ujiyari

UPSC & STATE PCS CURRENT AFFAIRS · UJIYARI.COM

EDITORIAL ANALYSIS

The Double Helix: Why AI and Cyber Are No Longer Two Problems

THE HINDU

4 August 2026 · SECURITY & DEFENCE · SCIENCE & TECH · GS3 · GS2

CURATED & WRITTEN BY

Bharat Choudhary

UPSC Educator & Content Creator

linkedin.com/in/epicbharat

The Double Helix: Why AI and Cyber Are No Longer Two Problems

The Hindu 4 August 2026 GS3 GS2

INTERVIEW ANGLE

"If an AI system can independently discover zero-day vulnerabilities, the same capability that lets an attacker find them lets a defender find them first. What determines which side that capability favours, and is there any policy lever that shifts the balance toward defence?"

Every fact web-verified against primary sources

THE LIFT LINE

Zero Trust was built to keep out anyone who could not prove they belonged. It has no answer to an attacker that can.

WHY THIS EDITORIAL MATTERS FOR YOUR EXAM

Cybersecurity answers in GS3 tend to be lists: CERT-In, the National Cyber Security Policy 2013, critical information infrastructure, a gesture at capacity-building. This editorial supplies the mechanism layer, why AI changes the character of the threat rather than its volume, which converts a list into an argument. The generative-versus-agentic distinction and the Zero Trust critique are both new enough that most candidates will not have them.

GS Paper 3: Basics of cyber security; challenges to internal security through communication networks; role of media and social networking sites in internal security challenges; awareness in the fields of IT and computers.

GS Paper 2: Important international institutions and groupings, and the gaps in global governance architecture.

CONCEPT	MEANING	WHY IT IS TESTABLE
Generative AI	Systems that produce text, code or images in response to a prompt	The familiar category; useful as the contrast term
Agentic AI	Systems that pursue multi-step goals autonomously, taking actions without prompting at each step	Where the editorial locates the danger; the distinction is the examinable point
Zero Trust	A security model that grants no implicit trust and verifies every request regardless of network location	The current enterprise standard, and what the editorial argues is being structurally defeated
Zero-day vulnerability	A flaw unknown to the vendor, for which no patch exists	AI-assisted discovery inverts the discovery-to-patch relationship
Insider threat	Risk arising from an actor operating with legitimate access	The vector an autonomous agent with valid credentials occupies

BACKGROUND AND CONTEXT

The author is **M.K. Narayanan**, former National Security Adviser and former Governor of West Bengal, writing in The Hindu on 4 August 2026.

The organising claim is that AI has crossed a threshold. So long as AI was an efficiency mechanism, adopted where it improved a process, its failures were contained to that process. Once AI becomes an **infrastructure mandate**, embedded in power dispatch, financial settlement, clinical decision support and telecommunications routing, an AI failure is an infrastructure failure. India is directly exposed here. **CERT-In handled about 29.44 lakh incidents in 2025**, up roughly 44 per cent on the 20.41 lakh of 2024, and the **World Economic Forum's Global Risks Report 2026 ranks cybersecurity as India's single largest national risk**, ahead of economic downturn, climate disaster and armed conflict. Meanwhile the country's power grids, banking, health systems and telecom have been substantially digitised over a short period.

THE CORE ARGUMENT / ISSUE

The three multiplications

AXIS	WHAT AI CHANGES	WHY THE EXISTING DEFENCE FAILS
Autonomy	Malware adapts to the defensive environment it encounters	Signature-based detection matches known patterns; self-modifying code has no stable pattern
Accessibility	Capability once requiring state resources becomes available to non-state actors	Threat models built around a small number of identifiable state adversaries misdescribe the threat population
Speed	Attack sequences compress below human decision cycles	Any defence with a human in the loop is, by construction, too slow

The third is the one with the uncomfortable implication. If the attack completes faster than a human can respond, the only viable defence is also autonomous, which means delegating consequential security decisions to machines. That is a governance problem dressed as a technical one.

Why the Zero Trust critique is the sharpest claim

Zero Trust replaced the older perimeter model, in which anything inside the network was trusted. Its principle is that trust is never implicit and every request is verified regardless of origin. Against an external attacker attempting to appear internal, this works well.

An **autonomous malicious agent** presents a different problem. It operates with credentials that are genuinely valid, having been obtained through phishing, compromise or misconfiguration, and it can behave in ways statistically indistinguishable from an authorised user because it can model what authorised behaviour looks like. Verification succeeds, because there is nothing wrong with the credential. The attacker is not defeating the verification; it is satisfying it.

This is why the editorial's framing matters, and it is also where the framing should be handled with care. Narayanan's claim is not that Zero Trust is badly implemented or under-funded, but that its foundational assumption, that verification distinguishes the legitimate from the illegitimate, does not hold against an adversary that can be verified and is illegitimate.

The security literature agrees on the **mechanism** and is more divided on the **conclusion**. It broadly accepts that an autonomous agent occupies the insider-threat vector that credential checking cannot reach; the working consensus is that every autonomous agent should be treated as an insider threat, and the number of non-human identities in enterprise environments now vastly exceeds the number of human ones. But the prevailing prescription is to **extend** Zero Trust to non-human principals, through per-agent identity, short-

lived and narrowly scoped credentials, runtime authorisation and continuous behavioural monitoring, rather than to abandon it. Narayanan's own remedy, notably, is **"Zero Trust 2.0"**, which is an upgrade rather than a replacement.

A candidate should therefore present the strong form as the columnist's position and the extension response as the field's, because the difference between them is exactly the kind of distinction an examiner rewards.

The vulnerability-discovery inversion

Vulnerability management rests on an implicit race. Researchers and vendors find flaws and issue patches; attackers find flaws and exploit them. The system is tolerable because discovery is slow, expensive and requires scarce expertise.

AI systems capable of independently discovering zero-days remove the scarcity constraint. If discovery becomes cheap and fast while patching remains slow, because patching requires vendor action, testing, distribution and, in critical infrastructure, scheduled downtime, then the stock of exploitable unpatched vulnerabilities grows structurally rather than incidentally.

The governance vacuum

DOMAIN	GOVERNING ARCHITECTURE
Nuclear weapons	NPT, CTBT, bilateral arms-control treaties, IAEA safeguards
Chemical weapons	Chemical Weapons Convention, with OPCW verification
Biological weapons	Biological Weapons Convention, **without a verification protocol**; negotiations on one collapsed in 2001
AI-enabled offensive cyber operations	**Thin, fragmented and largely non-binding on offensive use**

The accurate statement is not that nothing exists, but that what exists is thin and that **India is outside all of it**. That is a sharper and more useful point than a blanket vacuum claim.

INSTRUMENT	STATUS	INDIA
Budapest Convention on Cybercrime (2001)	Binding on parties; India has declined since 2001, on sovereignty and data-sharing grounds and because it was drafted without Indian participation	**Not a party**
UN Convention against Cybercrime (opened for signature at Hanoi, 25 October 2025)	Around 65 to 70 states signed on the opening day	**Has not signed**
Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law (adopted May 2024, opened for signature September 2024)	The **first legally binding international AI treaty**; signed by the United States, United Kingdom, European Union, Japan, Canada, Australia and others	**Not a signatory**
Lethal autonomous weapons systems	Informal meetings under the Convention on Certain Conventional Weapons since 2014, a Group of Governmental Experts since 2016; no binding instrument	Participant, no binding obligation
State behaviour in cyberspace	UN GGE norms of 2015, voluntary; the Open-Ended Working Group concluded in 2025 with a successor permanent mechanism	Participant, voluntary only

None of these constrains AI-enabled offensive operations specifically, and the acceleration of the United States and China capability race is proceeding without one. Narayanan's own formulation is that international law here is in a fledgling state without unified enforcement, which is more precise than saying it does not exist.

HOW TO THINK ABOUT THIS (ANALYTICAL FRAME)

When a defensive system fails, distinguish failures of implementation from failures of assumption, because the two require entirely different responses. An implementation failure is met with more funding, better training, wider deployment. An assumption failure is met by redesigning the model, and no amount of investment in the existing model helps. The diagnostic question is: does the attack succeed despite the defence working correctly? If it does, the assumption is wrong. Apply this to fraud detection, border management, financial regulation and epidemiological surveillance alike; systems under pressure reliably respond to assumption failures by funding implementation, because that response is available and the other is not.

THE DIAGRAM IN WORDS

Picture a building with a guard at every door who checks identification against a register. That is Zero Trust, and it is a genuine improvement on the older arrangement of one guard at the front gate and free movement inside. Now picture a visitor who holds a genuine pass, issued to a real employee, and who has studied enough of that employee's routine to walk the corridors in a way nobody finds odd. Every guard checks the pass. Every check passes. The building's security is functioning exactly as designed, and the intruder is inside. The problem is not that the guards are lazy or too few. It is that the building's security was built on the belief that identification distinguishes friend from adversary, and that belief has stopped being true.

WAY FORWARD

1. **Preserve meaningful human control over military decision-making**, and press for international norms on autonomous weapons and AI-enabled operations in a domain that presently has none.
2. **Mandate pre-deployment safety testing** for AI systems entering critical infrastructure, on the same logic that governs pharmaceutical and aviation certification.
3. **Build indigenous AI and semiconductor capability**, on the reasoning that a country that imports its computational substrate cannot fully secure it.
4. **Fund cybersecurity research and skilled manpower**, since the binding constraint on Indian cyber defence is people rather than policy documents.
5. **Harden critical infrastructure specifically**, treating power grids, banking, health systems and telecom as a distinct protection category rather than as ordinary enterprise networks.
6. **Address algorithmic radicalisation as a security matter**, since personalised feed distortion is a societal-security threat that sits outside the conventional cyber portfolio.

PYQ LINKAGE AND PRACTICE

UPSC has tested cyber security, critical information infrastructure, social media and internal security, and the challenges of emerging technologies across recent GS3 cycles. This editorial supplies the mechanism that connects the AI questions to the cyber questions, which most answers currently treat as separate syllabus entries.

Practice question: "The convergence of artificial intelligence and cyber capability has rendered prevailing defensive architectures obsolete at the level of assumption rather than implementation." Examine this claim, and assess what institutional and international responses it implies for India. (250 words, 15 marks)

Interview angle: If an AI system can independently discover zero-day vulnerabilities, the same capability that lets an attacker find them lets a defender find them first. What determines which side that capability favours, and is there any policy lever that shifts the balance toward defence?

Sources: The Hindu, CERT-In, Ministry of Electronics and Information Technology

KEY ARGUMENTS AT A GLANCE

Artificial intelligence and cyber are no longer separable security domains but two intertwined strands of one threat architecture, each amplifying the other; because AI has crossed from being an efficiency tool to an infrastructure mandate, every AI weakness is now a national-infrastructure weakness, and Narayanan's sharpest claim is that the prevailing defensive paradigm fails at the level of its assumptions rather than its resourcing, though the security literature treats this as a case for extending that paradigm rather than abandoning it.

✓ SUPPORTING

- AI acts as a force multiplier for offensive cyber capability in three distinct ways: it enables autonomous malware that adapts to the defensive environment it encounters, which defeats signature-based detection; it lowers the technical barrier to entry so that non-state actors can run campaigns that previously required state resources; and it compresses attack timelines below the speed of human response.
- The critical distinction is between generative AI, which responds to prompts by producing text or code, and agentic AI, which pursues multi-step goals autonomously. Malicious autonomous agents undermine Zero Trust architectures by exploiting legitimate credentials and behaving like authorised insiders, which places them in the insider-threat vector that credential verification was never designed to reach; the prevailing response in the field is to extend Zero Trust to non-human identities through per-agent identity, short-lived scoped credentials and runtime authorisation, which Narayanan himself calls Zero Trust 2.0.
- AI systems capable of independently discovering zero-day vulnerabilities constitute a new class of strategic capability, because discovery can now outpace patching; and unlike nuclear or chemical weapons, AI-enabled offensive operations are governed only by thin and largely non-binding instruments, and India stands outside the three that exist, the Budapest Convention, the UN Convention against Cybercrime opened at Hanoi in October 2025, and the Council of Europe Framework Convention on Artificial Intelligence of 2024, while the United States and China accelerate a capability race.

⚠ COUNTER

The same capabilities are available to defenders, and asymmetry claims in cybersecurity have a long history of being overstated: AI-driven anomaly detection, automated patching and behavioural analytics scale defence in ways signature-based tools never could, and a well-

resourced defender with better data about its own network has advantages an attacker does not. The honest position is that AI raises the ceiling on both offence and defence, and which side gains depends on institutional capacity and investment rather than on anything intrinsic to the technology.

→ **WAY FORWARD**

Maintain meaningful human control over military decision-making, build international norms on autonomous weapons and AI-enabled operations where treaty architecture does not yet exist, mandate rigorous pre-deployment safety testing, and for India specifically develop indigenous AI and semiconductor capability, fund cybersecurity research, build skilled manpower, and harden the critical infrastructure, power grids, banking, health systems and telecom, that digitisation has exposed.

MAINS ANSWER FRAMEWORK

QUESTION

Examine the proposition that artificial intelligence and cybersecurity have converged into a single threat domain, and assess the adequacy of India's institutional preparedness for AI-enabled offensive cyber operations. (250 words)

INTRODUCTION

The conventional treatment of artificial intelligence and cybersecurity as adjacent but separate policy domains has become untenable. They now function as two strands of a single threat architecture, in which AI supplies capability to cyber operations and cyber vulnerability supplies the attack surface for AI systems, and the convergence is significant because AI has moved from being an efficiency mechanism to an infrastructure mandate, which means every weakness in an AI system is now a weakness in national infrastructure.

BODY

AI multiplies offensive cyber capability along three axes. First, autonomy: malware that adapts to the defensive environment it encounters defeats signature-based detection, which works by matching known patterns and therefore cannot recognise code that rewrites itself.

Second, accessibility: capabilities that once required state-level resources, sustained reconnaissance, custom tooling, skilled operators, are now available to non-state actors, which changes the threat population rather than merely the threat level. Third, speed: attack sequences that compress below the interval a human analyst needs to observe, decide and respond effectively remove the human from the defensive loop.

The sharpest claim concerns architecture. Zero Trust, the prevailing enterprise security model, assumes no implicit trust and verifies every request, which is robust against an attacker impersonating an outsider. An autonomous malicious agent operating with legitimate credentials and behaving in ways statistically consistent with an authorised user is not impersonating an outsider; it aggravates the insider-threat vector, which Zero Trust was never designed to solve. The defensive paradigm is therefore obsolete at the level of assumption rather than execution.

Two further concerns follow: AI systems that independently discover zero-day vulnerabilities invert the discovery-to-patching relationship on which vulnerability management depends, and algorithmic radicalisation through personalised feed distortion converts recommendation systems into a societal-security problem. Governing all of this is a treaty vacuum, since nuclear and chemical weapons have regimes and AI-enabled offensive operations have none, while the United States and China accelerate a capability race.

CONCLUSION

The response has to operate at two levels. Internationally, norms on meaningful human control and on autonomous weapons need building where no treaty architecture exists, alongside mandatory pre-

deployment safety testing.

Domestically, India, among the most-targeted countries for cyber attack and with power grids, banking, health systems and telecom now substantially digitised, needs indigenous AI and semiconductor capability, funded cybersecurity research and skilled manpower, on the reasoning that a country that imports its computational substrate cannot secure it.

RELATED DAILY ARTICLES

4 Aug Current Affairs Today, August 4, 2026

3 Aug Current Affairs Today, August 3, 2026

3 Aug Who Pays for Free? The Proposal to Bring Back MDR on...

1 Aug Current Affairs Today, August 1, 2026

CURATED & WRITTEN BY

Bharat Choudhary

UPSC Educator & Content Creator

linkedin.com/in/epicbharat

📖 Read Full Article on Ujiyari →

https://ujiyari.com/editorials/2026/08/th-ai-cyber-double-helix-security-threats-2026/