks, which seek to exploit vulnerabilities in AI-driven decision-making processes, can have profound consequences, particularly in critical domains such as transportation and defence.[206] Furthermore, AI-driven cyberattacks have introduced new levels of stealth and adaptability, as perpetrators increasingly use generative AI to craft malware capable of evading detection.[207]

BOX 22.

What is an adversarial attack on an AI system?

An adversarial attack on an AI system is a technique used by malicious actors to trick artificial intelligence into making mistakes. AI models rely on patterns in data to make decisions, but an attacker can manipulate these patterns in subtle ways to confuse the system.[208] For example, it might slightly alter an image, a piece of text or even a voice recording in a way that looks normal to humans but causes the AI to misinterpret it. In cybersecurity, adversarial attacks can be used to bypass AI-powered security systems, evade malware detection or trick facial-recognition software. In the event of future widespread deployment of AI systems in high-stakes areas like finance, defence and healthcare, adversarial attacks could pose serious risks to national security and economic stability and could put individuals at risk. This makes it crucial to develop AI models that can resist such manipulations.

3.1.3. AI-powered cyberdefenses

Third, **AI also holds potential to improve cybersecurity defences**. AI-powered systems already enhance the skill, speed and knowledge of defenders by automating threat analysis, incident response and monitoring processes.[209] AI-driven security analytics can also help detect previously unknown threats by analysing massive data sets to identify suspicious activities.[210] For example, AI-based defences have proven effective in fending off cyberattacks during the conflict in Ukraine, where proactive identification and blocking of previously unknown malicious code reportedly played a pivotal role in preventing large-scale disruptions across several critical sectors.[211]

BOX 23.

How AI strengthens cyberdefences

AI is already being used to enhance cybersecurity by automating and accelerating defensive measures. The key ways in which AI may contribute to cyberdefence include:

- **Automated Threat Analysis:** AI can process vast amounts of data in real time, identifying patterns and detecting anomalies that may indicate cyberthreats.
- **Rapid Incident Response:** AI can automatically take action to block attacks, reducing damage before humans even notice the problem.
- **Advanced Threat Detection:** AI helps security teams discover cyberthreats that have not previously been seen by looking for strange patterns in data.[212]
- **Proactive Malware Defence:** AI systems can identify and block malicious code before it executes, even if it has never been seen before.
- **Boosting Cyberdefence in Conflict Zones:** AI-based defences have been deployed in the context of armed conflict to successfully counter cyberattacks, preventing large-scale disruptions across critical sectors.

These advancements suggest that AI, when combined with cloud-based analysis and threat intelligence from individual devices, could eventually tilt the balance in favour of some cyberdefenders. However, the d**eployment of AI in cybersecurity requires significant resources, energy and infrastructure, including vast data sets and computational power**.[213] This may limit accessibility of these defensive measures to a handful of well-resourced organizations and exacerbate existing inequalities in the cybersecurity landscape. This cybersecurity gap may raise concerns around equity in cybersecurity capabilities and could leave smaller organizations and some developing states ever more vulnerable to sophisticated cyberthreats over time.

On the other hand, too great a dependency on automated defences may leave organizations vulnerable in the event of an adversarial cyberattack that targets the AI-based security systems themselves. **Threat actors are reported to be developing techniques to manipulate AI decision-making models**, creating a significant risk for security applications that rely exclusively on machine learning algorithms.[214] Furthermore, the **opacity of AI algorithms may complicate the detection and mitigation of adversarial AI attacks**, leading some actors to call for international normative frameworks to prohibit cyberattacks targeting AI systems,[215] particular in high-stakes domains such as transportation and defence, where AI tampering could cost lives.[216]

3.1.4. AI–cyber nexus

To capture the potential uses of AI in the field of cybersecurity, UNIDIR developed an **ICT Intrusion Path framework that maps out the progress**ion of cyber intrusions based on their position relative to the targeted network's security boundary. The framework consists of three layers:

a. **Outside the perimeter:** The stage where attackers gather intelligence from publicly accessible sources, such as social media, open databases and Dark Web forums, and develop malicious tools and techniques to prepare for an intrusion
b. **On the perimeter:** The security boundary at which attackers attempt to gain unauthorized access by exploiting vulnerabilities, bypassing firewalls or leveraging deceptive tactics like phishing
c. **Inside the perimeter:** Where attackers have already breached the network and can move laterally, steal administrator-level privileges and manipulate critical systems

AI can be used by both attackers and defenders across all three layers of the ICT Intrusion Path (see Table 1). Understanding these layers can help cybersecurity professionals anticipate and respond to threats at different stages of a cyberattack.

TABLE 1.

The impact of AI on the ICT intrusion path[217]

LAYER	AI USE BY ATTACKERS	AI USE BY DEFENDERS
Outside the Perimeter	AI can automate reconnaissance by analysing vast amounts of public and Dark Web data, identifying potential vulnerabilities and crafting targeted phishing campaigns. AI can also assist in malware development, such as polymorphic malware, which evolves to evade detection.	AI-driven threat intelligence systems can monitor open-source intelligence (OSINT), Dark Web forums and network traffic to detect early-stage reconnaissance activities. AI can also enhance predictive analytics to pre-emptively identify potential attack vectors.
On the Perimeter	AI can generate highly personalized and deceptive spear phishing emails and deepfake content to deceive users. AI-enhanced malware can dynamically adjust its attack vectors to exploit detected vulnerabilities in firewalls and intrusion-detection systems.	AI can strengthen perimeter security through enhanced email filtering, anomaly detection and automated penetration testing. AI can also support adaptive authentication mechanisms that detect and prevent unauthorized access attempts in real time.
Inside the Perimeter	AI can facilitate lateral movement by automating credential theft, privilege escalation and evasion of security controls. AI-driven malware can modify itself in real time to avoid detection, while AI-powered exfiltration techniques can optimize data theft and system disruption.	By autonomously analysing network behaviour for suspicious activities, AI can rapidly detect and respond to intrusions. AI-powered endpoint-protection solutions can identify and mitigate insider threats, while AI-enhanced forensic analysis can accelerate post-incident investigations.

In the OEWG on ICT security, states have acknowledged that AI, like other emerging technologies, is inherently neutral but presents both opportunities and risks for international security. **States have agreed that AI's evolving capabilities could introduce new vulnerabilities in the ICT landscape**, enhance the speed and precision of malicious cyber activities, and increase the volume and impact of cyberattacks. They have **expressed particular concern regarding the safety and security of AI systems**, as well as the integrity of data used to train AI models in the context of ICT security. While AI has the potential to strengthen cybersecurity

by improving incident-response times and network resilience, states have cautioned that it could also be leveraged to amplify cyberthreats, escalate cascading effects and cause unintended harm. In response, **states underscored at the OEWG the importance of deepening understanding of AI-related risks**, implementing robust security measures across the AI life cycle, and ensuring that AI technologies are harnessed for peaceful purposes, in line with shared international security interests.[218]

BOX 24.

2025 meeting of the United Nations Security Council to discuss AI-based cyberthreats

In 2025, the misuse of artificial intelligence in cyber operations was prominently highlighted as a concern that could have negative impacts on the international community's efforts to maintain international peace and security. During a September 2025 meeting of the Security Council, the **United Nations Secretary-General** warned that AI-powered cyberthreats have could escalate tensions between states. The Secretary-General emphasized the urgent need for global cooperation to establish safeguards and prevent the weaponization of AI in cyberspace, highlighting that AI-enabled cyberattacks could rapidly disrupt critical infrastructure, that AI-driven manipulation of information poses risks of escalation and diplomatic crises, and that transparency, global regulatory frameworks, and strengthened national capacities are essential to deter AI-generated deception and uphold international peace and security.[219]

Looking ahead, **AI is set to play a central role in future cybersecurity**. While its potential to augment human capabilities and address cybersecurity workforce shortages through automation is promising, achieving widespread benefits will require investments in infrastructure, international collaboration and equitable access to AI-driven tools. Policymakers may also need to consider ethical concerns, biased AI systems and the risks associated with AI misuse and adversarial attacks on critical AI-systems to ensure that advancements in AI strengthen global cybersecurity resilience without exacerbating existing vulnerabilities and introducing new ones.

3.2. Cybersecurity and quantum computing

In 2024, the **United Nations designated 2025 as the International Year of Quantum**, recognizing the profound impact that quantum computing is poised to have on global security and technological innovation.[220] This decision is timely, as quantum technologies, grounded in the foundational principles of quantum mechanics, are poised to revolutionize an expansive range of sectors, with cybersecurity at the forefront.[221] By harnessing the principles of quantum mechanics, quantum computers have the potential to solve complex problems that are beyond the reach of classical computers, **unlocking new advancements in secure communications and data processing**.

BOX 25.

How quantum threats work

Quantum computing operates on principles of quantum mechanics that allow it to perform calculations at speeds unattainable by classical computers. This capability enables quantum systems to solve the mathematical problems – such as factoring large prime numbers – that underpin almost all modern cryptographic algorithms. Once operational, quantum computers could decrypt secure communications, leaving sensitive information and critical infrastructure systems vulnerable.[222]

However, while quantum-driven advancements in secure communications and threat detection could enhance digital security for some, the ability of quantum computers to break current encryption standards may present a significant challenge for many. **Widely used asymmetric encryption methods, which protect all types of today's communications – from national defence to financial transactions – could become obsolete within years rather than decades**. This disruption could erode trust in global cybersecurity frameworks, compromise sensitive diplomatic and intelligence communications, and expose critical infrastructure and financial systems to unprecedented vulnerabilities. Additionally, the potentially **uneven pace of quantum development among states raises concerns about technological asymmetry**, where states or non-state actors with early access to quantum capabilities may gain a strategic advantage, potentially fuelling new geopolitical tensions and exacerbating existing cybersecurity gaps on a global scale.

One of the most pressing concerns arising from quantum computing's rapid development is the **"harvest now, decrypt later"** strategy – allegedly already initiated by some threat actors.[223] While large-scale quantum computers are not yet widely available, this threat is not hypothetical. Malicious actors are believed to be actively intercepting and storing encrypted data today, with the intention of decrypting it once quantum technology matures. Industry experts estimate that quantum computers could break widely used cryptographic systems within the next **5–30 years**, with a **50–70 per cent chance** that this could happen as soon as the next 5 years.[224] Given the high stakes, governments and industries must act pre-emptively to develop and implement post-quantum cryptographic (PQC) solutions before current encryption methods become obsolete.

BOX 26.

What is a "Harvest Now, Decrypt Later" cyberattack?

A "Harvest Now, Decrypt Later" (HNDL) attack is a long-term cyber espionage strategy in which attackers intercept and store encrypted data today with the expectation that they will be able to decrypt it in the future using more advanced computational capabilities, particularly quantum computing.[225] State actors and cybercriminals engaging in HNDL attacks may target encrypted government communications, financial transactions and sensitive national defence data, aiming to gain future access to classified or commercially valuable information.

There is also concern about the implications of quantum computing for **economic stability and technology supply chains**. Some states are actively working to develop **sovereign quantum capabilities** from a fear that dependence on foreign quantum infrastructure could create security vulnerabilities and **increase reliance on external technology providers**.[226] Additionally, the **high cost and complexity of quantum research and development** may widen the technological divide between states, limiting equitable access to its benefits. Concerns are also emerging over the **security of quantum supply chains**, as disruptions in the production of key quantum components – such as specialized hardware and rare materials – could create new vulnerabilities. In response, UNIDIR has initiated dialogues[227] that focus on the implications of quantum technologies for global security and **inclusive and responsible quantum development** to ensure that technological advancements do not exacerbate geopolitical divides or introduce new security risks.[228]

In anticipation of quantum breakthroughs that loom large on the horizon, the **cybersecurity community is pivoting towards post-quantum cryptography**: development of encryption algorithms that are resistant to quantum attacks. Governments and organizations worldwide are currently exploring and investing in quantum-safe technologies. For instance, the European Union has integrated quantum technologies into its cyber policy goals through initiatives such as the European Quantum Communication Infrastructure (EuroQCI) and the Quantum Flagship programme.[229] Similarly, some major technology firms are accelerating their transition to PQC, setting the stage for a migration towards quantum-resilient systems.[230]

Apart from new risks, **quantum computing also introduces opportunities to enhance cybersecurity**. Quantum sensors and communication technologies may soon offer heightened security for information transfer and the ability to detect cyberthreats in real time.[231] For example, **quantum key distribution (QKD)** provides a resilient method of secure communication that already being deployed in critical sectors by states and private entities.[232] With **quantum communication** promising ultra-secure information transfer, several states are already migrating sensitive military communications to **quantum-secured networks** to guard against cyber espionage.[233]

BOX 27.

What is quantum key distribution?

Quantum key distribution (QKD) is a cutting-edge method for securing communications using quantum mechanics. Unlike traditional encryption, QKD relies on particles of light (photons) to generate a secret encryption key. If a hacker tries to intercept the key, quantum laws ensure that any eavesdropping disrupts the system and alerts the sender and the receiver. This makes QKD highly secure and resistant to cyberthreats. As quantum computers advance, QKD is emerging as a key solution for protecting financial transactions, government communications and critical infrastructure from future cyberthreats.[234]

However, as states race to achieve quantum breakthroughs, the **risk of technological inequality and geopolitical tensions grows**, potentially leading to **new security dilemmas**. Limited access to quantum technology and infrastructure may concentrate significant security advantages in the hands of a few, leaving many smaller and under-resourced states and organizations behind.[235] This disparity could have far-reaching implications for trust and stability in the international security landscape with some experts warning that while quantum technology may create absolute security for a few, it could leave others facing "absolute insecurity".[236] As quantum technologies continue to evolve, their transformative potential should be balanced with proactive measures to mitigate risks. Investments in transition to PQC, QKD and collaborative international frameworks are essential to harness these technologies for the greater good while minimizing their potential for human harm and disruptions to international stability.[237]

BOX 28.

Example of the impact of quantum threats on international peace and security

A hypothetical yet plausible scenario involves a malicious actor capturing encrypted diplomatic communications through with the intent to decrypt them in the future using quantum computing (HNDL). If successful, this could expose sensitive negotiations, disrupt diplomatic relations and escalate geopolitical tensions. For instance, confidential communications intercepted during critical peace negotiations could be exploited to undermine trust and derail international efforts to resolve conflicts.

At the **OEWG on ICT security**, states have recognized that quantum computing, like other emerging technologies, offers significant development opportunities but also introduces potential security risks. **States have acknowledged that quantum advancements could create new vulnerabilities in ICT security**, particularly by increasing the speed and effectiveness of malicious cyber activities and compromising traditional encryption methods. **They have also emphasized the need to better understand quantum-related risks and to**

develop security measures throughout a technology's life cycle in order to mitigate the potential impact of quantum technologies on ICT security. The international community has further agreed that it is in the interest of all states to **promote the peaceful use of quantum technologies**, ensuring that their benefits are harnessed while minimizing the security risks they may pose to international stability.[238]

Beyond considering advancements in quantum and AI separately, the potential future **convergence of AI and quantum computing technologies could also transform the field of cybersecurity** by enabling the development of artificial general intelligence (AGI).[239] Such AGI systems may take autonomous decision-making in both defensive and offensive cyber operations. For example, AGI could process vast data sets and execute complex algorithms at unprecedented speeds, potentially allowing **AGI to identify vulnerabilities and deploy defensive measures or counterattacks without human intervention**. Such a shift may raise concerns about maintaining meaningful human control over such systems, as their ability to operate independently could lead to unpredictable and potentially uncontrollable outcomes.

To ensure that AGI remains aligned with human values and strategic oversight, states may consider the establishment of robust ethical frameworks and control mechanisms to prevent unintended escalations in cyberspace. As these two disruptive technologies evolve in tandem, states may also consider how to **ensure that discussions on cybersecurity, AI and quantum do not evolve in isolation**. This would allow states to assess their potential combined impacts to ensure that safeguards and governance mechanisms evolve in step with technological advancements to prevent unintended escalatory risks in cyberspace.

Working Group on Security in Use of Information and Communications Technologies, February 2025, New York. Credit: UN Photo / Manuel Elías.

Conclusions

The contemporary cybersecurity landscape presents both urgent challenges and opportunities for advancing international peace and security. As malicious cyber activities grow in sophistication and scale, they may increasingly threaten critical infrastructure, destabilize economies and strain international relations. These threats can have far-reaching implications for international stability, with cascading impacts on public safety, governance and global trust. Addressing these multifaceted challenges requires a united and proactive approach, grounded in international cooperation, technological innovation and a commitment to collective security.

A key opportunity lies in leveraging the role of the United Nations to strengthen global governance in cyberspace. In 2025, Member States decided to establish the *Global Mechanism on developments in the field of ICTs in the context of international security and advancing responsible State behaviour in the use of ICTs*, with its organizational session to be convened in early 2026. As the Mechanism is operationalized, Member States may wish to consider how its working methods and the thematic focus of its dedicated working groups can be aligned with the evolving realities of the cyberthreat landscape. By reflecting current risks and challenges, the mechanism can more effectively coordinate responses to emerging threats, support the advancement of the United Nations Framework for Responsible State Behaviour in Cyberspace and build capacity. Looking ahead, ensuring that the Global Mechanism remains attuned to practical threat dynamics will also be essential to fostering trust, advancing responsible state conduct and contributing meaningfully to international cyber stability.

The potential future impacts of cyberthreats on international peace and security cannot be overstated. In 2025, disruptions to critical infrastructure, financial systems and democratic processes already posed significant risks to global stability. Misattribution or escalation resulting from cyber incidents could deepen geopolitical tensions, while the exploitation of emerging technologies, such as AI and quantum computing, amplifies the complexity of these challenges. Without coordinated action, the international community risks losing the foundations of trust, cooperation and security that underpin peace in the digital age.

In taking forward the establishment of the Global Mechanism and strengthening international cooperation, Member States have an opportunity to address the multifaceted nature of contemporary cybersecurity challenges and contribute to maintaining an open, peaceful, stable and accessible cyberspace for all.

Going forward, sustained investment in education, capacity-building and emerging technologies will be critical to ensuring equitable access to cybersecurity resources and safeguarding the digital domain. Advancing respect for existing international norms, alongside fostering innovative public–private partnerships, can further bolster collective cyber resilience. In advancing the establishment of a permanent United Nations mechanism and reinforcing global cooperation, states have a unique opportunity to address the complexities of cybersecurity at a global level whilst ensuring that cyberspace remains open, peaceful, stable and accessible for all.

Endnotes

1 UNIDIR Security and Technology Programme, "2024 Cyber Stability Conference: Unpacking Cyber Threats to International Peace and Security", 29 February–1 March 2024, https://unidir.org/event/2024-cyber-stability-conference-unpacking-cyber-threats-to-international-peace-and-security.

2 UNIDIR, "2024 Cyber Stability Conference: Unpacking Cyber Threats to International Peace and Security", X, 29 February 2024, https://x.com/UNIDIR/status/1763694205140865253.

3 Celine Rosak, "2024 in Review: Cyber Threats and the Fight to Secure Critical Infrastructure", Xpage Security, 18 December 2024, https://xage.com/blog/cyber-attack-news-2024-attacks-on-critical-infrastructure.

4 KnowBe4, "Cyber Attacks on Infrastructure: The New Geopolitical Weapon", December 2024, https://www.knowbe4.com/hubfs/Global-Infrastructure-Report-2024_EN_US.pdf?hsLang=en-us.

5 Open-ended working group on Information and Communication Technology (ICT), 2nd meeting, 20th substantive session, 17–21 February 2025, https://webtv.un.org/en/asset/k1n/k1nuiidprr.

6 US Cybersecurity and Infrastructure Security Agency, "PRC State-Sponsored Actors Compromise and Maintain Persistent Access to U.S. Critical Infrastructure", 7 February 2024, https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-038a.

7 Juliet Skingsley, "Cyber-Rattling: Can 'Pre-Positioning' in Cyberspace Amount to a Threat of the Use of Force under Article 2(4) of the United Nations Charter?", *Journal on the Use of Force and International Law*, vol. 11, no. 1–2 (18 October 2024), https://www.tandfonline.com/doi/full/10.1080/20531702.2024.2413791.

8 Celine Rosak, "2024 in Review".

9 Gordon James, "11 Recent Cyber-Attacks on the Water and Wastewater Sector", Wisdaim, 13 October 2024, https://wisdiam.com/publications/recent-cyber-attacks-water-wastewater/.

10 UNIDIR Security and Technology Programme, "2024 Cyber Stability Conference".

11 Gaglina Antova, "Why Ransomware on Hospitals Is One of the Greatest Dangers of Our Time", *Forbes*, 18 April 2022, https://www.forbes.com/councils/forbestechcouncil/2022/04/14/why-ransomware-on-hospitals-is-one-of-the-greatest-dangers-of-our-time/.

12 Kristian McCann, "How Hackers Are Hitting Healthcare via Their Supply Chain", *Cyber Magazine*, 15 October 2024, https://cybermagazine.com/articles/cyber-attacks-threaten-healthcare-supply-chains.

13 US Environmental Protection Agency, "EPA Cybersecurity for the Water Sector", December 2024, https://www.epa.gov/waterresilience/epa-cybersecurity-water-sector.

14 Jonathan Reed, "Cyberattack on American Water: A Warning to Critical Infrastructure", IBM, 3 November 2024, https://securityintelligence.com/news/cyberattack-on-american-water-warning-critical-infrastructure.

15 Thales Report, "2024 Cloud Security Study: Boom Times for the Cloud, Is Security Ready?", Thales Group, 25 June 2024, https://www.thalesgroup.com/en/worldwide/defence-and-security/press_release/cloud-resources-have-become-biggest-targets.

16 Federico Mantellassi and Giacomo Persi Paoli, *Cloud Computing and International Security: Risks, Opportunities and Governance Challenges* (Geneva: UNIDIR, 2024), https://unidir.org/publication/cloud-computing-and-international-security-risks-opportunities-and-governance-challenges/.

17 Triskele Labs, "Cloud Cyber Attacks: The Latest Cloud Computing Security Issues", Triskele Labs Blog, 2024, https://www.triskelelabs.com/blog/cloud-cyber-attacks-the-latest-cloud-computing-security-issues.

18 ArcServe, "7 Most Infamous Cloud Security Breaches", ArcServe Blog, 20 December 2023, https://www.arcserve.com/blog/7-most-infamous-cloud-security-breaches.

19 Walter Peters, "Cyberattacks on Satellites: An Underestimated Political Threat", London School of Economics and Political Science, 2024, https://www.lse.ac.uk/ideas/projects/space-policy/publications/Cyberattacks-on-Satellites.

20 Shaun Waterman, "Experts Outline the Growing Threat of Cyber Attacks Against Space Systems", Via Satellite, 18 November 2024, https://www.satellitetoday.com/cybersecurity/2024/11/18/a-frightening-future-experts-outline-the-growing-threat-of-cyber-attacks-against-space-systems/.

21 R. Roberts et al., "Stellar Safeguards: How organizations can protect space assets from cyberthreats", Deloitte Center for Government Insights, 29 August 2024, https://www2.deloitte.com/us/en/insights/industry/public-sector/defending-against-cyber-threats-space-systems.html.

22 Viasat Inc., "KA-SAT Network cyber attack overview", 30 March 2022, https://news.viasat.com/blog/corporate/ka-sat-network-cyber-attack-overview.

23 CyberPeace Institute Cyber Conflict Tracker, "Case Study: Viasat", CyberPeace Institute, 2024, https://cyberconflicts.cyberpeaceinstitute.org/law-and-policy/cases/viasat.

24 Nicola Smith, "Australia on Red Alert for Surge in High-Impact Sabotage Warns Top Spy", The Nightly, 20 February 2025, https://thenightly.com.au/politics/australia/australia-on-red-alert-for-surge-in-high-impact-sabotage-warns-top-spy-c-17786702.

25 Dina Temple-Raston, "Neuberger: Defining Espionage vs. Pre-Positioning for Attacks is Key to Battling State Actors", Recorded Future News, 15 February 2024, https://therecord.media/volt-typhoon-china-defining-espionage-pre-positioning-neuberger-munich.

26 Morgan Demboski and Brent Eskridge, "Cyber Attacks on the Power Grid", IronNet Threat Research, 19 May 2022, https://www.ironnet.com/blog/cyber-attacks-on-the-power-grid.

27 Surbhi Misra and David Shepardson, "AT&T, Verizon Targeted by Salt Typhoon Cyberespionage Operation", Reuters, 29 December 2024, https://www.reuters.com/technology/cybersecurity/chinese-salt-typhoon-cyberespionage-targets-att-networks-secure-carrier-says-2024-12-29/.

28 Maxime A. and Livia Tibirna, "The Transportation Sector Cyber Threat Overview", Sekoia Threat Research and Intelligence, 12 September 2023, https://blog.sekoia.io/the-transportation-sector-cyber-threat-overview/.

29 Skingsley, "Cyber-Rattling".

30 United Nations, General Assembly, Open-ended Working Group on security of and in the use of ICTs, Third Annual Progress Report, A/79/214, July 2024, paragraphs 14–17, https://docs.un.org/en/A/79/214.

31 United Nations, A/79/214, paragraphs 14–16.

32 United Nations Office for Disarmament Affairs, *New Agenda for Peace* (New York: United Nations, July 2023), https://www.un.org/sites/un2.un.org/files/our-common-agenda-policy-brief-new-agenda-for-peace-en.pdf.

33 United Nations International Computing Center, "Cybersecurity Threat Landscape Report 2023", May 2024, https://www.unicc.org/wp-content/uploads/2024/11/2023-Cyber-Threat-Landscape-Report-v2.pdf.

34 United Nations Development Programme, "UNDP Investigates Cyber-Security Incident", UNDP News Centre, 16 April 2024, https://www.undp.org/speeches/undp-investigates-cyber-security-incident.

35 Tilman Rodenhauser, Balthasar Staehelin and Massimo Marelli, "Safeguarding Humanitarian Organizations from Digital Threats", Humanitarian Law and Policy, 13 October 2022, https://blogs.icrc.org/law-and-policy/2022/10/13/safeguarding-humanitarian-organizations-from-digital-threats/.

36 International Committee of the Red Cross (ICRC), "ICRC Cyber-Attack: Sharing Our Analysis", 16 February 2022, https://www.icrc.org/en/document/icrc-cyber-attack-analysis.

37 CyberPeace Institute, "Humanity Under Attack", Humanitarian Cybersecurity Center, 2024, https://cyberpeaceinstitute.org/humanitarian-cybersecurity-center.

38 European Commission, "Critical Infrastructure Resilience at EU-level", EU Migration and Home Affairs, 11 September 2025, https://home-affairs.ec.europa.eu/policies/internal-security/counter-terrorism-and-radicalisation/protection/critical-infrastructure-resilience-eu-level_en.

39 Public Safety Canada, "Guidance on Essential Services and Functions in Canada During the COVID-19 Pandemic", 2024, September 2020, https://www.publicsafety.gc.ca/cnt/ntnl-scrt/crtcl-nfrstrctr/esf-sfe-en.aspx.

40 US Federal Emergency Management Agency, "Coronavirus (COVID-19) Pandemic Guidance for Private Nonprofit Organizations", US Department of Homeland Security, 2 April 2020, https://www.fema.gov/fact-sheet/coronavirus-covid-19-pandemic-private-nonprofit-organizations; Hana Driss, "The Effects of COVID-19 Pandemic on Humanitarian Operations in Jordan", SIT Digital Collections, December 2020, https://digitalcollections.sit.edu/cgi/viewcontent.cgi?article=4270&context=capstones.

41 International Committee of the Red Cross (ICRC), "Sophisticated Cyber-Attack Targets Red Cross Red Crescent Data on 500,000 People", News Release, 19 January 2022, https://www.icrc.org/en/document/sophisticated-cyber-attack-targets-red-cross-red-crescent-data-500000-people.

42 United Nations, General Assembly, Open-ended Working Group on security of and in the use of ICTs, Final Report, A/80/257, July 2025, paragraph 21, https://docs.un.org/en/A/80/257.

43 United Nations, Security Council, resolution 2730 (2024), 24 May 2024, https://digitallibrary.un.org/record/4049572?ln=en&v=pdf.

44 International Conference of the Red Cross and Red Crescent, "Resolution 34IC/24/R2 on Protecting civilians and other protected persons and objects against the potential human cost of ICT activities during armed conflict", 28–31 October 2024, https://rcrcconference.org/app/uploads/2024/11/34IC_R2-ICT-EN.pdf.

45 Google, "Cybercrime: A Multifaceted National Security Threat", Annual Google Threat Intelligence Report, February 2025, https://services.google.com/fh/files/misc/cybercrime-multifaceted-national-security-threat.pdf.

46 Prithwish Ganguli, "The Rise of Cybercrime-as-a-Service: Implications and Countermeasures", SSRN Research Paper, 15 September 2024, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4959188.

47 Interpol, "African Cyberthreat Assessment Report 2024", Interpol African Cybercrime Operations Desk, April 2024, https://www.interpol.int/content/download/21048/file/AJFOC_Africa Cyberthreat Assessment Report_2024.pdf.

48 David Kasabji, "Ransomware-as-a-Service: An Infamously Lucrative Business Model", Conscia Threat Intelligence Blog, 18 May 2022, https://conscia.com/blog/ransomware-as-a-service-an-infamously-lucrative-business-model/.

49 Cloudfare, "What is ransomware-as-a-service (RaaS)?", Cloudfare Security Blog, March 2025, https://www.cloudflare.com/en-gb/learning/security/ransomware/ransomware-as-a-service.

50 See Kumar Ritesh, "Who's Buying and Selling Ransomware Kits on the Dark Web", *CyberCrime Magazine*, 13 March 2021, https://cybersecurityventures.com/whos-buying-and-selling-ransomware-kits-on-the-dark-web/.

51 Iman Vakilinia and Shamik Sengupta, "Vulnerability Market as a Public-Good Auction with Privacy Preservation", *Computers and Security*, vol. 93, no. 1 (June 2020), https://www.sciencedirect.com/science/article/abs/pii/S0167404820300924.

52 Rand Corporation, "Black Markets for Hackers Are Increasingly Sophisticated, Specialized and Maturing", *News Release*, 25 March 2024, https://www.rand.org/news/press/2014/03/25.html.

53 See for example the estimates on cybercrime-related financial losses by the World Economic Forum, https://www.weforum.org/stories/2023/01/global-rules-crack-down-cybercrime/; Cybersecurity Ventures, https://cybersecurityventures.com/hackerpocalypse-cybercrime-report-2016/; and Statista, https://www.statista.com/forecasts/1280009/cost-cybercrime-worldwide.

54 See the latest economic data in International Monetary Fund, "World Economic Outlook Database", https://www.imf.org/en/Publications/WEO/weo-database/2024/October/weo-report.

55 Microsoft Corporation, "Microsoft Digital Defense Report 2024", Microsoft Threat Intelligence Center, November 2024, https://www.microsoft.com/en-us/security/security-insider/intelligence-reports/microsoft-digital-defense-report-2024.

56 Sangfor Technologies, "Ransomware Attacks 2024: A Look Back at the Top Ransomware Headlines", Sangfor Security Blog, 19 January 2025, https://www.sangfor.com/blog/cybersecurity/ransomware-attacks-2024-top-ransomware-headlines.

57 Mariusz Michalowski, "50+ Ransomware Statistics for 2025", Spacelift Security Blog, 28 July 2025, https://spacelift.io/blog/ransomware-statistics.

58 SentinelOne, "What is Double-Extortion Ransomware", Cybersecurity Threat Intelligence Blog, August 2025, https://www.sentinelone.com/cybersecurity-101/threat-intelligence/what-is-double-extortion.

59 Adi Bleih, "Ransomware Annual Report 2024", Cyberint, 13 January 2025, https://cyberint.com/blog/research/ransomware-annual-report-2024/.

60 Antova, "Why Ransomware on Hospitals Is One of the Greatest Dangers of Our Time".

61 Chainalysis, "The 2024 Crypto Crime Report: The Latest Trends on Ransomware, Scams, Hacking, and More", Chainalysis Threat Intelligence, February 2025, https://go.chainalysis.com/crypto-crime-2024.html.

62 Chainalysis, "2025 Crypto Crime Mid-year Update: Stolen Funds Surge", Chainalysis Crime Blog, 17 July 2025, https://www.chainalysis.com/blog/2025-crypto-crime-mid-year-update.

63 Jonathan Greig, "Ransomware Gang Threatens to 'Overthrow' New Costa Rica Government, Raises Demand to $20 million", Recorded Futures News, 16 May 2022, https://therecord.media/ransomware-gang-threatens-to-overthrow-new-costa-rica-government-raises-demand-to-20-million.

64 Interpol, "Financial Fraud assessment: A Global Threat Boosted by Technology", Interpol Financial Fraud Unit, 11 March 2024, https://www.interpol.int/en/News-and-Events/News/2024/INTERPOL-Financial-Fraud-assessment-A-global-threat-boosted-by-technology.

65 The United Nations Office on Drugs and Crime has highlighted the growth of "pig butchering" scams in South East Asia, where criminals use generative AI and deepfakes to defraud victims, leading to an estimated $75 billion in losses. See United Nations Office on Drugs and Crime (UNODC), *Transnational Organized Crime and the Convergence of Cyber-Enabled Fraud, Underground Banking and Technological Innovation in Southeast Asia: A Shifting Threat Landscape* (Vienna: UNODC, October 2024), https://www.unodc.org/roseap/uploads/documents/Publications/2024/TOC_Convergence_Report_2024.pdf.

66 Spenser Feingold and Johnny Wood, "'Pig-Butchering' Scams on the Rise as Technology Amplifies Financial Fraud", World Economic Forum Center for Cybersecurity, 10 April 2024, https://www.weforum.org/stories/2024/04/interpol-financial-fraud-scams-cybercrime.

67 Younghoo Lee and Ben Gelman, "The Dark Side of AI: Large-Scale Scam Campaigns Made Possible by Generative AI", Sophos Threat Intelligence Update, 27 November 2023, https://news.sophos.com/en-us/2023/11/27/the-dark-side-of-ai-large-scale-scam-campaigns-made-possible-by-generative-ai.

68 Julia Dickson and Lauren Burke Preputnik, "Cyber Scamming Goes Global: Unveiling Southeast Asia's High-Tech Fraud Factories", Centre for Strategic and International Studies, 12 December 2024, https://www.csis.org/analysis/cyber-scamming-goes-global-unveiling-southeast-asias-high-tech-fraud-factories.

69 Office of the United Nations High Commissioner for Human Rights, "Hundreds of Thousands Trafficked to Work as Online Scammers in SE Asia, says UN Report", Press Release, 29 August 2023, https://www.ohchr.org/en/press-releases/2023/08/hundreds-thousands-trafficked-work-online-scammers-se-asia-says-un-report.

70 Juan H, "Dark Web: Lifecycle of Stolen Credentials Explored", PreyProject Threat Detection Blog, 26 February 2024, https://preyproject.com/blog/lifecycle-stolen-credentials-dark-web.

71 Leila Goldstein, "Thousands Rescued from Illegal Scam Compounds in Myanmar as Thailand Launches Huge Crackdown", *The Guardian*, 19 February 2025, https://www.theguardian.com/world/2025/feb/19/myanmar-scam-call-centre-compound-rescues-thailand-crackdown.

72 Rebecca Ratcliffe, "Tens of Thousands Could be Held in Illegal Scam Compounds in Myanmar, Thai Police General Says", *The Guardian*, 21 February 2025, https://www.theguardian.com/world/2025/feb/21/myanmar-scam-call-centre-compounds.

73 United Nations Office of Counter-Terrorism (UNOCT) and United Nations Interregional Crime and Justice Research Institute (UNICRI), "Beneath the Surface: Terrorist and Violent Extremist Use of the Dark Web and Cybercrime as a Service for Cyber-Attacks",28 June 2024, https://digitallibrary.un.org/record/4063352?ln=en&v=pdf.

74 Group of Seven (G7), "Ransomware Annex Statement", Adopted by G7 Member States as a part of Broader Statement on Digital Payments, 13 October 2020, https://home.treasury.gov/system/files/136/G7-Ransomware-Annex-10132020_Final.pdf.

75 United Nations, A/80/257, paragraph 24.

76 See the UNODC website for more details on the UN Convention Against Cybercrime and the Convention Signing Ceremony in Hanoi in October 2025, https://www.unodc.org/unodc/en/cybercrime/convention/home.html.

77 United Nations, General Assembly, resolution 79/243, "Countering the Use of Information and Communication Technologies for Criminal and Terrorist Purposes", 31 December 2024, https://docs.un.org/A/RES/79/243.

78 Sonatype, "State of the Supply Chain Security: 10th Anniversary Report", Sonatype Threat Intelligence Report Series, October 2024, https://www.sonatype.com/state-of-the-software-supply-chain/introduction.

79 Jonathan Munshaw, "Are Hardware Supply Chain Attacks 'Cyber Attacks?'", Cisco Talos Threat Source Newsletter, 26 September 2024, https://blog.talosintelligence.com/threat-source-newsletter-sept-26-2024.

80 VNG Cloud, "10 Types of Cloud Computing Attacks", VNG Cloud Tech Blog, 3 April 2023, https://vngcloud.vn/blog/muoi-hinh-thuc-tan-cong-tren-dien-toan-dam-may.

81 Zac Amos,"Why Software Update Can Lead to Cyberattacks", Hackernoon Blog, 27 August 2024, https://hackernoon.com/why-software-updates-can-lead-to-cyberattacks-and-what-to-do.

82 Fortra's Alert Logic Staff, "Why Software Updates Are Critical for Cybersecurity", Alert Logic Security Update, 26 July 2021, https://www.alertlogic.com/blog/why-software-updates-are-critical-for-cybersecurity/.

83 Microsoft Corporation, "Microsoft Digital Defense Report 2024".

84 Oscar Collins, "Pirated Software Presents New Cybersecurity Risks for Small Business Owners", *Cybersecurity Magazine*, February 2024, https://www.uscybersecurity.net/pirated-software-presents-new-cybersecurity-risks-for-small-business-owners.

85 Eva Dou and Gerrit De Vynck, "Pagers Attack Brings to Life Long-Feared Supply Chain Threat", *Washington Post*, 19 September 2024, https://www.washingtonpost.com/technology/2024/09/19/hezbollah-pager-attack-supply-chain/.

86 Ari Hawkins and Joseph Gedeon, "Middle East Pager Attacks Ignite Fear of Supply Chain Warfare", Politico, 19 September 2024, https://www.politico.com/news/2024/09/19/pager-attacks-supply-chain-warfare-00180136.

87 David E. Sanger, "A New Era in Sabotage: Turning Ordinary Devices Into Grenades on a Mass Scale", *New York Times*, 19 September 2024, https://www.nytimes.com/2024/09/19/us/politics/israel-hezbollah-pager-attacks.html.

88 European Agency for Cybersecurity (ENISA), "Understanding the Increase in Supply Chain Security Attacks", Press Release, 29 July 2021, https://www.enisa.europa.eu/news/enisa-news/understanding-the-increase-in-supply-chain-security-attacks.

89 Sonatype, "Scale: State of the Supply Chain Security: 10th Anniversary Report", Sonatype Threat Intelligence Report Series, October 2024, https://www.sonatype.com/state-of-the-software-supply-chain/2024/scale.

90 Thales Group Report indicates that cloud resources have become the primary targets for cyberattacks, with SaaS applications (31%), cloud storage (30%) and cloud management infrastructure (26%) being the most affected categories. Thales, "Cloud Resources Have Become Biggest Targets for Cyberattacks", 25 June 2024, https://cpl.thalesgroup.com/about-us/newsroom/cloud-resources-biggest-cyberattack-targets.

91 Alexander Liskin et al., "Story of the Year: Global IT Outages and Supply Chain Attacks", Kaspersky Security Bulletin, 9 December 2024, https://securelist.com/ksb-story-of-the-year-2024/114883.

92 James Coker, "Threat Actor Breaches Snowflake Customers", *Infosecurity Magazine*, 11 June 2024, https://www.infosecurity-magazine.com/news/threat-actor-breaches-snowflake.

93 See for example statements made by the Governments of the United States, https://therecord.media/volt-typhoon-china-defining-espionage-pre-positioning-neuberger-munich; and China, https://thehackernews.com/2024/10/china-accuses-us-of-fabricating-volt.html.

94 Beth Maundrill, "Supply Chain Attacks Top Cyber Threat for 2030", InfoSecurity Europe Blog, 17 May 2024, https://www.infosecurityeurope.com/en-gb/blog/threat-vectors/supply-chain-attacks-cyber-threat.html.

95 Liz Young and Heather Haddon, "Retailers Hit by Ransomware Attack on Tech Provider", Wall Street Journal, 25 November 2024, https://www.wsj.com/articles/starbucks-other-retailers-hit-by-ransomware-attack-on-tech-provider-98314528.

96 McCann, "How Hackers Are Hitting Healthcare via Their Supply Chain".

97 KnowBe4, "Cyber Attacks on Infrastructure".

98 James Rundle, "Fintech Company Finastra, Used by the Largest Banks, Discloses Hack", *Wall Street Journal*, 21 November 2024, https://www.wsj.com/articles/fintech-company-finastra-used-by-the-largest-banks-discloses-hack-ef5a575d.

99 Lamont Atkins et al., "Derisking Emerging Technologies in Financial Services", McKinsey and Company Risk Assessment Report, 11 March 2024, https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/the-cyber-clock-is-ticking-derisking-emerging-technologies-in-financial-services.

100 Security Scorecard, "97% of Leading U.S. Banks Impacted by Third-Party Data Breachers in 2024", Security Scorecard Threat Intel Report, 12 December 2024, https://securityscorecard.com/company/press/security-scorecard-threat-intel-report-97-of-leading-u-s-banks-impacted-by-third-party-data-breaches-in-2024.

101 Graham Fraser, "CrowdStrike: What was the Impact of the Global IT Outage", BBC News, 24 September 2024, https://www.bbc.com/news/articles/cr54m92ermgo.

102 US Department of Defense (DoD), "Defense Industrial Base Cybersecurity Strategy 2024", DoD Office of Publication and Security Review, 21 March 2024, https://dodcio.defense.gov/Portals/0/Documents/Library/DIB-CS-Strategy.pdf.

103 Center for Strategic and International Studies (CSIS), "Significant Cyber Incidents Since 2006", CSIS Strategic Technologies Programme Report, June 2025, https://www.csis.org/programs/strategic-technologies-program/significant-cyber-incidents.

104 KnowBe4, "Cyber Attacks on Infrastructure".

105 Heather Wishart-Smith, "The Semiconductor Crisis: Addressing Chip Shortages and Security", *Forbes*, 19 July 2024, https://www.forbes.com/sites/heatherwishartsmith/2024/07/19/the-semiconductor-crisis-addressing-chip-shortages-and-security.

106 United Nations, A/80/257, paragraph 23.

107 United Nations, A/80/257, paragraph 34.

108 United Nations, A/79/214, Annex A.

109 United Nations, A/79/214, Annex A.

110 UNIDIR Security and Technology Programme, "2024 Cyber Stability Conference".

111 OpenAI, "Influence and Cyber Operations: An Update", OpenAI Threat Intelligence Report, October 2024, https://cdn.openai.com/threat-intelligence-reports/influence-and-cyber-operations-an-update_October-2024.pdf.

112 Sam Stockwell, "AI-Enabled Influence Operations: Threat Analysis of the 2024 UK and European Elections", Centre for Emerging Technology and Security Briefing Paper, 19 September 2024, https://cetas.turing.ac.uk/publications/ai-enabled-influence-operations-threat-analysis-2024-uk-and-european-elections.

113 Randolf Carr and Paula Kohler, "AI-pocalypse Now? Disinformation, AI, and the Super Election Year", Munich Security Conference Analysis, 4 October 2024, https://securityconference.org/en/publications/analyses/ai-pocalypse-disinformation-super-election-year.

114 Tiffany Hsu, Stuart Thompson and Steven Lee Myers, "Elections and Disinformation Are Colliding Like Never Before in 2024", *New York Times*, 9 January 2024, https://www.nytimes.com/2024/01/09/business/media/election-disinformation-2024.html.

115 Ali Swenson and Kelvin Chan, "Election Disinformation Takes a Big Leap with AI Being Used to Deceive Worldwide", Associated Press, 14 March 2024, https://apnews.com/article/artificial-intelligence-elections-disinformation-chatgpt-bc283e7426402f0b4baa7df280a4c3fd.

116 Stockwell, "AI-Enabled Influence Operations".

117 A study analysing 27 viral AI-enabled disinformation cases in elections (16 in the United Kingdom and 11 in the European Union) found that, while these cases did not alter election outcomes, they raised concerns about deepfakes inciting hate and causing confusion over content authenticity. See Stockwell, "AI-Enabled Influence Operations".

118 OpenAI, "Influence and Cyber Operations".

119 Sean Cordey, "Cyber Influence Operations: An Overview and Comparative Analysis", Cyber Defense Project (CDP), Center for Security Studies (CSS), ETH Zurich, October 2019, https://css.ethz.ch/content/dam/ethz/special-interest/gess/cis/center-for-securities-studies/pdfs/Cyber-Reports-2019-10-CyberInfluence.pdf.

120 US Attorney's Office, District of Columbia, "Three Cyber Actors Indicted for 'Hack-and-Leak' Operation Designed to Influence the 2024 U.S. Presidential Election", Press Release, 27 September 2024, https://www.justice.gov/usao-dc/pr/three-irgc-cyber-actors-indicted-hack-and-leak-operation-designed-influence-2024-us.

121 Reza Rafati, "Hack and Leak Crime", Threat Intelligence Lab, 18 September 2024, https://threatintelligencelab.com/blog/hack-and-leak-crime.

122 Charles Owen-Jakcson, "Social Engineering in the Era of Generative AI: Predictions for 2024", IBM Security Intelligence Blog, 9 May 2024, https://securityintelligence.com/articles/social-engineering-generative-ai-2024-predictions.

123 Jack Goodman and Mohanad Hashim, "AI: Voice Cloning Tech Emerges in Sudan Civil War", BBC News, 5 October 2023, https://www.bbc.com/news/world-africa-66987869.

124 Bobby Allyn, "Deepfake Video of Zelenskyy Could Be 'Tip of the Iceberg' in Info War, Experts Warn", NPR News, 16 March 2022, https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia.

125 Sarah Cahlan, "How Misinformation Helped Spark an Attempted Coup in Gabon", *Washington Post*, 13 February 2020, https://www.washingtonpost.com/politics/2020/02/13/how-sick-president-suspect-video-helped-sparked-an-attempted-coup-gabon/.

126 Josh A. Goldstein et al., "How Persuasive is AI-Generated Propaganda?", *PNAS Nexus Journal*, vol. 3, no. 2 (February 2024), https://academic.oup.com/pnasnexus/article/3/2/pgae034/7610937.

127 Julius Endert, "Generative AI is the Ultimate Disinformation Amplifier", DW Akademie, 17 March 2024, https://akademie.dw.com/en/generative-ai-is-the-ultimate-disinformation-amplifier/a-68593890.

128 OpenAI reported disrupting over 20 deceptive networks worldwide attempting to misuse AI models for influence operations in 2024. See OpenAI, "Influence and Cyber Operations".

129 Resecurity, "Global Malicious Activity Targeting Elections is Skyrocketing", Resecurity Cyber Threat Intelligence Report, 12 February 2024, https://www.resecurity.com/blog/article/global-malicious-activity-targeting-elections-is-skyrocketing.

130 Joao Tome, "Global Elections in 2024: Internet Traffic and Cyber Threat Trends", Cloudfare Blog, 23 December 2024, https://blog.cloudflare.com/elections-2024-internet.

131 Mike Wendling, "Whirlwind of Misinformation Sows Distrust Ahead of US Election Day", BBC News, 4 November 2024, https://www.bbc.com/news/articles/czj7eex29r3o.

132 Derek B. Johnson and Aj Vincens, "Cyberattack Hits Georgia County at Center of Voting Software Breach", Cyberscoop News, 26 April 2024, https://cyberscoop.com/cyberattack-hits-georgia-county-at-center-of-voting-software-breach.

133 Robert Booth, "META Says it has Taken Down about 20 Covert Influence Operations in 2024", *The Guardian*, 3 December 2024, https://www.theguardian.com/technology/2024/dec/03/meta-says-it-has-taken-down-about-20-covert-influence-operations-in-2024.

134 Helena Junqueira, "Digital Extremism: How Algorithms Feed the Politics of Polarisation", Ipsos Synthesio Flair Collection, 17 October 2023, https://www.ipsos.com/en/flair-collection/digital-extremism-how-algorithms-feed-politics-polarisation.

135 Salsa Della Guitara Putri et al., "Echo Chambers and Algorithmic Bias: The Homogenization of Online Culture in a Smart Society", SHS Web of Conferences 202, 05001, December 2024, https://www.shs-conferences.org/articles/shsconf/pdf/2024/22/shsconf_icense2024_05001.pdf.

136 International Committee of the Red Cross, "Addressing Harmful Information in Conflict Settings: A Response Framework for Humanitarian Organizations", 30 January 2025, https://www.icrc.org/en/publication/addressing-harmful-information-conflict-settings-response-framework-humanitarian.

137 World Health Organization (WHO), "Disinformation and Public Health", WHO Information Bulletin, 6 February 2024, https://www.who.int/news-room/questions-and-answers/item/disinformation-and-public-health.

138 Samantha Bradshaw, "Disinformation and Identity-Based Violence", Stanley Center for Peace and Security, 1 November 2024, https://reliefweb.int/report/myanmar/disinformation-and-identity-based-violence.

139 Tilman Rodenhauser and Samit D'Cunha, "Foghorns of War: IHL and Information Operations During Armed Conflict", Humanitarian Law and Policy, 12 October 2023, https://blogs.icrc.org/law-and-policy/2023/10/12/foghorns-of-war-ihl-and-information-operations-during-armed-conflict.

140 Katyanna Quach, "Adobe Sells Fake AI-Generated Israel–Hamas War Images – Then the News Ran Them as Real", The Register, 8 November 2023, https://www.theregister.com/2023/11/08/adobe_ai_israel_hamas_war_pics/.

141 United Nations, A/80/257, paragraph 18.

142 ICRC statement, https://www.icrc.org/en/un-oewg-cyber-threats-7th-meeting-statement; UN Secretary-General's remarks to the Security Council's High-Level Debate on 20 June, 2024 on "Maintenance of International Peace and Security: Addressing Evolving Threats in Cyberspace", https://www.un.org/sg/en/content/sg/statements/2024-06-20/secretary-generals-remarks-the-security-councils-high-level-debate-maintenance-of-international-peace-and-security-addressing-evolving-threats-cyberspace.

143 See Member States statements on threats in the Open-ended Working Group on security of and in the use of ICTs 2021-2025, 8th and 9th substantive sessions.

144 US Cyber and Infrastructure Security Agency (CISA), "Cybersecurity Toolkit and Resources to Protect Elections", CISA-JCDC Joint Advisory, Updated 2025, https://www.cisa.gov/cybersecurity-toolkit-and-resources-protect-elections.

145 European Commission, "Strategic Communication and Countering Foreign Information Manipulation and Interference", Standard Eurobarometer 102, October 2024, https://commission.europa.eu/topics/countering-information-manipulation_en.

146 Susan Gonzales, "AI Literacy and the New Digital Divide: A Global Call for Action", United Nations Educational, Scientific and Cultural Organization (UNESCO), 23 September 2025, https://www.unesco.org/en/articles/ai-literacy-and-new-digital-divide-global-call-action.

147 World ealth Organization, "Ageism in Artificial Intelligence for Health", Policy Brief, 9 February 2022, https://www.who.int/publications/i/item/9789240040793.

148 See for example the UNIDIR Women in AI Fellowship programme, https://unidir.org/women-ai/.

149 Shimona Mohan and Dongyoun Cho, "Gender and Lethal Autonomous Weapons Systems", UNIDIR, 26 August 2024, https://unidir.org/publication/gender-and-lethal-autonomous-weapons-systems/.

150 Alistair Knott and Dino Pedreschi, "Human, or Human-Like? Transparency for AI-Generated Content", OECD AI Policy Observatory, 4 December 2023, https://oecd.ai/en/wonk/human-or-human-like-transparency-for-ai-generated-content.

151 Raj Sharma, "Why Transparency is Key to Unlocking AI's Full Potential", World Economic Forum Blog, 2 January 2025, https://www.weforum.org/stories/2025/01/why-transparency-key-to-unlocking-ai-full-potential.

152 Mantellassi and Persi Paoli, *Cloud Computing and International Security*.

153 Stephanie Pell, "Private-Sector Cyber Defense in Armed Conflict", Lawfare, 1 December 2022, https://www.lawfaremedia.org/article/private-sector-cyber-defense-armed-conflict.

154 Alex Horton and Serhii Korolchuk and Eva Dou, "Russia's Starlink Terminals Help Power Its Advance in Ukraine", *Washington Post*, 12 October 2024, https://www.washingtonpost.com/world/2024/10/12/starlink-russia-ukraine-elon-musk.

155 Bill Brenner, "How Active Adversaries Divide Labor to More Effectively Target Victims", CyberRisk Alliance Blog, 19 February 2024, https://www.scworld.com/resource/how-active-adversaries-divide-labor-to-more-effectively-target-victims.

156 Rand Corporation, "Black Markets for Hackers Are Increasingly Sophisticated, Specialized and Maturing", News Release, 25 March 2014, https://www.rand.org/news/press/2014/03/25.html.

157 See the reported average time-window for exploitation of disclosed ICT vulnerabilities in Microsoft Digital Defense Reports of 2022 and 2024 for comparison.

158 Deen Hans, "Understanding and Mitigating Distributed Denial of Service (DDoS) Attacks", Deimos Cloud Security Blog, 20 September 2023, https://www.deimos.io/blog-posts/understanding-and-mitigating-distributed-denial-of-service-ddos-attacks.

159 Thales, "Website Defacement Attack", Thales Cybersecurity Learning Center, n.d., https://www.imperva.com/learn/application-security/website-defacement-attack.

160 Adam Weiss, "How Hackers Are Using AI to Launch Smarter Phishing Campaigns", *Atlantic*, 12 November 2024, https://tomorrowsoffice.com/blog/how-hackers-are-using-ai-to-launch-smarter-phishing-campaigns.

161 KeepNet, "Ukraine's Cyber Army: A New Front in the Conflict with Russia", KeepNet Security Blog, 17 January 2024, https://keepnetlabs.com/blog/ukraine-assembles-an-it-army.

162 Cyble, "Russian Hacktivists Increasingly Tamper with Energy and Water System Controls", Cyble Blog Post, 6 December 2024, https://cyble.com/blog/russian-hacktivists-target-energy-and-water-infrastructure/.

163 Security Solutions, "Record-Breaking DDoS Attack Linked to Pro-Palestinian Group – Radware", Security Solutions Cybersecurity Blog, 12 August 2024, https://www.securitysolutionsmedia.com/2024/08/12/record-breaking-ddos-attack-linked-to-pro-palestinian-group-radware.

164 Aiden Render-Katolik, "The IT Army of Ukraine", Center for Strategic and International Studies, 15 August 2023, https://www.csis.org/blogs/strategic-technologies-blog/it-army-ukraine.

165 Ian Phillips and Vladimir Isachenkov, "Putin: Russia Doesn't Hack But "Patriotic" Individuals Might", Associated Press, 1 June 2017, https://apnews.com/article/281464d38ee54c6ca5bf573978e8ee91.

166 Hannah Beech, "The Tale of One Patriotic Cyberwarrior", *Time*, 13 February 2013, https://world.time.com/2013/02/21/chinas-red-hackers-the-tale-of-one-patriotic-cyberwarrior/.

167 Franklin D. Kramer, "The Sixth Domain: The Role of the Private Sector in Warfare", Atlantic Council, 4 October 2023, https://www.atlanticcouncil.org/in-depth-research-reports/report/the-sixth-domain-the-role-of-the-private-sector-in-warfare/#cyber.

168 Michael Hill, "How Security Vendors are Aiding Ukraine", CSO, 22 May 2022, https://www.csoonline.com/article/572163/how-security-vendors-are-aiding-ukraine.html.

169 Brad Smith, "Extending Our Vital Technology Support for Ukraine", Microsoft Corporation, On the Issues Blog, 3 November 2022, https://blogs.microsoft.com/on-the-issues/2022/11/03/our-tech-support-ukraine.

170 Horton et al.,"Russia's Starlink Terminals Help Power Its Advance in Ukraine".

171 Kim Zetter, "Security Firms Aiding Ukraine During War Could Be Considered Participants in Conflict", Zero Day Newsletter, 7 December 2022, https://www.zetter-zeroday.com/security-firms-aiding-ukraine-during/.

172 Pell, "Private-Sector Cyber Defense in Armed Conflict".

173 Matt Pollard, Fauve Kurnadi and Coline Beytout-Lamarque, "What Private Businesses Need to Know About International Humanitarian Law", ICRC Humanitarian Law and Policy Blog, 26 November 2024, https://blogs.icrc.org/law-and-policy/2024/11/26/what-private-businesses-need-to-know-about-international-humanitarian-law.

174 Pell, "Private-Sector Cyber Defense in Armed Conflict".

175 Heide Moore and Dan Roberts, "AP Twitter Hack Causes Panic on Wall Street and Send Dow Plunging", *The Guardian*, 23 April 2013, https://www.theguardian.com/business/2013/apr/23/ap-tweet-hack-wall-street-freefall.

176 Luke Harding and Charles Arthur, "Syrian Electronic Army: Assad's Cyber Warriors", *The Guardian*, 30 April 2013, https://www.theguardian.com/technology/2013/apr/29/hacking-guardian-syria-background.

177 Kubo Macak and Mauro Vignati, "Civilianization of Digital Operations: A Risky Trend", Lawfare, 5 April 2023, https://www.lawfaremedia.org/article/civilianization-digital-operations-risky-trend.

178 John Leyden, "US Government Warns Script Kiddies to Stay Out of Cyberwar", The Register, 13 February 2023, https://www.theregister.com/2003/02/13/us_gov_warns_script_kiddies/.

179 Joe Tidy, "Meet the Hacker Armies on Ukraine's cyber front line", BBC News, 15 April 2023, https://www.bbc.com/news/technology-65250356.

180 Wieslaw Gozdziewicz, "Militias, Volunteer Corps, Levee en Masse or Simply Civilians Directly Participating in Hostilities? Certain Views on the Legal Status of 'Cyberwarriors' under Law of Armed Conflict", *European Cybersecurity Journal*, vol. 2, no. 2 (February 2016), https://www.2017.cybersecforum.eu/files/2016/12/ecj_vol2_issue2_w.gozdziewicz_militias_volunteer_corps_levee_en_masse_or_simply_civilians_directly_participating_in_hostilities.pdf.

181 Macak and Vignati, "Civilianization of Digital Operations".

182 Blackberry researchers note an increase in the outsourcing of cyber espionage to mercenary APT groups. This may allow states to protect themselves from being identified, effectively hiding behind proxy attackers. While. in theory, anyone can hire a mercenary APT, the more sophisticated APT mercenary groups typically prefer high-profile, well-paying customers. See Anna Zhadan, "The Curious Case of Cyber Warriors: Backing Nation States in Cyberwarfare", Cybernews, 28 September 2022, https://cybernews.com/editorial/cyber-warriors-backing-states-in-cyberwarfare/.

183 Csaba Krasznay, "Bridging the Gap: Private Sector's Vital Role in Military Cyber Defense", White Hat IT Security Blog, 3 January 2024, https://whitehat.eu/bridging-the-gap-private-sectors-vital-role-in-military-cyber-defense.

184 Macak and Vignati, "Civilianization of Digital Operations".

185 US Cybersecurity and Critical Infrastructure Security Agency (CISA), "Iran-based Cyber Actors Enabling Ransomware Attacks on US Organizations", Joint CISA and FBI Advisory, 28 August 2024, https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-241a.

186 Natalia Krapiva and Anastasya Zhyrmont, "Civil Society in Latvia, Lithuania, and Poland Targeted with Pegasus Spyware", Access Now News Update, 30 May 2024, https://www.accessnow.org/publication/civil-society-in-exile-pegasus/.

187 United Nations, Security Council, Report of the Panel of Experts established pursuant to resolution 1874 (2009), S/2023/171, 7 March 2023, https://docs.un.org/S/2023/171.

188 United Nations, A/79/214, paragraphs 11, 16.

189 Vibhu Mishra, "UN Chief Warms of 'Cyber Mercenaries' Amid Spike in Weaponising Digital Tools", United Nations News, 20 June 2024, https://news.un.org/en/story/2024/06/1151266.

190 British Foreign, Commonwealth and Development Office (FCDO), "The Pall Mall Process Declaration: Tackling Proliferation and Irresponsible Use of Commercial Cyber Intrusion Capabilities", FCDO Policy Paper, 28 February 2025, https://www.gov.uk/government/publications/the-pall-mall-process-declaration-tackling-proliferation-and-irresponsible-use-of-commercial-cyber-intrusion-capabilities.

191 Cybersecurity Tech Accord, "Cyber Mercenaries: An Old Business Model, a Modern Threat: Cybersecurity Tech Accord Principles Limiting Offensive Operations In Cyberspace", Tech Accord News, 27 March 2023, https://cybertechaccord.org/uploads/prod/2023/03/Cyber-mercenary-principles_Tech-Accord_032723_FINAL.pdf.

192 Timan Rodenhouser and Mauro Vignati, "8 Rules for 'Civilian Hackers' and 4 Obligations for States to Restrain Them", ICRC Humanitarian Law and Policy Blog, 4 October 2023, https://blogs.icrc.org/law-and-policy/2023/10/04/8-rules-civilian-hackers-war-4-obligations-states-restrain-them/.

193 Kaspersky Institute, "Cybersecurity in the AI Era: How the Threat Landscape Evolved in 2023", Press Release, 11 December 2023, https://www.kaspersky.com/about/press-releases/cybersecurity-in-the-ai-era-how-the-threat-landscape-evolved-in-2023.

194 Derek Manky and Gil Baram, "Beyond Phishing: Exploring the Rise of AI-enabled Cybercrime", Barkeley University Center for Long-Term Cybersecurity, January 2025, https://cltc.berkeley.edu/2025/01/16/beyond-phishing-exploring-the-rise-of-ai-enabled-cybercrime.

195 Catherine Bolgar, "Microsoft's Majorana 1 Chip Carves New Path for Quantum Computing", Microsoft Corporation, 19 February 2025, https://news.microsoft.com/source/features/innovation/microsofts-majorana-1-chip-carves-new-path-for-quantum-computing.

196 See for example the Open-ended Working Group on security of and in the use of ICTs, 10th substantive session, 21 February 2025, https://webtv.un.org/en/asset/k15/k15n8qgoli.

197 Lucia Stanham, "AI-Powered Cyberattacks", CrowdStrike Cybersecurity 101: The Fundamentals of Cybersecurity, 16 January 2025, https://www.crowdstrike.com/en-us/cybersecurity-101/cyberattacks/ai-powered-cyberattacks.

198 British National Cyber Security Center (NCSC), "NCSC Warns of Widening AI Gap Between Cyber Threats and Defense Capabilities", NCSC News, 16 October 2024, https://www.ncsc.gov.uk/news/ncsc-warns-widening-gap-between-cyber-threats-and-defence-capabilities.

199 Vasil Michev, "Cyber Attack Vectors in Microsoft 365: Detect and Prevent Entry", Core View Security Blog, 20 September 2024, https://www.coreview.com/blog/the-anatomy-of-a-microsoft-365-hack-part-1-entry.

200 Abdulaziz Almaslukh, "AI Could Empower and Proliferate Social Engineering Cyberattacks", World Economic Forum, 25 October 2024, https://www.weforum.org/stories/2024/10/ai-agents-in-cybersecurity-the-augmented-risks-we-all-need-to-know-about.

201 Dan Milmo and Alex Hern, "AI Will Make Scam Emails Look Genuine, UK Cybersecurity Agency Warns", *The Guardian*, 24 January 2024, https://www.theguardian.com/technology/2024/jan/24/ai-scam-emails-uk-cybersecurity-agency-phishing.

202 Christine Barry, "5 Ways Cybercriminals Are Using AI: Malware Generation", Barracuda Security Blog, 16 April 2024, https://blog.barracuda.com/2024/04/16/5-ways-cybercriminals-are-using-ai--malware-generation.

203 Jeff Sims, "EyeSpy Proof-of-Concept: A Cognitive Threat Agent", HYAS Labs, 1 August 2023, https://www.hyas.com/blog/eyespy-proof-of-concept.

204 Dena De Angelo, "The Dark Side of AI in Cybersecurity: AI-Generated Malware", Palo Alto Networks, 15 May 2024, https://www.paloaltonetworks.com/blog/2024/05/ai-generated-malware/.

205 See the Mitre Atlas Machine Learning Cyber Attack Matrix for possible attack vectors of adversarial attacks on AI systems, https://atlas.mitre.org/matrices/ATLAS.

206 Ioana Puscas, "AI and International Security: Understanding the Risks and Paving the Path for Confidence-Building Measures", UNIDIR Security and Technology Programme, 12 October 2023, https://unidir.org/wp-content/uploads/2023/10/UNIDIR_AI-international-security_understanding_risks_paving_the_path_for_confidence_building_measures.pdf.

207 Microsoft Corporation, "Microsoft Digital Defense Report 2024".

208 Lewis Birch, "AI Under Attack: Six Key Adversarial Attacks and Their Consequences", 20 September 2025, https://mindgard.ai/blog/ai-under-attack-six-key-adversarial-attacks-and-their-consequences.

209 Iqbal H. Sarker et al., "Explainable AI for Cybersecurity, Automation, Intelligence and Trustworthiness: Methods, Taxonomy, Challenges and Prospects", *ICT Express*, vol. 10, no. 4 (August 2024), https://www.sciencedirect.com/science/article/pii/S2405959524000572?via%3Dihub.

210 Samuele Dominioni, "Exploring the AI–ICT Security Nexus", UNIDIR Security and Technology Programme, 5 December 2024, https://unidir.org/wp-content/uploads/2024/12/UNIDIR_ICT_Intrusion_Path_A4_Final.pdf.

211 Microsoft Corporation, "Governments Face Unprecedented Cyber Threats: AI Emerges as the Ultimate Defense to Cybercrime", CEE Multi-Country News Center, 28 January 2025, https://news.microsoft.com/en-cee/2025/01/28/governments-face-unprecedented-cyber-threats-ai-emerges-as-the-ultimate-defense-to-cybercrime/.

212 Terralogic, "The State of AI in Cybersecurity: How AI will Impact the Cyber Threat Landscape in 2025", Terralogic Blog, December 2024, https://terralogic.com/ai-in-cybersecurity/.

213 International Energy Agency, "AI and Energy Security", World Energy Outlook Special Report, January 2025, https://www.iea.org/reports/energy-and-ai/ai-and-energy-security.

214 Puscas, "AI and International Security".

215 Marcus Comiter, "Attacking Artificial Intelligence: AI's Security Vulnerability and What Policymakers Can do About It", Belfer Center for Science and International Affairs, Harvard Kennedy School Report, August 2019, https://www.belfercenter.org/publication/AttackingAI.

216 DeepMind Google Report, "Evaluating Potential Cybersecurity Threats of Advanced AI", Google DeepMind, January 2024, https://deepmind.google/discover/blog/evaluating-potential-cybersecurity-threats-of-advanced-ai.

217 Dominioni, "Exploring the AI–ICT Security Nexus".

218 United Nations, A/80/257, paragraph 26.

219 United Nations Secretariat, *Humanity's Fate Cannot Be Left to Algorithm,' Warns Secretary-General in Security Council Debate on Artificial Intelligence*, Secretary General Office Press Release, 25 September 2025, https://press.un.org/en/2025/sgsm22830.doc.htm.

220 United Nations Educational, Scientific and Cultural Organisation (UNESCO), "International Year of Quantum Science and Technology", 7 June 2024, https://quantum2025.org/.

221 Zhanna L. Malekos Smith and Giacomo Persi Paoli, "Quantum Technology, Peace and Security: A Primer", UNIDIR Security and Technology Programme, 21 November 2024, https://unidir.org/wp-content/uploads/2024/11/UNIDIR_quantum_technology.pdf.

222 Malekos Smith and Persi Paoli, "Quantum Technology, Peace and Security".

223 United Nations, General Assembly, "Current Developments in Science and Technology and Their Potential Impact on International Security and Disarmament Efforts", Report of the Secretary-General, A/79/224, 23 July 2024, https://documents.un.org/doc/undoc/gen/n24/218/85/pdf/n2421885.pdf.

224 Rachel Hall, "Microsoft unveils chip it says could bring quantum computing within years", *The Guardian*, 19 February 2025, https://www.theguardian.com/technology/2025/feb/19/topoconductor-chip-quantum-computing-topological-qubits-microsoft.

225 On2IT, "Harvest Now, Decrypt Later: Preparing for Quantum Computing Threats", Zero Trust Innovators Security Bulletin, n.d., https://on2it.net/preparing-for-quantum-computing-threats/.

226 Malekos Smith and Persi Paoli, "Quantum Technology, Peace and Security".

227 See, for example, the UNIDIR event on "Quantum Technologies and their Implications for International Peace and Security", https://unidir.org/event/multi-stakeholder-dialogue-on-quantum; or UNIDIR "2024 Innovations Dialogue: Quantum Technologies", https://unidir.org/event/2024-innovation-dialogue-quantum-technologies-and-their-implications-for-international-peace-and-security/.

228 Dongyoun Cho, "2024 Innovations Dialogue: Quantum Technologies and Their Implications for International Peace and Security – Conference Summary Report", UNIDIR Security and Technology Programme, December 2024, https://unidir.org/wp-content/uploads/2024/12/UNIDIR_Innovations_Dialogue_2024.pdf.

229 European Commission, "European Quantum Communication Infrastructure – EuroQCI", December 2024, https://digital-strategy.ec.europa.eu/en/policies/european-quantum-communication-infrastructure-euroqci.

230 Surbhi Singh, "Microsoft and Apple Advance Post-Quantum Cryptography Support in Upcoming OS Releases", Encryption Consulting Blog, 15 June 2025, https://www.encryptionconsulting.com/microsoft-and-apple-advance-post-quantum-cryptography-support-in-upcoming-os-releases/.

231 Kelly Simon, Henderson Tim and Adebis Hertiage Samuel, "Quantum Computing and Cybersecurity: Emerging Threats and Strategic Opportunities", ResearchGate: Computer Science and Engineering, March 2025, https://www.researchgate.net/publication/390095755_Quantum_Computing_and_Cybersecurity_Emerging_Threats_and_Strategic_Opportunities.

232 Jennifer Lavoie, Samantha Smoak and Jamie Moody, "JPMorgan Chase, Toshiba and Ciena Build the First Quantum Key Distribution Network Used to Secure Mission-Critical Blockchain Application", Toshiba Press Release, 17 February 2022, https://www.toshiba.eu/solutions/quantum/news/jpmorgan-chase-toshiba-and-ciena-build-the-first-quantum-key-distribution-network-used-to-secure-mission-critical-blockchain-application/.

233 Edward Parker, "US-Allied Militaries Must Prepare for the Quantum Threat to Cryptography", RAND Corporation, 6 June 2025, https://www.rand.org/pubs/commentary/2025/06/us-allied-militaries-must-prepare-for-the-quantum-threat.html.

234 Fortinet, "Quantum Key Distribution (QKD)", Fortinet Cyber Glossary, n.d., https://www.fortinet.com/resources/cyberglossary/quantum-key-distribution.

235 Argyri Panezi, "The Security Stakes in the Global Quantum Race", Just Security, 15 July 2025, https://www.justsecurity.org/116473/security-stakes-global-quantum-race/.

236 Nik Faiz Ruzman, "Quantum Threat to Cryptopgraphy: Experts Warn of an Extinction-Level Event in Digital Trust", Cybersecurity Asia, 2 October 2024, https://cybersecurityasia.net/quantum-threat-to-cryptography-experts-warn/.

237 Scott Buchholz et al., "The Realist's Guide to Quantum Technology and National Security", Deloitte, 6 February 2020, https://www.deloitte.com/us/en/insights/industry/government-public-sector-services/the-impact-of-quantum-technology-on-national-security.html.

238 United Nations, A/80/257, paragraph 26.

239 WIS@key, "Cybersecurity in a Post-Quantum AI Era: What Happens When you Mix Generative AI with QUantum Computing", Davos 2025 WIS&key panel discussion, January 2025, https://www.wisekey.com/davos25/quantumpanel/.

@unidir
/unidir
/un_disarmresearch
/unidirgeneva
/unidir

Palais des Nations
1211 Geneva, Switzerland

WWW.UNIDIR.ORG