BEYOND FIREWALLS:

THE MYTH-REALITY DIVIDE IN CYBER SECURITY

Editor-in-chief

Prof. (Dr.) Nirmal Kanti Chakrabarti

Prof. (Dr.) Paresh Kumar Acharya

Edited by

Ms. Lamiya Sultana.

Dr. Chandrima Chakbarorty.

Ms. Ankita Mukherjee

BEYOND FIREWALLS 2026

ABOUT THE BOOK

Beyond Firewalls: The Myth–Reality Divide in Cyber Security challenges the narrow view that cybersecurity is limited to tools like firewalls and antivirus software. It highlights how modern threats—such as phishing, ransomware, and social engineering—often exploit human vulnerabilities rather than technical flaws.

The book explores the "myth–reality divide" between perceived security and actual resilience. It argues that true cybersecurity depends on people, processes, and policies, not just technology.

Emerging risks like AI-driven attacks and cybercrime networks are also examined, especially in rapidly digitizing regions like India. Emphasizing strategies like Zero Trust and continuous monitoring, the book advocates a proactive approach. Overall, it presents cybersecurity as an ongoing process requiring awareness, adaptability, and responsibility.

9-78-81-993286-5-1

Department of School of Legal Studies, Swami Vivekananda University,

Barrackpore

Edition: First Edition, 2025

Publisher: Swami Vivekananda University Press

Telinipara, Barasat–Barrackpore Road

Bara Kanthalia, West Bengal – 700121, India

Website:

https://www.swamivivekanandauniversitypress.com

ISBN: 978-81-993286-3-1

Cover Designed By: Multimedia Team, Swami Vivekananda

University

Disclaimer:

The views, interpretations, and conclusions expressed in this book are those of the author and do not necessarily reflect the views of the publisher or the institution. The publisher assumes no responsibility for any errors or omissions or for any consequences arising from the use of the information contained in this book.

Price: 350/-

BEYOND FIREWALLS
THE MYTH-REALITY DIVIDE IN CYBER SECURITY

SWAMI VIVEKANANDA UNIVERSITY
SCHOOL OF LEGAL STUDIES

Editor-in-Chief
Prof. (Dr.) Nirmal Kanti Chakrabarti
Prof. (Dr.) Paresh Kumar Acharya

Edited by
Lamiya Sultana
Dr. Chandrima Chakraborty
Ankita Mukherjee

Swami Vivekananda University, Kolkata (Institutional Publisher) Published by the Swami Vivekananda University (Institutional Publisher), Kolkata-700121, West Bengal, India ISBN: 978-81-993286-3-1 First Edition: March, 2026, Publisher: Swami Vivekananda University (Institutional Publisher), Kolkata-700121,

Preface

In an age where digital systems underpin nearly every aspect of human life, cybersecurity has emerged as both a shield and a paradox. *Beyond Firewalls: The Myth–Reality Divide in Cyber Security* is an attempt to navigate this paradox—where perception often diverges sharply from practice, and where confidence in protection can mask profound vulnerabilities.

The popular imagination tends to reduce cybersecurity to visible defenses: firewalls, antivirus software, and encryption tools. While these are important, they represent only a fraction of a far more complex ecosystem. This book challenges the comforting myth that technology alone can secure us. It argues instead that cybersecurity is as much about human behavior, institutional culture, governance structures, and evolving threat landscapes as it is about code and hardware.

Across its chapters, this work explores the gap between what individuals, organizations, and governments believe about cybersecurity and what actually unfolds in practice. Why do sophisticated systems still fail? How do human errors and cognitive biases undermine advanced defenses? Why do policy frameworks often lag behind technological change? These questions form the core of the inquiry.

This book is not written solely for technical experts. It seeks to engage students, policymakers, educators, and general readers who wish to understand cybersecurity beyond jargon and surface-level narratives. By bridging theoretical insights with real-world examples, it aims to demystify complex concepts while encouraging critical reflection on widely held assumptions.

Importantly, *Beyond Firewalls* also emphasizes that cybersecurity is not merely a technical challenge but a societal one. Trust, accountability, ethics, and resilience are central themes that run throughout the discussion. In a world increasingly shaped by data and digital interdependence, security cannot be an afterthought—it must be an integrated and conscious practice.

This preface sets the stage for a journey that moves beyond simplistic binaries of safety and risk, strength and weakness, myth and reality. Instead, it invites readers to embrace nuance and complexity, recognizing that true cybersecurity lies not in absolute protection, but in continuous adaptation, awareness, and informed engagement.

Ultimately, this book hopes to inspire a shift in perspective—from seeing cybersecurity as a static barrier to understanding it as a dynamic, shared responsibility.

Ms. Lamiya Sultana
Dr. Chandrima Chakraborty
Ms. Ankita Mukherjee

CONTENTS

Chapter – I

FOUNDATIONS OF CYBER SECURITY:

CONCEPTS, DEFINITIONS AND CORE PRINCIPLES

Moumita De, State Aided College Teacher, Department of Human Rights, Ramakrishna Sarada Mission Vivekananda Vidyabhavan College and Ph.D. Research Scholar, Adamas University.

Abstract:

Cyber security is the discipline dedicated to protecting digital systems, networks, programs, and data from unauthorized access, misuse, disruption, or destruction. As societies increasingly rely on interconnected technologies, the need for structured security foundations has become critical. The foundations of cyber security are built upon clear concepts, precise definitions, and established core principles that guide protective measures across digital environments. Central to this field is the understanding of threats, vulnerabilities, risks, and controls, which together shape effective defense strategies. The Confidentiality, Integrity, and Availability (CIA) triad serves as a primary model, ensuring that information is accessible only to authorized users, remains accurate and unaltered, and is available when required. Additional principles such as authentication, authorization, accountability, and non-repudiation strengthen trust within systems. Risk management, governance frameworks, and security policies further support structured protection efforts. By integrating technical controls with organizational practices and human awareness, cyber security establishes a resilient defense posture. A strong foundation in these core concepts enables institutions and individuals to anticipate emerging threats, minimize potential damage, and maintain the reliability and stability of digital infrastructures in an evolving technological landscape.

Keywords: interconnected technologies, organizational practices and human awareness, cyber security.

Introduction

In an era defined by digital transformation, cybersecurity has emerged as a central concern for governments, corporations, educational institutions, and individuals. The rapid expansion of internet connectivity and reliance on digital systems have created a landscape where cyber threats can inflict economic loss, privacy violations, reputational damage, and even national insecurity.

Cybersecurity, as a discipline, addresses the protection of systems and information from unauthorized access, damage, or disruption.[1]

Although the concept may seem simple, cybersecurity encompasses a complex array of technologies, processes, principles, and human behaviours. As such, a clear conceptual and definitional foundation is essential to understanding its practice and challenges. This paper synthesizes core ideas from academic literature, standards bodies, and industry frameworks to provide a comprehensive overview of the foundations of cybersecurity.

Core Concepts in Cybersecurity

Several foundational concepts underpin cybersecurity theory and practice. These concepts help structure understanding and guide implementation of security measures.

1. **Confidentiality, Integrity and Availability (CIA Triad)**

The CIA Triad remains the fundamental conceptual model in cybersecurity:

- Confidentiality: Ensures that sensitive information is accessed only by authorized entities. Mechanisms such as access controls, encryption, and authentication help enforce confidentiality.

- Integrity: Protects information from unauthorized modification, ensuring that data remains accurate and trustworthy. Hashing, digital signatures, and auditing are common integrity tools.

- Availability: Ensures that systems and information are accessible to authorized users when needed. Redundancy, fault tolerance, and disaster recovery strategies promote availability.[2]

Together, these three principles provide a foundation for evaluating and implementing security controls.

2. **Authentication, Authorization and Accounting (AAA)**

AAA is a model governing how user access is controlled:

- Authentication verifies the identity of a user or system component, typically through passwords, biometrics, or cryptographic keys.

- Authorization determines what an authenticated user is permitted to do.

---

[1] Clark DD and Landau S, *Cybersecurity: Toward an Interdisciplinary Approach* (Oxford University Press 2018).

[2] Stallings W, *Effective Cybersecurity: A Guide to Using Best Practices and Standards* (Addison-Wesley 2017).

- Accounting records actions performed by users, enabling auditing and accountability.[3]

AAA is central to access management in secure systems.

3. Defense in Depth

Defense in depth refers to the use of multiple, redundant security measures that protect a system at various level.[4] Rather than relying on a single control, a layered strategy increases resilience; if one control fails, others remain to mitigate risk.

Threat Landscape

Understanding threats is critical to designing effective cybersecurity measures. Threats vary in intent, sophistication, and impact.

Types of Threat Actors

Threat actors include:

- Malicious external actors: Cybercriminals, hacktivists, and state-sponsored groups who seek financial gain, political influence, or disruption.
- Insiders: Employees or contractors who intentionally or inadvertently compromise security.
- Automated threats: Bots and malware that operate without direct human control once released.[5]

Each actor type poses distinct challenges and requires tailored defences.

Common Attack Vectors

Some widely encountered attack vectors include:

- Malware: Software designed to disrupt, damage, or gain unauthorized access.
- Phishing: Social engineering attacks that deceive users into revealing credentials or executing harmful actions.
- Denial of Service (DoS): Attempts to overwhelm systems, making them unavailable to legitimate users.

---

[3] Zwicky ED, Cooper S and Chapman DB, *Building Internet Firewalls* (O'Reilly Media 2000).

[4] Clark DD and Landau S, *Cybersecurity: Toward an Interdisciplinary Approach* (Oxford University Press 2018).
[5] Ibid.

- Man-in-the-middle attacks: Intercepting communications between two parties to eavesdrop or manipulate data.
- Zero-day exploits: Attacks that exploit previously unknown vulnerabilities (Sutton, Greene & Amini, 2018).[6]

Principles of Cybersecurity

While foundational concepts explain what cybersecurity aims to protect, core principles describe how security should be structured and implemented.

1. **Risk Management**

Cybersecurity is fundamentally about managing risk. Risk is defined as the potential for loss when a threat exploits a vulnerability (ISO/IEC, 2018). A risk management approach involves:

- Risk identification: Cataloging assets, threats, and vulnerabilities.
- Risk assessment: Evaluating the likelihood and impact of potential events.
- Risk mitigation: Implementing controls to reduce risk to acceptable levels.
- Monitoring and review: Continuously assessing control effectiveness.

Frameworks such as NIST's Risk Management Framework provide structured approaches for organizations[7].

2. **Security by Design and Default**

Security by design means incorporating security principles into systems from the outset rather than as an afterthought. Systems should be built with robust defaults, minimizing insecure configurations and requiring explicit action to relax security controls.[8] This principle reduces the likelihood that systems are deployed with vulnerabilities.

3. **Accountability and Auditing**

[6] Casey E, *Digital Evidence and Computer Crime* (Academic Press 2011).

[7] National Institute of Standards and Technology, *Framework for Improving Critical Infrastructure Cybersecurity* (2021).

[8] Lipner S, *Security Design and Deployment in a Large Enterprise* (Microsoft Research 2005).

Effective cybersecurity requires mechanisms to track and verify actions. Auditing and logging ensure that activities can be reviewed, investigated, and traced to responsible parties. These mechanisms support compliance and deter negligent or malicious behaviour.

4. Resilience and Recovery

Despite defences, breaches may occur. Resilience—the ability to maintain operations during and after an incident—is therefore crucial. Recovery planning includes backups, incident response teams, and continuity of operations plans. These capabilities reduce downtime and data loss.

Security Controls and Technologies

A variety of security controls and technologies implement cybersecurity principles, ranging from technical tools to policy frameworks.

1. Cryptography

Cryptography protects data confidentiality and integrity through mathematical algorithms. Encryption transforms readable data into an unreadable format without a decryption key. Public key infrastructure (PKI) enables secure key exchange and digital signatures, which verify data origin and authenticity.[9]

2. Firewalls and Intrusion Detection Systems

Firewalls filter network traffic according to defined rules, preventing unauthorized access. Intrusion Detection Systems (IDS) and Intrusion Prevention Systems (IPS) monitor for suspicious activity and can alert or respond automatically.[10]

3. Endpoint Protection

Endpoint protection tools secure devices such as laptops and mobile phones. Antivirus, anti-malware, and endpoint detection and response (EDR) software detect, block, and remediate threats across user devices.[11]

4. Identity and Access Management (IAM)

IAM systems control how users authenticate and access resources. Features often include single sign-on, multi-factor authentication, and role-based access control, increasing both convenience and security.

Challenges and Future Directions

Cybersecurity continues to face evolving challenges.

[9] Hadnagy C, *Social Engineering: The Science of Human Hacking* (Wiley 2018).
[10] Sutton M, Greene A and Amini P, *Fuzzing: Brute Force Vulnerability Discovery* (Addison-Wesley 2018).
[11] Jaquith A, *Security Metrics: Replacing Fear, Uncertainty, and Doubt* (Addison-Wesley 2007).

1. Rapid Technological Change

Technological innovation—such as cloud computing, the Internet of Things (IoT), and automation—increases security complexity. These technologies expand attack surfaces and require new protective strategies.[12]

2. Skill Shortages

Demand for skilled cybersecurity professionals outpaces supply. Educational pathways, certifications, and workforce development efforts aim to address this gap but progress remains incremental.[13]

3. Global Collaboration

Cyber threats often cross-national boundaries, necessitating international cooperation. Agreements on norms of behaviour, information sharing, and joint response mechanisms enhance collective security

Foundational dimensions of cyber security.

1. Governance and Strategic Alignment

Cybersecurity is not solely a technical function; it is a governance issue that must align with organizational strategy. Governance refers to the structures, policies, and decision-making processes that ensure cybersecurity objectives support broader institutional goals. According to Von Solms and Van Niekerk (2013), cybersecurity governance integrates risk management, compliance, and operational controls into executive-level oversight. Without strategic alignment, cybersecurity initiatives often become reactive rather than preventive.[14]

Board-level accountability has increasingly become essential. Senior leadership must define risk tolerance levels, allocate resources, and ensure regulatory compliance. The integration of cybersecurity into enterprise governance frameworks strengthens resilience and reduces fragmented security practices.[15]

2. Cybersecurity Frameworks and Standards

Foundational cybersecurity practices are often structured around recognized frameworks and standards. These frameworks provide systematic approaches to managing risk and implementing controls. For example, the National Institute of Standards and Technology (NIST) Cybersecurity

[12] Conti M, Dehghantanha A, Franke K and Watson S, 'Internet of Things Security and Forensics: Challenges and Opportunities' (2018) 78 *Future Generation Computer Systems*.
[13] Ibid.
[14] Von Solms R and Van Niekerk J, 'From Information Security to Cyber Security' (2013) 38 *Computers & Security*.
[15] Ibid.

Framework organizes security activities into five core functions: Identify, Protect, Detect, Respond, and Recover (NIST, 2021). This lifecycle approach reinforces that cybersecurity is continuous rather than static.[16]

Similarly, the ISO/IEC 27001 standard provides requirements for establishing, implementing, and maintaining an Information Security Management System (ISMS) (ISO/IEC, 2018). The ISMS model emphasizes continuous improvement through a "Plan–Do–Check–Act" cycle, promoting adaptability to evolving threats.[17] These frameworks do not prescribe identical solutions but provide structured guidance adaptable to various industries and organizational sizes.

3. Cybersecurity Architecture and System Design

Secure architecture forms the structural backbone of cybersecurity. Security architecture refers to the design principles and models that define how security controls are integrated within systems. A well-designed architecture ensures that protection mechanisms are embedded into networks, applications, and data flows rather than added after vulnerabilities appear.[18]

One foundational architectural model is segmentation, which divides networks into smaller, isolated sections to limit the spread of attacks. Another is zero trust architecture, which assumes no user or device should be trusted automatically, even if inside the network perimeter (Rose et al., 2020). Under zero trust principles, continuous verification replaces implicit trust.[19]

Secure coding practices also contribute to strong architecture. Developers must anticipate misuse scenarios and design applications that minimize exploitable vulnerabilities. This proactive mindset aligns with the principle of security by design.

4. Incident Response and Digital Forensics

Even with strong preventive controls, security incidents are inevitable. Incident response is therefore a core cybersecurity function. An incident response plan outlines procedures for identifying, containing, eradicating, and recovering from cyber events.

According to Casey (2011), digital forensics plays a critical role in incident response by preserving and analysing digital evidence. Proper forensic procedures ensure that evidence remains admissible in legal proceedings and supports accurate root-cause analysis.[20]

---

[16] National Institute of Standards and Technology, *Framework for Improving Critical Infrastructure Cybersecurity* (2021).

[17] Ibid.

[18] International Organization for Standardization, *ISO/IEC 27000:2018 Information Security Management Systems — Overview and Vocabulary* (ISO 2018).

[19] Craigen D, Diakun-Thibault N and Purse R, 'Defining Cybersecurity' (2014) 4(10) *Technology Innovation Management Review.*

[20] Casey E, *Digital Evidence and Computer Crime* (Academic Press 2011).

A structured incident response capability minimizes operational disruption, reduces financial loss, and strengthens future defensive strategies through lessons learned.

5. Cybersecurity Metrics and Performance Measurement

Measuring cybersecurity effectiveness remains a complex challenge. Unlike physical security, success in cybersecurity is often defined by the absence of incidents. However, quantitative and qualitative metrics can help organizations evaluate their posture.

Jaquith (2007) argues that meaningful security metrics should be consistent, inexpensive to gather, and aligned with business risk. Examples include mean time to detect (MTTD), mean time to respond (MTTR), patch management timelines, and user training completion rates.[21]

Metrics serve two primary purposes: demonstrating accountability and guiding improvement. Without measurement, organizations cannot determine whether security investments produce tangible benefits.

6. Supply Chain and Third-Party Risk

Modern organizations rely heavily on external vendors, cloud providers, and software suppliers. This interconnected ecosystem introduces supply chain risks, where vulnerabilities in third-party systems may expose core organizational assets. Boyson emphasizes that supply chain cybersecurity requires contractual controls, vendor risk assessments, and continuous monitoring. Organizations must evaluate third-party security maturity before integrating services or products into their environment.[22]

Recent high-profile breaches have demonstrated that attackers often exploit weaker suppliers to gain indirect access to larger targets. Therefore, cybersecurity foundations must extend beyond internal boundaries.[23]

Cybersecurity and Privacy Integration

Cybersecurity and privacy, while distinct, are deeply interconnected. Cybersecurity protects systems and data from unauthorized access, while privacy ensures that personal information is collected, processed, and stored responsibly.

Privacy-by-design principles require that data minimization, purpose limitation, and user consent mechanisms be integrated into system architecture (Cavoukian, 2011). Strong cybersecurity

[21] Jaquith A, *Security Metrics: Replacing Fear, Uncertainty, and Doubt* (Addison-Wesley 2007).

[22] Sharma B and Sood SK, 'Fundamentals of Cybersecurity' (2020) 50 *Journal of Information Security and Applications.*

[23] Sandhu R and Samarati P, 'Access Control: Principles and Practice' (1994) *IEEE Communications Magazine.*

safeguards support privacy compliance, but privacy governance also demands transparency and accountability in data usage practices.

Balancing operational needs with individual rights is a central ethical challenge in digital environments.

Ethical Hacking and Security Testing

Ethical hacking, also known as penetration testing, is a proactive strategy used to identify vulnerabilities before malicious actors exploit them. Security professionals simulate attacks under controlled conditions to evaluate system resilience. Penetration testing aligns with the principle of continuous improvement. By identifying weaknesses early, organizations can strengthen controls and reduce exposure. According to Allen (2012), structured testing methodologies improve both technical defences and organizational awareness. Ethical hacking reinforces the concept that security must be regularly evaluated rather than assumed effective.

Education and Capacity Building

Sustainable cybersecurity depends on education at multiple levels. Academic programs, professional certifications, and organizational training initiatives contribute to capacity building. Cybersecurity education must address technical competencies, legal awareness, and ethical reasoning.[24] National strategies increasingly emphasize workforce development to close the cybersecurity skills gap.[25] A well-trained workforce enhances both prevention and response capabilities. Education also extends to public awareness, empowering individuals to practice safe digital behaviour in everyday life.[26]

Conclusion

Cybersecurity is a multifaceted discipline rooted in protecting digital systems and data from unauthorized access, exploitation, and disruption. Its foundations rest on concepts such as confidentiality, integrity, and availability, and core principles including risk management, security by design, accountability, and resilience. Though grounded in technology, successful cybersecurity integrates human behaviour, organizational culture, legal compliance, and ethical practice.[27]

As digital systems continue to evolve and intertwine with every aspect of modern life, the need for robust cybersecurity grows. A comprehensive foundation that blends theory, practice, governance,

[24] Craigen D, Diakun-Thibault N and Purse R, 'Defining Cybersecurity' (2014) 4(10) *Technology Innovation Management Review.*

[25] Von Solms R and Van Niekerk J, 'From Information Security to Cyber Security' (2013) 38 *Computers & Security.*

[26] International Organization for Standardization, *ISO/IEC 27000:2018 Information Security Management Systems — Overview and Vocabulary* (ISO 2018).

[27] Cavoukian A, *Privacy by Design: The 7 Foundational Principles* (Information and Privacy Commissioner of Ontario 2011).

and awareness equips organizations and individuals to navigate an uncertain digital landscape with confidence and resilience.

Expanding the foundation of cybersecurity requires moving beyond technical controls toward an integrated framework that includes governance, architecture, metrics, privacy, supply chain management, and workforce development. Cybersecurity is not a single technology or product but a structured discipline built upon risk management, strategic oversight, and continuous adaptation.

As digital ecosystems become more complex, foundational principles must evolve to address interconnected systems, globalized threats, and regulatory expectations. A comprehensive understanding of these additional dimensions strengthens the theoretical and practical grounding of cybersecurity as an academic and professional field.

Chapter – II

From Threats to Trust: Securing the Digital World

Sonia Das, Assistant Professor, Department of Political Science, St. Xavier's College, Maharo, Dumka.

Abstract

This chapter reinvents cybersecurity roots by leaving the technical definitions and fixed protection models behind. It claims that the source of protection does not only reside in the technological safeguards but also in the socio-technical governance, institutional design, and human behaviour. Modern cyber threats are working in integrated digital ecosystems, whereby states, non-governmental organisations and individuals are all sharing complementary responsibilities. The long-standing ideas of traditional principles of protecting confidentiality, integrity and availability are no longer applicable in such settings to describe protection building or maintenance.

The chapter establishes a critical framework that broadens the fundamental principles to encompass trust, accountability, adaptability, and resilience. Trust is discussed as a structural state which is formed by architecture, control, and social requirements. The analysis of accountability relates to the problem of attribution, legal responsibility, and the problem of governance gaps in transnational cyber activities. Adaptability is given as an active capacity that would allow institutions to move in response to uncertainty and swift changes in technology. The conceptualisation of resilience is the ability to withstand the attacks and in addition, absorb the disruption, recover effectively, and reshape systems to respond to changing risks.

The impact of the organisational culture, economic interests, and political interests on the process of security outcomes is also given attention. The chapter puts emphasis on the fact that systemic vulnerabilities and human factors frequently cause more vulnerabilities compared to individual technical vulnerabilities. It would offer a progressive concept of cybersecurity based on adaptive governance and inclusive digital ecosystems by applying the risk management theory, governance studies, and ethical analysis. This redefinition is in line with other contemporary academic debates that cybersecurity should be viewed as a socio-technical and institutional problem, as opposed to that of only a technological problem. The chapter thus adds to a polished theoretical base of the future research, policies, and strategies in the field of cybersecurity.

Keywords: organisational culture, economic interests, accountability, adaptability, and resilience.

Introduction: Reframing the Roots of Protection

Cybersecurity should be perceived as a dynamic issue of governance, which will not be limited to purely technical control systems. The conventional cybersecurity strategies are centred on perimeter defence, firewalls and intrusion detection systems. Nonetheless, these methods are

required, but inadequate due to the growing complexity of threats that are linked up, along with the wider socio-economic environment in which digital systems are used in. Modern cyberattacks use not only technical vulnerabilities but also socio-technical, which is based on the organisational culture, human behaviour and operational decision-making. The constraints of a narrow technical model are noted by research, and it is found that the systems that cannot unify both social and organisational dynamics are still vulnerable despite the high level of technical protection[28].

This chapter provides the claim that a range of core tenets of cybersecurity needs to change by including governance, trust, accountability, and resilience into its rationale. Governance models construct security due to institutional structures, cross-sector co-operation and regulatory policy as opposed to being merely technology changes. Governance perspective is an acceptance that actors like states, private organisations, and individuals have shared (distributed) roles to secure digital ecosystems, particularly with the growing pattern of interdependence between infrastructures[29].

Conceptual Reorientation of Cyber Security Foundations

From Technical Defence to Socio-Technical Governance

The history of cybersecurity has been a more profound change in isolated technical control systems to socio-technical models of governance that are aware of interdependencies between technology, organisations and society. The classical security models that are based on perimeter defences like firewalls and antivirus software were based on the premise that the source of risks was external to a well-delineated border. However, the current digital worlds are dynamic, networked and can be permeable, i.e. it misses external demarcation of threats, but may occur within the socio-technical systems themselves. Scholars allege that cybersecurity will have to be redefined to help in tackling this complexity by combining social, organisational, economic, and political factors with technical interventions[30].

28 Ceur, Towards operationalizing cyber resilience - a socio-technical analytical framework (2026) Accessed on 6th March 2026: https://ceur-ws.org/Vol-4134/paper7.pdf

29 Hoong, Yang, and Davar Rezania. "Navigating cybersecurity governance: the influence of opportunity structures in socio-technical transitions for small and medium enterprises." *Computers & Security* 142 (2024): 103852.

30 Ceur, Towards operationalizing cyber resilience - a socio-technical analytical framework (2026) Accessed on 6th March 2026: https://ceur-ws.org/Vol-4134/paper7.pdf

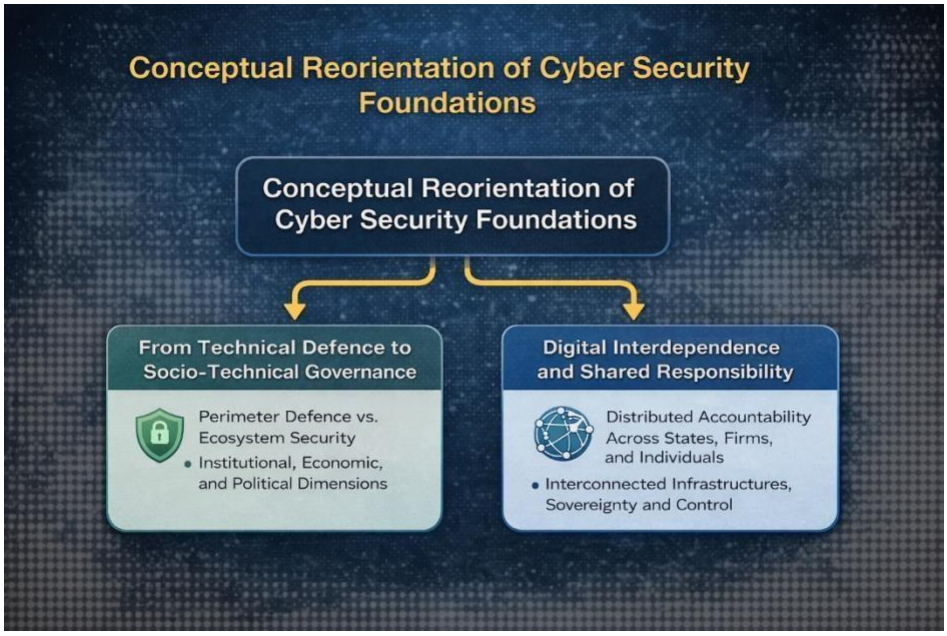

Figure 1: Cyber security conceptual reorientation

(Source: Self-Created)

A socio-technical system of governance is aware that technology and human actors share the same success in terms of security. Indicatively, cyber resilience models incorporate adaptive capabilities, including anticipation, recovery and learning in the system design with the focus placed on the co-optimisation of technical and human factors. This way of considering security is both as avoiding breaches and as the capacity to continue operation during periods of stress, a feature which in many old technical models is frequently ignored[31]. In addition, the models of governance emphasise institutional and regulatory frameworks that affect the way organisations distribute resources, agenda-setting, and reaction to changing threats. Investing decisions and risk acceptance are influenced by institutional influences, economic factors and legal rules, i.e. cybersecurity efficiency is as much dependent on governance decisions as on technological initiatives.

Adaptive governance also means that rulemaking and implementation of policies involve various stakeholders, such as the government, firms, and civil society[32]. This democratic governance is indicative of the fact of inter-relating digital ecosystems, where a failure in any part can spill over into widespread social or economic instabilities.

Digital Interdependence and Shared Responsibility

The concept of responsibility in cybersecurity goes further than the individual organisations to the networks of actors, as critical digital infrastructures are becoming more and more linked. Digital

31 Ceur, Towards operationalizing cyber resilience - a socio-technical analytical framework (2026) Accessed on 6th March 2026: https://ceur-ws.org/Vol-4134/paper7.pdf

32 Melaku, Henock Mulugeta. "A dynamic and adaptive cybersecurity governance framework." *Journal of Cybersecurity and Privacy* 3, no. 3 (2023): 327-350.

interdependence implies that a failure in one of the subsystems can be system-wide, so collective defence is necessary. Here, accountability is shared between the states, private businesses, and individual users[33]. Everyone has a contribution towards the process of establishing digital security and resilience. The research in cyberspace regulation confirms that the indefinite presence of security gaps exists because regulatory environments are disjointed and enforcement systems are inconsistent without integrated policy frameworks and multistakeholder interactions[34].

The notion of distributed accountability also transforms the conceptions of sovereignty and control in cyberspace. The conventional concept of national sovereignty presupposes territorial states and a central government. Conversely, infrastructures based on networks cross these boundaries, making it challenging to have state control and creating cooperative rules of governance.

Reconstructing Core Principles Beyond Traditional Models

Limits of Classical Security Principles

The classical concept of security that is based on the model of the static perimeter defence and deterministic control logics is becoming unsuitable in cloud-native, AI-enhanced, and platform ecosystem environments. The traditional "castle and moat" design presupposes definite borders and prescribed areas of trust, but digital space in the modern world has no boundaries. The cloud infrastructures are distributed across distributed services, microservices, and hybrid architecture, and there is a challenge to adhere to the static policies determined by network location[35]. The classical models rely on established rules and signatures that do not work with dynamic threats and new attack patterns that utilise AI-assisted exploitation and intra-service movement. Recurring vulnerability reporting and rule sets to older network arrangements will fail to keep up with the continuous integration and deployment, scale out, and on-demand resource provisioning found in cloud environments. Such incompatibility creates loopholes that are manipulated by advanced persistent threats to cause breaches that are not detected until a lot of damage has been done. With AI-based systems, attackers can tamper with the models of learning or can feed the models with poisoned data as a tactic to avoid detection.

33 Abrahams, Temitayo Oluwaseun, Oluwatoyin Ajoke Farayola, Simon Kaggwa, Prisca Ugomma Uwaoma, Azeez Olanipekun Hassan, and Samuel Onimisi Dawodu. "Cybersecurity awareness and education programs: a review of employee engagement and accountability." *Computer Science & IT Research Journal* 5, no. 1 (2024): 100-119.

34 Qudus, Lawal. "Cybersecurity governance: Strengthening policy frameworks to address global cybercrime and data privacy challenges." *International Journal of Science and Research Archive* 14, no. 1 (2025): 1146-1163.

35 Anh, Nguyen Hoang. "Hybrid cloud migration strategies: balancing flexibility, security, and cost in a multicloud environment." *Transactions on Machine Learning, Artificial Intelligence, and Advanced Intelligent Systems* 14, no. 10 (2024): 14-26.

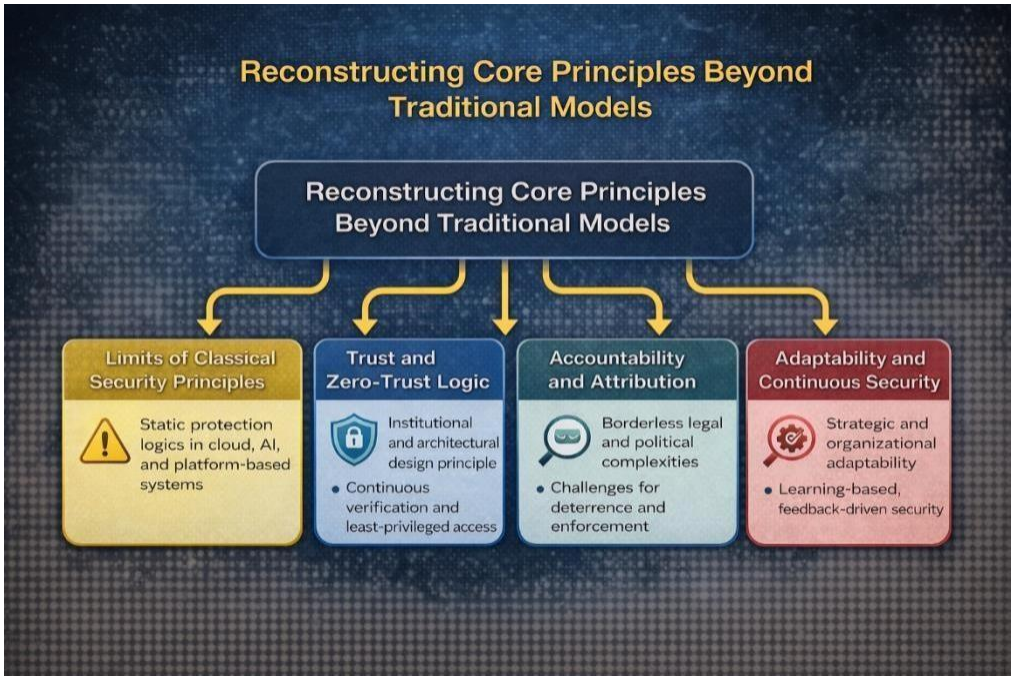

Figure 2: Reconstructing cybersecurity core principles

(Source: Self-Created)

Trust and Zero-Trust Logic as Governance Mechanisms

The implicit trust of traditional cybersecurity has been founded on the idea that users and devices within a perimeter are safe by default. By way of contrast, Zero Trust logic represents a principle of governance in which there is no automatic granting of trust but rather constant checking with respect to identity, context, and behaviour. Zero Trust philosophy, known as never trust, always verify, is becoming known as a strategic concept in the governance of security policies in digitally diverse environments, particularly in cloud and hybrid environments where former boundaries are melted away. Zero Trust must have continuous authentication of access, the least rights of access, and dynamic policy resolution of each access, irrespective of its source. This strategy changes the concept of trust not as a fixed quality but as an institutional and architectural design choice that is embedded in all the interactions. It drives organisations to the assumption that any element, user, or relationship can be hacked and implement policy choices that decrease implicit trust on the network. In this manner, Zero Trust minimises attack points and limits the horizontal movement, which third-party traditional models tend to overlook, providing vast access once the initial perimeter defence is compromised[36].

Accountability and Attribution in a Borderless Environment

The concept of accountability in cybersecurity requires the ability to trace malicious practices back to a recognised actor, and this is difficult due to the borderless and obscured nature of cyberspace.

36 Nvlpubs,: NIST Special Publication 800-207 (2020) https://nvlpubs.nist.gov/nistpubs/specialpublications/NIST.SP.800-207.pdf

Attribution, the process of identifying the origin of a particular cyber incident, is technically, legally, and politically difficult due to the use of anonymity, proxies, and the spread of infrastructure across the world[37]. The current cyber activities tend to cross borders, are nonstate activities and are proxies sponsored by states, and hence the lines of responsibility are blurred.

Adaptability and Continuous Security

Adaptability in cyber security focuses on strategic and organisational capability to adjust defence mechanisms to emerge threats. When the enemies keep on innovating, then the use of static controls becomes inadequate. Adaptive security is a dynamic control of policies driven by feedback, real-time observation and data analytics learnt by observing the behaviour of the system and threat information[38]. This model empowers systems to identify concept drift in patterns of threat and optimally retrain defensive mechanisms as environments evolve, ensuring that the detection accuracy remains and the resilience is maintained. Constant security is just a constant exposure to threats, constant presence of anomalies, and dynamism in policy implementation. There are examples of how learning-based frameworks can be more effective than just constant rule sets. Adaptive security conceptualises threats and defence as a coevolution process, where risk contextualised decision-making and intelligence integration across environments is a constant process[39].

Human Agency, Incentives and Systemic Vulnerability

Organisational Behaviour and Security Culture

Human behaviour in organisations is a core factor that determines the results of cybersecurity since it frequently influences the success or failure of these technical systems. Organisational culture implies a set of beliefs, values, and behaviours that affect the perception and relations of employees to cybersecurity policies and practices[40]. Empirical evidence indicates that when an organisation has a good security culture, meaning employee perception of security as part of their job and mission, compliance is generated and behaviour risks associated with decision making biases or shortcuts are minimised. To illustrate, workers can disregard security controls by responding to them as heavyweight or not suitable to the productivity requirements, thus contributing to the

37 Gul, Seema, Wasmiya Malik, and Gohar Masood Qureshi. "Cybersecurity And Sovereignty: The Role Of International Law In Governing State Behaviour In Cyberspace." *Policy Journal of Social Science Review* 3, no. 5 (2025): 121-135.

38 Farooq, M. and Khan, M.H., 2024. AI-driven network security: innovations in dynamic threat adaptation and time series analysis for proactive cyber defense. *Int. J. Wirel. Microwave Technol.(IJWMT)*, *14*(2), pp.17-26.

39 Iyera, CA Vishwanathan, and Nilesh P. Gokhaleb. "Artificial Intelligence, Deepfakes, and Corporate Intelligence: A Systematic Literature Review on Ethical and Strategic Implications for Business Transparency and Decision-Making." *Transforming Financial Management with AI, BI, and Data-Driven Decision Making* (2026): 67.

40 de Bruin, Marten, and Konstantinos Mersinas. "Individual and Contextual Variables of Cyber Security Behaviour--An empirical analysis of national culture, industry, organisation, and individual variables of (in) secure human behaviour." *arXiv preprint arXiv:2405.16215* (2024).

heightening vulnerability. This is symptomatic of behavioural risk when convenience of preference or mental shortcuts are put ahead of security. Preparing resilience through such an investment in adaptive security awareness programmes and the integration of security norms into everyday actions makes secure behaviour a matter of practice instead of choice[41].

Organisational security culture is strongly determined by leadership. Leaders with cyber security as their priorities instil a sophisticated environment whereby secure behaviour is acknowledged, justified, and aligned towards the business goals. On the other hand, leaders who consider cyber security an IT problem but not a strategic priority make employees have lax approaches to complying. Active leadership that sets expectations, incorporates security into the performance systems, and journalizes secure behaviour will contribute to increased risk awareness and a decrease in security fatigue, the level of exhaustion due to the continuous pressure to practise security that results in appropriation and slackness. Human factors research points to the fact that security culture can never be elicited solely through policy, but it needs to be engaged, trained and reinforced with leadership and governance arrangements.[42]

Cognitive biases that are associated with behavioural risks include overconfidence, routine behaviour and underestimation of threats. Such biases affect such decisions such as reusing passwords or disregarding suspicious warnings and reveal the weaknesses in the system.[43]

41 Iyer, Shankar Subramanian, and Brinitha Raji. "Cybersecurity culture and organizational resilience: A humancentered approach to digital risk management." *American Journal of Industrial and Business Management* 15, no. 5 (2025): 748-766.

42 Iyer, Shankar Subramanian, and Brinitha Raji. "Cybersecurity culture and organizational resilience: A humancentered approach to digital risk management." *American Journal of Industrial and Business Management* 15, no. 5 (2025): 748-766.

43 Darley, Hanah-Marie. "Dangers of succumbing to bias in cyber security: An evaluation of the impact of cognitive biases on threat assessments and cyber security strategies." *Cyber Security: A Peer-Reviewed Journal* 6, no. 3 (2023): 211-219.

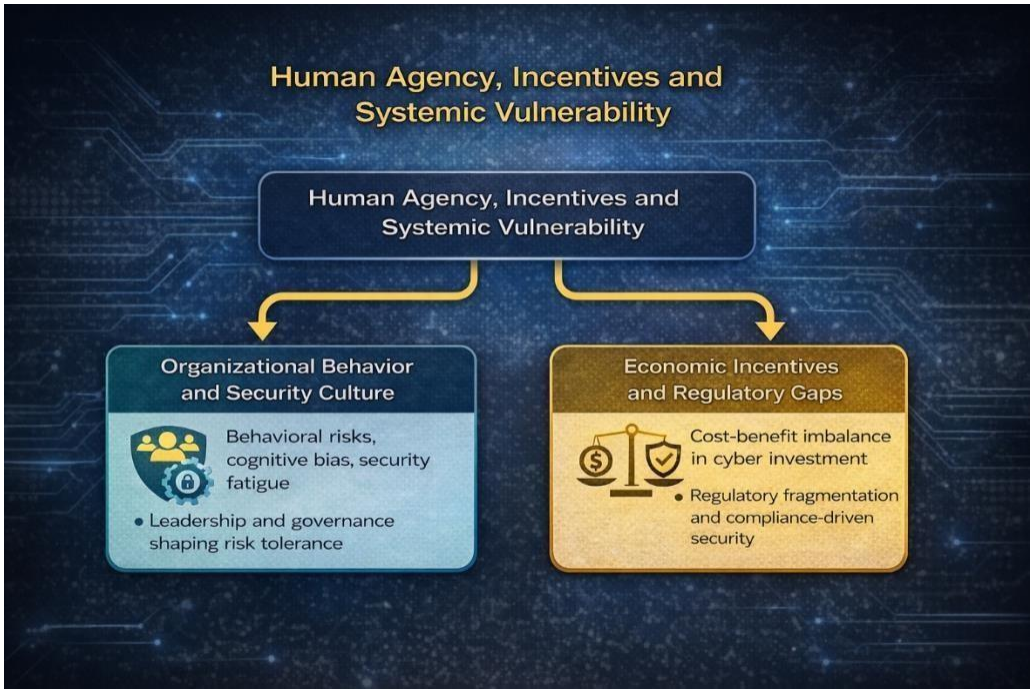

Figure 3: Human agency in cyber security dynamics
(Source: Self-Creaed)

Economic Incentives and Regulatory Gaps

Investments in cyber security are typically influenced by economic reasons and legislative conditions that may contribute to or create roadblocks in terms of organisational devotion to sound protective endeavours. The essence of the cost-benefit imbalance in the decision-making on cyber security lies in the fact that organisations consider the expense of security investments now and possible losses that are not present yet but might happen due to a breach[44].

The regulatory frameworks affect the organisational behaviour to determine the incentives towards compliance and investment. Formulated regulations can influence risk management practises to improve through creating obligations towards incident reporting, risk assessments, and minimum-security standards. Nonetheless, regulatory fragmentation through practice across jurisdictions and industries usually gives rise to unequal expectations and compliance costs, which are relatively more burdensome to smaller organisations. Disjointed regulatory systems may cause either duplication of needs or lack of enforcement, thus decreasing the overall efficiency of the policy incentives. Such mesquite may undermine group cyber resilience by enabling opponents to take advantage of jurisdictions that have looser enforcement.[45]

44 Shaikh, Faheem Ahmed, and Mikko Siponen. "Information security risk assessments following cybersecurity breaches: The mediating role of top management attention to cybersecurity." *Computers & Security* 124 (2023):

102974.

45 Iyer, Shankar Subramanian, and Brinitha Raji. "Cybersecurity culture and organizational resilience: A humancentered approach to digital risk management." *American Journal of Industrial and Business Management* 15, no. 5 (2025): 748-766.

Risk, Resilience and the Evolution of Protection Logic

Risk Management Under Uncertainty

Cyber security risk management commonly uses probabilistic models, estimating probability and the effects of threats, which are constrained in difficult digital systems (where uncertainty is common, interdependence is possible, and enemy behaviours are unpredictable). Conventional probabilistic models assume that previous data may be relied upon to forecast any form of risk, yet cyber threats change fast, and the previous trends cannot affect the study of the fresh attack vectors[46]. Additionally, the major variables like zero-day vulnerabilities, developing technology and human factors cause in-depth uncertainty that is a challenge to the statistical prediction and deterministic risk assumptions. Existing studies on cyber governance point to the fact that the risk management approach is usually deployed by organisations to regulate uncertainty but is potentially inadequate due to the tendency to view risk as a predictable aspect, as opposed to an outcome of socio-technical interactions[20].

Complex digital environments have cascading failures and systemic risk, where a failure in some component of a network spreads to connected components and has an amplified impact compared to its original cause. Cascading failures. Unlike in non-networked systems, in a networked system, a failure in one component raises the risk of failure of a second component[47]. The large-scale effects of risk are amplified in critical infrastructures like energy grids, finance, or supply chains, which initiates the small domino effect of isolated probabilistic models in that interdependency is not considered. The systemic risk arises not due to individual vulnerabilities but because of the collective behaviour of networked elements, and should be treated in a holistic manner, not in a fractured risk analysis manner.

Resilience as a Foundational Principle

Cyber resilience can be defined as the capacity of a system to receive, adapt to, and recuperate negatively affecting cyber events without disrupting vital operations and restoring operations rapidly. From a literary perspective, resilience is the ability to sustain critical activities in the presence of disruption and to reestablish the quality of the system after the event, thus resilience shifts the emphasis of prevention into an extended functionality in the face of stress.

---

46 Mallick, Md Abu Imran, and Rishab Nath. "Navigating the cyber security landscape: A comprehensive review of cyber-attacks, emerging trends, and recent developments." *World Scientific News* 190, no. 1 (2024): 1-69. [20] Abrardi, Laura, Stefano Comino, and Sasha Grassini. "The economics of cyber risk: a survey of the literature." *Journal of Industrial and Business Economics* (2025): 1-35.

47 Karch, Benjamin, Titus Gray, and Michael T. Rowland. *Field Programmable Gate Array-Based Reactor Protection Systems and Potential for Inclusion of Secure Elements to Improve Cybersecurity*. No. SAND--202414666R. Sandia National Laboratories (SNL-NM), Albuquerque, NM (United States), 2024.

The fundamental idea of resilience is the capacity to respond to the largest of shocks with only minor harm to the system (absorptive capacity) and rearranging itself to fit the shifting threat environments (adaptive capacity). Companies that develop resilience as an organisational philosophy embrace pre-emptive threat detection, redundancy and recovery mechanisms that facilitate a rapid response and learning[48].

Ethical and Strategic Foundations for the Future

Security, Privacy and Civil Liberties

Growth in cybersecurity capabilities has led to a heightening confrontation between the national security interests and the safeguarding of privacy and civil liberties. An increase in defensive capacity is supported by surveillance technologies, data retention requirements, and algorithmic surveillance tools, though it raises the issue of proportionality and the rights of the individual. Unmonitored security operations are a threat to social confidence and democratic legitimacy[49].

Transparency, Proportionality and Digital Solidarity

New ideas of cyber regulation in the future focus on transparency, proportionality, and digital solidarity. Exposure embodies effective communication concerning security practices, risks and accountability measures. Proportionality is used so that the protection measures are commensurate with the scope and characteristics of threats. Digital solidarity represents the shared digital security responsibility between involved states, the real actors, and civil society[50]. Governance structures that bring on board different stakeholders can empower legitimacy, propagate confidence, and equity in securing similar administration results in interdependent societies.

Conclusion: Re-Rooting Protection in Adaptive Governance

By considering socio-technical complexity that affects cybersecurity as an adaptive governance issue, this chapter has relegated the notion of technical control in efforts to resolve the issue. It has demonstrated that shared protection logics are not enough in the interconnected digital ecosystems. Enhanced general principles - trust, accountability, adaptability and resiliency, offer a stronger model of dealing with systemic risk and uncertainty. Both human behaviour and economic motivations, and regulatory frameworks have a great impact on vulnerability and response capacity.

48 Qudus, Lawal. "Resilient systems: building secure cyber-physical infrastructure for critical industries against emerging threats." *Int J Res Publ Rev* 6, no. 1 (2025): 3330-46.

49 Dowling, Melissa-Ellen. "Cyber information operations: Cambridge Analytica's challenge to democratic legitimacy." *Journal of Cyber Policy* 7, no. 2 (2022): 230-248.

50 Villani, Susanna. "The Cyber Solidarity Act: Framework and Perspectives for the New EU-Wide Cybersecurity Solidarity Mechanism Under the EU Legal System." *European Journal of Risk Regulation* 16, no. 2 (2025): 485497.

Chapter – III

Gendered Cyber Insecurity: Between Legal Architecture and Lived Reality

Fazle Wakil, Research Scholar, Department of International Relations, Jadavpur University and

Anneshya Sanyal, Research Scholar, School of International Relations and Strategic Studies, Jadavpur University

Abstract

Cybersecurity, a foundation stone of the digital age, affects individuals across all demographics, yet its connection with gender is often unnoticed. The surge in importance of cybersecurity has not diminished the ever-present obstacles related to gender equality and the underrepresentation of women in this field. From online harassment to cyber-stalking, from cyber-bullying to hate speech, from threats to image-based abuse, from revenge pornography to discrimination, women and other marginalised genders experience heightened risk of cybersecurity. The gendered approach to cyber security is interconnected with the other aspects; complexities and needs of people based on gender, religion, race and sexual orientation. Gender-based violence is not a new phenomenon, but its migration to the digital world has magnified its scale and impact. The digital age has transformed how people interact, communicate, and conduct business. While the internet has provided numerous benefits, it has also given rise to new forms of crimes, particularly cybercrimes against women. These crimes not only invade the privacy of women but also threaten their safety, dignity, and mental well-being.

The rapid revolution in the IT industry, with a cornerstone of Artificial Intelligence, has affected the lives of individuals at large. People are more connected and engaged with each other on online platforms compared to physical interactions. These developments have assisted in some instances, but put privacy in danger. Women fear giving any of their details on any online platform due to the lack of security and privacy. Gender and cybersecurity are two correlated terms which have gained popularity in recent years. Privacy and security are still questioned in this era. Despite global progress and increased representation of women across various fields, this issue remains a major setback, limiting potential advancements in cybersecurity as a crucial societal factor. This imbalance deprives the field of talent and weakens its capacity to address new threats. In the field of cybersecurity specifically, according to figures from the International Information Systems Security Certification Consortium, that are based on a study carried out in 2019, women who work in cybersecurity represent only 24% of the total workforce globally.

Cybercrime against women has emerged as a critical concern within the United Nations system, particularly through the work of the International Telecommunication Union (ITU). Since its establishment in 1865, the ITU has addressed evolving technological risks from telegraphy to the internet age recognizing cybersecurity as foundational to a connected society. The proliferation of malware, cyber terrorism, and organized digital crime has intensified vulnerabilities, disproportionately affecting women online. The rise of cybercrimes against women in India has

also prompted the development of a multi-layered legal framework. The Information Technology Act, 2000 criminalizes hacking, identity theft, and electronic stalking, while enabling specialized cybercrime investigation cells. The Indian Penal Code has been amended to address voyeurism, cyberstalking, and non-consensual dissemination of intimate material. Additionally, the Protection of Women from Domestic Violence Act, 2005 extends protection to online harassment within abusive relationships. Although this framework is normatively robust, enforcement deficits, limited institutional capacity, and low awareness among women continue to constrain effective access to justice and meaningful digital protection.

Despite this layered statutory architecture, the paper argues that the gap between law on the books and law in action remains significant. The existence of multiple statutes and institutional mechanisms inevitably raises a critical question: have these legal interventions meaningfully reduced cybercrimes against women, or are such offences continuing to escalate in scale? The available trends suggest that reported cases are rising year after year which partly due to increased awareness and reporting, but also because digital dependence has deepened across every sphere of life. In an era where education, employment, governance, finance, and even intimate relationships are mediated through technological platforms, disengagement from digital systems is neither feasible nor desirable. This structural dependence renders women simultaneously visible, accessible, and vulnerable in unprecedented ways. Also while dealing against cybercrime as a whole, the inquiry must move beyond statutory adequacy to assess whether institutions possess technological capacity, trained personnel, and gender-sensitive protocols, while critically examining accountability when violations arise externally or within power structures, complicating the assumption of the state as an unequivocal protective actor.

This paper thus moves beyond formal legal analysis to examine patterns of incidence, enforcement capacity, infrastructural preparedness, and systemic accountability. By situating cybercrime within broader structures of digital governance and gendered power relations, it seeks to evaluate whether women's digital safety is substantively secured or remains an aspirational promise within an increasingly technologized society. Finally, the paper evaluates the effectiveness of existing cybersecurity frameworks themselves. Cybersecurity often operates reactively rather than pre-emptively, and rapid technological innovation frequently outpaces regulatory adaptation. The study therefore assesses whether current cybersecurity mechanisms function as meaningful deterrents or merely as post-facto response systems.

Keywords: *Privacy, Information Technology Act, Gendered Cyber Violence, Digital Governance, Institutional Capacity, Women's Online Safety, International Telecommunication Union*

Introduction – Gendering Cybersecurity

The expansion of the internet and digital technologies has fundamentally transformed the architecture of contemporary social life. Digital platforms now shape the way individuals communicate, access information, conduct economic activities, and participate in political and cultural processes. From education and governance to commerce and interpersonal interaction, the digital sphere has become deeply embedded within everyday life. Within this rapidly evolving

technological landscape, cybersecurity has emerged as a critical mechanism for safeguarding digital infrastructures and protecting individuals from malicious attacks, data breaches, and unauthorized access.[51] In its broadest sense, cybersecurity refers to the set of technological, legal, and institutional measures designed to protect computers, networks, servers, devices, and data from digital threats. However, while the digital revolution has generated unprecedented opportunities for connectivity and empowerment, it has also produced new vulnerabilities that affect users in unequal ways. The risks associated with cyber threats do not operate in a social vacuum. Instead, they intersect with existing structures of inequality embedded within societies. Among the groups most affected by these vulnerabilities are women, who frequently encounter disproportionate exposure to online harassment, privacy violations, and digital exploitation. Gender biases, patriarchal social norms, and structural inequalities often shape women's experiences in digital environments, making cybersecurity not only a technical issue but also a deeply social and political one.[52] As the digital domain increasingly functions as a site of social interaction, economic participation, and political expression, the absence of gender-sensitive cybersecurity frameworks raise important concerns regarding safety, equality, and access.

The relationship between gender and cybersecurity has therefore gained increasing scholarly and policy attention in recent years. Digital spaces often replicate and, at times, intensify existing social hierarchies. Women's participation in online environments is frequently shaped by cultural expectations, discriminatory attitudes, and institutional limitations that influence their access to technology as well as their ability to use it safely.[53] A gendered understanding of cybersecurity recognizes that digital threats are experienced differently across social groups. It draws attention to the ways in which vulnerabilities are influenced by intersecting identities such as gender, race, religion, sexual orientation, and socio-economic status. Consequently, addressing cybersecurity challenges requires a broader analytical lens that acknowledges how online harms disproportionately affect certain communities.[54]

Online harassment represents one of the most visible manifestations of gendered insecurity within digital environments. As digital communication platforms expand, they have simultaneously become spaces where abusive and discriminatory behaviour can proliferate with relative anonymity. The Dart Centre for Journalism and Trauma defines online harassment as any unwanted verbal or non-verbal behaviour that occurs in digital spaces and violates the dignity of an individual by creating a hostile, degrading, or offensive environment. This definition highlights the psychological and emotional consequences of cyber abuse, demonstrating that online harassment is not merely an inconvenience but a serious violation that can affect an individual's well-being, reputation, and sense of safety.[55] Women frequently encounter a range of abusive practices in online environments, including threats, trolling, hate speech, doxxing, cyberstalking, and other forms of targeted harassment. These practices are often rooted in misogynistic attitudes that seek to intimidate or silence women who participate in public discourse. Social media platforms, online

[51] Wall DS, *Cybercrime: The Transformation of Crime in the Information Age* (Polity Press 2007)
[52] Yar M and Steinmetz KF, *Cybercrime and Society* (3rd edn, Sage Publications 2019)
[53] Kshetri N, *Cybercrime and Cybersecurity in the Global South* (Palgrave Macmillan 2013)
[54] Singer PW and Friedman A, *Cybersecurity and Cyberwar: What Everyone Needs to Know* (Oxford University Press 2014)
[55] Ibid.

forums, and digital communication channels have increasingly become arenas where gender-based hostility manifests through coordinated attacks and derogatory commentary. For many women, navigating digital spaces involves constant negotiation between the benefits of online participation and the risks of exposure to harassment.[56]

The broader context of cybercrime further underscores the significance of these challenges. Cybercrime has become a global phenomenon, affecting millions of individuals each year. Statistical estimates suggest that cybercrime occurs at an alarming frequency worldwide, with new victims emerging every few seconds. These figures illustrate the scale of the problem and highlight how deeply cyber risks have become embedded within modern societies. As the information technology sector continues to evolve particularly with the integration of artificial intelligence, machine learning, and data-driven algorithms the digital ecosystem has become increasingly complex.[57] Individuals now rely on digital platforms not only for communication but also for financial transactions, professional activities, and personal relationships. This rapid technological transformation has led to a paradoxical situation in which increased connectivity coexists with heightened vulnerability. While digital technologies enable individuals to build networks, share knowledge, and access opportunities, they also expose users to surveillance, data exploitation, and identity-related risks. Personal information shared online can be stored, analysed, and potentially misused by a range of actors, including corporations, governments, and cybercriminals. Privacy concerns have therefore become central to contemporary debates about digital governance and technological ethics.[58]

The tension between technological advancement and privacy protection has long been recognized. In 1999, Scott McNealy, the then Chief Executive Officer of Sun Microsystems, famously remarked that "privacy is dead—get over it." Although the remark was controversial, it reflected a growing recognition that digital technologies were fundamentally altering traditional notions of privacy. Over the past two decades, the proliferation of social media platforms, mobile applications, and data-driven services has intensified these concerns. The widespread collection and monetization of personal data have blurred the boundaries between public and private life, raising important questions about who controls digital information and how it is used. For women navigating digital spaces, these developments often translate into heightened anxieties about safety and personal security.[59] Many women remain cautious about sharing personal information online due to fears of harassment, identity theft, stalking, or misuse of their data. The lack of effective institutional safeguards further amplifies these concerns. In many contexts, women perceive digital platforms as environments where privacy protections are fragile and where abusive behaviour may occur with limited accountability. As a result, the digital sphere can simultaneously function as a site of empowerment and a space of vulnerability. Research increasingly indicates that women encounter cybersecurity threats more frequently than men, particularly in the form of gender-based

[56] Rid T, *Cyber War Will Not Take Place* (Oxford University Press 2013)
[57] Rid T, *Cyber War Will Not Take Place* (Oxford University Press 2013)

[58] Brenner SW, *Cybercrime and the Law: Challenges, Issues and Outcomes* (Northeastern University Press 2012)

[59] Ibid.

harassment.[60] One of the key factors contributing to this disparity is the absence of gender-responsive policy frameworks within cybersecurity governance. Traditional cybersecurity approaches have primarily focused on protecting technological infrastructures such as networks, databases, and information systems. While such protection is essential, it often overlooks the human dimension of cyber threats, including the ways in which social inequalities shape digital risks.

The neglect of gender-specific concerns within cybersecurity policy is reflective of broader patterns in which issues related to privacy, safety, and bodily autonomy are marginalized within discussions of technological development. Debates about equality and justice frequently emphasize political representation or economic participation, yet the digital security of women remains insufficiently addressed.[61] In an era in which large segments of social interaction occur online, the absence of robust protections against digital harassment has significant implications for women's ability to participate freely in digital environments. Another factor that exacerbates women's vulnerability in digital spaces is the persistent digital literacy gap. Access to technological knowledge and cybersecurity awareness remains unevenly distributed across societies. In many regions, women face barriers to technological education and digital training, limiting their ability to recognize and respond to cyber threats.[62] Without adequate knowledge of privacy settings, secure communication practices, or data protection strategies, users may unknowingly expose themselves to exploitation. Addressing this gap requires targeted initiatives aimed at strengthening digital literacy, particularly among women and marginalized communities.

The gender disparities evident in contemporary digital spaces are also rooted in historical developments within the technology sector itself. Contrary to common perceptions, women played an important role in the early development of computing. During the early decades of computing history, programming was often regarded as routine clerical work and was performed largely by women. In fact, computing was sometimes described as a "pink-collar profession," reflecting the significant presence of women within the field. Female programmers and mathematicians made substantial contributions to technological innovation, particularly during the mid-twentieth century. Women's participation was especially visible during the Second World War, when female codebreakers and analysts played critical roles in decoding encrypted communications and supporting military intelligence operations. Their analytical skills and technical expertise were essential to wartime cryptographic efforts.[63] Despite these contributions, the professionalization of the technology sector gradually marginalized women's roles within computing. As computing became associated with higher levels of prestige, economic power, and technical specialization, gender stereotypes increasingly shaped hiring practices and workplace cultures. Women's expertise was often undervalued, and many were discouraged from pursuing long-term careers in science and technology. Over time, this shift contributed to the gender imbalance that continues to

[60] Singer PW and Friedman A, *Cybersecurity and Cyberwar: What Everyone Needs to Know* (Oxford University Press 2014)

[61] Brenner SW, *Cybercrime and the Law: Challenges, Issues and Outcomes* (Northeastern University Press 2012)

[62] Ibid.

[63] Deibert R, *Black Code: Surveillance, Privacy and the Dark Side of the Internet* (Signal Books 2013)

characterize the technology sector today. The cybersecurity industry, in particular, remains heavily male-dominated. Global estimates suggest that women constitute only a small proportion of the cybersecurity workforce. Although this figure has gradually increased over time, the gender gap remains substantial.

The limited representation of women within cybersecurity has important consequences for the development of digital security frameworks. Diversity within technological fields is crucial for addressing complex challenges, as different perspectives and experiences contribute to more comprehensive solutions. When women are underrepresented in decision-making roles, the specific vulnerabilities they face may remain overlooked within policy and technological design. The absence of female role models and mentors also discourages younger generations of women from pursuing careers in cybersecurity, perpetuating a cycle of underrepresentation. At the same time, the expansion of digital platforms has intensified various forms of cyber violence directed at women.[64] Online harassment, misogynistic abuse, and targeted hate campaigns have become increasingly visible in social media environments. Surveys conducted in different regions suggest that a significant proportion of women have experienced some form of online harassment. Female journalists, activists, academics, and public figures often face coordinated digital attacks designed to silence their voices and discourage participation in public debate.

Cyberstalking represents another significant dimension of gendered cyber insecurity. This form of abuse involves persistent monitoring or harassment through digital technologies, including unauthorized access to personal accounts, spyware installations, and location tracking through GPS-enabled devices. In many cases, cyberstalking occurs within intimate relationships, where abusive partners use technology as a tool of control. Digital surveillance can restrict victims' autonomy, undermine their sense of safety, and escalate into real-world threats. Non-consensual dissemination of intimate images has also emerged as a pervasive form of cyber violence. The unauthorized sharing of private photographs or videos—often referred to as revenge pornography—can have devastating consequences for victims. Women subjected to such violations frequently experience emotional trauma, social stigma, and reputational damage. In certain contexts, victims may face exclusion from professional or social networks, and in extreme cases they may even encounter rejection from their own families. Despite the severity of these harms, legal frameworks in many jurisdictions remain inadequate for addressing such offences effectively. Another increasingly prevalent practice is doxxing, which involves the public release of personal information such as home addresses, contact details, or workplace information without consent. By exposing victims' private data, perpetrators create conditions that facilitate harassment, intimidation, and physical threats. Women who advocate for gender rights or challenge patriarchal norms are particularly vulnerable to such attacks, which are often intended to silence dissent and discourage activism.

These patterns of abuse reveal how patriarchal norms and gender biases are reproduced within digital environments. Online misogyny often reflects broader cultural attitudes that seek to police women's behaviour and restrict their participation in public spaces. The normalization of such behaviour within digital cultures can discourage victims from reporting abuse, particularly when

[64] Nissenbaum H, *Privacy in Context: Technology, Policy and the Integrity of Social Life* (Stanford University Press 2010)

institutional responses are slow or ineffective. The consequences of cyber violence extend beyond immediate emotional harm. Persistent harassment can affect women's professional opportunities, economic independence, and willingness to engage in digital platforms that are increasingly essential for career development. Female entrepreneurs, journalists, and professionals often rely on digital tools for networking, communication, and business promotion. However, exposure to cyber harassment, identity theft, or privacy violations may discourage them from maintaining an active online presence. Women in developing regions face additional vulnerabilities due to limited access to technological resources and cybersecurity awareness. Structural inequalities in education, income, and digital infrastructure often restrict women's ability to benefit from technological advancements while simultaneously increasing their exposure to cyber risks. In such contexts, the digital divide intersects with gender inequality, reinforcing patterns of marginalization.[65] These dynamics highlight the importance of examining cybersecurity not only as a technological challenge but also as a social phenomenon shaped by power relations, institutional structures, and cultural norms. Understanding the gendered dimensions of cyber insecurity requires a broader analytical framework that situates digital threats within the context of structural inequality and social transformation. Such an approach opens the way for deeper exploration of how online violence operates, how institutional responses have evolved, and how gender-sensitive cybersecurity frameworks can be developed to address these emerging challenges in the contemporary digital era.[66]

Cybersecurity Governance and the Evolving Landscape of Global Digital Threats

The rapid digitalization of contemporary societies has transformed data, digital infrastructures, and interconnected networks into central components of economic activity, governance, and everyday life. As organizations, governments, and individuals increasingly rely on complex technological systems to conduct operations, the importance of cybersecurity governance has grown substantially.[67] Cybersecurity governance broadly refers to the institutional frameworks, policies, and decision-making processes through which organizations manage cyber risks, protect digital assets, and ensure compliance with legal and regulatory standards. It functions as the strategic backbone of cybersecurity management by establishing accountability structures, clarifying responsibilities, and aligning security practices with broader organizational objectives. In an era characterized by accelerating digital transformation, cybersecurity governance has become indispensable for maintaining operational continuity, safeguarding sensitive information, and strengthening resilience against increasingly sophisticated cyber threats. At its core, cybersecurity governance provides a systematic approach to identifying, assessing, and mitigating cyber risks. It encompasses the development of policies, oversight mechanisms, and risk management strategies that enable organizations to protect their digital infrastructure and respond effectively to emerging threats. Without effective governance structures, cybersecurity initiatives often remain fragmented, leaving critical gaps that can be exploited by malicious actors. By defining clear roles and responsibilities and integrating cybersecurity into broader organizational strategies,

[65] Floridi L, *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality* (Oxford University Press 2014)

[66] Floridi L, *The Fourth Revolution: How the Infosphere Is Reshaping Human Reality* (Oxford University Press 2014)

[67] MacKinnon R, *Consent of the Networked: The Worldwide Struggle for Internet Freedom* (Basic Books 2012)

governance frameworks enable institutions to adopt proactive rather than reactive approaches to digital risk management. In doing so, they help organizations anticipate vulnerabilities, strengthen defence mechanisms, and ensure that cybersecurity remains aligned with institutional priorities and regulatory requirements.

The significance of cybersecurity governance has become more pronounced as the digital landscape continues to expand in scale and complexity. The increasing reliance on interconnected systems, cloud-based services, and digital platforms has created a highly integrated technological ecosystem in which vulnerabilities in one component can rapidly cascade across multiple networks. In such an environment, cyberattacks can have far-reaching consequences, affecting not only individual organizations but also entire industries and national infrastructures. Effective governance frameworks therefore play a crucial role in coordinating security strategies, ensuring regulatory compliance, and fostering collaboration among stakeholders responsible for protecting digital systems.[68] One of the primary functions of cybersecurity governance is to address the growing risks associated with cybercrime and data privacy violations. The absence of coherent governance structures often results in inconsistent security policies and inadequate risk management practices, creating opportunities for cybercriminals to exploit weaknesses in digital systems. By establishing standardized procedures for risk assessment, incident response, and compliance monitoring, governance frameworks help organizations minimize exposure to cyber threats while maintaining transparency and accountability. Furthermore, effective governance strengthens trust between organizations and their stakeholders by demonstrating a commitment to safeguarding sensitive information and protecting user privacy.[69] The expanding digital ecosystem has simultaneously intensified concerns regarding the protection of personal and organizational data. Data has become one of the most valuable assets in the modern economy, driving innovation, enabling personalized services, and facilitating strategic decision-making. However, the increasing volume and sensitivity of data collected by institutions also amplify the risks associated with unauthorized access, misuse, and breaches. Data protection has therefore emerged as a central concern within cybersecurity governance, requiring robust frameworks capable of balancing security, privacy, and technological innovation.[70]

The contemporary cybersecurity landscape is marked by a growing number of sophisticated and large-scale cyber incidents that reveal the vulnerabilities inherent in interconnected digital infrastructures. High-profile attacks in recent years have underscored the systemic risks associated with weak cybersecurity governance. Incidents such as the SolarWinds supply chain breach and the ransomware attack on the Colonial Pipeline illustrate how cyber threats can disrupt critical infrastructure, compromise sensitive information, and inflict substantial financial and reputational damage.[71] These attacks not only exposed weaknesses in organizational security practices but also highlighted the broader implications of cyber vulnerabilities for national economies and public trust in digital systems. The economic consequences of cybercrime are particularly striking. Global estimates indicate that cybercrime imposes trillions of dollars in annual losses on the global

[68] MacKinnon R, *Consent of the Networked: The Worldwide Struggle for Internet Freedom* (Basic Books 2012)
[69] Deibert R, *Black Code: Surveillance, Privacy and the Dark Side of the Internet* (Signal Books 2013)
[70] Ibid.
[71] Nissenbaum H, *Privacy in Context: Technology, Policy and the Integrity of Social Life* (Stanford University Press 2010)

economy, with projections suggesting that these costs will continue to rise as digitalization expands. Financial losses associated with cyberattacks include direct theft, ransom payments, operational disruptions, and the costs of recovery and remediation. However, the implications of cyber incidents extend far beyond immediate financial damage.[72] Data breaches and service disruptions can undermine public confidence in digital technologies, discourage technological adoption, and hinder innovation within digital economies. Cyberattacks targeting essential services further illustrate the societal implications of inadequate cybersecurity governance. Attacks on healthcare systems, for instance, can disrupt medical services and jeopardize patient safety by disabling critical information systems. The ransomware attack on Ireland's Health Service Executive demonstrated how cyber incidents can interfere with healthcare delivery, forcing hospitals to cancel procedures and delay treatments. Such incidents highlight the vulnerability of critical infrastructure sectors that rely heavily on digital systems for operational efficiency and service provision.[73]

Beyond economic and societal impacts, cybersecurity governance has also become a matter of national security. State-sponsored cyber operations have increasingly targeted government institutions, defence systems, and electoral processes, raising concerns about the use of cyberspace as a domain of geopolitical competition. Cyberattacks directed at energy infrastructure, communication networks, and financial systems can destabilize entire regions by disrupting essential services and undermining public confidence in national institutions. The coordinated cyberattacks on Ukraine's power grid during the mid-2010s demonstrated how cyber warfare can directly affect critical infrastructure, leaving millions without electricity and highlighting the strategic potential of cyber operations in modern conflicts.[74] Within this evolving threat landscape, cybercrime itself has become increasingly sophisticated and diversified. Among the most prevalent forms of cyber threats are ransomware attacks, phishing campaigns, and supply chain compromises, each exploiting different vulnerabilities within digital ecosystems. Ransomware attacks have expanded dramatically in scale and complexity, targeting organizations across sectors ranging from healthcare and education to energy and finance. These attacks involve malicious software that encrypts a victim's data and demands payment for its release. The global impact of ransomware was dramatically illustrated by the WannaCry attack in 2017, which affected hundreds of thousands of computer systems across multiple countries and disrupted critical services worldwide. In recent years, ransomware tactics have evolved to include "double extortion," where attackers not only encrypt data but also threaten to publicly release sensitive information if ransom demands are not met.

Phishing attacks remain another highly effective method used by cybercriminals to infiltrate organizational networks. These attacks exploit human vulnerabilities rather than technical weaknesses, manipulating individuals into revealing confidential information such as login credentials or financial details. The increasing sophistication of phishing techniques—including spear phishing and business email compromise schemes—has made detection and prevention more

[72] Crete-Nishihata M and Deibert R (eds), *Access Contested: Security, Identity and Resistance in Asian Cyberspace* (MIT Press 2012)
[73] Ibid.
[74] Greenberg A, *Sandworm: A New Era of Cyberwar and the Hunt for the Kremlin's Most Dangerous Hackers* (Doubleday 2019)

challenging for organizations. Because phishing often relies on deception rather than technical intrusion, addressing such threats requires a combination of technological safeguards and user awareness initiatives. Supply chain attacks have emerged as particularly concerning threats within interconnected digital ecosystems. By targeting third-party vendors or service providers, attackers can infiltrate otherwise secure organizations and gain access to sensitive networks. The SolarWinds breach represents a striking example of how vulnerabilities within supply chains can be exploited to compromise thousands of organizations simultaneously, including government agencies and major corporations.[75] Such incidents illustrate how the interconnected nature of modern digital infrastructures creates systemic risks that extend far beyond individual institutions. Advanced persistent threats and state-sponsored cyber operations represent another dimension of contemporary cyber risk. These threats are typically characterized by long-term, stealthy infiltration of targeted systems, often aimed at extracting valuable intelligence or compromising strategic assets. Operations such as the Stuxnet attack on Iranian nuclear facilities demonstrated the potential of cyber tools to achieve strategic objectives traditionally associated with conventional military operations. Similarly, various state-linked hacking groups have been accused of conducting prolonged campaigns targeting intellectual property, industrial secrets, and government communications across multiple countries. Alongside these threats, the challenges of protecting data privacy have become increasingly complex. Data breaches have become a recurring feature of the digital age, exposing sensitive personal information such as financial records, health data, and identification details.[76] Large-scale breaches affecting social media platforms and technology companies have demonstrated how vulnerabilities in data protection practices can expose millions of users to identity theft, fraud, and other forms of exploitation. Beyond the immediate consequences for affected individuals, such incidents also erode public trust in digital platforms and raise questions about the ethical responsibilities of organizations that collect and store vast amounts of personal data. Concerns about surveillance further complicate the relationship between cybersecurity governance and data privacy. Governments often justify extensive surveillance programs as necessary for national security and law enforcement purposes. However, these initiatives frequently raise concerns about the balance between security and individual privacy rights. Revelations regarding mass surveillance activities have intensified debates about the limits of governmental authority and the potential misuse of digital monitoring technologies. Similarly, corporate data collection practices—particularly those used for targeted advertising and behavioural profiling—have raised ethical questions about consent, transparency, and the commercialization of personal information.[77]

The issue of informed consent is particularly significant in the context of digital data governance. Many users unknowingly agree to extensive data collection practices through complex terms of service agreements that are rarely read or fully understood. High-profile controversies involving the misuse of personal data for political or commercial purposes have highlighted the consequences of weak consent mechanisms and inadequate oversight. These developments have prompted policymakers in several jurisdictions to introduce regulatory frameworks aimed at

[75] Brown I and Marsden C, *Regulating Code: Good Governance and Better Regulation in the Information Age* (MIT Press 2013)
[76] Ibid.
[77] Crete-Nishihata M and Deibert R (eds), *Access Contested: Security, Identity and Resistance in Asian Cyberspace* (MIT Press 2012)

strengthening data protection and empowering individuals with greater control over their information. Among the most influential regulatory initiatives are the European Union's General Data Protection Regulation (GDPR) and the California Consumer Privacy Act (CCPA). These frameworks have sought to enhance privacy protections by establishing stricter requirements for data collection, processing, and storage. The GDPR, in particular, has established a global benchmark for data governance by emphasizing principles such as user consent, data minimization, and the right to erasure. Its extraterritorial application has compelled organizations around the world to adopt stronger data protection measures when handling information belonging to European citizens. Similarly, the CCPA has expanded consumer rights in the United States by granting individuals greater control over how their personal data is accessed and used by corporations.[78] Despite these advances, the global cybersecurity governance landscape remains fragmented and uneven. One of the most significant challenges arises from the cross-border nature of cybercrime. Cyberattacks frequently involve perpetrators, victims, and digital infrastructures located in different jurisdictions, complicating efforts to investigate and prosecute offenders. Jurisdictional boundaries, differences in legal definitions of cybercrime, and the absence of harmonized regulatory frameworks often allow cybercriminals to exploit gaps in international enforcement mechanisms. Attacks such as the global spread of the NotPetya malware demonstrated how cyber incidents can affect organizations across dozens of countries while remaining difficult to attribute or prosecute effectively.[79]

International cooperation initiatives have attempted to address these challenges by promoting collaborative approaches to cybersecurity governance. Agreements such as the Budapest Convention on Cybercrime provide frameworks for information sharing and coordinated law enforcement efforts. However, the effectiveness of such initiatives is limited by uneven participation and differing national priorities. Several major countries have not ratified these agreements, citing concerns about sovereignty and fairness in the development of international legal frameworks. As a result, global efforts to combat cybercrime remain constrained by geopolitical tensions and disparities in technological capabilities among nations. Emerging technologies such as artificial intelligence and the Internet of Things further complicate the governance landscape.[80] AI systems rely on large volumes of data to function effectively, raising concerns about transparency, bias, and accountability in automated decision-making processes. Meanwhile, the proliferation of IoT devices from smart home appliances to industrial sensors has introduced new vulnerabilities into digital networks. Many IoT devices are produced with minimal security features, making them attractive targets for cybercriminals seeking to exploit weaknesses in distributed networks. Incidents such as the Mirai botnet attack demonstrated how insecure IoT devices can be harnessed to launch large-scale distributed denial-of-service attacks capable of disrupting major internet services. These developments reveal significant gaps in existing cybersecurity governance frameworks. Many current policies focus primarily on reactive measures such as breach notifications and penalties rather than proactive strategies aimed at building resilience and preventing attacks. Small and medium-sized enterprises, in particular, often lack the

[78] Rid T, *Cyber War Will Not Take Place* (Oxford University Press 2013)
[79] Ibid.
[80] Brown I and Marsden C, *Regulating Code: Good Governance and Better Regulation in the Information Age* (MIT Press 2013)

resources required to implement comprehensive cybersecurity programs, creating vulnerabilities that can be exploited by attackers. At the same time, regulatory fragmentation across jurisdictions continues to complicate compliance efforts for organizations operating in global digital markets.[81]

The growing nuances of cyber threats, the rapid evolution of digital technologies, and the persistent fragmentation of governance frameworks collectively underscore the need for more adaptive and coordinated approaches to cybersecurity governance. Addressing these challenges requires not only stronger regulatory frameworks but also greater international cooperation, improved technological safeguards, and sustained investment in capacity-building initiatives. Within this evolving context, a deeper examination of governance mechanisms, policy frameworks, and institutional responses becomes essential for understanding how digital security can be strengthened in an increasingly interconnected world. Cybercrime against women has increasingly emerged as a significant concern within global governance frameworks, particularly within the institutional architecture of the United Nations. As digital technologies expand across societies, the risks associated with cybercrime, online harassment, and digital exploitation have grown in scale and complexity. Women, in particular, experience disproportionate exposure to cyber threats due to existing social inequalities, gender-based discrimination, and structural barriers that shape their access to digital spaces. International organizations have therefore begun to recognize the gendered dimensions of cybersecurity and the urgent need to develop institutional responses that address online violence, digital exclusion, and technological vulnerabilities affecting women.

International and National Responses to Cybercrime Against Women

Within the United Nations system, the International Telecommunication Union (ITU) has played a pivotal role in addressing evolving technological risks. Established in 1865, the ITU is one of the oldest international organizations and has historically been responsible for coordinating global telecommunication standards and policies. Over time, its mandate has expanded to encompass broader issues related to digital governance and cybersecurity. As communication technologies evolved from telegraph networks to the contemporary internet-based ecosystem, the ITU increasingly acknowledged that cybersecurity is a foundational component of a safe and connected global society. The proliferation of malware, cyber espionage, cyber terrorism, and organized digital crime has intensified concerns about the security of digital infrastructures. These developments have not only affected governments and institutions but have also created new vulnerabilities for individuals navigating online environments. Women are often disproportionately affected by such threats, particularly in the form of cyber harassment, identity theft, online stalking, and the non-consensual dissemination of personal content. Complementing the technological focus of the ITU, the United Nations Entity for Gender Equality and the Empowerment of Women (UN Women), established in 2010, has played a central role in addressing the gendered implications of digital insecurity. UN Women operates as the primary United Nations body dedicated to advancing gender equality, promoting women's rights, and supporting global efforts to eliminate violence against women. The organization also engages with issues affecting marginalized communities, including violence against LGBTQ+ individuals and gender minorities, particularly within digital spaces. Through its governance structure—

[81] Ibid.

comprising an executive board representing different geographic regions—UN Women operates within the broader framework of the United Nations Charter to promote inclusive policies and gender-responsive strategies. In recent years, the organization has expanded its focus to include the intersection between gender, peace, and cybersecurity. Recognizing the growing role of digital technologies in shaping social and political life, UN Women initiated the "Women, Peace and Cybersecurity" project in 2021, aimed at enhancing women's participation in cybersecurity governance and promoting safer digital environments. Such initiatives reflect a broader acknowledgment that cybersecurity must incorporate gender-sensitive perspectives in order to address the distinct challenges faced by women in digital spaces. In 2024, UN Women further strengthened these efforts by launching a free e-learning platform designed to equip women with knowledge and practical skills relevant to the digital age. By providing accessible training resources, the initiative seeks to bridge the digital literacy gap and empower women to navigate technological environments with greater confidence and security. Alongside international efforts, national governments have also begun developing legal frameworks to address cybercrime against women.

In India, the rise in cyber offenses targeting women has prompted the development of a multi-layered legislative architecture aimed at regulating digital conduct and protecting victims. The Information Technology Act, 2000 constitutes the primary legal instrument governing cyber offenses in India. It criminalizes activities such as hacking, identity theft, unauthorized access to digital systems, and electronic stalking. The Act also provides for the establishment of specialized cybercrime investigation cells that are tasked with addressing digital offenses more effectively. In addition to the IT Act, provisions within the Indian Penal Code (IPC) have been amended to address emerging forms of cyber-enabled violence. Legal provisions now cover offenses such as voyeurism, cyberstalking, and the non-consensual dissemination of intimate images, commonly referred to as "revenge pornography."[82] The Protection of Women from Domestic Violence Act, 2005 further extends protection to cases where digital technologies are used within abusive relationships to harass, threaten, or monitor victims. These legal instruments collectively reflect an effort to adapt existing criminal justice frameworks to the realities of the digital era. Despite the existence of this normative legal structure, challenges related to enforcement, institutional capacity, and public awareness continue to limit its effectiveness. Many victims remain unaware of the legal protections available to them, while law enforcement agencies often face resource constraints and technical limitations when investigating cyber offenses. The gap between legislative provisions and their implementation therefore remains a significant obstacle to ensuring meaningful digital protection for women. Statistical trends further highlight the growing scale of the problem. Reports indicate that cybercrimes against women in India have increased steadily in recent years. In 2022 alone, the National Commission for Women reportedly received over 31,000 complaints related to crimes against women, including online harassment and digital abuse. Among these offenses, the non-consensual circulation of intimate images—commonly described as revenge pornography—has become one of the most disturbing manifestations of cyber violence.

[82] Crete-Nishihata M and Deibert R (eds), *Access Contested: Security, Identity and Resistance in Asian Cyberspace* (MIT Press 2012)

Such crimes not only violate personal privacy but also expose victims to severe psychological distress, social stigma, and reputational harm.

Indian law contains several provisions intended to address these forms of digital abuse. Section 72 of the Information Technology Act, 2000 criminalizes the unauthorized disclosure of personal information obtained through digital means. Any individual who discloses or publishes such information without consent may face imprisonment of up to three years, a monetary fine that may extend to five lakh rupees, or both. The provision emphasizes the protection of individual privacy and recognizes the harm caused by the misuse of personal data. The constitutional dimension of privacy protection was further strengthened in 2017 when the Supreme Court of India delivered its landmark judgment in *Justice K.S. Puttaswamy v. Union of India*. In this historic ruling, a nine-judge constitutional bench affirmed that the right to privacy is a fundamental right under Article 21 of the Constitution, which guarantees the right to life and personal liberty.[83] The judgment marked a transformative moment in Indian constitutional jurisprudence by recognizing privacy as an essential component of human dignity and autonomy in the digital age. However, despite these legal and constitutional advancements, the persistence of cybercrime across Indian states illustrates the continuing challenges of translating legal protections into effective safeguards. Rapid digitalization, uneven access to cybersecurity awareness, and persistent gender inequalities contribute to an environment in which women remain vulnerable to online abuse. Consequently, addressing cybercrime against women requires not only stronger enforcement mechanisms but also sustained institutional efforts to enhance digital literacy, strengthen investigative capacity, and promote gender-sensitive approaches within cybersecurity governance.[84]

Conclusion

Despite the existence of an extensive statutory and institutional framework addressing cybercrime against women, the practical effectiveness of these mechanisms remains subject to critical scrutiny. The present analysis suggests that a significant gap persists between the normative strength of legal provisions and their operational implementation. India possesses a layered legal architecture that includes the Information Technology Act, 2000, relevant provisions of the Indian Penal Code, and specialized statutes such as the Protection of Women from Domestic Violence Act, 2005 and the POCSO Act, 2012. In addition, constitutional jurisprudence has reinforced the importance of privacy and dignity through the recognition of the right to privacy as a fundamental right in *Justice K.S. Puttaswamy v. Union of India (2017)*. Together, these instruments create a comprehensive legal framework that formally recognizes the seriousness of cyber violence and the need to safeguard women in digital spaces. However, the persistence and growing visibility of cyber offences indicate that the existence of legal protections alone does not necessarily translate into effective prevention or redress.

[83] Zuboff S, *The Age of Surveillance Capitalism* (PublicAffairs 2019)

[84] Brown I and Marsden C, *Regulating Code: Good Governance and Better Regulation in the Information Age* (MIT Press 2013)

A closer examination reveals that the disparity between "law on the books" and "law in action" remains a defining challenge within India's cybersecurity governance framework. While legislative provisions criminalize acts such as identity theft, cyberstalking, voyeurism, and the non-consensual dissemination of intimate images, enforcement mechanisms often struggle to keep pace with the rapidly evolving nature of digital threats. Law enforcement agencies frequently encounter technical limitations, shortages of trained personnel, and jurisdictional complications when investigating cyber offences. The transnational nature of digital crimes further complicates prosecution, as perpetrators may operate across multiple jurisdictions while exploiting anonymity and technological sophistication. Consequently, victims often face procedural delays, difficulties in evidence collection, and limited institutional support, which collectively weaken the deterrent value of existing legal provisions. Statistical trends further underscore the complexity of the problem. Reported cases of cybercrime against women have increased steadily in recent years, reflecting a pattern that is simultaneously encouraging and concerning. On one hand, the rise in reported complaints may indicate growing awareness among women about their legal rights and greater willingness to approach institutional mechanisms such as the National Commission for Women or specialized cybercrime cells. On the other hand, the upward trajectory of these cases also suggests that cyber offences themselves are expanding in scale as digital technologies become more deeply embedded within everyday life. The proliferation of social media platforms, online financial services, remote work systems, and digital communication tools has transformed the internet into an indispensable infrastructure for modern living. In such a context, complete disengagement from digital environments is neither practical nor desirable. Women must participate in these spaces for education, employment, social interaction, and civic engagement. Yet this very participation simultaneously increases their visibility and accessibility within digital networks, often exposing them to new forms of vulnerability and exploitation.

This structural dependence on digital systems therefore creates a paradoxical condition. While digital technologies offer unprecedented opportunities for empowerment, connectivity, and economic participation, they also expand the terrain on which gender-based violence can occur. Cyber harassment, stalking, doxxing, and non-consensual image dissemination illustrate how traditional forms of gendered violence have migrated into digital environments, often amplified by the speed, anonymity, and reach of online platforms. Addressing these challenges requires moving beyond a narrow focus on statutory adequacy toward a broader examination of institutional readiness and governance capacity. It is essential to assess whether law enforcement agencies possess the technological infrastructure, investigative expertise, and gender-sensitive protocols necessary to respond effectively to cyber offences. Equally important is the question of accountability. Violations may originate not only from individual offenders but also from systemic failures within institutions, regulatory frameworks, or digital platforms themselves. In such cases, the assumption of the state as an unequivocal protective actor becomes more complex and warrants careful scrutiny. Against this background, the chapter highlights the need for a multidimensional approach to cybersecurity governance that integrates legal, technological, and social perspectives. Ensuring meaningful digital safety for women requires sustained investment in institutional capacity-building, specialized cybercrime training for law enforcement agencies, and stronger coordination between regulatory bodies, technology companies, and civil society organizations. At the same time, initiatives aimed at strengthening digital literacy and cybersecurity awareness among women remain crucial for enabling individuals to recognize and respond to online threats.

Ultimately, it suggests that cybersecurity frameworks often function reactively rather than preventively. Regulatory mechanisms frequently emerge in response to technological disruptions rather than anticipating them. As digital innovation continues to accelerate, the gap between technological capability and regulatory adaptation may widen further unless proactive governance strategies are adopted. Evaluating the effectiveness of cybersecurity frameworks therefore requires asking whether they operate as genuine deterrents capable of preventing cyber offences or whether they primarily function as post-facto response systems activated only after harm has already occurred. This question forms a critical foundation for the subsequent analysis undertaken in this study, which seeks to examine how legal institutions, governance mechanisms, and technological systems can be strengthened to ensure that women's digital safety becomes a substantive reality rather than a merely aspirational promise in an increasingly technologized society.

Chapter – IV

Cybersecurity Risks in the Digitisation of Cultural Heritage in India: A Legal and Governance Analysis

Priyanka Das, Assistant Professor of Law, Swami Vivekananda University

Abstract

India is using digitization to both safeguard and promote the country's heritage through its cultural institutions like museums, manuscripts, archives, and other historical records. Digital forms of preservation are often considered a more secure and permanent way to preserve cultural assets, and many continue to subscribe to a larger myth associated with cybersecurity and how technology is assumed to be safe; therefore, this chapter will discuss if the current legal and governance framework created in India is sufficient to protect digitized cultural heritage from cyber-related threats. Most current literature discusses cultural heritage protection and cybersecurity governance in conjunction, but little existing analysis focuses directly on how current cybersecurity law applies to digital heritage institutions. This gap is important to note, as digital repositories become subject to various cyber threats, including data breaches, ransomware attacks, unplanned access via hacking, and weaknesses with the organizations that rely upon digital assets. This chapter uses doctrinal and analytical methods to review relevant laws related to this area of study, including but not limited to the IT Act, 2000, the Digital Personal Data Protection Bill 2023, and any other relevant heritage legislation and policy documents, along with secondary sources of information. By studying the intersection between digitization initiatives and cybersecurity regulatory initiatives, this study highlights the disparity between the intended benefits of digital preservation and the obstacles associated with cyber governance.

Keywords: Digital Preservation, Cybersecurity Law, Cultural Heritage Protection, Data Governance, Institutional Vulnerability.

Introduction

Cultural heritage is the artistic, scientific, or historical heritage of a culture or society. The term "cultural heritage" has various of definitions at both national and international level. However, it lacks any uniform definition. As per the UNESCO, "cultural heritage" includes "artefacts, monuments, a group of buildings and sites, museums that have a diversity of values including symbolic, historic, artistic, aesthetic, ethnological or anthropological, scientific and social significance. It includes tangible heritage, movable, immovable and underwater, intangible cultural heritage (ICH) embedded into cultural, and natural heritage artefacts, sites or monuments. The definition excludes ICH related to other cultural domains such as festivals, celebration etc. It

covers industrial heritage and cave paintings."[85] The preservation of the immense variety of heritage has long been viewed as a priority for India, where numerous cultures have existed throughout its history. Historically, preservation was limited to the physical preservation of monuments and artefacts via statutory methods. However, digitalisation has become an emerging model of preserving and promoting cultural heritage in recent years. There has been a greater commitment to digitalisation as a preservation method because digitisation allows cultural institutions to convert tangible materials into digital formats, which makes for greater access to cultural heritage, promotes research opportunities, and helps prevent the degradation of fragile resources.[86]

In addition, the creation of digital repositories and archives allows for a greater possibility of public engagement through the access to cultural heritage, as well as allowing for new ways to document and disseminate cultural heritage. The result of this belief is that many people now trust that digital preserving solutions for the conservation of heritage will be the most reliable and secure option to meeting the difficulties of heritage conservation in the future. This trust reflects the confidence most people have in technology systems as providing a safe and long-term method for completing the process of preserving cultural, historical, or scientific heritage.

However, the assumption that digitisation automatically guarantees security overlooks the vulnerabilities associated with digital infrastructures. Cultural institutions that rely on digital platforms are exposed to cyber risks such as data breaches, ransomware attacks, unauthorised access and systemic institutional weaknesses. When heritage resources are integrated into digital environments, they become subject to the same threats that affect other information systems. Any compromise of digital heritage may have consequences that extend beyond operational disruption and may affect cultural continuity and public trust.

The Indian government has enacted legislation to deal with cyber issues. The legislation currently in place includes the Information Technology Act, of 2000 and the Digital Personal Data Protection Act, of 2023. The two laws set up a general framework to regulate cybercrime, but there has been very little attention given to how they are to be applied specifically to institutions with digitized cultural heritage. Most of the available literature on the issue has treated the two areas of research i.e Cultural Heritage Protection and Cyber Security as separate and neither area has had much in the way of scholarly examination that touches on the intersection between the two.[87]

The purpose of this chapter is to determine if current cyber governance and the legal framework in place in India is adequate to protect digitized Cultural Heritage from the potential of being harmed by cyber related threats. In order to accomplish this, this chapter will incorporate a doctrinal and analytical methodology in order to conduct an examination of the applicable statutes, legislation governing Heritage Sites, policy documents, and applicable secondary academic sources. This chapter will examine the applied relationship between digitized Cultural Heritage, digitization initiatives and Cyber Security Regulation, and will highlight gaps between the declared

[85] UNESCO Institute for Statistics, *2009 UNESCO Framework for Cultural Statistics (FCS)* (UNESCO 2009).
[86] Press Information Bureau, 'Digitization of Cultural Heritage in India' (17 March 2025) https://www.pib.gov.in/PressReleasePage.aspx?PRID=2111884&reg=3&lang=2 accessed 2 March 2026.
[87] DLA Piper, 'Data Protection Laws of the World: India – Data Protection in India' https://www.dlapiperdataprotection.com/?t=law&c=IN accessed 2 March 2026.

intent of the policies and practices that are implemented to preserve Cultural Heritage Digitally and the challenges that exist in Nurturing Cyber Governance.

Digitisation of Cultural Heritage in India: Policy Context and Institutional Framework

India is working hard to create a digital representation of its cultural past. This is part of an initiative to create better government services and improve the uses of the internet. There has been a shift by many sources of cultural goods towards using computers to help preserve the information about their items, recording that information, and making it available to the public. The aims of these different agencies go beyond just needing to keep records, they want to provide access to their records, allow people to see them, and involve people once they have been made available.

Several successful projects have been supported by the Indian government to help digitise material.[88] All major projects have aimed to create online collections of items that could be used by everyone to learn about India's history. The national cultural institutions responsible for these projects include the National Archives of India, the Indira Gandhi National Centre for the Arts and various state museums. They have developed ways to scan and catalogue items digitally, so they can be accessed through various means online. The Digital India programme also emphasises the importance of electronic government and creates better digital infrastructure, which indirectly influences the cultural sector.

There are many benefits to digitisation. Digitisation aids in preserving delicate artefacts due to the decrease in physical handling of these items. Additionally, digitisation allows researchers, academics and the general public to have access to materials, regardless of geographic distance. Another benefit is that digitisation helps with documentation, and inventory management processes and therefore improves the administration efficiencies of institutions. Digital storage also permits duplication and backup, which potentially decreases the risk of irretrievable loss from natural disasters or deterioration of the physical object.

However, the digitisation of cultural assets results in a transformation of the cultural asset itself. Once in digital format, cultural heritage items are mere data format and stored on servers, networks and cloud systems. Consequently, protecting them will consist of information security functions rather than conservation functions. The risks associated with this changeover generally receive little or no attention. Cultural heritage institutions are generally underprepared for many digital security risks due to a lack of adequate cybersecurity infrastructure, or appropriate technical skill or financial resources to adequately manage digital threats.[89]

Also, third-party services are frequently engaged to support digitisation projects such as storage of data, software management, or cloud hosting, generating multiple futures for potential failure due to additional layers of vulnerability. Cultural heritage institutions may have weak contracts with third-party service providers, as well as poor security practices, and insufficient oversight leading

[88] Ministry of Culture, 'Digital Cultural Governance and Monitoring of Cultural Initiatives' (Press Information Bureau, 2 February 2026) https://www.pib.gov.in/PressReleseDetailm.aspx?PRID=2222115 accessed 2 March 2026.
[89] Samar I Bakhshi, 'Digitization and Digital Preservation of Cultural Heritage in India with Special Reference to IGNCA, New Delhi' (2016) 6(2) Asian Journal of Information Science and Technology 1, 2 https://ajist.co/index.php/ajist/article/view/134 accessed 2 March 2026.

to unauthorised access or use of digital data repositories. In some instances, cultural heritage institutions may not consider their digital collections to be valuable information assets, thereby limiting their investment in security measures.

Another important dimension concerns the value of cultural data. Digitised manuscripts, artefacts and archival records may contain sensitive historical information, indigenous knowledge systems or rare documentation. Such data may be attractive targets for cyber criminals, whether for financial gain, political motives or ideological reasons. Ransomware attacks, in particular, pose a serious threat because they can lock institutions out of their own collections unless payment is made. The loss or manipulation of cultural data may result in reputational damage and undermine public trust.

Policies governing digitalisation tend to put emphasis on growth/as well as access, with little thought given to Security arrangements. Most of the time, Innovation, and access are seen as high priority in relation to Technology, however, Security is viewed as secondary. Thus, there seems to be an implied belief that Digital Technologies, by virtue of being more advanced than other technologies, automatically provide Security. With the increased dependence on Digital Technology Infrastructures, one of the most important questions regarding Digital Heritage management will be whether or not the current legal and regulatory framework in India is sufficient for the Cybersecurity risks cultural institutions experience. In order to have an answer to this question, the appropriate statutes dealing with Cybercrime and Data Privacy should be reviewed and their relevance to the Heritage sector evaluated. This will be done in the following chapter.[90]

Cybersecurity Law in India and Its Applicability to Digital Heritage Institutions

Cybersecurity in India is primarily governed by legal regulation through the Information Technology Act of 2000, which was originally intended to enable e-commerce through the use of electronic signatures and provide a legal framework for e-businesses.[91] As it has evolved into a comprehensive piece of legislation, it now regulates various types of cyber-related misconduct, including cyber terrorism, as well as providing a legal framework for electronic records and electronic documents. The IT Act was amended to include provisions for data protection, victimisation and intermediary liability in 2008.

The IT Act includes several provisions that apply to digital cultural heritage institutions, including Section 43 that is unauthorised access to computer systems, theft of data and damage to computer systems and Section 66 that states about acts against computer systems that are done dishonestly or fraudulently. Both of these provisions would apply if a person accessed a digital archive or museum database without authorisation or altered, copied or deleted information from a digital collection of cultural heritage. Additionally, Section 66F or cyber terrorism could become applicable if an attack on a cultural heritage repository creates a threat to national security or public order.

[90] National Archives of India, 'Information Technology' https://nationalarchives.nic.in/Information-Technology accessed 2 March 2026.
[91] 'Understanding the Information Technology Act, 2000 in E-commerce' (Lloyd Law College, 24 June 2025) https://www.lloydlawcollege.edu.in/blog/it-act-2000-ecommerce-legal-framework.html accessed 2 March 2026.

The objective of protecting sensitive personal information and requiring reasonable security practices with respect to sensitive personal data are among the several areas in which these two statutes overlap. While earlier regulations around Section 43A of the Act provided a basis for understanding how to meet the standards of data protection, the Digital Personal Data Protection Act of 2023 has now significantly enhanced this regulation by establishing a clear, structured system of providing adequate safeguards for processing personal data, including specifying the responsibilities and rights of individual data fiduciaries and data principals, respectively.

Although many cultural heritage organizations do not routinely have access or use large volumes of personal data, some archives or historical records may still contain information that is identifiable to persons, groups or communities who are the subjects of that collection.[92] This will therefore create a legal obligation to comply with all applicable provisions of the Act as it specifically relates to data processing activities and records. Rather, the primary and only purpose of most cultural heritage organizations that hold archival materials is to preserve and disseminate historical records, not to collect and store data for commercial purposes. Thus, it is highly debatable whether the cybersecurity requirements established pursuant to the Digital Personal Data Protection Act, 2023 satisfy the cybersecurity needs of cultural heritage organizations.[93]

The corporate and regulatory statutory provisions in India are only one of many systems for governing digital Heritage organizations and their controls. The National Computer Emergency Response Team functions under the authority of the Information Technology Act, which includes guidelines for reporting and responding to incidents related to cyber security. Many of the organizations outlined above operate digital systems and may be subject to reporting obligations established in the guidelines from N-CERT; however, the current level of awareness and compliance in this sector is unclear. The significant limitation within this existing framework is its general applicability. The Information Technology Act and Digital Personal Data Protection Act were designed and implemented across all sectors of the economy. Neither specifically identifies Digitized Cultural Heritage organizations as critical information infrastructure nor does either provide cultural repository organizations with sector-specific guidelines for implementing standards that enable the preservation of these types of collections.

This absence ultimately creates a situation where Digitized Cultural Heritage organizations must interpret the broad statutory provisions and statutory obligations on their own, without having access to meaningful assistance from the statutes. Moreover, cybersecurity governance requires more than just adherence to legal obligations. These organizations typically operate with limited technical capacity and financial means; therefore, even when statutory obligations exist, it may not be possible for them to properly implement those obligations. The difference between the formal protections provided by law and the institutional capacity to adhere to the law demonstrates a structural flaw within the existing systems.[94]

---

[92] T.K. Gireesh Kumar, 'Digital meets culture: towards a national portal for aggregating Indian cultural heritage' (2026) Journal of Cultural Heritage Management and Sustainable Development 1, 5 https://doi.org/10.1108/JCHMSD-02-2024-0031 accessed 2 March 2026.

[93] DLA Piper (n 4).

[94] Ministry of Electronics and Information Technology, 'Digital India Programme' https://www.digitalindia.gov.in/ accessed 2 March 2026.

In summary, an analysis of the extent to which heritage specific laws and policies complement or contradict the wider cyber security framework must occur following this assessment. This will require that the intersection of cyber regulation with heritage protection laws in India.

Heritage Protection Laws and the Regulatory Disconnect in the Digital Context

Historically, the legal protection of India's cultural heritage has been aimed at preserving tangible monuments, sites and artefacts. The main legislation regarding the protection of ancient monuments is the Ancient Monuments and Archaeological Sites and Remains Act 1958. This Act primarily regulates and preserves monuments classified as being of 'national importance'. Another piece of legislation regarding antiquities is the Antiquities and Art Treasures Act 1972, which governs the export and trade in antiquities and art objects, amongst other things. Both pieces of legislation emphasise physical protection and regulatory oversight in the safeguarding of material heritage.

These pieces of legislation, while still important for protecting material heritage, are applicable only to the material heritage that existed before the rise of the internet and online technology. None of these statutes take into account the effects of digitisation and do not seek to provide for protection against hacking, data manipulation or the digital destruction of records. Once material culture has been converted to a digital format, it is no longer clearly covered by traditional heritage statutes.

The regulatory gap becomes apparent when we consider the holdings of archival and manuscript collections. Archival institutions, such as the National Archives, and public libraries operate from administrative structures and frameworks of rules, but these rules are usually more about custodial or facility responsibility than about securing the information that resides there.[95] Although many policy efforts have encouraged these institutions to digitize their manuscript and archival collections, there are currently few cybersecurity standards specifically tailored to their digital holdings.

This disconnection points to a larger systemic problem within the administrative framework for managing digital heritage. Heritage law and cybersecurity law exist in parallel with respect to the protection of cultural property versus addressing digital systems. Digital heritage occupies a hybrid reality that spans both areas of law. Digital heritage is, by nature, both a cultural good and, therefore, a data resource. The lack of explicit legal recognition of digital heritage creates an incomplete and fragmented protection of this unique type of property.

In addition to ownership and authenticity, there are many other issues related to cultural materials in a digital environment including their copying, altering and redistributing. Questions about intellectual property rights and the moral rights of the original creator of the work are common; as are ethical questions about the integrity of digital reproductions. Although copyright law provides some protection from infringement, copyright law does not provide protections related to the broader security risks associated with unauthorised access or systemic cyberattacks on repositories.

[95] National Archives of India (n 7).

In India, since cultural institutions are not generally classified as critical information infrastructure under the Indian cybersecurity policy, therefore cultural materials do not get the level of protection provided by the sector focus for critical infrastructure sectors such as finance, energy and telecommunications. Because damage to these critical infrastructure sectors has more immediate economic and security implications, the greater priority is to ensuring that they are able to operate without interruption or threat. Cultural heritage continues to be a significant part of a country's identity and helps to establish historical continuity, but they are often not perceived as being strategic targets for cyber-attacks. The perception of risk in this area may contribute to insufficient investment in cyber resilience for the cultural sector.

This regulatory disconnect operates on several levels: the statutory level, the institutional coordination level, and the level of government priorities. While within the area of heritage legislation there is not a requirement to conduct a digital risk assessment, there is also not any way for the cyber-security legislation to provide any protective measures for the heritage sector. Consequently, a regulatory framework that is designed to deal with cyber threats in general but fails to take into account or provide for the cultural significance of the information being protected.

Cybersecurity Risks and Governance Challenges in Digital Heritage

Acting as a bridge between digital and physical collections, cultural heritage institutions are digitising their manuscripts, artefacts, and archival records so that they are stored as digital assets in networked systems.[96] Even though improving access to collections and preserving collections through digitisation are invaluable benefits, digitisation puts cultural institutions at a higher risk of cybersecurity. A major cybersecurity concern facing cultural institutions is unauthorised access to a database that could allow a hacker to copy or change the digitised collection. This could put the authenticity and the integrity of the digitised collection at risk. Unlike physical damage, digital evidence of manipulation might not be readily apparent.

This creates long-term risks associated with the distortion of the historical record. Apart from the risk of unauthorised access, ransomware attacks present another serious concern.[97] If an attacker is able to encrypt digital repositories, institutions may have no way of recovering access to the digitised collections. If an institution does not have secure backups in place, this circumstance may result in the permanent loss of the digitised collection and could also result in reputational damage to the institution. Data that is inappropriately accessed can also create potential liability exposure for institutions if the archives contain personal or sensitive information.

Lastly, cultural institutions often operate with limited financial and technological resources. Portfolio risk may limit the priority given to cybersecurity at many museums and archives. Cultural institutions are often reliant on third party service providers for solutions related to cloud storage and database management, creating further complications when it comes to establishing responsibility and liability in the event of a breach.

[96] T.K. Gireesh Kumar, 'Digital meets culture: towards a national portal for aggregating Indian cultural heritage' (2026) Journal of Cultural Heritage Management and Sustainable Development 1, 3.
[97] CERT-In, 'Notes on Ransomware Prevention and Response' (2023) https://www.cert-in.org.in accessed 2 March 2026.

Conclusion and Recommendations

The digitisation of India's cultural heritage is vital in preserving it, making it more accessible and allowing distribution of knowledge democratically. Digital methods also expand the audiences that can engage with historical materials and limit the physical deterioration of fragile items.[98] However, moving from traditional methods of preserving items to digital preservation involves a set of new vulnerabilities, which cannot be solved simply by advancing technology. As demonstrated through the analysis provided in this chapter, India has some comprehensive cybersecurity legislation; for example, the information technology Act 2000 and the Digital Personal Data Protection Act 2023 provide some protections for all sectors of industry.

However, neither the IT Act nor the Digital Personal Data Protection Act were written with the unique needs of institutions with digitally preserved cultural heritage in mind. In contrast, laws to support the preservation of cultural heritage institutions primarily address the physical preservation of cultural artifacts and do not adequately address the risks associated with the digital preservation of these items. This gap in laws exposes cultural repositories to the potential threat of losing access to cultural heritage through cyber-related incidents.[99]

Several approaches can be explored to close this divide. Firstly, guidelines providing policy direction for the digital heritage sector should be created, incorporating cybersecurity standards within strategies for cultural preservation. Secondly, heritage institutions must follow a set of minimum security standards, including routine evaluations of risk, secure backup systems for data, and clear channels to report incidents related to cybersecurity. Thirdly, building capacity will provide staff within heritage institutions with a stronger technical knowledge of cyber-related hazards, as well as raise their level of awareness in relation to these risks.

Finally, it is essential that there is improved coordination between cultural bodies and cybersecurity agencies to clarify regulations and enhance institutional preparedness. Digitisation alone cannot be relied upon as the sole solution to the challenges of preservation; it needs to be backed by sufficient systems of governance that acknowledge digital heritage as an asset that is both cultural and informational; protecting it calls for more than regulatory compliance institutions must demonstrate their ongoing commitment and co-ordinate their efforts through policies that all government entities support.

[98] Sudipta Shee, 'Digital preservation of cultural heritage in India: A digital age' (2025) 7(1) International Journal of Humanities and Education Research 260.
[99] DLA Piper (n 4).

Chapter – V

Artificial Intelligence in Banking and Ombudsman Services: Addressing Cybersecurity Challenges While Ensuring Justice

Dr. Souvik Dhar, Assistant Professor, School of Law, Brainware University

Abstract

This paper explores the transformative role of Artificial Intelligence (AI) and technology in banking and Ombudsman services. The research questions focus on identifying risks and proposing mitigation strategies for fair AI adoption. A qualitative methodology was adopted, analyzing legal statutes, literature, and expert opinions. The thematic analysis categorized risks into legal, technological, ethical, and operational domains. The paper discussed AI's integration into banking, highlighting customer support, fraud detection, personalized services, back-end automation, compliance, and cybersecurity. AI-driven chatbots and virtual assistants provide round-the-clock customer support, enhancing service efficiency and customer satisfaction. Fraud detection systems can analyze transactions in real-time, preventing financial crimes effectively. Personalized banking through automated advisors and hyper-personalization improves customer engagement. Back-end automation accelerates operations, reduces errors, and lowers costs. AI also streamlines regulatory reporting and strengthens cybersecurity defenses. These innovations collectively make banking faster, safer, and more customer-centric. The paper also critically examines risks associated with deploying AI in Banking Ombudsman services. It identifies legal challenges such as threats to natural justice and lack of explainability in AI decisions. Technological issues include bias, unreliability, and cybersecurity vulnerabilities. Ethical and institutional risks involve dehumanization, accountability gaps, and erosion of public trust. Regulatory gaps exist due to the absence of specific AI guidelines, risking exclusion of marginalized groups. Results indicate that while AI enhances efficiency, it also risks violating constitutional principles and data privacy. Biases inherited from training data can lead to discriminatory outcomes. Errors and cyber threats compromise decision reliability and data security. The lack of clear liability frameworks complicates accountability. The discussion emphasizes the need for human oversight, ethical guidelines, and regulatory reforms. Strengthening cybersecurity and establishing sector-specific regulations are crucial. Responsible deployment requires balancing innovation with safeguards. Developing AI and ensuring inclusive access are vital. Implementing legal and ethical frameworks will foster trust and justice. AI can significantly improve banking and Ombudsman services if risks are managed effectively. This paper aims to support policymakers, regulators, and financial institutions in fostering trustworthy AI systems. Ultimately, responsible AI use can promote equitable, efficient, and secure banking and dispute resolution services.

Keywords: Artificial Intelligence; Banking system; Grievance redressal; Ombudsman; policy framework; cybersecurity.

Introduction

The fast-paced evolution of technology, particularly Artificial Intelligence (AI), has significantly impacted the growth of various sectors around the world, with the banking sector being cited as one of the sectors that have been significantly impacted. The convergence of AI and the banking sector has, in effect, transformed the conventional banking system, resulting in increased efficiency, customer service, and security.[100] In the context of the country, the banking sector is undergoing a transformation, with the evolution of technology being the major contributor. Moreover, the Reserve Bank of India (RBI), which is the apex bank of the country, has been at the forefront in the promotion of the digital banking system, particularly the adoption of AI, through various measures.[101] Some of the measures include the introduction of AI-based innovations, including the use of chatbots, virtual assistants, biometric systems, etc.[102] These innovations have significantly impacted the growth of the banking system, resulting in the provision of faster, more efficient, and customer-centric services. These innovations have significantly impacted the growth of the banking system, resulting in the provision of faster, more efficient, and customer-centric services. Moreover, the evolution of technology has also had an indirect positive impact on the growth of the banking system, resulting in the promotion of financial inclusion.

The present study aims to investigate the application of AI in the banking sector and Ombudsman services, highlighting the legal, ethical, technological, and operational issues arising therefrom. While AI has the advantage of offering a plethora of benefits, it also poses a number of risks, particularly in the context of cybersecurity, data privacy, bias, and transparency, among others. The application of AI in critical sectors, such as dispute resolution and customer grievances, necessitates a high level of oversight to avoid the infringement of constitutional rights, particularly the right to privacy, and the principles of natural justice.[103] The Banking Ombudsman Scheme, introduced by the RBI, is a vital instrument in the resolution of customer grievances against banks. Traditionally, the Banking Ombudsman Scheme has relied on manual and paper-based documentation. However, with the emergence of AI and digital technologies, there is a tremendous opportunity to transform the existing system, with a view to increasing the efficiency, transparency, and accessibility of the system, although the application of AI in quasi-judicial and administrative functions poses a number of issues, which are critical and require due consideration.

From the academic and legal viewpoints, there is an immediate need for exploring the interface of AI, banking law, and the constitutional framework, especially in the context of the Indian scenario. The landmark judgment in Justice K.S. Puttaswamy v. Union of India[104] recognized the right to

[100] Shyam Sunder Agrawal, Ninu Rose and K. Prabhu Sahai, 'The fintech revolution: AI's role in disrupting traditional banking and financial services' (2024) 7(1) Decision Making: Applications in Management and Engineering 243.
[101] P Kumar and Rishi Manrai, 'A study on accelerating digital financial inclusion for positioning India through AI-enabled banking services' (2024) 6(4) International Journal for Multidisciplinary Research.
[102] S Agarwal, B Agarwal, and R Gupta, "Chatbots and virtual assistants: a bibliometric analysis," Library Hi Tech, vol. 40, no. 4, 2022, pp. 1013-1030.
[103] D. Leslie, C. Burr, M. Aitken, J. Cowls, M. Katell, and M. Briggs, 'Artificial intelligence, human rights, democracy, and the rule of law: a primer', arXiv preprint arXiv:2104.04147, 2021.
[104] Justice K.S. Puttaswamy v. Union of India, AIR 2018 SC (SUPP) 1841.

privacy as an integral aspect of the fundamental right under Article 21 of the Indian Constitution. This underlines the need for the legal framework to ensure the handling of data and the use of AI in the context of dispute resolution in the banking sector.

In the current paper, the legal, ethical, and operational aspects of the integration of AI in the banking sector and the Ombudsman schemes in the Indian context are explored. The focus of the paper is to explore the scope of the use of AI in the context of dispute resolution in the banking sector, while ensuring the constitutional rights of the parties involved in the dispute resolution process.

Research Objectives

This paper aims to give an exhaustive overview of the use of AI in the banking system, including the Ombudsman service, in India. This study aims to critically analyze the myriad implications of the use of AI in the banking sector and the Ombudsman schemes in the Indian legal and regulatory scenario. The objectives of the study are as follows:

1. Analyse the existing legal and regulatory scenario in the context of the use of AI in Indian banking services: This would involve the critical examination of the laws, guidelines, and judicial pronouncements applicable to the use of AI in the context of data privacy and dispute resolution laws and the constitutional provisions in the Indian context.

2. Evaluate the Banking Ombudsman Scheme in the context of the use of AI in the digital age.

3. Identify and critically examine the legal and ethical issues that are linked to the integration of AI in the banking system of dispute resolution.

4. Propose the legal and policy changes that are required to ensure the responsible use of AI in the banking system, including the Ombudsman service.

5. Examine the need for the incorporation of cybersecurity in AI-based banking systems.

Materials and Methods

The methodology that is being followed in this study is doctrinal, qualitative, based on legal analysis, interpretive approaches, and comparative perspectives. The aim is to arrive at an in-depth understanding of the legal and ethical principles that govern the use of AI in the context of banking, along with the dispute resolution mechanisms, particularly in the context of India.

Primary Data Sources

1. Statutes and Regulations:

The basic legal instruments would include the Reserve Bank of India Act, 1934, Banking Regulation Act, 1949, Information Technology Act, 2000, and the Digital Personal Data Protection Act, 2023. These acts would provide the legal guidelines for banking operations, data protection, and digital dispute resolution. RBI guidelines, circulars, and notifications, such as RBI's Master

Directions on 'Digital Payment Security Controls' and 'Integrated Ombudsman Scheme 2021,' would be critically analyzed to understand the current scenario in the country.

2. Judicial Decisions:

Important cases like the Justice K.S. Puttaswamy v. Union of India[105] would be the guiding principles for understanding the constitutional aspects of the right to privacy. Important cases would include Indian Medical Association v. Union of India[106], Shreya Singhal v. Union of India,[107]etc. These cases would be analyzed to understand the jurisprudence of the matter and the changing dynamics of the use of AI in the country.

3. Official Reports and Policy Documents:

Reports from the RBI, NITI Aayog, OECD, and other international agencies would provide the necessary context for understanding the changing dynamics of the use of AI in the country. RBI reports on fintech, digital banking, and the use of AI would be highly instrumental in understanding the perspectives of the RBI.

4. Case Law and International Jurisprudence:

The study would include the regulations of other nations like the European Union's regulations under the GDPR and AI Act, Singapore's regulations on the use of AI, and the UK's regulatory sandbox initiatives.

Secondary Data Sources

1. Academic Literature:

We will use journal articles, commentaries, and research papers to shed lighter and challenge the underlying principles of law, the technology issues, and the ethical issues involved.

2. Legal Commentaries and Textbooks:

We will use commentaries and textbooks on banking law, data privacy, and AI ethics to support the underlying principles of law and legal discussions.

Research Techniques

1) Legal and Doctrinal Analysis:

The core of the work will involve the interpretation of the legal provisions to understand the implications of the legal principles in the context of AI in banking and grievance redressal. This will involve an in-depth study of the legal provisions with a critical eye.

[105] Justice K.S. Puttaswamy v. Union of India, AIR 2018 SC (SUPP) 1841
[106] Indian Medical Association vs Union Of India & Ors., AIR 2011 SUPREME COURT 2365
[107] Shreya Singhal vs U.O.I, AIR 2015 SUPREME COURT 1523.

2) Comparative Analysis:

The study will also cover the legal frameworks in other jurisdictions to identify the best practices that could be adopted in the Indian legal system.

3) Critical and Thematic Analysis:

The study will cover the risks of bias, opacity, and breaches in the context of AI in banking and grievance redressal, along with the other issues that are connected with the use of AI. This will cover the implications of the use of AI in the context of the legal principles of fairness, transparency, etc.

4) Policy and Regulatory Framework Development:

The study will cover the development of the policy in the context of the legal provisions with the help of the doctrinal study and the comparative study.

5) Interdisciplinary Approach:

The study will cover the different aspects of the use of AI in banking and grievance redressal with the help of the different disciplines.

Limitations

In the context of the nascent stage of legislation related to the field of AI in Indian law, the study relies upon existing laws, judicial decisions, and international best practices. The dynamic nature of the technology and the law in the field of AI implies that the study may require updating in the near future with the enactment of fresh laws. Further, empirical data in the field is limited, which underlines the theoretical nature of the study.

Literature Review

Today, AI has become a mainstay in the lives of various sectors, with banking leading the way. In the past few years, there has been a lot of research and expert opinion on the role of AI in the banking sector, how it is changing the way banking is done, and how it is speeding things up, making it safer, and providing a more personalized service, though there are also risks and challenges that come with the adoption of AI. In this paper, a literature review on the role of AI in the banking sector will be presented, highlighting the benefits, the challenges, the legal and ethical issues, and the role of AI in dispute resolution and customer service.

Benefits of AI in Banking

Many experts are in agreement that AI has a lot to offer to the banking sector. AI is able to help banks sort through large amounts of data in a fast and accurate manner.[108] Moreover, AI is able to help detect fraud in a timely manner, thereby increasing the safety of the customer.[109]

Chatbots and virtual assistants are some of the most popular AI applications used in the banking sector. For example, chatbots are able to communicate with customers online and assist them with a number of issues. For example, banks like HDFC, SBI, and ICICI have their own chatbots, which are referred to as Eva, SIA, and iPal, respectively.[110] These chatbots are able to operate around the clock, which means that the customer is able to get assistance at any time of the day or night. For example, a customer is able to get the status of a loan application or check the account balance in a matter of a second, whereas in the past, the customer would have had to wait in a queue.[111]

Advantage of using AI in the banking sector is that it is able to provide the customer with a personalized banking experience. AI is able to help the customer plan their investments with the help of a robo-advisor, which is a device that is able to assist the customer in the creation of a plan to help them achieve their goals.[112] The robo-advisor is able to analyze the customer's goals, income, and risk tolerance and provide them with the best investment strategy. AI is able to read the customer's needs and provide them with the most appropriate services. If the customer is a fan of traveling, AI is able to provide them with a credit card that rewards them with miles.

Another advantage of the use of AI in the banking sector is that it is able to help the bank automate a number of back-end operations. The AI is able to help the bank open a number of accounts, verify the details of the customer, and verify the compliance of the customer with the rules and regulations of the bank.[113] For example, in case the customer wants to open an account with the bank or wants to take a loan, the AI is able to verify the details of the customer in a matter of a second.

The AI also helps the bank in following the rules that are set by the regulators. For example, the AI is able to scan the transactions in the bank to verify whether there is any money laundering or fraud activities going on in the bank. The AI is also able to help the bank in the preparation of the

[108] Suparna Biswas, Brant Carson, Violet Chung, Shwaitang Singh, and Renny Thomas, AI-bank of the future: Can banks meet the AI challenge (McKinsey & Company 2020) 28.
[109] Rahul Khurana, 'Fraud detection in ecommerce payment systems: The role of predictive AI in real-time transaction security and risk management', International Journal of Applied Machine Learning and Computational Intelligence, vol. 10, no. 6, 2020, pp. 1-32.
[110] Kanika Sachdeva, and Meenakshi Dhingra, 'Role of Bank Chatbots in Managing Business during Crisis: A Study of Customers' Perceptions of ICICI Bank and SBI Bank', in Building Resilience in Global Business During Crisis, pp. 38-57. Routledge India, 2024.
[111] Everestus Obinwanne Eze, and Adaora Darlingtina Odunukwe, 'On application of queuing models to customers management in banking system', American Research Journal of Bio Sciences, vol. 1, no. 2, 2015, pp. 14-20.
[112] P. Sironi, FinTech innovation: from robo-advisors to goal based investing and gamification. John Wiley & Sons, 2016.
[113] Abdel-Rahman Layla, and Yuli Andriansyah, 'The role of artificial intelligence in modern banking: an exploration of AI-driven approaches for enhanced fraud prevention, risk management, and regulatory compliance', Reviews of Contemporary Business Analytics, vol. 6, no. 1, 2023, pp. 110-132.

reports that are to be submitted to the regulators.[114] This way, the AI is able to ensure that the bank is within the law, which is very important in the current digital world.

Risks of Using AI

Despite the many advantages associated with AI, a significant number of researchers and experts also raise concerns about the possible disadvantages of AI. One of the main disadvantages of AI is the bias associated with it. AI learns and makes decisions based on the data available to it, and if the data is biased, it is possible that AI will also be biased in its decision-making process.[115] For example, if AI is trained with historical data indicating that certain groups of customers are less likely to obtain loans, AI may end up denying loans to those customers disproportionately in the future, thereby encouraging discrimination and unfair treatment of certain classes of customers.

The lack of transparency is also a major disadvantage of AI. AI, particularly deep learning, is a black box, and it is difficult to understand how AI makes decisions.[116] In banking and other areas such as dispute resolution, it is essential for decisions to be transparent so that customers or clients understand the rationale for the decision and, as such, are able to put their trust in AI. AI decisions, if they are not transparent, may also be deemed unfair and against the principles of natural justice.

The other major disadvantage of AI is the lack of data privacy associated with it. AI needs a significant amount of data to make decisions, and this data may include customers' personal information, which may be compromised if AI is mishandled, thereby damaging customers and creating a lack of trust in AI. Experts are of the opinion that data protection legislation is necessary to protect customers' data.

Cybersecurity is another challenge. As AI tools become more advanced, hackers also try to find ways to attack banking systems. Cyberattacks like hacking, phishing, or ransomware can cause huge losses. AI can help prevent some attacks, but it can also be used by hackers to create smarter attacks. Therefore, cybersecurity measures must keep up with AI advancements.

Regulatory gaps are also a problem. Many countries, including India, lack clear rules about how AI should be used in banking. Without proper laws, there is a risk that AI could be misused or cause harm. Experts suggest creating specific laws and guidelines to regulate AI use, especially in areas like dispute resolution.

AI in Dispute Resolution and Customer Service

The role of AI is also being increasingly used in the resolution of customer complaints and disputes. A significant research literature suggests that AI may facilitate the resolution of customer complaints and disputes by speeding up the process of handling customer complaints and

[114] Jon Truby, Rafael Brown, and Andrew Dahdal, 'Banking on AI: Mandating a proactive approach to AI regulation in the financial sector', Law and Financial Markets Review, vol. 14, no. 2, 2020, pp. 110-120.
[115] Adheesh Kadiresan, Yuvraj Baweja, and Obi Ogbanufe, 'Bias in AI-based decision-making', in Bridging Human Intelligence and Artificial Intelligence, pp. 275-285. Cham: Springer International Publishing, 2022.
[116] Arun Rai, 'Explainable AI: From black box to glass box', Journal of the Academy of Marketing Science, vol. 48, no. 1, 2020, pp. 137-141.

disputes.[117] This can be done by categorization of complaints, prioritization of urgent complaints, and even providing possible solutions for customer complaints and disputes. Accordingly, the time taken to resolve customer complaints and disputes is reduced, and customer satisfaction is improved.

Some concerns remain with regards to issues of fairness and justice in the resolution of customer complaints and disputes, such as those in the Indian context, where the Banking Ombudsman Scheme is available for customers to raise complaints against banks. Earlier, the process of resolving customer complaints and disputes is manual, where human officials listen to customer complaints and disputes and make decisions. The incorporation of AI may facilitate the resolution of customer complaints and disputes by aiding in the categorization of customer complaints, providing initial responses, and even understanding patterns in customer complaints and disputes.

However, some researchers suggest that AI should not be exclusively used for the resolution of customer complaints and disputes,[118] as it does not understand human emotions and context, as observed in the resolution of disputes, where human relationships and context play a major role in the resolution of customer complaints and disputes. Secondly, fairness and absence of bias in AI decisions also remain a concern, and AI should be used as a supporting tool for human decision-makers rather than exclusively using AI for the resolution of customer complaints and disputes.[119]

Legal and Ethical Issues

Many researchers stress the importance of the protection of constitutional rights, especially the right to privacy under Article 21 of the Indian Constitution. The judgment in the landmark case of Justice K.S. Puttaswamy v. Union of India[120] recognized the right to privacy as a fundamental right. This places the responsibility upon the banks and the regulatory agencies to ensure that the use of AI applications complies with the right to privacy by ensuring the use of personal information with the consent of the parties involved. Therefore, the use of AI must comply with the principles of the right to privacy.

Another issue that needs to be considered is the issue of transparency and interpretability. The ability of clients to understand the decisions made by the AI system needs to be considered. This is because the use of an uninterpretable decision made by the AI system would be in contravention of the principles of natural justice, which emphasize the right to be heard in the decision-making process. Ethical issues in the use of AI are many, especially the issue of discrimination. Many researchers stress the importance of auditing the use of AI in order to ensure that the use of the technology is fair and non-discriminatory. This is because the use of AI technology in the financial sector could unjustly discriminate against certain classes of people in society, thereby exacerbating the social inequality that already exists in society.

[117] Sonali Chadha, and Sid Sirisukha, 'Empowering Indian banks–AI powered dispute resolution for better customer service', ICL Journal, 2024, p. 85.
[118] Muhammad Ali Malik, 'The Effect of AI Powered Sentiment Analysis on Consumer Complaint Resolution in E-commerce: A Comparative Study of Human Vs AI Meditation Support', PhD diss., Business Studies, 2024.
[119] Moses Alabi. "Ethical implications of AI: bias, fairness, and transparency." Computer Science and Engineering (2024).
[120] Justice K.S. Puttaswamy v. Union of India, AIR 2018 SC (SUPP) 1841.

Another issue that needs to be considered in the use of AI in the financial sector is the issue of accountability. This is because the error in the decision made by the AI system could be attributed to many parties, including the developer of the system, the financial institution, and the regulatory agency. Therefore, the development of the use of the technology needs to be well understood in order to ensure the use of the technology in the financial sector.

International Perspectives and Best Practices

There are many countries working on developing regulatory guidelines and standards for AI systems. The European Union, for example, has introduced the AI Act, which categorizes AI systems based on risk and sets rules for high-risk AI applications, including those used in justice and finance systems.[121] Singapore has developed its national AI governance framework, which prioritizes transparency, fairness, and safety.[122]

India can take inspiration from these international developments. There are experts who have recommended developing local laws and ethics for AI systems used in banking, with the intention of ensuring that technology is used for the greater good of all citizens and their rights are protected.

Results and Discussion

Drawing from extensive literature, legal frameworks, case laws, and expert opinions, the discussion here explores how AI transforms banking operations, enhances dispute resolution, and raises critical legal and ethical issues. The discussion seeks to emphasize the critical point that, even as the use of AI in the banking sector enhances efficiency, accessibility, and personalization, it presents critical risks that need to be regulated properly.

The study indicates that AI has redefined the operations of the banking sector in many ways, resulting in critical efficiencies in the operations of the sector. In particular, the study discloses that AI tools, such as chatbots and virtual assistants, are integral components of the contemporary banking sector.[123]

Enhanced Customer Support and Accessibility

One of the most prominent benefits of AI is its ability to provide round-the-clock customer service through chatbots and virtual assistants. Banks such as SBI, ICICI, and HDFC have developed their own AI-powered chatbots—SIA, iPal, and Eva—that respond instantly to customer queries. These tools handle a large volume of routine questions, such as checking account balances, transaction history, or loan status, significantly reducing waiting times and improving customer satisfaction.

[121] Anat Keller, Clara Martins Pereira, and Martinho Lucas Pires, 'The European Union's approach to artificial intelligence and the challenge of financial systemic risk', in Multidisciplinary perspectives on artificial intelligence and the law, pp. 415–439. Cham: Springer International Publishing, 2023.
[122] Philip L. Frana, 'Assessing Smart Nation Singapore as an international model for AI responsibility', International Journal on Responsibility, 7, no. 1 (2024): 2.
[123] Margherita Mori, 'AI-powered virtual assistants in the realms of banking and financial services', (2021).

The deployment of AI in customer support has made banking services more accessible, especially in a vast and diverse country like India.[124] Customers in remote areas or with limited digital literacy can receive assistance via multilingual AI interfaces, ensuring inclusivity. AI-based virtual assistants also help overcome language barriers, catering to India's linguistic diversity.

Personalization and Financial Advisory

AI enhances customer engagement through hyper-personalization. Robo-advisors analyze individual financial data, preferences, and risk profiles to recommend tailored investment options and savings plans. This democratizes wealth management, making sophisticated financial advice available to a broader audience at a lower cost.

Banks leverage AI to offer targeted marketing, customized credit offers, and personalized product suggestions.[125] This not only improves customer experience but also results in increased loyalty and retention. Such personalized services foster a more inclusive financial environment by catering to diverse customer needs.

Fraud Detection and Risk Management

AI systems excel at analyzing large datasets for detecting suspicious activities and potential fraud in real-time.[126] For instance, AI algorithms monitor transactions for anomalies, such as unusual login locations or sudden large withdrawals, and alert banks immediately. This proactive approach enhances security, reduces losses, and builds customer trust.

In credit risk assessment, AI models evaluate multiple data points—social media activity, transaction history, and behavioral patterns—to make faster and more accurate lending decisions. This facilitates financial inclusion by extending credit to individuals with limited credit histories, provided the AI models are designed fairly.

Automation of Back-End Processes and Compliance

AI automates routine back-end operations such as account opening, verification, and compliance checks, significantly reducing operational costs and human errors.[127] Robotic Process Automation (RPA) handles repetitive tasks efficiently, freeing human resources for more complex decision-making.[128] Furthermore, AI supports regulatory compliance by automating the monitoring of

[124] Khusboo Mittal, 'Leveraging Artificial Intelligence to enhance customer service in cooperative banks in India', (2025).
[125] Chandrima Bhattacharya, and Manish Sinha, 'The role of artificial intelligence in banking for leveraging customer experience', Australasian Accounting, Business and Finance Journal, 16, no. 5 (2022).
[126] N. Al, Faisal, Janifer Nahar, Niger Sultana, and Abdul Awal Mintoo, 'Fraud detection in banking leveraging AI to identify and prevent fraudulent activities in real-time', Journal of Machine Learning, Data Engineering and Data Science, 1, no. 01 (2024), pp. 181–197.
[127] Robert N. Boute, Joren Gijsbrechts, and Jan A. Van Mieghem, 'Digital lean operations: Smart automation and artificial intelligence in financial services', in Innovative Technology at the Interface of Finance and Operations: Volume I, pp. 175–188. Cham: Springer International Publishing, 2021.
[128] C. Vijai, and M. Mariyappan, 'Robotic Process Automation (RPA) in human resource functions', Advances In Management, no. 16 (2023), p. 3.

transactions for anti-money laundering (AML) and combating the financing of terrorism (CFT).[129] AI-driven systems generate timely alerts, ensuring banks adhere to legal requirements and avoid penalties.

AI enhances cybersecurity by continuously monitoring network activity, identifying threats, and responding to cyberattacks swiftly.[130] Biometric authentication, facial recognition, and fingerprint scans powered by AI improve security and reduce fraud, making digital banking safer.

Challenges of AI Integration

Despite the numerous benefits, the research identifies significant challenges and risks associated with AI in banking, especially concerning legal, ethical, and operational aspects.

Algorithmic Bias and Discrimination

One of the most critical concerns is bias inherited from training data. AI systems, if not properly designed, can perpetuate existing social biases, leading to unfair treatment of certain groups. For example, AI models used for credit scoring might unfairly deny loans to marginalized communities if historical data reflects discriminatory practices. Research highlights those biases in AI can violate principles of equality and non-discrimination enshrined in constitutional and legal frameworks. Regular audits, diverse data training, and transparent algorithms are necessary to mitigate this risk.

Lack of Transparency

Many advanced AI models, particularly deep learning algorithms, operate as "black boxes," making their decision-making processes opaque. Customers and regulators demand explainability, especially in sensitive areas like dispute resolution and credit approval.

The inability to explain AI decisions can violate the principles of natural justice, which require that decisions affecting individuals' rights be transparent and justifiable. Courts and regulators are increasingly scrutinizing the explainability of AI systems, emphasizing the need for human oversight and interpretable models.

Data Privacy and Security Concerns

AI systems require vast amounts of personal and financial data, raising concerns about privacy breaches and misuse. The recent case of data breaches and cyberattacks illustrates the vulnerability of banking systems to malicious actors.

India's evolving data protection laws, such as the Digital Personal Data Protection Act, 2023, aim to address these concerns. However, the lack of comprehensive legal safeguards and enforcement

[129] Georgios Pavlidis, 'Deploying artificial intelligence for anti-money laundering and asset recovery: the dawn of a new era', Journal of Money Laundering Control, vol. 26, no. 7 (2023), pp. 155–166.
[130] Goutham Sunkara, 'AI-driven cybersecurity: Advancing intelligent threat detection and adaptive network security in the era of sophisticated cyber attacks', Well Testing Journal, vol. 31, no. 1 (2022), pp. 185–198.

mechanisms remains a challenge, which could undermine customer trust and violate constitutional rights like privacy under Article 21.

Regulatory Gaps and Lack of Legal Frameworks

The research finds that India lacks specific regulations governing AI use in banking and dispute resolution. Existing laws are inadequate to address issues such as liability for AI errors, accountability, and standards for ethical AI deployment.

International models like the European Union's proposed AI Act and Singapore's AI governance framework provide useful references. India needs to develop sector-specific guidelines and oversight bodies to ensure responsible AI adoption.

Cybersecurity Threats

As AI becomes more integrated into banking systems, cybercriminals also exploit vulnerabilities. AI-driven cyberattacks, such as ransomware and phishing, threaten the security of customer data and financial assets.

While AI can strengthen cybersecurity, over-reliance on automated systems without robust safeguards can lead to vulnerabilities.[131] Continuous updates, threat intelligence, and human oversight are essential to counteract evolving cyber threats.

Digital Divide and Inclusivity

The deployment of AI-based systems might unintentionally exclude marginalized groups lacking digital literacy, internet access, or regional language support. This digital divide undermines the goal of financial inclusion and equitable access to grievance redressal mechanisms.[132] AI interfaces need to be designed with inclusivity in mind, incorporating regional languages, accessible formats, and offline options.

Judicial and Legal Developments

The judiciary in India has played a vital role in shaping the legal landscape related to AI, privacy, and digital evidence. Landmark cases like Justice K.S. Puttaswamy affirm the constitutional right to privacy, emphasizing that data collection and processing by AI systems must respect individual rights.[133]

Courts have increasingly recognized that AI decisions must be fair, transparent, and accountable. Judicial decisions have also clarified that electronic evidence, such as transaction logs and chatbot

[131] Khang, Alex, ed., AI-Powered Cybersecurity for Banking and Finance: How to Enhance Security, Protect Data, and Prevent Attacks (CRC Press, 2025).

[132] Abhijit Ghosh and Ankita Bhatia, 'Bridging the Digital Divide: Leveraging Technological Innovations for Inclusive and Sustainable Finance', Financial Innovation for Global Sustainability (2025), pp. 305–336.
[133] Justice K.S. Puttaswamy v. Union of India, AIR 2018 SC (SUPP) 1841.

transcripts, must be supported by proper certification to be admissible in courts, which is crucial for AI-driven dispute resolution.

The courts are also examining the scope of natural justice in AI-based decisions, emphasizing that decisions must be explainable and provide individuals with a fair opportunity to be heard. As AI becomes more prevalent, judicial scrutiny will intensify to ensure that technological innovations do not violate constitutional principles.

Recommendations

The research indicates that AI's future in Indian banking and dispute resolution is promising but requires careful regulation and oversight. Policymakers should develop comprehensive legal frameworks that address AI-specific issues such as liability, bias, explainability, and data privacy. Establishing regulatory sandboxes, as seen in the UK and Singapore, can facilitate experimentation with AI while ensuring compliance with safety standards. International best practices suggest that AI governance should be based on principles of transparency, fairness, accountability, and inclusivity. Regular audits, impact assessments, and stakeholder engagement are necessary to monitor AI systems' performance and address emerging risks. Building awareness among consumers about their rights and the role of AI in banking will foster trust and confidence. Finally, a hybrid model combining AI automation with human oversight is recommended. Human judgment is vital in complex disputes, ensuring that decisions are empathetic, fair, and legally sound.

Conclusion

Research and expert opinions agree that AI has great potential to transform banking and dispute resolution in India. Its benefits include faster services, better fraud detection, personalized advice, and improved regulatory compliance. However, risks related to bias, transparency, privacy, and accountability must be carefully managed. Creating clear laws, ethical standards, and oversight mechanisms is crucial for responsible AI use. Regular audits, transparency in decision-making, and human oversight are necessary to ensure fairness and justice. International best practices can serve as models for India to develop a balanced approach that harnesses AI's power while protecting citizens' rights. Overall, the literature emphasizes that AI can bring many benefits to banking and dispute resolution if used responsibly. The focus should be on creating an inclusive, fair, and safe environment where technology enhances trust and confidence in the financial system.

The integration of AI into banking and Ombudsman services offers immense potential to improve efficiency, accessibility, and customer satisfaction. However, the associated risks—bias, lack of transparency, data privacy, and legal gaps—must be addressed proactively. The legal and regulatory environment must evolve to keep pace with technological innovations, balancing progress with constitutional safeguards. As India advances in AI adoption, the focus should be on responsible innovation, ensuring that AI serves justice, promotes financial inclusion, and upholds individual rights. The future of AI in banking depends on creating a resilient, transparent, and inclusive ecosystem that leverages technology's benefits while safeguarding against its risks.

Chapter – VI

Digital Afterlives: Cyber Vulnerability and Social Belonging

Adrija Nath, Independent researcher, Jadavpur University

Abstract

In contemporary digital societies, cybersecurity is generally grounded as a technical problem, so much so that breaches are measured in financial losses, identity theft is reduced to credit score damage, and phishing is treated as a failure of awareness or digital literacy. Yet for those who experience hacking, impersonation, data breaches, or sudden deplatforming, the entire experience does not end when passwords are reset or accounts are restored. Instead, it initiates what this paper calls as a *digital afterlife* which is a chronic socio-psychological condition in which identity, their sense of belonging, and trust are reshaped by conditions of cyber vulnerability. This study moves the focus away from technical fixes to people's real-life experiences. It argues that cyber victimisation is not just about financial loss, but a deep disruption to a person's social life and sense of stability. Using in-depth interviews with people who have gone through hacking, identity theft, phishing scams, or being removed from online platforms, this study places these experiences within wider sociological ideas such as stigma, trauma, social identity, and trust in institutions. Drawing on Erving Goffman's idea of a "spoiled identity," the paper shows th1at many victims begin to blame themselves after such incidents. They often feel ashamed and describe themselves as careless, naïve, or lacking digital skills. These experiences show that it is not just personal embarrassment at play, but also a broader cultural belief that online safety is primarily an individual's responsibility. As a result, being a victim of cybercrime becomes a stigmatised condition, something people often keep hidden, rarely speak about openly, and one that is closely tied to judgments about a person's competence in the digital age.

The paper also tries to understands cyber victimisation as a kind of identity trauma. Unlike ordinary property crimes, digital breaches unsettle the line between who a person is and how they appear online. When someone's email is hacked, their social media is impersonated, or their account is suspended without any explanation. Their digital self, which in today's world is very central to work, relationships, and public life, becomes fragile and exposed. Participants describe feeling detached from themselves and exposed, almost like as if they have lost control over their own story and identity. Especially in cases of suspension of profiles, people describe a feeling of being erased from public life. Losing access to a platform is seen as a social exclusion from spaces where their friendships, work opportunities, and political voice are rooted.

The study also looks into how cyber vulnerability changes people's trust in institutions. Many victims describe a second layer of harm when they face unclear and unsupportive customer service systems, and a sense of bureaucratic indifference. Banks, online platforms, and government agencies often come across as distant, impersonal, and hard to deal with. As a result, people's trust in these institutions is often shaken. The effects of a breach therefore go beyond personal anxiety,

leading to a wider sense of doubt and mistrust toward the systems and institutions that govern digital life.

The study also reflects that the long-term effects of cyber incidents are not experienced equally. People with limited digital skills, insecure jobs and thus people who are marginalised socially often face stronger stigma and take much longer to recover. For instance, the freelance workers and content creators, when removed from a platform means endangering their livelihood. For women and people from marginalised gender identities, hacking and impersonation often bring extra reputational damage and moral scrutiny. Cyber vulnerability therefore works through the already existing social inequalities, making already precarious lives more unstable. Using ideas about social identity and trust in institutions, the paper argues that digital harm is deeply social and relational. By focusing on people's life experiences, the study questions mainstream cybersecurity approaches that mainly gives more importance to prevention and financial damage, while overlooking the difficult social recovery individuals go through after a digital violation.

In terms of method, the study uses narrative analysis to follow how people tell the story of a cyber incident over time, from the initial shock and confusion to the longer-term ways they cope, with it. The digital aftermath is marked by a sense of tension, people still have to rely on digital systems in everyday life, yet they trust them much less than before. In the end, the paper argues that as more of our social, economic, and civic lives move onto digital platforms, cyber vulnerability should be understood as a serious form of social harm, not just a technical risk.

Keywords: cyber vulnerability, cyber victimization, extra reputational damage and moral scrutiny.

Introduction

Digital technologies now shape much of everyday social life. Emails structure work routines in offices and institutions. Messaging platforms sustain friendships and maintain contact across distance. Social media profiles often function as public markers of identity where individuals present themselves to others. For many people, the internet is therefore not a separate space that exists outside social life. It has become one of the main arenas where relationships are maintained, reputations are built, and opportunities emerge. In this sense, digital presence is closely tied to how individuals are seen and recognised by others. Despite the growing importance of digital spaces in everyday life, discussions around cybersecurity are still framed largely as technical matters. When incidents such as hacking, phishing scams, identity theft, deplatforming are discussed, the focus usually falls on data protection, financial damage, or the need for better digital awareness. Institutional advice often centres on practical precautions such as creating stronger passwords or being more careful while using online platforms. While these measures are important, they also shape how cyber incidents are understood. They come to appear as problems that can be fixed once accounts are recovered or financial losses are managed, leaving little room to consider how such experiences might affect people beyond the immediate moment of the breach.

Yet these ways of thinking about cyber incidents leave out an important part of the story: how such events are actually experienced by the people who go through them. For someone whose account

has been hacked or whose profile has suddenly been taken down, the experience rarely ends when the technical issue is fixed.Many people continue to feel uneasy about their digital presence for some time afterwards. Some speak of embarrassment, especially when others become aware of what happened. Others describe a sense of exposure, as if something personal had been briefly taken out of their control. It is also common for people to begin questioning their own abilities with digital systems, wondering whether they missed some warning sign or made a mistake along the way. For a while after the incident, even routine actions online can feel slightly different. Activities that were once automatic may now be approached with caution.

This chapter approaches cyber incidents from this experiential perspective. Instead of looking at the technical side of hacking or data breaches, the focus here is on how individuals themselves understand and recount what happened to them. What matters is not only the incident, but what follows it. How do people make sense of such moments? How do they respond to the sudden awareness of vulnerability? And how does the experience slowly reshape the way they relate to the digital systems that are part of their everyday lives?

To describe this prolonged aftermath, the chapter introduces the idea of a *digital afterlife*. The phrase refers to the state in which people continue to live with the effects of a cyber breach even after the immediate disruption seems to have been resolved. When someone's account is hacked, their identity is impersonated, or they fall prey to cybercrimes, the event often leaves behind traces that linger for some time. People begin to see their digital identity differently. Their trust in the systems that organise online life may also change.

The chapter explores three central questions. First, how do individuals narrate their experiences of cyber victimisation? Second, in what ways do such incidents influence their sense of identity and belonging within digital spaces? Third, how do these experiences affect trust in institutions such as banks, online platforms, and other systems that organise digital life?

This chapter combines primary and secondary sources to explore how people experience cyber harm in everyday life. The primary data comes from qualitative interviews with individuals who had faced incidents such as hacked accounts, impersonation, cyber bullying or loss of access to digital platforms. These conversations focused on how participants understood and coped with these experiences in their daily lives. Alongside this, they also chapter draws on existing academic literature on digital identity, trust, and inequality. Bringing together personal accounts and scholarly work helped situate these experiences within the framework of social realities of life in a growing digital world.

The digital self and networked identity

Digital technologies have slowly become part of how people show themselves to others and stay connected in everyday life. Over time, these spaces begin to hold traces of everyday interactions such as messages, photographs, exchanges, payments, and memories. Because of this, digital accounts rarely feel separate from the individual who uses them. They become one of the ways through which people are recognised and remembered by others. Our digital presence today therefore, is much more than just a technical profile attached to a platform. It has become a part of how individuals see themselves and how they are seen by others.

Long before the rise of digital platforms, scholars such as *Erving Goffman*[134]suggested that everyday life involves a constant process of presenting oneself to others. People adjust how they speak, behave, and appear depending on the situation and the audience they face. A classroom, a workplace, or a gathering with friends each call for slightly different forms of self-presentation. Digital platforms extend this process into new settings. A profile page, a carefully chosen photograph, or a short biography now becomes part of how individuals introduce themselves to wider audiences. Even seemingly small choices like what to share and what to hide can shape how others perceive a person online. At the same time, these digital environments differ from earlier social settings as well. Interactions that once disappeared after the moment passed, now leave behind records that may remain visible for years. Messages, photographs, documents and posts accumulate over time, forming a kind of archive of a person's digital presence. This awareness often influences how people manage their online behaviour, encouraging a more careful form of self-presentation.

The growing importance of digital networks has also changed the spaces through which relationships are maintained. *Manuel Castell*[135]*s* describes contemporary societies as being increasingly organised through networks of information and communication technologies. Work relationships, friendships, public conversations, debates, shopping, transactions frequently unfold within these digital networks. Participation in such spaces is often necessary for staying relevant and connected with the world around us.

Anthony Giddens[136] too has offered a useful way of thinking about these developments through his idea of the "reflexive projection of the self." According to Giddens, individuals continuously reflect upon and shape their own life stories as they move through different experiences. Identity develops gradually through the choices people make, and the narratives they construct about themselves. Digital platforms have become one of the places now, where this process unfolds.

As a result, the boundary between online and offline life often becomes difficult to draw clearly.

Several participants in this study have described how closely their digital presence had become tied to their sense of everyday stability. One respondent, a freelance photographer in her late twenties, explained that most of her professional communication took place through a single email account and an online portfolio linked to social media. When she temporarily lost access to the account following a hacking attempt, she described the experience as deeply unsettling.

"It wasn't only about the account," she reflected during the interview. *"For a few days it felt like I had disappeared from the place where my work and contacts actually exist."*

Sherry Turkle's[137] work on digital life offers a helpful way of thinking about how these spaces gradually become part of how people understand themselves. In her writings on life with

134 Erving Goffman, *The Presentation of Self in Everyday Life* (Penguin Books 1959).

135 Manuel Castells, *The Rise of the Network Society* (2nd edn, Wiley-Blackwell 2010)
136 Anthony Giddens, *Modernity and Self-Identity: Self and Society in the Late Modern Age* (Polity Press 1991)
137 Sherry Turkle, *Life on the Screen: Identity in the Age of the Internet* (Simon & Schuster 1995).

computers and the internet, Turkle suggests that digital environments create new spaces where individuals can express and sometimes even experiment with different parts of their identity. Through conversations, posts, photographs, or messages, people begin to share fragments of their everyday lives online. Over time, these fragments form a visible presence through which others come to know them. Turkle points out that this process is not only about communication across distance. Online spaces also allow individuals to find communities that share similar interests, experiences, or concerns. Someone who feels unheard or unnoticed in one setting may find recognition in another digital space. And any disruption to this side of our lives gives meaning to what we can call a *digital afterlife* which is like a lingering condition that can follow a cyber incident long after the immediate problem appears to have been resolved.

Experiences of cyber victimisation are often accompanied by feelings that go beyond the immediate shock of the incident. Alongside confusion or anxiety, many individuals also describe a quieter emotion that surfaces soon afterwards that is embarrassment. Even when people recognise that they have been targeted by sophisticated scams or hacking attempts, they frequently speak as though the event reflects some personal mistake on their part. In conversations about cyber incidents, it is not uncommon to hear individuals describe themselves as having been "*careless*," "*naïve*," and "*not careful enough*." The language people use to recount these experiences often carries a tone of self-blame.

One way to understand this reaction is through Goffman's [138] discussion of stigma. He described stigma as a situation in which certain experiences or attributes come to be seen as discrediting within a particular social setting. When this happens, individuals may begin to feel that something about them has been judged negatively by others. They may try to conceal the experience or even internalise the judgement themselves. In such situations, the issue is not simply the event that took place, but the social meaning that is attached to it.

Cyber victimisation can easily take on this kind of meaning. In an environment where digital competence is increasingly taken for granted, people are often expected to know how to protect themselves online. Advice about cybersecurity regularly emphasises individual vigilance for instance recognising suspicious emails, avoiding unknown links, managing passwords carefully and report suspicious calls. While such guidance is important, it also creates a subtle expectation that responsible users should be able to avoid these risks. When something does go wrong, the incident can therefore feel less like a random misfortune and more like a personal lapse.

One participant in this study, a postgraduate student, recalled how uncomfortable she felt telling her friends that her email account had been compromised after she responded to a phishing message. Although she later realised that the email had been carefully designed to appear legitimate, her first reaction was embarrassment. During the interview she explained*:*

"I kept thinking I should have noticed something was wrong. It made me feel like I had been careless, even though the message looked very real at the time."

[138] Erving Goffman, *Stigma: Notes on the Management of Spoiled Identity* (Prentice-Hall 1963).

Another participant described a similar hesitation when discussing a social media account that had been temporarily taken over by someone impersonating him. Even after the account was recovered, he admitted that he had not told many people about the incident.

"It felt strange to talk about," he said. *"You start wondering what others will think. Maybe they will assume you weren't paying attention or you are a technological handicap"*

These reactions they show how cyber incidents can quickly become framed as individual failures. Instead of viewing themselves simply as victims of a crime, individuals may begin to interpret the event as evidence that they were not sufficiently cautious or knowledgeable. In Goffman's terms, the experience can threaten a person's sense of competence and produce what he described as a "*spoiled identity*." The individual may feel that something about their social image has been damaged, even if the incident itself was beyond their control. This sense of embarrassment often leads people to limit how openly they speak about what happened. Some respondents mentioned that they only told a small circle of close friends, while others avoided discussing the incident altogether. The reluctance to share these experiences is shaped partly by the fear of judgement.

One respondent, a young professional working in a technology-related field, spoke about how difficult it felt to admit that she had been the target of a phishing attempt that briefly compromised her account.

"*Because I work in a field where everyone is supposed to be digitally aware, it felt almost worse,"* she explained. *"I kept thinking that people would expect me to know better."*

Her reaction reveals how expectations surrounding digital competence can heighten this sense of personal failure when something goes wrong.

News reports and public campaigns frequently emphasise caution and awareness, sometimes implying that those who fall victim to scams may have ignored obvious warning signs. Although such messages aim to encourage safer online practices, they can unintentionally reinforce the belief that victims are responsible for their own misfortune. This cultural expectation is what places us in a difficult position making it an intersectional issue.

Identity Trauma and the Fragility of the Digital Self

When a cyber incident takes place, the first concern is often practical that is recovering the account securly and preventing further damage. Yet for many individuals, the experience does not end here. Over time, the incident can begin to affect how people think about their own digital presence and their sense of control over it. Digital accounts often hold much more than access to a platform and because of this, losing control over such an account can feel upsetting in ways that extend beyond technical control.

Anthony Giddens[139] understands this reaction through the idea of ontological security. He uses this term to describe the sense of stability that individuals rely upon to move through everyday life

[139] Anthony Giddens, *The Consequences of Modernity* (Polity Press 1990).

with confidence. People develop routines and expectations that allow them to trust that the world around them will remain reasonably predictable. This sense of continuity is not always something individuals consciously think about; rather, it quietly supports the feeling that everyday life makes sense and remains under some degree of control. Digital systems have gradually become part of these everyday routines. Checking email, responding to messages, or logging into accounts or making online purchases or payments often happens almost automatically. Over time, these activities create a sense of familiarity with the digital environments people rely upon. Whenever there is a breach in this environment without warning, that sense of familiarity can be deranged. The experience can make individuals aware that something they once trusted can suddenly become uncertain.

Participants in this study when asked about this described this moment of disruption quite vividly. One respondent recalled discovering that someone had briefly gained access to her social media account and had begun sending messages to people in her contact list. Although the situation was resolved within a few hours, she described the experience as –

"For those few hours, I kept thinking that someone else was speaking as me. It felt strange to imagine that people might read those messages and think they came from me."

Such experiences point to the fragile boundary between a person and the digital representations through which others encounter them. When someone impersonates an account or gains access to a personal profile, that boundary becomes dimmed. The account continues to exist in public view, but the individual behind it may no longer have control over what appears there. In that moment, a digital identity can begin to feel detached from the person it represents.

Kai Erikson's[140] work on trauma provides another way to think about this sense of disruption. Erikson described trauma not simply as a response to physical harm but as an experience that shaken a person's sense of continuity and belonging. Trauma, creates a feeling that the familiar order of things has been disturbed. While cyber incidents differ greatly from the kinds of disasters Erikson originally wrote about, the underlying idea of disruption can still be useful. For some individuals, losing control over their digital identity produces a moment of disorientation. Accounts that once felt like stable extensions of their everyday life begin to appear fragile and uncertain.

Another participant described this experience after his professional email account was briefly compromised. Although no significant damage occurred, he remembered feeling uneasy even after the account was secured again. *"I kept thinking about how much of my work life is in that account,"* he said.

"If someone can step into that space, even for a short time, it makes you realise how exposed things actually are."

[140] Kai Erikson, *Everything in Its Path: Destruction of Community in the Buffalo Creek Flood* (Simon & Schuster 1976).

Account suspension can produce a different but related feeling. In these cases, individuals do not lose control because someone else takes over their identity. Instead, they find themselves suddenly removed from a digital space where they previously had an active presence. One respondent explained that losing access to a platform where she had spent years sharing photographs and writings made her feel as though a small part of her social life had been abruptly cut off.

"It was strange," she reflected.

"It wasn't just about posting pictures. It felt like a place where my life had been unfolding had suddenly closed its doors."

These reactions they show how digital identities today, have become intertwined with everyday social interactions. Some participants also described a lingering awareness of vulnerability after the incident had passed. Even once accounts were recovered or secured, the earlier sense of certainty rarely returned in the same form. Routine activities such as logging in or sharing information began to carry a new layer of caution.

Institutional Trust and Bureaucratic Indifference

When people become victims of cyber incidents, the harm rarely ends with the initial breach. The experience often extends into a quieter struggle that involves navigating institutions that are supposed to help. Banks, online platform, customer support systems, and cybercrime authorities become the next points of contact. In theory, these organisations exist to restore order and protect individuals. In practice, many victims encounter something very different. Automated messages, complicated procedures, delays in follow ups and distant forms of communication t feel impersonal and slow. What begins as a technical problem quickly becomes a crisis of trust.

Trust plays a central role in how modern digital environments function. Individuals believe the systems managing their online activities are reliable. The idea of trust explored by *Niklas Luhmann*[141] helps explain this dynamic. Luhmann suggested that trust reduces complexity in everyday life. In a world filled with uncertainty, people rely on trusted institutions and systems so that they do not have to question every interaction or decision. Without this reliance, ordinary activities would become overwhelming. Digital infrastructures depend heavily on this kind of trust. Most users do not fully understand the technical mechanisms behind online platforms. Instead, they assume that these systems will work as promised. The presence of passwords, verification processes, and security alerts gives the impression that protection is built into the structure of the system. If things run smoothly, this trust remains largely invisible.

Problems arise when the system fails. A hacked account, fraudulent transaction, stolen identity or cyber bullying disrupts the quiet confidence people place in these systems. Suddenly, what once felt routine becomes uncertain. At this moment, victims often turn to institutions for assistance. The expectation is simple that is someone will listen, investigate, and help restore what has been lost. However, the reality is quite different.

[141] Niklas Luhmann, *Trust and Power* (John Wiley & Sons 1979)

Let us consider the experience of contacting customer support after an online account has been compromised. Many platforms rely heavily on automated systems to manage large volumes of complaints. Victims are frequently directed through layers of help pages and chatbots before they reach an actual person, if they reach one at all. While these systems are designed for efficiency, they can feel deeply frustrating for someone who is already dealing with anxiety and confusion. The language of automated responses often sounds neutral and procedural, yet to victims it can appear dismissive or indifferent. A similar experience can emerge in interactions with banks following digital financial fraud. Victims may spend hours explaining their situation, submitting documentation, and waiting for internal investigations. During this period, uncertainty lingers. The individual is left wondering whether their money will be recovered, whether the institution believes their account of events, and whether the system truly prioritises their protection. Even when banks eventually resolve the issue, the process itself can leave a lingering sense of vulnerability.

The concept of trust in modern systems developed by *Anthony Giddens*[142] provides a useful way to understand this experience. Giddens argued that modern societies rely on what he called *"abstract systems"* which is a large, complex networks of expertise and technology that people depend on without fully understanding how they operate. For examples, we can take financial institutions, technological infrastructures, digital platforms and online payments. People interact with these systems daily, trusting that they are managed by competent experts. Cyber incidents expose the fragile nature of this reliance.

When something goes wrong, individuals suddenly confront the distance between themselves and the systems they depend on. The result is a feeling of disconnection. Victims may feel that their personal crisis has been absorbed into an impersonal administrative process. Further, interactions with cybercrime authorities can reinforce this sense of distance. Victims often describe feeling as though their experiences have been reduced to case numbers within a system that moves slowly and communicates little.

This bureaucratic character of institutional response can be better understood through the work of Max Weber[143]. Weber described bureaucracy as a rational organisational structure designed to maximise efficiency through rules, procedures, and hierarchies. Such systems are meant to produce consistency and fairness by treating cases according to standardised guidelines rather than personal judgement.

Yet bureaucracy has a paradoxical side. While its structured procedures are meant to ensure fairness, they can also produce emotional distance. Individuals facing urgent or distressing situations may feel that the system responds too slowly or too rigidly. The rules that keep institutions organised can also make them appear indifferent to the personal realities behind each case. This phenomenon can be described as secondary victimisation. The initial attack may involve financial loss, identity misuse, or the invasion of personal information. The institutional response,

[142] Anthony Giddens, *The Consequences of Modernity* (Polity Press 1990) 27–29
[143] Max Weber, *Economy and Society: An Outline of Interpretive Sociology* (University of California Press 1978).

however, can unintentionally deepen the sense of vulnerability. Victims who struggle to obtain clear answers or timely support may begin to question whether the systems they trusted are truly capable of protecting them.

Inequality and Differential Cyber Vulnerability

Cyber harm is almost never socially neutral. While digital spaces often present themselves as open and universal, the risks and consequences attached to cyber incidents are quite unevenly distributed. People's social positions, play a significant role in determining how deeply they are affected. For some individuals, a cyber incident may be frustrating but manageable. For others, it can threaten their livelihood, reputation, or sense of safety. To understanding this uneven landscape it's important to grasp the of *digital divide*. This concept is not only about who has access to the internet but also about who feels confident accessing the digital systems. Individuals with strong digital literacy often recognise warning signs such as suspicious login alerts or phishing attempts. They may also know how to recover accounts, report incidents, or secure their information quickly. But those users who are less familiar with digital technologies often struggle to interpret what has happened or how to respond.

One participant, a 52-year-old tailoring service provider who had recently begun accepting online orders, described how difficult it was to deal with a phishing scam that compromised her messaging account:

"My customers started telling me they received messages asking for advance payments. I didn't even know someone had taken control of my account. My son had to sit with me and figure out what was happening."

Her experience reflects the uncertainty that can accompany limited digital familiarity. The breach itself was stressful, but the process of understanding the problem felt equally overwhelming. Instead of acting independently, she relied on younger family members to learn the technical aspects of account recovery.

Economic vulnerability also shapes how cyber incidents unfold. Freelancers and gig workers often depend entirely on digital platforms for visibility, payments, and client relationships. When these platforms malfunction or accounts become inaccessible, the consequences are immediate. One participant who worked as a freelance translator described how a security flag temporarily froze her payment account:

"The platform said they were reviewing my account for suspicious activity. During that time, I couldn't withdraw anything. I had already completed projects, but the money was stuck there. For two weeks I was just waiting and hoping they would clear it."

For her, the delay was not merely inconvenient. As someone without a fixed salary, access to that payment was crucial for covering monthly expenses. The platform's automated security measures, though designed to prevent fraud, ended up magnifying the precariousness of her situation.

Gendered expectations too influence how cyber harm is experienced. Several women participants spoke about the reputational anxieties that followed digital incidents. The fear was not simply about technical loss but about how others might see them. A postgraduate student described the distress she felt after discovering that someone had circulated edited screenshots of her social media posts in a private online group:

"They changed the captions and made it look like I had written things I never said. Even after I explained it was fake, I kept worrying about who had seen it."

In her case, the harm came less from the technological act itself and more from the social consequences attached to being the secondary sex. For individuals with marginalised gender identities, cyber incidents often intersect with the already present experiences of discrimination. Digital platforms may offer opportunities for visibility and connection, yet they can also become spaces where harassment and targeted attacks occur.

One interviewee, who identifies as a transgender man and runs a small online art page, described receiving repeated attempts by strangers to access his account:

"At first I thought it was random hacking attempts. But then I realised people were also sending messages saying things like 'people like you shouldn't be online.' That's when it started feeling personal."

The experience blurred the boundary between a technical threat and social hostility. What might have been let go as routine hacking attempts in another context became a question linked to his identity.

Another participant, a young non-binary student, spoke about the vulnerability of relying on online communities for emotional support:

"Most of my support system is online. When someone reported my account and it got temporarily restricted, I felt cut off from everyone who understood me."

For them, the platform was not just a communication tool but a crucial space for belonging. Losing access, even briefly, created a sense of isolation. These real life experiences reveal how cyber vulnerability is closely tied to broader social inequalities.

Living in the Digital Afterlife

Cyber incidents are often spoken about as - something goes wrong, the account is recovered, and life continues. But as we have understood for people the experience does not end when the technical part of it is resolved. The breach leaves behind a residual thought that their digital identity can be accessed, breached, altered, impersonated, and taken away without warning. The repercussion becomes a kind of *digital afterlife*. The event itself may pass, but the feeling remains.

For many victims, the experience begins with *shock*. The first moment of discovering that something is wrong rarely feels clear or immediate. Strange notifications, unfamiliar posts or unexpected transactions create confusion. Digital accounts usually feel stable and predictable, so

the idea that someone else may be inside them is difficult to process. Once the reality of the breach becomes clearer, the emotional response often shifts inward. The second stage is quite often marked by *self-blame*. Instead of focusing only on the attacker or the system failure, individuals begin questioning their own behaviour. They wonder whether they could have done things differently. Even when cyber incidents are complex and difficult to prevent, victims often feel that they should have been more careful. This quiet sense of embarrassment can make people hesitant to talk openly about what happened.

After this comes a more practical but often frustrating phase which is the *struggle to regain control.* Individuals must deal with these platforms in order to recover their accounts or secure their information. While these processes are designed to protect users, they can feel slow and impersonal. The breach itself may have taken minutes, but the effort to repair its consequences can last much longer. Even after access is restored, many people find that their relationship with digital spaces has changed. A period of *social withdrawal* often follows. This does not always mean abandoning online platforms altogether, but participation becomes more cautious. People post less or think twice before engaging online. Activities that once felt effortless begin to carry a small sense of risk.

Over time, most individuals reach a stage of *adaptation.* They develop new habits which could be having stronger passwords, additional security settings, or more careful digital routines. These adjustments help restore a degree of stability, but they don't quite bring back the original sense of trust. Instead, digital life continues with a constant awareness that things can go wrong. In such a way, cyber harm often becomes less like a temporary interruption and more like a *chronic condition.*

Conclusion

Cyber incidents are often treated as technical mishaps. Yet the experiences discussed in this chapter suggest that the damage does not stay within the realm of tech. When a digital account is compromised, what is shaken is not just a platform but the person attached to it. Digital profiles today hold fragments of our identity. When these spaces are breached, the boundary between the self and its online representation suddenly feels brittle, leaving individuals feeling exposed or oddly detached from versions of themselves circulating online. At the same time, cyber incidents quietly reshape belonging. Because friendships, work, communities, and services now exist largely on platforms, losing access even briefly can feel like being locked out of a room where social life continues without you. The result is sometimes withdrawal, with individuals becoming more cautious about what they share and how they participate online from then on. These experiences also affect trust in institutions. Cyber vulnerability also tends to mirror existing inequalities- freelancers dependent on platforms risk losing income, individuals facing social scrutiny risk reputational damage, the differential experiences that comes with belonging from different social locations and digitally less literate users struggle more to regain control. For some, a breach is just an inconvenience but for many, it can deeply interrupt our everyday life.

These patterns reveal that cyber vulnerability is no longer just a technical risk but a social condition. As more of identity, work and relationships move onto digital platforms, cyber harm increasingly reflects broader structures of power, inequality, and institutional trust. In a world where so much of life unfolds online, protecting digital security is no longer just about better passwords or stronger firewalls rather it is about recognising cyber vulnerability as a form of *structural social harm* embedded within the very systems that organise contemporary social life.

Chapter – VII

Offence–Defence Asymmetry in the Age of Artificial Intelligence: Implications for the Cyber Threat Landscape

Apala Ghosh, PhD Research Scholar, Department of International Relations, Jadavpur University

Abstract

The swift integration of artificial intelligence (AI) into digital architecture is generally considered a transformative process in the area of cybersecurity. However, much of the current literature on the topic considers AI either as a technology that amplifies cyber capabilities or as a new form of governance challenge that demands regulation. This chapter presents a different case: that AI is best conceptualized as a structural variable that alters the balance of offence and defence in cyberspace. By using the theoretical framework of International Relations (IR) theory, specifically the literature on offence-defence asymmetry and security dilemmas, this chapter contends that the current form of AI systems aggravates the pre-existing imbalance between offence and defence in cyberspace, thereby exacerbating the instability of the global cyber threat environment. The offence-defence balance approach, which has traditionally been used to assess conventional and nuclear security, examines the prevailing technological and strategic conditions of a given environment to determine whether it is more advantageous to launch an attack or defend against one. When the costs of offence are lower than those of defence, states are likely to pursue pre-emptive strategies, the threat of escalation rises, and deterrence policy becomes unreliable. Cyberspace has long been considered an offence-preferred environment, where attackers need only find one vulnerability to breach the system, while defenders must protect a complex and constantly evolving target, where attribution is slow and uncertain, and where offensive capabilities can be easily copied at a low cost. The pre-AI cyber environment is already characterized by asymmetrical vulnerability. Cyber operations are different from traditional warfare in that they leverage systemic interconnectivity, civilian infrastructure, and dual-use technology. The defensive effort is high due to the expansive and constantly updated nature of digital networks, while the offensive community, whether state or non-state, enjoys the benefits of obscurity and anonymity. These characteristics establish what can be termed a persistent offence bias in cyberspace. The chapter will establish this baseline in order to evaluate the impact of AI on the balance and in which direction. Second, the chapter will explore how particular AI capabilities increase the reach of the offensive community. AI-enabled vulnerability scanning, code generation, and machine learning-driven reconnaissance decrease the technical skill set and time previously required for sophisticated cyber operations. Generative tools facilitate scalable and context-specific social engineering attacks, reducing the traditional trade-off between credibility and scalability in phishing and fraud. Adaptive malware can iterate independently in response to defensive efforts, improving the resilience and longevity of intrusions. Taken together, these trends reduce the barriers to entry for malicious actors, accelerate attack cycles, and improve the strategic reach of relatively resource-constrained actors.

Notably, the chapter will show how AI interacts with existing cyber structures to increase the flexibility of the offense. Third, the chapter will critically evaluate whether defensive AI mitigates these trends. Advocates claim that machine learning improves anomaly detection, threat intelligence analysis, and real-time response capabilities. While these tools improve defensive monitoring, they are subject to structural and institutional constraints. Defensive AI systems require high-quality training data, produce a high rate of false positives, and require human direction within legal and administrative frameworks. Furthermore, defensive infrastructures are typically situated in complex public-private ecosystems, which impede coordination and response. By contrast, the offensive community enjoys fewer normative or institutional constraints. The imbalance is not only based on technical capability but also on organisational and regulatory environments that shape the use of AI. On the basis of this comparative analysis, the chapter goes on to examine the wider implications of international security. If AI further expands the offence-defence gap in cyberspace, several implications follow. First, the risk of escalation could rise as the condensed time frames for detection and response leave less time for decision-makers to deliberate. Second, deterrence becomes more problematic as the delay between attacks and attribution trails the rapid execution of AI-enabled attacks. Third, the proliferation of superior capabilities to non-state actors further cloud state-centric visions of cyber stability. The chapter thus places AI in the context of debates on strategic stability, contending that cybersecurity must be understood not only as a technical challenge but also as a space of evolving power relations, and explores structural avenues for mitigating AI-driven asymmetry through architectural reform, collective defence coordination, improved attribution, and normative adaptation.The aim is to offer a framework by which policymakers and scholars might assess future technological developments without recourse to either alarmism or technological optimism.

Keywords - Offense–Defence Asymmetry, Artificial Intelligence (AI), Cybersecurity, Cyber Conflict, Strategic Stability, Emerging Cyber Threats.

Introduction

The integration of artificial intelligence (AI) into digital infrastructure has been broadly described as a transformative moment for the field of cybersecurity. However, the current state of debate on this issue tends to vacillate between a vision of technological utopianism and a more alarmist vision of the future, without properly locating the phenomenon of AI within the existing body of literature on international security analysis. This chapter will contend that AI should not be considered simply a new “emergent risk” or “governance challenge” for cybersecurity. Rather, it should be considered a new “structural variable” that alters the balance of offence and defence in cyberspace.

The balance of offence and defence is a foundational concept in the field of International Relations (IR) theory, and it seeks to assess the prevailing strategic and technological conditions in a given international system as to whether they are more conducive to offensive or defensive action. When a system is characterised by a structural advantage for offence, this tends to increase the likelihood of escalation and undermine the stability of deterrence. Cyberspace has long been a system that is characterised by offence-preferential traits such as asymmetrical vulnerability, the difficulty of attribution, and the systemic interconnectedness of digital networks.

This chapter will argue that AI tends to further exacerbate these asymmetries by reducing the costs of innovation, increasing the tempo of operations, and diffusing advanced capabilities beyond the traditional state actor.

The main argument that will be made in this chapter is threefold. First, cyberspace is already a system that is characterised by a structural advantage for offence. Second, the current state of AI technology tends to increase this imbalance in a manner that is more pronounced on the side of offence than on the side of defence. Third, this has a number of implications for strategic stability, escalation, and the diffusion of power in international politics.

By locating the analysis of AI within the existing body of literature on offence-defence theory and applying it to the phenomenon of AI-enabled cyber operations, this chapter hopes to provide a more structured explanation of how technological change tends to alter the strategic equilibrium of international politics.

Section I

Reconsidering the Offence–Defence Balance in Cyberspace

The incorporation of artificial intelligence into digital systems has sparked an increasing amount of commentary speculating about the potential for transformation in the nature of cyber warfare. However, much of this literature is conducted without specifying the theoretical framework in which the potential for transformation is being measured. In order for AI to be considered a material factor, it must be considered not only as an innovation in technology but also as a factor within pre-existing frameworks of international security studies. Of these, the offence-defence balance is a particularly useful point of departure[144]. Rather than asking whether AI is "dangerous" or "useful," the offence-defence balance asks a more systematic question: does AI change the relative ease of attack and defence in cyberspace?

The offence-defence balance has long been used in the conventional and nuclear contexts. At its heart is a simple but profound observation: the balance between offence and defence determines the strategic incentives for conflict, arms racing, and strategic miscalculation. When offence is perceived to be superior, states are more likely to pursue pre-emptive strategies and to view defensive preparations as threatening. When defence is superior, the price of aggression increases and deterrence stabilizes strategic interaction. The balance itself is not simply a descriptive tool for military capability but shapes strategic behaviour through perception and expectation.

Applying this framework to cyberspace demands a nuanced conceptual translation. In contrast to territorial space, cyberspace is not geographically bounded but rather defined by its networked architecture. Occupation does not involve control of territory but rather access to systems and data. The relevant strategic question is therefore whether it is more structurally feasible to gain access to and exploit these networks than to comprehensively protect them. By this measure, the offence in cyberspace is simply gaining unauthorized access, obtaining or manipulating information, or disrupting its functionality. Defence is simply preventing unauthorized access, detecting intrusion, and restoring system integrity.

[144] Ben Buchanan, *The Cybersecurity Dilemma: Hacking, Trust and Fear Between Nations* (Oxford University Press 2017).

Even before the advent of more sophisticated AI systems, cyberspace has shown characteristics that make defensive primacy difficult. Digital systems are complex and interconnected. Software development is a layered process, involving third-party libraries and constant updates. Each layer is potentially vulnerable. Because digital systems are constantly evolving, the defender must deal with an ever-expanding attack surface. The defender's problem is therefore systemic and ongoing. The attacker, on the other hand, needs only one point of vulnerability. This is not a guarantee of success, but it certainly establishes an asymmetry.

The balance of offence and defence in cyberspace is also affected by information asymmetry. Attribution of cyber-attacks is often difficult and delayed. Unlike conventional attacks, which often betray their origin through force projection, cyber-attacks can be routed through multiple sources and masked by technology. This makes deterrence credibility difficult to establish. If retaliation is dependent on successful attribution, and attribution is difficult or delayed, the cost of the offence is perceived to be lower. Structural ambiguity therefore combines with vulnerability to tip the balance in favour of the offence.

Nevertheless, the existence of an offence bias in cyberspace does not necessarily mean that innovation in defence is a futile endeavour. The balance between offence and defence is not static. Encryption protocols, multi-factor authentication, intrusion detection systems, and international norms have all changed the character of cyber engagement over time. The challenge is not to assert that cyberspace is necessarily offensive in nature, but to assess how technological change shapes relative costs.

Artificial intelligence must be placed in the context of this constantly shifting balance of power. AI systems improve pattern recognition, facilitate analysis, and support adaptive responses. These characteristics are dual-use. They can improve anomaly detection and defensive monitoring; they can also support automated reconnaissance, exploit development, and deception[145]. The strategic effects of AI depend on whether it upends the ratio of offensive to defensive effort. The role of effort must not be dismissed. In offence-defence theory, the key variable is not capability but cost. A technology that lowers the cost of performing offensive actions more than it lowers the cost of maintaining defence will shift the balance even if both sides are better off in absolute terms. Marginal change, not absolute change, is the key to structural advantage. In cyberspace, cost is more than just money. It includes time, expertise, coordination, and institutional constraint. Traditional cyber offence has required technical expertise and careful reconnaissance. Cyber defence requires constant monitoring, organizational coordination, and compliance with legal and regulatory requirements. In assessing the effects of AI, the key question is whether it changes these cost factors symmetrically.

A further conceptual clarification is needed on the role of technology and perception. The offence-defence balance of power theory highlights that what matters is not just material advantage but perceived advantage. If states perceive that AI-enabled cyber offence is gaining the upper hand, they may choose to counter by building their own cyber offence. These actions can create competitive spirals even in the absence of conclusive evidence that cyber offence has the

[145] Rebecca Slayton, 'What Is the Cyber Offense-Defense Balance? Conceptions, Causes, and Assessment' (2017) 41 *International Security* 72.

advantage. In this way, AI can shape strategic stability through expectation effects as much as through capability effects.

Cyberspace also brings about complexities that do not exist in the traditional realm. Offensive and defensive capabilities can be the same. Research used for vulnerability identification for defensive patching can be used for offensive exploitation. Intelligence gathering for network defence is similar to intelligence gathering for attack. The dual-use nature of cyber capability makes it difficult to distinguish between the two, which were traditionally viewed as more distinct by offence-defence theory. AI makes this even more difficult because the machine learning models used for defensive anomaly detection can have the same underlying architecture as those used for offensive pattern recognition[146].

Moreover, cyberspace is inhabited not only by nation-states but also by corporations, criminal groups, and loosely connected collectives. The offence-defence balance thus applies at various levels of analysis. A technology that benefits non-state offensive actors can upset interstate relations indirectly. The spread of capability makes it difficult to apply traditional deterrence models, which rely on the existence of specific state-based adversaries. It is important to note that these complexities do not make the theory of offence and defence redundant. Rather, they highlight its versatility. The theory is a useful tool for examining the relationship between new technologies, incentives, and vulnerabilities. Rather than dwelling on the sensational implications of autonomous cyberwar, the offence-defence approach focuses on relative effort, scalability, and institutional constraints. The analytical argument made in this chapter is that artificial intelligence co-constitutes existing cyber asymmetries in ways that systematically favour offence. This argument is not based on technological determinism. Artificial intelligence does not necessarily enable attackers relative to defenders. Instead, it amplifies existing structural characteristics that already disadvantage defenders. Artificial intelligence decreases the marginal cost of reconnaissance, facilitates scalable deception, and accelerates adaptive exploitation, thereby reducing the marginal cost of offensive activities. Defensive systems are aided by artificial intelligence-assisted detection and response, but they remain limited by complexity and organizational factors.

The applicability of this argument extends beyond the technical realm of cybersecurity. If AI further exacerbates the offence-defence imbalance in cyberspace, it carries implications for strategic thinking among states[147]. Perceived vulnerability may lead to pre-emptive strategies, expanded offensive capabilities, and heightened suspicion. Conversely, if defensive innovation does not keep pace, confidence in digital security may erode, with attendant economic and political consequences.

Section II

The Structural Foundations of Cyber Offence Advantage

Any hypothesis about the role of artificial intelligence in changing the balance of offence and defence in cyberspace must begin with a description of the underlying distribution of advantage. Otherwise, AI may be seen as the source of instability in itself, rather than a catalyst for the

[146] Henry Farrell and Abraham L Newman, 'Weaponized Interdependence: How Global Economic Networks Shape State Coercion' (2019) 44 *International Security* 42.
[147] Martin C Libicki, *Cyber Deterrence and Cyberwar* (RAND Corporation 2009).

underlying conditions. The strategic nature of cyberspace was established well before the advent of massive machine learning systems. The architecture of cyberspace, its economic model, and its governance regime have created a space in which the superiority of defence has always been impossible to achieve.

The first structural basis for offence superiority is the architecture of digital systems themselves. Today's networks are not point solutions or self-contained systems. They are complex ecosystems that include operating systems, applications, firmware, third-party libraries, cloud infrastructure, and human interfaces. Each layer of the ecosystem introduces code, and each piece of code introduces the possibility of error. The size of today's digital systems ensures that vulnerabilities are not exceptional but statistical certainties. In such an environment, the defender's job is not simply to defeat an attack but to sustain systemic integrity through a constantly shifting technological landscape.

This is an architectural truth that creates a deep asymmetry of effort. Defence must be comprehensive. Offence only needs to be selective. An attacker needs to find one vulnerability that has not been noticed; a defender has to protect thousands of possible points of entry. The asymmetry is not merely one of degree but of logic. Failure for defence can occur through a single mistake; failure for offence does not preclude another attack. Since the cost of multiple attacks is low, persistence becomes a feasible strategy for attack. The complexity of digital systems thus embeds asymmetry at the level of system design.

The second basis for offence superiority is found in the economics of software development and cybersecurity investment. Cyber systems are built in a competitive market environment that favours innovation, speed, and functionality. Security is often an afterthought. Patching vulnerabilities takes time and effort that does not produce direct revenue. For many actors in the private sector, cybersecurity is a cost centre rather than a strategic investment. This economic incentive structure creates an uneven defensive posture across sectors and borders.

Attacks are driven by a different economic model. The development of exploits can be directly monetized through ransomware, data exfiltration, or denial-of-service attacks. Black markets enable the trading of vulnerabilities and malware[148]. Once developed, exploit code can be repurposed and adapted at a low additional cost. The commoditization of offensive capabilities reduces barriers to entry and allows capability to be widely distributed outside of elite technical circles. In this way, the economics of the cyber domain favour successful exploitation over successful prevention.

Third, the attribution challenge undermines the credibility of deterrence. In traditional military conflict, tracing the source of attack is often an immediate or physically traceable process based on evidence of force projection. Cyber-attacks can be masked through proxy servers, compromised hosts, and routing techniques. Attribution is often a probabilistic process rather than a certainty. Even when technical analysis points to a likely source, political factors can complicate public attribution and retaliation.

[148] Michael N Schmitt (ed), *Tallinn Manual 2.0 on the International Law Applicable to Cyber Operations* (Cambridge University Press 2017).

Deterrence requires a credible threat of retaliation. When attribution is uncertain or delayed, there is a risk of miscalculation in response. This generates strategic ambiguity. Such ambiguity can be exploited by offensive actors who operate below the threshold likely to provoke escalation. Espionage, data manipulation, and limited disruption can be conducted in ways that test defensive resolve without triggering overt confrontation. Structural uncertainty therefore reduces the expected cost of certain forms of cyber aggression.

Temporal dynamics further reinforce asymmetry. Cyber intrusions are often gradual and may remain undetected for extended periods. During this time, attackers conduct reconnaissance, map networks, and prepare for disruption. Detection frequently occurs only after significant compromise has already taken place. The delay between intrusion and response provides attackers with informational advantage. Even when ultimately effective, defensive measures are reactive.

The defender's temporal burden is one of constant vigilance. Monitoring, updating, auditing, and responding require sustained organisational effort. Offensive actors can choose the timing of engagement strategically and are not required to maintain continuous visibility. This asymmetry of engagement places structural strain on defensive institutions.

A further imbalance arises from the diffusion of capability among both state and non-state actors. Unlike nuclear or advanced conventional weapons, cyber capabilities do not depend on large-scale industrial infrastructure. Skilled individuals, small groups, and criminal enterprises can develop impactful tools. The dissemination of knowledge through online platforms and open-source repositories accelerates this process. As capability spreads, the number of potential adversaries expands. Defensive actors must therefore guard against a diverse spectrum of threats, ranging from state-sponsored units to opportunistic criminal networks.

This diversity complicates strategic stability. Traditional deterrence models assume identifiable adversaries with centralised command structures. In cyberspace, potential actors include loosely organised networks whose motivations may not align with geopolitical rationality. Even if states achieve mutual deterrence, non-state actors may operate outside those constraints. The structural consequence is increased unpredictability.

Institutional and normative constraints further shape asymmetry. Defensive cybersecurity, particularly in democratic contexts, is embedded within legal frameworks governing surveillance, data retention, and automated response. Security measures must be balanced against privacy and civil liberties. Offensive actors, especially criminal organisations and authoritarian regimes, often face fewer normative constraints. This divergence affects the speed and intensity with which capabilities can be deployed. While defenders operate within compliance regimes, attackers iterate with comparatively fewer restrictions.

Interdependence compounds these structural factors. Digital infrastructure supports financial systems, healthcare networks, energy grids, and public administration. Civilian and military domains are interconnected. A vulnerability in one sector can cascade across others. The interconnectedness that enhances economic efficiency simultaneously amplifies systemic risk. Offensive actors benefit from this condition, as disruption in a single node can generate disproportionate consequences.

Taken together, these structural features—architectural complexity, economic incentive misalignment, attribution difficulty, temporal asymmetry, diffusion of capability, institutional constraint, and systemic interdependence—have produced a cyber domain in which defence faces intrinsic burdens. This does not imply the inevitability of offensive dominance; defensive innovation has strengthened resilience in numerous contexts. Nevertheless, the underlying asymmetry endures.

Establishing this baseline is essential for evaluating the impact of artificial intelligence. In a structurally neutral domain, the dual-use character of AI might generate symmetrical enhancement. In a domain already inclined toward offence, however, technologies that reduce operational cost or accelerate exploitation are likely to magnify imbalance. The analytical task that follows is therefore to determine whether AI interacts with these structural conditions in ways that amplify offensive leverage more than defensive resilience, and whether such amplification reshapes strategic incentives at the international level.

Section III

Artificial Intelligence as an Asymmetry Multiplier

Artificial intelligence does not operate in a neutral strategic environment. Instead, it is superimposed upon a cyber environment already marked by architectural vulnerability, economic misalignment, attribution uncertainty, and temporal imbalance. The relevant question, therefore, is not whether AI enhances capability on both sides—it clearly does—but how it interacts with these pre-existing asymmetries to alter their scale or character. The argument advanced here is that AI functions as an asymmetry multiplier[149]. It does not create structural imbalance; rather, it amplifies it by reducing the marginal effort required for offensive action more substantially than it reduces the systemic burden of defence.

The concept of marginal effort is central to this claim. In strategic analysis, absolute increases in capability are less significant than relative reductions in cost. If both offence and defence become more sophisticated, the balance remains stable unless the cost structures shift unevenly. Artificial intelligence lowers specific categories of effort—time, labour, expertise, and iterative testing—associated with offensive cyber operations. Although defensive actors also benefit from automation, the structural constraints identified earlier limit the degree to which those benefits can counterbalance offensive acceleration.

One of the most consequential developments introduced by AI is the automation of analytical labour. Historically, cyber reconnaissance required extensive system mapping, code examination, and behavioural analysis. Machine learning models excel at identifying patterns across vast datasets. When applied offensively, they can analyse code repositories, network metadata, and publicly available information at scales beyond manual capacity. The effect is not merely increased speed but reduced marginal cost. Expanding analysis to additional targets does not require proportional increases in personnel. Computational scaling substitutes for human scaling.

This development directly interacts with the architectural complexity of digital systems. Given that vulnerabilities are statistically inevitable in large-scale infrastructures, lowering the cost of

[149] Thomas Rid, *Cyber War Will Not Take Place* (Hurst 2013).

discovering them produces disproportionate strategic effects. The defender's obligation remains comprehensive: all potential weaknesses must be addressed. The attacker's obligation is selective: a single viable point of entry is sufficient. AI reduces the cost of selectivity. As search becomes faster and less resource-intensive, the likelihood of identifying exploitable weaknesses rises without equivalent growth in expenditure. The asymmetry embedded in system design thereby becomes more readily exploitable.

Generative AI produces a comparable shift in the domain of social engineering. Cyber intrusions frequently rely not only on technical exploits but also on manipulation of human behaviour. Previously, a trade-off existed between scale and specificity: highly tailored deception required labour-intensive preparation, while mass campaigns sacrificed credibility. Generative language models collapse this trade-off by enabling contextually tailored communication at negligible marginal cost[150]. Offensive actors can replicate institutional tone, incorporate contextual references, and adapt language to cultural settings without extensive human intervention.

The strategic consequence is enhanced scalability. When credible deception can be reproduced algorithmically, the volume of high-quality intrusion attempts increases. Defensive filtering systems confront dynamic variation rather than static templates. Machine-generated diversity challenges detection mechanisms reliant on known patterns. While defensive AI can analyse linguistic anomalies, it must minimise disruption to legitimate communication. Excessive sensitivity generates operational friction and economic cost. Offensive actors, by contrast, can tolerate failed attempts; unsuccessful experimentation carries limited penalty. This asymmetry in tolerance for error reinforces the broader asymmetry in cost reduction.

AI-driven iterative adaptation further widens the imbalance. Defensive cybersecurity often depends on identifying malicious signatures and deploying patches, processes historically marked by temporal lag. AI-enabled offensive systems can compress adaptation cycles by generating behavioural variants in response to detection. An offensive model can iteratively adjust until it identifies configurations that evade monitoring. Defensive responses, however, remain embedded within institutional processes of validation, coordination, and network-wide deployment. The obligation to maintain system stability constrains the speed of adaptation. Offence benefits from agility, whereas defence bears responsibility for continuity.

Lowered expertise thresholds also reshape the strategic landscape. AI-assisted coding and vulnerability analysis tools reduce technical barriers to conducting sophisticated cyber operations. While skilled actors remain important, the pool of potential participants expands as marginal expertise requirements decline. Capability diffusion accelerates. More actors can attempt intrusion with less specialised training. Defensive institutions cannot equivalently reduce expertise thresholds without increasing systemic risk; effective security monitoring demands precision and oversight. The asymmetry between experimentation and reliability becomes more pronounced as AI tools proliferate.

Institutional context magnifies these dynamics. Offensive actors—particularly criminal organisations and loosely regulated entities—can deploy AI tools rapidly and without extensive

[150] Brandon Valeriano and Ryan C Maness, *Cyber War versus Cyber Realities: Cyber Conflict in the International System* (Oxford University Press 2015).

oversight. They are not constrained by privacy regulations, procurement procedures, or public accountability. Defensive institutions, especially in democratic states and corporate settings, must integrate AI within regulatory and ethical frameworks. Concerns regarding transparency, bias, and legality mediate implementation. These governance requirements slow deployment and limit operational aggressiveness. Even where defensive AI offers technical advantages, its use is filtered through institutional friction.

It would be analytically unsound to claim that AI renders defence ineffective. Machine learning enhances anomaly detection, integrates threat intelligence across datasets, and supports predictive monitoring. In certain contexts, defensive automation significantly strengthens resilience. Nevertheless, these gains operate within a domain defined by comprehensive obligation. Defence must secure entire networks, ensure continuity of service, and minimise false positives that disrupt legitimate activity. Offence can concentrate on specific vulnerabilities, accept failure, and capitalise opportunistically on success. AI reduces the cost of opportunism more sharply than it reduces the burden of comprehensive protection.

The cumulative effect is a relative widening of offence–defence asymmetry in cyberspace. Both sides acquire enhanced tools, yet the structural characteristics of digital infrastructure enable offensive gains to translate more readily into scalable advantage. AI amplifies the selective logic of exploitation more effectively than it mitigates the systemic logic of protection.

The implications extend beyond technical cybersecurity. A perceived expansion of offensive advantage reshapes strategic incentives[151]. States that interpret increased vulnerability may invest more heavily in offensive cyber capabilities, anticipating the need for pre-emption or retaliation. The belief that AI-enabled intrusion is becoming easier can stimulate competitive dynamics even where defensive capabilities improve concurrently. As offensive operations become cheaper and more scalable, the threshold for initiating limited cyber actions may decline, rendering strategic interaction more fluid and potentially more volatile.

Artificial intelligence thus operates as a multiplier of pre-existing imbalance. It does not transform cyberspace into an entirely new strategic domain, nor does it negate defensive innovation. Rather, it alters the marginal calculus of effort in ways that deepen structural asymmetry. Appreciating this dynamic is essential for evaluating its broader consequences for escalation, deterrence, and strategic stability.

Section IV

Artificial Intelligence, Escalation Dynamics, and Strategic Stability

If artificial intelligence operates as an asymmetry multiplier in cyberspace, its relevance extends beyond questions of technical exposure. An imbalance between offence and defence matters because it reshapes strategic incentives. Classical security theory suggests that when offence is perceived to hold the advantage, actors experience greater pressure to adopt pre-emptive postures, to engage in competitive capability accumulation, and to interpret uncertainty as threatening. The effects are probabilistic rather than automatic: imbalance heightens the risk of instability. In the

[151] Erik Gartzke, 'The Myth of Cyberwar: Bringing War in Cyberspace Back Down to Earth' (2013) 38 *International Security* 41.

cyber domain, AI-induced shifts in relative advantage generate analogous concerns, albeit within a non-kinetic environment in which escalation manifests through disruption rather than territorial seizure.

Escalation in cyberspace differs in important respects from conventional military escalation. Cyber operations frequently remain below the threshold of armed conflict, encompassing espionage, intellectual property expropriation, infrastructure interference, and influence activities. Because such operations can be calibrated in intensity and obscured through ambiguous attribution, they enable incremental probing. Artificial intelligence reinforces this incremental dynamic by lowering the cost and increasing the speed of limited actions. As small-scale intrusions become cheaper to conduct, their frequency may increase, rendering strategic rivalry more persistent and less episodic.

A central concern lies in the compression of decision-making time. AI accelerates reconnaissance, exploit development, and adaptive evasion, thereby narrowing the window for detection and assessment. Defensive institutions may confront pressure to respond rapidly to ambiguous indicators[152]. In contexts where attribution remains uncertain, accelerated response elevates the risk of miscalculation. The conjunction of speed and uncertainty is inherently destabilising. Strategic stability depends in part on deliberation—the time required to evaluate intent, determine proportionality, and coordinate response. AI-driven acceleration reduces this temporal buffer.

At the same time, the destabilising implications of speed require careful qualification. Cyber operations rarely generate immediate physical destruction, and many are reversible or containable. This has led some observers to argue that cyberspace possesses stabilising characteristics, permitting signalling without catastrophic escalation. Artificial intelligence complicates this assessment in two respects. First, by enhancing scalability, it expands the potential magnitude of disruption. Second, by diffusing capability across a broader array of actors, it increases the number of potential sources of instability. Stability grounded in restraint among major powers becomes more fragile when non-state actors can initiate cascading effects.

Deterrence in cyberspace is already characterised by ambiguity. Effective retaliation depends upon credible attribution and proportionate response. If AI widens offence–defence asymmetry, it may weaken deterrence by reinforcing perceptions that defensive systems are increasingly penetrable. States that perceive persistent vulnerability in their critical infrastructure, despite defensive investment, may pursue alternative strategies, including forward defence or expanded offensive postures. Such approaches risk normalising sustained cyber engagement rather than confining competition to discrete episodes.

Perception plays a decisive role in this process. Offence–defence theory emphasises that actors respond not only to objective capabilities but also to perceived advantage. If policymakers interpret AI-enhanced cyber capabilities as decisively favouring offence, they may prioritise offensive development accordingly. This perception can generate competitive spirals even when defensive capabilities improve in parallel. The security dilemma intensifies when defensive research into

[152] Joseph S Nye, 'Deterrence and Dissuasion in Cyberspace' (2017) 41 *International Security* 44.

vulnerabilities is indistinguishable from offensive preparation. AI amplifies this ambiguity, as similar machine learning techniques may underpin both intrusion and detection.

Threshold ambiguity further complicates escalation dynamics. In conventional deterrence theory, escalation thresholds are often associated with physical destruction or territorial violation. In cyberspace, such thresholds are less clearly defined. AI-enabled operations may permit increasingly intrusive or disruptive activity while remaining below established markers of armed conflict[153]. The cumulative effect of sustained low-level intrusions may erode trust and heighten suspicion without provoking overt retaliation. An environment of continual probing fosters strategic fatigue and reactive positioning.

It would, however, be analytically reductive to conclude that widening asymmetry necessarily produces instability. Defensive adaptation is ongoing. As AI-enabled attacks grow more sophisticated, defensive institutions may reorganise, introduce redundancy, and cultivate more resilient architectures. Strategic stability does not require invulnerability; it requires predictability and credible signalling. Should major powers develop shared norms regarding acceptable cyber conduct, AI-enhanced capabilities may be integrated into managed competition rather than uncontrolled escalation.

Yet the normative dimension remains underdeveloped. International agreements governing cyber operations are limited and frequently non-binding. The pace of AI development exceeds diplomatic consensus. In the absence of clearly articulated rules, states interpret one another's capabilities through worst-case assumptions, amplifying perceived imbalance. Uncertainty regarding AI's future trajectory—its autonomy, adaptability, and scalability—contributes to strategic anxiety.

An additional source of instability arises from the interaction between cyber and conventional domains. As digital systems underpin military command, logistics, and communications, AI-enhanced cyber intrusion into these networks could affect crisis management. If states believe their command-and-control infrastructure is vulnerable to AI-driven disruption, they may perceive incentives to act early in crises rather than risk incapacitation. This logic echoes concerns from nuclear strategy regarding vulnerability and pre-emption, although the modalities and stakes differ.

AI's influence on strategic stability thus operates through multiple interconnected pathways: acceleration of operational tempo, diffusion of capability, amplification of perception-driven insecurity, and deeper integration of digital infrastructure with critical systems. None of these dynamics makes escalation inevitable[154]. Each, however, increases the complexity of deterrence and the difficulty of sustaining equilibrium.

A widening offence–defence asymmetry does not necessarily foreshadow catastrophic cyber conflict. Rather, it points toward a strategic environment characterised by persistent contestation, compressed response windows, and heightened suspicion. Artificial intelligence intensifies the

[153] Miles Brundage and others, *The Malicious Use of Artificial Intelligence: Forecasting, Prevention, and Mitigation* (Future of Humanity Institute 2018).

[154] Lucas Kello, *The Virtual Weapon and International Order* (Yale University Press 2017).

structural conditions under which miscalculation becomes more probable, particularly when combined with ambiguous attribution and indeterminate thresholds.

Ultimately, the consequences of these developments depend not solely on technological innovation but also on political interpretation and institutional adaptation. Strategic stability in an AI-enhanced cyber domain will hinge as much on norm development and governance as on technical capability. Nevertheless, if AI continues to reduce the marginal cost of offensive cyber operations more rapidly than it alleviates the systemic burden of defence, the underlying asymmetry will continue to exert pressure on established deterrence frameworks.

Section V

Rebalancing the Offence–Defence Equation in an AI-Enhanced Cyber Domain

If artificial intelligence deepens offence–defence asymmetry by lowering the marginal cost of offensive cyber operations more than that of systemic defence, then effective responses must address structural incentives rather than rely on isolated technical remedies. The challenge identified in this chapter is not merely the proliferation of AI-enabled threats, but the transformation of relative cost structures within cyberspace. Rebalancing therefore requires measures that raise the cost of offensive exploitation, reduce the burden of defensive maintenance, or recalibrate the strategic expectations that sustain competitive escalation.

The most foundational avenue of rebalancing lies in architectural reform. Cyberspace's offence bias is rooted in complexity and the retrospective addition of security to systems not designed with resilience as a primary objective. Artificial intelligence intensifies this condition by automating vulnerability discovery across expansive infrastructures. A structural response must therefore intervene at the design stage. Secure-by-design standards, mandatory code auditing, and liability regimes that sanction negligent software practices would elevate baseline security. By decreasing the density of exploitable vulnerabilities, such reforms increase the cost of offensive discovery even within an AI-augmented environment. The aim is not to eliminate vulnerability entirely, but to narrow the asymmetry generated by systemic fragility.

A second pathway involves redistributing the defensive burden. At present, defenders carry comprehensive responsibility for safeguarding complex networks, while attackers exploit isolated weaknesses. Collective defence mechanisms—shared threat intelligence platforms, synchronised patch management, and cross-sector response coordination—can reduce duplication and lower marginal defensive costs. AI-based detection systems perform more effectively when trained on aggregated datasets rather than siloed organisational logs. In this respect, cooperation functions as a defensive multiplier, counterbalancing the offensive multiplier effect produced by AI.

Yet technical coordination alone cannot rectify structural incentives. The persistence of asymmetry is reinforced by limitations in deterrence, particularly those stemming from attribution uncertainty. Improvements in forensic AI that enhance attribution reliability may strengthen deterrent credibility. Although perfect certainty is unattainable, reducing ambiguity alters strategic calculation. If offensive actors anticipate more consistent identification and coordinated response, the perceived cost of aggression increases. Enhanced attribution thus operates as an indirect mechanism for elevating offensive risk.

Legal and regulatory reform also has structural implications. Current market incentives in software production do not fully internalise security externalities. Clearer liability standards for severe security failures would shift costs from victims to producers, altering investment incentives. When firms face tangible repercussions for inadequate safeguards, preventive investment becomes economically rational. Over time, strengthened baseline resilience diminishes the exploitable surface available to AI-assisted reconnaissance.

International norm development addresses the perceptual dimension of asymmetry. Offence–defence theory underscores the influence of expectations on strategic behaviour. If states perceive AI-enhanced cyber capabilities as unconstrained and readily deployable, they are more likely to invest pre-emptively in offensive instruments. Norms delineating protected targets—such as critical civilian infrastructure—can stabilise expectations even amid technological change. While norms do not eliminate capability, they restrict its legitimate application, thereby moderating worst-case assumptions.

The governance of AI dissemination presents an additional structural consideration. Open distribution of advanced generative or code-producing models lowers expertise thresholds broadly, including for malicious actors. Conversely, excessive restriction may impede defensive innovation. A calibrated approach—incorporating safety testing, red-teaming protocols, and monitored deployment—can introduce friction against misuse without suppressing legitimate research. Such mechanisms moderate the pace at which offensive marginal costs decline.

Rebalancing does not necessitate the complete eradication of offensive advantage. In international politics, perfect equilibrium is rarely attainable. The objective is to prevent widening asymmetry from destabilising strategic expectations. Investments in resilience, redundancy, and rapid recovery capacity can diminish the strategic payoff of intrusion. Systems designed to degrade gradually rather than fail catastrophically reduce the expected benefit of attack, thereby lowering incentive even where technical penetration remains feasible.

Adaptation must also occur at the doctrinal level. Cyber defence should be integrated into comprehensive strategic planning rather than treated as a peripheral technical function. Clear articulation of red lines, proportional response frameworks, and escalation management mechanisms can mitigate ambiguity. As artificial intelligence compresses operational tempo, doctrine must compensate by clarifying response logic in advance. Predictable signalling reduces uncertainty and constrains escalation dynamics.

The central analytical insight is that AI-induced asymmetry is neither fixed nor inevitable. It arises from the interaction between technological capability and structural context. By reforming architecture, strengthening collective defence, enhancing attribution, realigning economic incentives, developing norms, and clarifying doctrine, actors can influence the trajectory of the offence–defence balance. Although the effectiveness of such measures will vary across political systems, their shared objective is to reshape relative marginal costs rather than merely respond to discrete threats. Rebalancing is therefore a process of structural adaptation. Artificial intelligence magnifies existing asymmetries; deliberate institutional innovation can mitigate that amplification. Whether such adaptation can keep pace with technological advancement remains uncertain. Absent structural intervention, however, the logic of asymmetry outlined in this chapter is likely to continue shaping the evolution of the cyber threat environment.

Conclusion

This chapter has examined whether artificial intelligence alters the offence–defence balance in cyberspace. Rather than approaching AI as a discrete technological breakthrough, the analysis has situated it within the enduring structural conditions that define the cyber domain. The central argument has been that cyberspace exhibited offence-leaning asymmetry prior to the diffusion of advanced AI systems, and that AI deepens this asymmetry by reducing the marginal cost of offensive action more substantially than it alleviates the systemic burden of defence. The structural baseline is therefore decisive. Digital infrastructures are complex, interdependent, and in constant evolution. Defenders carry comprehensive responsibility for securing systems in which vulnerabilities are statistically unavoidable. Attackers require only selective success. Attribution remains uncertain, deterrence constrained, and capability diffusion ongoing. Artificial intelligence enters this environment as a force multiplier. By automating reconnaissance, enabling scalable deception, accelerating adaptive exploitation, and lowering expertise thresholds, AI reinforces the selective logic of intrusion. Although defensive actors also benefit from automation, their gains are circumscribed by institutional constraints, reliability requirements, and the imperative of systemic continuity. The outcome is a widening of relative advantage rather than a balanced transformation. AI magnifies pre-existing imbalance because the underlying architecture of digital systems remains vulnerability-dense and economically misaligned. Rebalancing, therefore, does not entail suppressing AI development but reshaping the structural context within which it operates. Secure-by-design principles, collective defence coordination, improved attribution mechanisms, realigned economic liability, normative boundary-setting, and doctrinal clarification all operate upon the cost structures that underpin the offence–defence balance. By raising the expected cost or diminishing the anticipated payoff of offensive cyber operations, such measures can temper the asymmetry intensified by AI. The strategic trajectory of cyberspace will depend less on the intrinsic sophistication of AI systems than on the institutional adaptations that accompany their integration. Offence–defence theory underscores the dynamic nature of balance: technological innovation may shift relative advantage, but governance, coordination, and perception can recalibrate it. Should states interpret AI-enhanced intrusion as an uncontested advantage and respond primarily through competitive offensive expansion, instability is likely to deepen. Conversely, if adaptation prioritises structural resilience and calibrated normative frameworks, asymmetry may be moderated even amid rapid technological change. Artificial intelligence does not convert cyberspace into an entirely new strategic arena. Rather, it sharpens and accelerates existing characteristics. It hastens processes that were previously slower, scales activities that were once limited, and diffuses capabilities that were formerly concentrated. Whether this sharpening results in sustained instability or in managed competition will depend on the extent to which political actors confront the structural foundations of imbalance rather than merely its technological expressions. The offence–defence balance in cyberspace is not a fixed conclusion but a continually evolving equation. AI has altered its variables. The eventual outcome will hinge on the effectiveness of institutional responses to that alteration.

Chapter - VIII

State and Sovereignty in Cyber Space: Government Frameworks, Laws, and Security Tools

Aungshuman Ghosh, Advocate Supreme Court of India & Visiting Faculty Department of Legal Studies Swami Vivekananda University, Barrackpore, West Bengal.

Abstract

The emergence of cyberspace as a domain of geopolitical contestation has fundamentally disrupted conventional notions of territorial sovereignty and state authority. Unlike physical domains where boundaries are cartographically defined and legally recognized, cyberspace operates as a borderless, decentralized, and asymmetric environment — one in which the traditional Westphalian framework of state sovereignty faces its most formidable challenge yet. This chapter examines the intricate relationship between state power and cyberspace, interrogating how governments across the world have responded to the imperatives of digital governance through the development of legal frameworks, regulatory architectures, and technological security tools.

The chapter begins by deconstructing the myth that cyberspace exists beyond the reach of sovereign authority. While early internet discourse celebrated its emancipatory and ungovernable character, the reality of contemporary cyberspace reflects an increasingly fragmented landscape — one shaped by competing national interests, divergent regulatory philosophies, and the growing assertion of digital sovereignty by state actors. From China's "Great Firewall" and Russia's sovereign internet legislation to the European Union's General Data Protection Regulation (GDPR) and India's evolving data protection jurisprudence under the Digital Personal Data Protection Act, 2023, states have demonstrated both the will and the capacity to territorialize the digital domain.

Central to this analysis is a critical assessment of the international legal frameworks governing state conduct in cyberspace. The chapter evaluates the applicability of existing instruments of public international law — including the UN Charter, customary international law, and the Tallinn Manual 2.0 — to cyber operations, exposing significant normative gaps and interpretive ambiguities that states and non-state actors routinely exploit. The absence of a binding multilateral treaty on cybersecurity, combined with the attribution challenge inherent to cyber incidents, continues to undermine accountability and collective security in the digital realm.

The chapter further explores the domestic legal responses of key jurisdictions, analyzing how national cybersecurity laws, cybercrime legislation, and critical information infrastructure protection regimes function as instruments of both security and control. Particular attention is given to the tension between state surveillance imperatives and the fundamental rights of

citizens — a tension that legislation such as India's Information Technology Act, 2000 and its amendments, the United Kingdom's Investigatory Powers Act, 2016, and the United States' Cybersecurity Information Sharing Act, 2015 have only partially resolved. The chapter argues that effective cybersecurity governance cannot be reduced to a technical exercise in deploying firewalls, encryption protocols, or intrusion detection systems; it is, at its core, a normative and institutional challenge demanding coherent legal architecture, inter-agency coordination, and meaningful democratic oversight.

The chapter concludes by proposing a multi-stakeholder model of cyber governance that reconciles state sovereignty with the transnational character of cyberspace. It advocates for enhanced international cooperation, capacity-building in developing nations, and the embedding of human rights principles within national cybersecurity strategies. Ultimately, the chapter contends that the myth of the impenetrable digital fortress — whether technological or legislative must yield to a more honest and resilient conception of cyber governance, one grounded in accountability, adaptability, and the rule of law.

Keywords: Cyber Sovereignty, Digital Governance, Cybersecurity Law, International Law, Data Protection, State Surveillance, Critical Infrastructure, Cyber Warfare, Information Technology Act, Tallinn Manual.

Key Outcomes

- Researching how state sovereignty developed through borderless cyberspace and digital domains shows its current state.
- The study examines international legal frameworks which control cyber operations to determine how these laws apply to state activities.
- The study analyzes domestic cybersecurity laws which exist in major jurisdictions such as India the United States the European Union and China.
- An examination involving the degree of state-security requirements that may infringe on fundamental human rights in cyberspace governance.
- An analysis into how security technologies or systems interact with prevailing legal and regulatory regimes.

Introduction: Sovereignty in the Digital Age

The concept of state sovereignty is rooted in the Westphalia Treaty (1648), which lays out a definition of sovereignty in terms of state action, which is made exclusive within its walls. The classic concept of "sovereignty of the state"[155] holds that a country is supreme in terms of both the citizens and the inhabitants of the physical territory, undisturbed morally or physically from

[155] Arora, S. (2023). Digital Personal Data Protection Act, 2023: A critical analysis of exemptions and enforcement mechanisms. *Indian Journal of Law and Technology*, 19(2), 145-172.

outside forces. In the cyberspace paradigm that has been introduced to undermine and alter the concept of territoriality, cyberspace is based on a network system that extends beyond physical boundaries and is devoted to global (data; aural, oral) communications. Specifically, the Internet has created great ambiguity in the manner people considered sovereignty when words like "sovereign"[156] evaporated, while some positions were creating questions about the probability of extant sovereignty structures ensuring some control over online operations, and there is a strong case for some urgency in considering novel structures of power and governance.[157] In the early years of the Internet, most discussions were full of a kind of techno-utopianism-a viewing of cyberspace as a wild, interstice, a digital common beyond the reach of any State, no matter how authoritative. From the "Declaration of the Independence of Cyberspace"[158] proposed by John Perry Barlow, came the idea that digital networks constitute an area that can slip outside the sovereignty of nation-states. That romantic viewpoint is totally prescription today on account of the stress these past twenty years have put on the governments. Basically, throughout the world governments have displayed both desire and ability regarding the employment of technology and law to capture the control over Cyberspace and to in fact disintegrate the once lofty dream of a worldwide digital network into a group of disjoined national online territories.[159]

Cyber sovereignty assertion has occurred through different methods which include technical infrastructure controls and legal frameworks and surveillance systems and international coordination efforts. China's internet filtering system and Russia's internet sovereignty system both demonstrate Cyber sovereignty through their operations.

Conceptualizing Cyber Sovereignty: Theoretical Foundations:

The Westphalian Challenge

In order from minor to major, the Westphalian concept of sovereignty is framed upon the notions of territorial integrity, political autonomy, and legal equality among independent states. None of these three can be realized in the digital domain. The constitution of digital territory presently finds itself taking a wholly digital-oriented form that comprises a pattern of meshed networks as well as sets of servers and data flows enforced by user information across different legal domains. In a very practical setting, an email will take just a couple of seconds to transmit from one server to another located in various places around the world, providing itself many entangled problems about which state has the right to monitor, control, regulate, etc., the transaction. [160]

---

[156] Arora, S. (2023). Digital Personal Data Protection Act, 2023: A critical analysis of exemptions and enforcement mechanisms. *Indian Journal of Law and Technology*, 19(2), 145-172.

[157] Kuehl, D. T. (2009). From cyberspace to cyberpower: Defining the problem. In F. D. Kramer, S. H. Starr, & L. K. Wentz (Eds.), *Cyberpower and National Security* (pp. 24-42). National Defense University Press.

[158] Arora, S. (2023). Digital Personal Data Protection Act, 2023: A critical analysis of exemptions and enforcement mechanisms. Indian Journal of Law and Technology, 19(2), 145-172.

[159] Goldsmith, J., & Wu, T. (2006). *Who Controls the Internet? Illusions of a Borderless World*. Oxford University Press.

[160] Brenner, S. W. (2007). At light speed: Attribution and response to cybercrime/terrorism/warfare. *Journal of Criminal Law and Criminology*, 97(2), 379- 475.

The necessary algorithm for this process to make political independence challenging is to form a web with its infrastructure spreading throughout all systems. No state, on the grounds of a state-wide economy, can ever completely become independent of cyberspace. There is an interface between national networks international protocols and domain name systems, while the physical cyber infrastructure functions from one legal territory to another. Thus, the internet control mechanism quite naturally works on simplistic lines as a multi-stakeholder system rather than being state-led to isolate internet resources and root servers and global routing systems. Such a mechanism virtually binds together all territorial statuses, hugely contrasting with the rights-of-sovereignty that are given to states on land.[161]

Third-party assessor discusses that the concept of Westphalian sovereignty of international equality is now diminishing due to technological disparities in cyber capabilities of different national states.[162] A plethora of technical capacities, its extensive know-how - whether about the good sophistication, security, offensive cyber warfare, and what have you-needed contribute considerably to the degree to which one of the countries is able to tag along, with the US and China or Russia of established giant cyber powers, or Israel running with high-tech cyber abilities, whereas the relatively low-cyber-capacities within individual member states could all be derived from the European Union. Some contrast to this, however, exists in the sense that most low-income countries lack even rudimentary cybersecurity measures. These differences in capability between countries mean that sometimes one country may just end up controlling another distant territory with the state remaining quite penetrable from beyond against international threats or internal, for that matter.[163]

Competing Models of Digital Governance

Two contrasting narratives are brought up currently around the very contemporary topic of cybersecurity. One points towards a liberal democracy-like model which offers a power- sharing platform in governance between many entities, be they governments, private sector, NGOs, and tech industry. The system could then be aligned like CSC group IANA or Internet Engineering Group around the three fundamental principles: public access; systems operability; and constraint of state power.[164]

Another model, the complete central authority or the government with some underlying authority, states that the cyberspace in any nation should come under the complete control of the state as if it were an indispensable part of the territory, with no scope for trade-off on grounds of national security or police integration and internet censorship. China's argument on "cyber sovereignty"[165] is probably the most complete implementation of such a model. According to the

[161] Mueller, M. L. (2010). *Networks and States: The Global Politics of Internet Governance*. MIT Press.

[162] Ibid.

[163] Goldsmith, J., & Wu, T. (2006). *Who Controls the Internet? Illusions of a Borderless World*. Oxford University Press.

[164] DeNardis, L. (2014). *The Global War for Internet Governance*. Yale University Press.

[165] Ibid.

statement of the government of China, countries as states are entitled, and in the matter of policy commitment, obligated to manage their Internet intern systems in compliance with national laws and culture, while there is an added dimension that sees criticism from outsiders as an intrusion into sovereign rights.[166]

The European Union develops its third model which establishes a system for government control that protects rights through its strong legal frameworks which govern both state and private sector operations. The General Data Protection Regulation (GDPR) serves as a regulatory sovereignty example because it enables EU legal standards to reach worldwide through its global enforcement power. The EU governance model requires effective control systems which include strong legal frameworks to limit state surveillance and corporate data operations while ensuring democratic oversight and fundamental rights protection.[167]

Table 1: Comparative Models of Cyber Governance

Governance Model	Primary Actors	Key Principles	Representative Examples
Multi-stakeholder	States, Private Sector, Civil Society, Technical Community	Openness, Minimal Intervention, Distributed Authority	ICANN, Internet Governance Forum, Early Internet Architecture
State-Centric Sovereignty	National Governments	Territorial Control, Regulatory Authority, Security Primacy	China's Great Firewall, Russia's Sovereign Internet Law, Iranian National Network
Regulatory Sovereignty	State Institutions with Rights Constraints	Legal Frameworks, Rights Protection, Democratic Accountability	EU GDPR, German NetzDG, French Data Protection Authority
Hybrid Models	Mixed State and Private Coordination	Pragmatic Balance, National Security with Limited Openness	India's IT Act, UK Investigatory Powers Act, Singapore Cybersecurity Act

[166] Creemers, R. (2020). China's conception of cyber sovereignty: Rhetoric and realization. In *Governing Cyberspace: Behavior, Power, and Diplomacy* (pp. 107- 145). Rowman & Littlefield.

[167] Bradford, A. (2020). *The Brussels Effect: How the European Union Rules the World*. Oxford University Press.

International Legal Frameworks: Norms, Gaps, and Ambiguities:

Applicability of Public International Law to Cyberspace

Scholars and policymakers have ongoing discussions about how existing public international law applies to activities in cyberspace. The United Nations Group of Governmental Experts (UN GGE) on Developments in the Field of Information and Telecommunications in the Context of International Security reached its 2013 and 2015 conclusions about state behavior in cyberspace by stating that international law, together with the UN Charter, governs state activities in cyberspace. The agreement established a major achievement because it proved that states cannot treat cyberspace as an area without laws which protects them from their duty to follow international treaties.[168]

International law applicability remains an area now needing further attention. It does nothing to help with the ultimate question as to how legal rules become applied to cyberspace. The example of using force rendered illegal by the UN Charter, under Article 2(4), brings to light the difficulties of interpreting that signal. This prohibition is meant to prohibit kinetic military operations, but creates limitless uncertainties when applied to cyber operations. So, should a more active breakdown in the functionality of critical infrastructure achieved by a distributed denial-of-service cyber attack be considered as a use of force? Likewise, what about cyber-spying, for instance by foreign intelligence agencies, wherever large-scale operation is concerned? On these questions, States are significantly divided, largely because of broader consensus as to where the threshold should be when assessing cyber activities as violations of sovereignty or acts of aggression.[169]
Test your understanding: In what way did poetry advance your understanding of love, death, or history in War of the Physicians? Stick to three examples from case materials supportive of your opinion, duly accepted by the community.

The Tallinn Manual Framework

There is no longer to wait to embellish the desktop with fascinating wallpapers by exploring your photographic abilities. By now, your picture frames are quite packed, but they prove inadequate. Even though unemployment forms are discussed by people, usually the hard-core job seekers fare well with handwriting superficial questions on resumes, the just conditions for face-to-face interviews vanish, and there is excessive weight on the question of what a person did.[170] However, by recording their portfolios online, several job seekers are entitling their fates to companies to evaluate their applications seriously through the information they supply directly. In terms of time, the concept should start with shots so less saleable they do not want them; [uploaders] may well have kept these images to this moment. Although the authors of the first paragraph point to endorsements of the Tallinn Manual, the text originated from an academic

[168] Tikk, E., & Kerttunen, M. (2017). The alleged demise of the UN GGE: An autopsy and eulogy. *Cyber Policy Institute Brief*, November 2017, 1-8.

[169] Roscini, M. (2014). *Cyber Operations and the Use of Force in International Law.* Oxford University Press.

[170] Ibid.

study conducted with no government intervention. This study was built on expert opinions rather than on established international practice. Such study has not been based on actual international customary law or an ensuing legal doctrine. Its judgment on contested issues such as: the point at which a cyber operation will be characterized as a use of force or as an infringement on sovereign rights is accepted but not uniformly accepted as a legal principle.[171] Though the Manual has policy relevance, it does not serve as an effective tool for understanding state behavior in cyberspace because there remains no among legal experts about what a given legal definition.

Domestic Legal Responses: National Cybersecurity Frameworks:

The Indian Framework: Information Technology Act and Beyond

India's cyberspace legal system developed significantly from 2000 because of the Information Technology Act implementation. The IT Act enables electronic commerce through digital transaction recognition which later required legal amendments to solve new cybersecurity threats and cybercrime issues and content management requirements. The 2008 amendments created new laws which addressed cyber terrorism and data protection and set intermediary liability standards because the government had started to view cyberspace as a legal area that needed complete control.[172]

The IT Act center section 69 gives central and state governments special rights to monitor and intercept all information which people send through computer systems to protect national security and public safety and international relations and territorial integrity. Civil rights defenders have raised major privacy issues about government monitoring systems because these systems depend on special rules which the Information Technology (Procedure and Safeguards for Interception, Monitoring and Decryption of Information) Rules 2009 establish for monitoring and protecting information.[173] (Bhatia, 2019).

The Digital Personal Data Protection Act 2023 stands as India's current attempt to create a complete framework for data protection rights. The Act establishes three data protection principles which require data minimization and purpose limitation and individual consent in addition to establishing the Data Protection Board as the nation's data protection authority. The Act imposes major restrictions on government surveillance activities because its main exemptions allow security agencies to operate without restrictions when they conduct operations for national security or public safety or territorial integrity. The government has chosen to protect national security and maintain flexible control over operations instead of protecting user privacy through this system of exemptions which has resulted in privacy complaints from users.

---

[171] Roscini, M. (2014). *Cyber Operations and the Use of Force in International Law*. Oxford University Press.

[172] Maini, T. S. (2012). The Information Technology Act, 2000: A legal framework for cyber law in India. *Indian Journal of Law and Technology*, 8(1), 89-112.

[173] Bhatia, R. (2019). Surveillance, privacy, and the Information Technology Act: Examining Section 69 and procedural safeguards. *Journal of Indian Law and Society*, 10(1), 34-58.

The United States Approach: Fragmented Regulation

The United States lacks a comprehensive federal cybersecurity law which matches the cybersecurity systems that other major jurisdictions have implemented. American cyber governance functions through a broken system which consists of specific industry rules and executive directives and optional security partnerships between public and private sectors. The Cybersecurity Information Sharing Act (CISA) of 2015 represents one of the few cross-sectoral federal initiatives which enable private companies and government bodies to exchange cybersecurity threat data. CISA fails to function as an all-encompassing governance instrument because its implementation depends on voluntary participation and the law provides only minimal liability coverage according to Watters.

Presidential Policy Directive 21 (PPD-21) on Critical Infrastructure Security and Resilience designates sixteen critical infrastructure sectors and establishes a framework for protecting these sectors from cyber threats through coordination between sector-specific agencies and private sector partners. The public-private partnership model serves as the American market- driven solution which requires minimal regulatory oversight while creating a stark contrast to standard regulatory methods that other jurisdictions employ. The Federal Information Security Modernization Act (FISMA) governs cybersecurity for federal agencies which need to do continuous monitoring and annual assessments and follow standardized security controls that use National Institute of Standards and Technology (NIST) guidelines.[174]

The Fourth Amendment of the United States Constitution protects citizens from government surveillance through its prohibition of unreasonable searches and seizures yet the extent of these protections remains disputed in relation to online activities. The Foreign Intelligence Surveillance Act (FISA) establishes procedures for electronic surveillance and physical searches related to foreign intelligence gathering while Section 702 of the FISA Amendments Act authorizes warrantless surveillance of non-U.[175]

The European Union: Rights-Based Regulatory Sovereignty

The European Union currently leads worldwide rights-based cyber governance because it developed extensive legal systems which regulate online activities while safeguarding essential human rights. The General Data Protection Regulation (GDPR), which entered into force in 2018, establishes stringent requirements for data processing, including principles of lawfulness, fairness, transparency, purpose limitation, data minimization, accuracy, storage limitation, integrity, and accountability. The GDPR applies extraterritoriality because it governs all organizations that handle personal data from EU residents, which extends European Union legal standards throughout the world according to Bradford.[176]

---

[174] Paganini, P. (2018). The evolution of U.S. federal cybersecurity policy: From FISMA to continuous monitoring. *International Journal of Cyber Warfare and Terrorism*, 8(3), 45-67.

[175] Ibid.

[176] Bradford, A. (2020). *The Brussels Effect: How the European Union Rules the World*. Oxford University Press.

The Network and Information Security (NIS) Directive, adopted in 2016 and updated as NIS2 in 2022, establishes cybersecurity requirements for operators of essential services and digital service providers. The directive requires member states to designate competent authorities, establish computer security incident response teams (CSIRTs), and implement national strategies for network and information security. The 2019 Cybersecurity Framework Act of the European Union gives EU Cybersecurity Agency (ENISA) added powers in setting up a European Certification Framework for cybersecurity-certification schemes. Certification guarantees that ICT products, services, and processes satisfy certain security requirements for trustworthiness
and integrity of the single market where applicable. The legislation is the EU's endeavor to use regulation to shape security standards worldwide and extend influence.

China's Comprehensive Control Architecture

The Chinese government has established its cyber sovereignty system which functions as the most advanced system for state control over online activities through its comprehensive digital presence.

The Chinese Cybersecurity Law which became effective in 2017 requires network operators to implement security measures which include conducting security assessments and storing data within national borders and working with government intelligence agencies and security organizations. The law defines critical information infrastructure as essential systems which protect approximately 80 percent of China's digital economy from security threats through government monitoring.[177] The Chinese government uses its cybersecurity framework to implement multiple regulatory systems which include monitoring online content and monitoring social credit systems and conducting full-scale surveillance operations. The Great Firewall functions as an advanced system which uses technical filtering methods and DNS manipulation techniques and deep packet inspection methods to prevent users from accessing foreign websites and services which the government considers dangerous for national security and public order. Social media platforms in the country must enforce strict content control measures which create heavy legal liability for them if they do not delete banned materials without delay. The Chinese government uses advanced technological systems for social control through its implementation of artificial intelligence and big data surveillance technologies which operate in Xinjiang and other areas.[178]

The Data Security Law and Personal Information Protection Law establish the complete legal structure which governs China's digital operations through their 2021 enactment. The laws create a framework for data classification which establishes rules for cross-border data transfers and specifies conditions under which government authorities can access information stored by private companies.[179]

---

[177] Creemers, R. (2020). China's conception of cyber sovereignty: Rhetoric and realization. In *Governing Cyberspace: Behavior, Power, and Diplomacy* (pp. 107- 145). Rowman & Littlefield.

[178] Mozur, P. (2019). One month, 500,000 face scans: How China is using AI to profile a minority. *The New York Times*, April 14, 2019.

[179] Ibid.

The framework shows China's intention to treat data as a national resource which the government will control while giving precedence to state security needs above individual and business rights. The current strategy uses data.

Table 2: Comparative National Cybersecurity Frameworks

Jurisdiction	Primary Legislation	Governance Approach	Key Features
India	IT Act 2000 (amended), Digital Personal Data Protection Act 2023	Centralized with security emphasis	Government interception powers, intermediary liability, broad security exemptions, emerging data protection
United States	CISA 2015, FISMA, Sector-specific regulations	Fragmented, public-private partnerships	Voluntary information sharing, critical infrastructure protection, Fourth Amendment constraints, sector-based regulation
European Union	GDPR 2018, NIS Directive/NIS2, Cybersecurity Act	Rights-based regulatory sovereignty	Extraterritorial application, comprehensive data protection, mandatory breach notification, certification schemes
China	Cybersecurity Law 2017, Data Security Law 2021, Personal Information Protection Law 2021	Comprehensive state control	Data localization, critical infrastructure security, content control, government access
			requirements, social credit integration
United Kingdom	Investigatory Powers Act 2016, Data Protection Act 2018, Network and Information Systems Regulations 2018	Balanced security and rights (post-Brexit hybrid)	Bulk interception powers, judicial authorization, incident reporting, GDPR adequacy maintenance

Security Tools and Technological Governance:

Technical Infrastructure and Control Mechanisms

The sovereignty of states to control their cyberspace depends on their capabilities to secure networks and manage their technology systems. States establish a framework of digital activity monitoring which includes technological systems that enable them to observe and control their online operations. The system provides two types of monitoring systems, with one being the lawful interception interface, which telecommunications networks must use and the other being advanced surveillance systems, which operate without public knowledge and democratic control.[180]

China's Great Firewall uses network filtering systems, which combine multiple technical strategies, including IP address blocking, DNS filtering, and redirection, URL filtering, and packet inspection, together with connection reset. States use these systems to block access to international websites and online services, while permitting users to connect to local services. Different jurisdictions show various degrees of technical development for these systems, as some countries use simple blocking methods, which users can bypass, while other countries use deep packet inspection and machine learning systems, which can detect and block encrypted data together with its circumvention tools.

Countries such as the United States, India, Australia, Spain, and the United Kingdom passed laws that enacted Telecommunications Assistance for Law Enforcement or CALEA to make telecommunications service providers build a system capable of lawful interception for government surveillance.[181] However, such countries had put certain regulations to incentivize end-to-end encryption and banned unnoticed crypto, condemned attempts to weaken any network design, or innovations for governments.

[180] Deibert, R., & Rohozinski, R. (2010). Liberation vs. control: The future of cyberspace. *Journal of Democracy*, 21(4), 43-57.

[181] Ibid.

Cyber Defense and Critical Infrastructure Protection

In the case of national cybersecurity strategies, attention to the protection of critical infrastructure as topically represented is now seeing several cyberattacks threatening the likes of all digital systems that support energy distribution, critical telecommunications networks, financial systems, water supplies, ground transportation infrastructures, and healthcare systems. Providing for the shortening of regulatory frameworks, the states have imposed baseline security measures, a requirement for incident reporting, and the provision for information sharing among private-sector operators in collaboration with government agencies.[182]

The elegance of the Cybersecurity Framework from the National Institute of Standards and Technology in the United States is a voluntary scheme that has been launched to voluntarily moderate the efforts to combat cybersecurity risks for critical infrastructure. It is based on the classification of its cybersecurity activities into five core functions: Identification, Protect, Detect, Response, and Recovery. It remains the most dominant standard in most sectors, while on a voluntary basis, because the standard asserts standardization presentations that are beneficial in business ventures involving large installations, exercising joint acceptance of them.[183] So, some concerted endeavor is most pressing for the intention to make possible the most viable arrangement of technical, procedural, and control mechanisms guaranteeing a likely assurance against intrusion vectors, devoid of a cumbersome, overly stringent tigress of imminent regulation that would be rendered obsolete too hastily by rapidly changing technology behaviors.

Having the Computer Emergency Response Teams (CERTs) or Computer Security Incident Response Teams (CSIRTs) solve cybersecurity incidents also resulted in working with the national and sector teams that are a part of these consortia. National CERTs operate in most countries to manage cyber incidents through their response coordination, threat intelligence sharing, and technical aid provision to government bodies and critical infrastructure operators. International coordination among CERTs through forums such as FIRST (Forum of Incident Response and Security Teams) enables international collaboration to combat cyber threats because effective cyber defense needs both national capabilities and international partnerships.[184]

---

[182] Lewis, J. A. (2019). *Critical Infrastructure Protection and Cyber Resilience*. Center for Strategic and International Studies.

[183] Ibid.

[184] Tøndel, I. A., Line, M. B., & Jaatun, M. G. (2014). Information security incident management: Current practice as reported in the literature. *Computers & Security*, 45, 42-57.

The Surveillance-Rights Tension: Democratic Accountability in Cyber Governance

The declaration of state control over cyberspace leads to conflicts between security needs and essential human rights which include privacy rights and freedom of speech and the right to fair legal proceedings. Democracy requires security systems to function within established constitutional boundaries which demand proper supervision and which should only be used for genuine government needs. The combination of complex technical systems for cyber surveillance together with hidden intelligence work and ineffective control systems prevents organizations from achieving their accountability standards according to Diffie and Landau.[185]

The Edward Snowden disclosures from 2013 revealed how the United States National Security Agency and its partner intelligence agencies operated extensive and sophisticated surveillance programs. The programs conducted bulk telephone metadata collection which enabled them to intercept internet traffic through their partnerships with telecommunications companies and they used commercial encryption system weaknesses to conduct their operations. The revelations created a global scandal which led to multiple countries implementing changes including USA FREEDOM Act in the United States and legal actions that resulted in the Schrems II decision by the European Court of Justice which rendered the EU US Privacy Shield agreement invalid according to Greenwald 2014.[186]

To achieve proper democratic control over cyber surveillance, multiple institutional structures need to be established. First, democratic legitimacy requires legal authorization through legislation instead of using executive power to establish limits on discretionary authority. Second, judicial oversight needs to include specific requirements for individualized

warrants which must meet the probable cause standard or an equivalent threshold, to prevent unauthorized surveillance. Third, all organizations need to meet legal security requirements when disclosing their surveillance activities and operations according to defined requirements.

Toward a Multi-Stakeholder Model: Reconciling Sovereignty with Transnational Governance

The existing conflict between territorial sovereignty and cyberspace's transnational nature needs more than state actions and technical solutions to achieve its resolution. Cyber governance requires multiple stakeholders to work together because states and international organizations and private sector entities and civil society organizations and technical communities need to cooperate. Stakeholders in governance frameworks possess separate abilities and requirements and viewpoints which need to be unified.[187]

Certainly, there are significant political DPI issues in international cybersecurity cooperation.

---

[185] Diffie, W., & Landau, S. (2007). *Privacy on the Line: The Politics of Wiretapping and Encryption*. MIT Press.

[186] Goldsmith, J., & Wu, T. (2006). *Who Controls the Internet? Illusions of a Borderless World*. Oxford University Press.

[187] DeNardis, L. (2014). *The Global War for Internet Governance*. Yale University Press.

Antagonism among the great powers turned on fundamental differences over norms governing the behavior of states in cyberspace, with authoritarian regimes speaking of state sovereignty and content control, while liberal democracies argue for openness and fundamental rights. At the international level, cybersecurity endeavor could not proceed, due to an irreconcilable normative position between two Chinese and Russian proposal for an International Code of Conduct for Information Security and the Western countries taking up the United Nation Group of Governmental Experts (UN GGE) meeting process which was to frame a voluntary, agreed-upon set of norms.[188]

Operational partnerships offer proof positively that cybersecurity cooperation has succeeded, in the face of the political impediments. Intergovernmental cooperation targeting the implementation of law enforcement in transnational cybercrime, inter alia, with Interpol cybercrime programs and mutual legal assistance frameworks, remains a possibility. ICANN and Internet Engineering Task Force (IETF) and the regional internet registries are technical bodies for coordination, which, through their capabilities of operational coordination, maintain the global internet infrastructure.[189] It thus may be observed that effective cyber governance is security-held by a formal operational partnership of an international nature, rather than the great many multilateral agreements.

Conclusion

The developing international system in cyberspace shifts the compartments. It is simply untenable to propose that cyberspace is ungovernable and therefore free from sovereignty. States have increasingly asserted their will to exert control over digital networks and thus raise legal and technological mechanisms and regulatory architectures that position the space closer to the national interest and political system; thus, governance systems across jurisdictions have a range of effectiveness and legitimacy, depending on their operational standpoint and doctrinal inconsistency.

Based on many widely accepted international norms, domestic forces may thus be useful in providing such guidelines for the behavior of states in cyberspace, with many needs to fill gaps and resolve ambiguities in frames of reference of rights regarding the "non-usage" of borders. Public international law," a term preferred by many supportive professors, is claimed to govern cyberspace operations everywhere while details such as points of threshold and proportionality in attribution raise yet more complex questions, not to mention the policy questions of how to respond to incidents.

The increasingly frequent collusion that states mount in the grey areas operating less than formal legal limits makes it difficult for everyone to anticipate the behavior of the said nations or for resulting actions to subsequently be condemned by the accepted norms of international law. Variably different domestic legal frameworks create varied governance models-some riotous,

[188] Tikk, E., & Kerttunen, M. (2017). The alleged demise of the UN GGE: An autopsy and eulogy. *Cyber Policy*

[189] Watters, P. (2016). The Cybersecurity Information Sharing Act: A legislative response to cyber threat proliferation. *Georgetown Journal of International Affairs*, 17(3), 125-138.

while others would be bordering on regulatory sovereignty or semi-principle-based regulatory sovereignty; where they seek a balance between state authority and private sector participation in various ways, each one putting the security vs rights argument first and then the economic efficiency vs equal access- to-innovation one.

The second concept (that is, equal access for all) runs parallel to the idea of ensuring all rights for all people and places; to what safeguards will the governance apparatus subscribe at the price of other rights? Vigorous debate about surveillance versus rights would now set out greener and more incipient paths that are not-and indeed never will be-hatch for established regulatory bodies. Thus, accounting for probable restraining conditions on surveillance by state institutions does require intense democratic oversight specimens accompanying judicial review and compliance with transparency demands on independent data protection.

Chapter – IX
Plagiarism and Cybersecurity in the Digital Age: A Cross-Level Study of Academic Integrity in Higher Education

Dr. Richa Chaurasia, Assistant Professor & Head, Department of Education, Saheed Anurup Chandra Mahavidyalaya, University of Calcutta

Abstract

Plagiarism has evolved significantly in the digital age, transforming from traditional cut, copy and paste based misconduct to technologically facilitated academic dishonesty. Simultaneously, cybersecurity threats in higher education institutions have intensified and alarming, at the same time exposing vulnerabilities related to intellectual property theft, unauthorized data access, and digital manipulation of academic manuscripts and content. The present study examines the intersection of plagiarism and cybersecurity across and among undergraduate (UG), postgraduate (PG), doctoral (PhD), and faculty levels. A cross-sectional survey of 200 participants was conducted to assess awareness, attitudes, experiences and expectations related to plagiarism and institutional cybersecurity mechanisms. The findings reveal high dependence on digital sources, moderate awareness of plagiarism policies, and limited understanding of cybersecurity frameworks about protecting academic data. Results indicate a statistically significant relationship between usage of digital tool and unintentional plagiarism practices. The study concludes that plagiarism must be addressed prominently as an ethical issue and also as a cybersecurity concern requiring technological, institutional, and behavioural interventions.

Keywords: plagiarism, cybersecurity, academic integrity, digital ethics, intellectual property, ethical issues.

Introduction

Plagiarism is traditionally defined as the act of presenting another individual's intellectual work ideas, words, data, or creative expressions as one's own without proper acknowledgment. Historically, plagiarism was confined to manual reproduction of texts. However, the rapid expansion of internet accessibility, cloud storage systems, and artificial intelligence tools has fundamentally transformed its scale and complexity. Higher education institutions increasingly rely on digital platforms for teaching, submission of assignments, peer collaboration, and research dissemination. While these advancements have democratized knowledge, they have also created opportunities for academic misconduct.

Digital plagiarism now includes copy-paste practices, paraphrasing software misuse, AI-generated text submission, and code replication.[190]

At the same time, universities face escalating cybersecurity threats. Data breaches, ransomware attacks, unauthorized database access, and intellectual property theft threaten academic ecosystems. The intersection between plagiarism and cybersecurity emerges when research data, unpublished theses, or proprietary academic content are accessed or misused through digital vulnerabilities.[191]

Plagiarism is no longer solely an ethical violation; it has become intertwined with digital security frameworks. Understanding this intersection is essential across UG, PG, and PhD levels, as well as among faculty who guide research integrity.[192]

This study aims to:

1. Examine the evolving nature of plagiarism in the digital environments.
2. Analyse cybersecurity and vulnerabilities that facilitate academic misconduct.
3. Assess awareness levels of students and faculty across UG, PG and PhD.
4. Provide recommendations based on evidence for strengthening academic integrity systems.

Review of Related Literature:

- Evolution of Plagiarism in Digital Contexts

1. Research indicates that internet proliferation has significantly increased plagiarism

rates among students. Digital accessibility reduces the perceived effort required to replicate academic material. Studies show that students often engage in unintentional plagiarism due to inadequate citation knowledge.[193]

2. AI-based writing systems have further complicated detection mechanisms. Automated

content generation tools blur the boundary between assistance and misconduct, challenging traditional definitions of originality.

- Cybersecurity in Higher Education

1. Universities hold large volumes of sensitive data such as unpublished research, patents,

[190] Bretag, T. (2016). Challenges in addressing plagiarism in education. Educational Integrity Journal, 12(2), 45–59.

[191] Park, C. (2003). In other (people's) words: Plagiarism by university students. Assessment & Evaluation in Higher Education, 28(5), 471–488.

[192] Ibid.

[193] Bretag, T. (2016). Challenges in addressing plagiarism in education. Educational Integrity Journal, 12(2), 45–59.

examination systems, and personal student records. According to cybersecurity frameworks proposed by the National Institute of Standards and Technology (NIST), US based standards agency that academic institutions often lack enterprise-level security infrastructure.[194]

2. Research published by IEEE organisation professional association for

technology emphasizes the growing trend of ransomware attacks targeting educational institutions. These attacks not only disrupt operations but also expose confidential academic material.[195]

- Detection Tools and Technological Safeguards

1. Plagiarism detection services such as Turnitin use similarity indexing and algorithmic comparison to detect content overlap.[196] However, these systems primarily address textual similarity rather than cybersecurity vulnerabilities like unauthorized file access.[197]

2. Studies suggest that integration of blockchain technology could secure academic

authorship records and timestamp intellectual property.[198]

Although plagiarism and cybersecurity have been individually studied by a large number of researchers but limited empirical research explores their intersection across academic levels. Therefore, the present study bridges that gap by combining theoretical analysis with survey-based evidence.

Theoretical background of Plagiarism and Cyber security

1. Deterrence Theory

Deterrence theory suggests individuals are less likely to commit misconduct when the perceived risk of detection and punishment is high. Plagiarism detection software acts as a deterrent mechanism. However, if the cybersecurity systems are weak there is a high opportunity of increase in misconduct.

2. Routine Activity Theory (Cyber Adaptation)

This theory posits that in digital academia crime occurs when three elements converge. They are: -

---

[194] Sharples, M., & Domingue, J. (2016). The blockchain and kudos: A distributed system for educational record. Journal of Learning Analytics, 3(3), 1–7.

[195] Ibid.

[196] Morthy, M.S.N. (2007) *Encyclopedia of Multimedia and Communication Technology (Vol-1)* Delhi: Arise Publishers and distributions

[197] National Institute of Standards and Technology. (2018). Framework for improving critical infrastructure cybersecurity.

[198] Bandhopadhya, A.K. (2008). *International Encyclopedia of Audio Visual Media (Vol- 1)*Delhi: Anmol Publication Pvt Ltd

- Motivated offender i.e. student or researcher,
- Suitable target i.e. digital academic content
- Absence of capable guardian i.e. cybersecurity & plagiarism systems[199]

The study concludes that weak digital guardianship increases risk of crime or misconduct.
3. Technology Acceptance Model (TAM)

TAM explains that perceived usefulness and ease of use influence adoption of detection software. Faculty acceptance significantly impacts institutional enforcement of plagiarism policies.

The study is based on following methodologies.
i) Research Design- A quantitative cross-sectional survey was conducted.
ii) Population – The population of the study will consist all the undergraduate students, post-graduate students, research scholars and faculty members of University of Calcutta and its affiliated colleges.[200]

iv) Sample - There are 100 participants as shown in detail in the table below which are selected from the population.[201]

Sl. No.	Types of participants	No. of participants
1	Undergraduate	40
2	Post- graduates	25
3	Research Scholars	20
4	Faculty Members	15
Total no. of participants		100

iii) Sampling Method

Convenience sampling method within the university setting is employed here.

iv) Tool

[199] Bandhopadhya, A.K. (2008). *International Encyclopedia of AudioVisual Media (Vol- 1)* Delhi: Anmol Publication Pvt Ltd

[200] Ibid.
[201] National Institute of Standards and Technology. (2018). Framework for improving critical infrastructure cybersecurity.

A self- made structured questionnaire containing short answer type, one word answer type, MCQ type and Yes or No type items constructed and administrated. The detailed structure of questions is given in the table below.

Sl. No.	Types of questions	No. of questions
1	Demographic questions	5
2	Likert-scale questions	15
3	Yes/No questions	5
Total no. of questions		25

v) Ethical Considerations

The study is conducted in an open environment where voluntary participation is encouraged and anonymity of the participants are maintained. However, the data collected strictly used for academic research only.

Data Analysis

The data collected are analysed through Descriptive statistics i.e. percentage analysis and Correlation analysis among the students, research scholars and faculty members based on three main points. They are -

- Digital tool use

- Plagiarism awareness

- Cybersecurity awareness

Results of the Survey

- Demographic Distribution

Sl. No.	Category of participants	Percentage of participants
1	Undergraduate	40%
2	Post- graduates	25%
3	Research Scholars	20%
4	Faculty Members	15%
TOTAL		100%

Findings of the study-

The key findings of the study are mentioned as follows.

A. Study of awareness of plagiarism policy: 78% of participants have keen awareness where 22% are partially aware.

B. Study of awareness of Cybersecurity Policy: 46% of participants are aware where 54% are unaware.[202]

C. Use of online sources without proper citation: 72% of participants uses online sources without proper citation and 18% of participants frequently omits where 10% never use any.

D. Use of AI tools: 64% of participants reported using AI assistance frequently among them undergraduate students have higher usage of 82%.[203]

E. Study of Confidence in Detection Tools: - 58% of participants have moderate confidence, 25% have high confidence and 17% low confidence.

Correlation Finding

A moderate positive correlation (r = .52) was observed between heavy digital tool usage and self-reported unintentional plagiarism.

Discussion

- The findings of the study suggest that while plagiarism awareness is relatively high among all whereas cybersecurity awareness remains significantly lower. This gap exposes institutions towards academic misconduct facilitated through digital systems.
- Undergraduate students show higher AI tool usage. Hence potentially increasing unintentional plagiarism risk among them.
- Faculty members show high confidence in detection systems that suggests institutional trust, yet awareness gaps among students weaken preventive impact too.
- The moderate correlation between digital tool use and plagiarism indicates that technology, without structured guidance, increases vulnerability.[204]

---

[202] Becta, (2004) British Educational Communications and Technology Agency. *A review of the Research literature on barriers to the uptake of ICT by teachers*

[203] Jain, M.E. and Agarwal, J.C. (2008) *Encyclopedia of Education (Vol-2)* Delhi: Anshan publishing house

[204] Anandan,K. & William Dharma Rja, B. (2010). *Educational Technology,* New Delhi: A.P.H. Publishing Corporation

Recommendations

1. There must be mandatory cybersecurity and academic integrity training in all levels of higher educational institutes for both students and faculties.
2. The integration of plagiarism detection into Learning Management Systems.
3. Introduction of Blockchain timestamping of theses and study materials.
4. There must be Two-factor authentication for academic submissions.
5. National academic cybersecurity frameworks should be developed.
6. AI-detection algorithm should be upgraded.

Conclusion

Therefore, it can be concluded that plagiarism in the digital era extends beyond the ethical misconduct and misuse. Hence, intersects with institutional cybersecurity vulnerabilities. In spite of availability of numerous detection tools, awareness and technological safeguards remain uneven across all academic levels both among students and teachers. Therefore, institutions must adopt a holistic strategy by combining education, technology and policy enforcement to protect intellectual integrity and minimise misconduct.

Suggestion for further studies

The present study has been conducted on a sample of students and teachers in the city Kolkata. This study may be replicated on school students and teachers and students and teachers of urban and rural areas, students and teachers of English medium as well as Hindi medium at different levels of education with different Boards of education. Studies may also be conducted to find out interaction or comparison between students and teachers of different universities, colleges or streams.[205]

[205] IEEE. (2020). Cybersecurity threats in higher education institutions. IEEE Security Review, 18(4), 22–30.

Chapter - X

Constitutionalism and Cyberspace Governance

Susmita Ghosh, Assistant Professor (Law), Swami Vivekananda University

Abstract

The expansion of digital technologies has fundamentally transformed the architecture of state power, raising critical questions about the applicability and resilience of constitutional principles in cyberspace governance. As governments increasingly rely on digital infrastructures for surveillance, data collection, cybersecurity, and public administration, the traditional understanding of constitutionalism - limited government, rule of law, separation of powers, and protection of fundamental rights- faces unprecedented challenges. Cyberspace operates across borders, often through algorithmic and automated systems, complicating established doctrines of territorial jurisdiction, accountability, and transparency. This chapter examines how constitutionalism adapts to digital governance, focusing on privacy, surveillance, data protection, intermediary regulation, and algorithmic decision-making. It analyses the recognition of informational privacy as a fundamental right in Justice K.S. Puttaswamy (Retd.) v. Union of India, which articulated the proportionality test as a constitutional standard for assessing state interference in digital contexts. The chapter further explores how statutory frameworks such as the Digital Personal Data Protection Act, 2023 seek to operationalize constitutional safeguards, while also raising concerns regarding state exemptions and oversight mechanisms. Comparative reference is made to the General Data Protection Regulation, which emphasizes accountability, data minimization, and individual autonomy.

The chapter argues that constitutionalism in cyberspace must evolve beyond traditional models to incorporate digital due process, algorithmic transparency, and robust judicial review. Courts play a crucial role in ensuring that technological governance remains subject to legality, necessity, and proportionality. Ultimately, the future of democratic governance depends on embedding constitutional values within digital infrastructures to prevent the normalization of unchecked surveillance and algorithmic control. Cyberspace, far from existing outside constitutional order, represents a new frontier where constitutional guarantees must be reaffirmed and reimagined.

Keywords: cyberspace, digital due process**,** data minimization, and individual autonomy.

Introduction

Constitutionalism is a foundational political philosophy that governs the relationship between authority and individual rights through a Constitution. Constitutionalism requires that all state actions are to be done within the limits of law by establishing the principles and framework of law that limit the power of the government.

The rapid development of digital technologies has essentially transformed the structure of the governance and relationship between the citizens and state. Governments today rely extensively on digital infrastructures for administration, public service delivery, national security, and regulation of

the social and economic activities. Cyberspace has therefore emerged as an important domain of governance where political authority, legal norms, and technological systems intersect.[206]

The constitutional frameworks were outlined at the time when administration was executed through territorial institutions and physical state machinery. The proliferation of AI, emergence of data driven and global communication platforms has fundamentally reshaped the traditional exercise of state power. The modern authorities are exercising unparalleled power to track communications, harvest personal data, and regulate digital spaces. This can also create threats to individual rights and democratic freedoms.

Constitutionalism, which emphasizes the limitation of governmental authority through the rule of law and protection of fundamental rights, plays a very important role in addressing these challenges. The principles of constitutional governance ensure that state actions remain accountable, transparent, and consistent with democratic values. In the digital era, constitutionalism must adapt to regulate new forms of power exercised through digital technologies.[207]

In contemporary constitutional jurisprudence the recognition of digital rights is an important aspect. Judiciary around the world recognize fundamental rights including privacy and freedom of speech and apply it to the digital world. By following this trend, the Indian Supreme Court in Justice ***K.S. Puttaswamy (Retd.) v. Union of India***[208] affirmed that right to privacy is a fundamental constitutional right and is a legal foundation for protecting digital autonomy and personal data.

Fundamental Rights in Cyberspace

The expansion of digital technologies has significantly influenced the scope and interpretation of fundamental rights. Rights that were traditionally exercised in physical spaces now operate within digital environments. The internet has become an essential platform for communication, political participation, economic activity, and social interaction.

One of the most important constitutional developments in this context is the recognition of the right to privacy in the digital sphere. The collection and processing of personal data by governments and private corporations raise serious concerns about surveillance, profiling, and misuse of information. Modern technologies enable the continuous monitoring of individuals through smartphones, online platforms, and digital identification systems.[209]

In India, the Supreme Court addressed these concerns in the landmark judgment of *Justice K.S. Puttaswamy (Retd.) v. Union of India,*[210] which recognized right to privacy as a fundamental right under Article 21 of the Constitution. The Court emphasized that informational privacy, including

[206] Andrew Murray, *Information Technology Law: The Law and Society* (Oxford University Press 2019).

[207] Jack M Balkin, *'The Future of Free Expression in a Digital Age'* (2017) 36 *Pepperdine Law Review* 427.

[208] (2017) 10 SCC 1

[209] Graham Greenleaf, *Asian Data Privacy Laws* (Oxford University Press 2014).

[210] (2017) 10 SCC 1

control over personal data, is an essential component of individual liberty and dignity. The judgment also introduced the proportionality test to evaluate state actions that interfere with privacy.

Freedom of speech and expression is another fundamental right significantly affected by cyberspace governance. Social media platforms and digital communication tools have created new opportunities for citizens to express opinions, mobilize political movements, and participate in democratic debates. However, the digital environment also presents challenges such as misinformation, hate speech, and online harassment.[211]

State Surveillance and Constitutional Safeguards

The expansion of digital technologies has dramatically increased the surveillance capabilities of modern states. Governments can now monitor electronic communications, track online activities, and analyze vast datasets using advanced analytical tools. These surveillance mechanisms are often justified as necessary for combating terrorism, cybercrime, and other threats to national security.[212]

However, ampel surveillance powers increase serious constitutional concerns for the protection of right to privacy and civil liberties. Many surveillance programs may enable governments to accumulate and reserve personal information about citizens without proper safeguards measures. These practices may threaten democracy and collapse the public confidence in state institutions.

Constitutionalism demands that surveillance activities be done within a clearly defined legal framework. State authorities must ensure that surveillance measures are conducted within legal limits, looking for a legitimate objective, and ensure the response is balanced. These requirements make sure that surveillance powers are not used arbitrarily or excessively.

For regulating the surveillance activity, the Judiciary plays an important role. Courts may assess the legality of surveillance laws and analyse whether they are against the constitutional rights. Autonomous bodies and parliamentary committees are there to monitor the implementation of surveillance programs to ensure transparency and accountability.

Another challenge relates to the use of encryption and digital security technologies. Encryption protects the privacy and security of digital communications, but governments often argue that encrypted systems may hinder law enforcement investigations. Balancing the need for strong encryption with the legitimate interests of national security remains a complex issue in cyberspace governance.[213] Ultimately, it should be ensured that effective constitutional safeguards must balance between protecting individual rights and entitle the state to address legitimate security concerns.

[211] Jack M Balkin, *'The Future of Free Expression in a Digital Age'* (2017).

[212] Andrew Murray, *Information Technology Law: The Law and Society* (Oxford University Press 2019)

[213] Lawrence Lessig, *Code and Other Laws of Cyberspace* (Basic Books 2006).

Data Protection and Digital Governance

The digital economy relies heavily on the collection and processing of personal data. Governments, corporations, and online platforms gather large volumes of information about individuals' activities, preferences, and communications. While such data can enhance innovation and economic growth, it also creates risks related to privacy, discrimination, and misuse of personal information.[214]

Laws on Data protection aim to regulate the collection, storage, and use of personal data. These laws establish rights for individuals and impose obligations on organizations that process personal information. Core principles of data protection typically include consent, transparency, purpose limitation, and accountability.[215]

In India, Digital Personal Data Protection Act, 2023 enacted on August 11, 2023, establishes a framework for processing digital personal data. The law requires free consent for data processing and mandates security mandates by data fiduciaries and imposed penalties for violation of laws.

Globally, the most influential data protection regime is the General Data Protection Regulation. It establishes strict rules regarding data processing, including the principles of data minimization, purpose limitation, and accountability.

Despite these developments, several challenges remain. One major concern relates to the exemptions granted to government agencies for reasons such as national security or public interest. While such exemptions may sometimes be necessary, they must be carefully regulated to prevent abuse of state power. [216]

Another challenge arises from the global nature of digital data flows. Personal data frequently moves across national borders, making it difficult for individual countries to enforce their regulatory frameworks. International cooperation and harmonization of data protection standards therefore play a crucial role in effective digital governance.[217]

Regulation of Digital Platforms and Emerging Constitutional Challenges

Digital platforms have become crucial to contemporary communication and economic activities. Social media networks, search engines, and online marketplaces influence how individuals access information, interact with one another, and participate in public discourse. These platforms often function as modern public forums where democratic debates take place.[218]

Governments have increasingly sought to regulate digital platforms to address issues such as misinformation, hate speech, cybercrime, and harmful content. Regulatory frameworks sometimes

[214] Graham Greenleaf, *Asian Data Privacy Laws* (Oxford University Press 2014).

[215] Christopher Kuner, *Transborder Data Flows and Data Privacy Law* (Oxford University Press 2013).

[216] Jack M. Balkin, "*The Future of Free Expression in a Digital Age*" (2017)

[217] Christopher Kuner, *Transborder Data Flows and Data Privacy Law* (Oxford University Press 2013).

[218] Andrew Murray, *Information Technology Law: The Law and Society* (Oxford University Press 2019).

focus on intermediary liability, which determines the extent to which online platforms are responsible for content posted by users.[219]

However, the rules and regulations for digital platforms are raising important constitutional questions. Excessive government dominance on online platforms may lead to censorship or risk of infringing fundamental right to freedom of speech. Conversely, complete hands-off approach can empower the harmful content to spread without accountability.

The challenge is in developing regulatory frameworks that balance freedom of expression with the need to maintain a safe and responsible digital environment. Transparency in content moderation policies and accountability mechanisms for technology companies are essential for protecting constitutional values.[220] Rapid technological developments also raise new constitutional challenges. Artificial intelligenc, biometric identification etc are increasingly used in public welfare services such as law enforcement, welfare distribution, and financial regulation. While these technologies may improve accuracy and efficiency, they may also create risks related to discrimination, lack of transparency, and degradation of due process. To mitigate these risks, constitutional frameworks must include principles such as algorithmic transparency, accountability, and human oversight. Governments make sure that technological innovation does not weaken the protection of fundamental rights.

Conclusion

The growth of digital technologies and cyberspace governance are facing one of the most significant challenges for contemporary constitutional systems. Governments now having unprecedented capabilities to monitor and regulated digital content personal data and communication, creating a need for new form of "digital Constitutionalism" to protect individual rights from state surveillance and control. While these developments lead to numerous benefits for economic growth and public administration, they also have some serious risks to individual rights and democratic values. Constitutionalism provides the framework necessary to regulate state power in the digital age and ensure that technological governance remains consistent with the rule of law. Protecting fundamental rights in cyberspace requires strong legal safeguards, effective data protection frameworks, transparent surveillance practices, and independent judicial oversight. Courts, legislatures, and regulatory institutions must work together to ensure that digital governance reflects constitutional principles.[221]

As emerging technologies continue to evolve, constitutional systems must adapt to address new challenges such as artificial intelligence, algorithmic decision-making, and cross-border data flows. Ultimately, cyberspace should not exist outside the constitutional order but must be governed by the same principles of legality, accountability, and respect for human rights that define democratic societies[222]

[219] Graham Greenleaf, *Asian Data Privacy* Laws (Oxford University Press 2014)

[220] Jack M. Balkin, "*The Future of Free Expression in a Digital Age*" (2017)

[221] Paul Bernal, *Internet Privacy Rights: Rights to Protect Autonomy* (Cambridge University Press 2014).

[222] Lawrence Lessig, *Code: And Other Laws of Cyberspace* (2nd edn, Basic Books 2006).

Chapter - XI

Illusion of Absolute Security: From Protection to Surveillance

Koyel Modak, Assistant Professor (Law), Swami Vivekananda University

Abstract

The promise of absolute security has become one of the most powerful narratives shaping contemporary governance, technology, and public imagination. In an age defined by cyber threats, terrorism, data breaches, and geopolitical instability, the demand for complete protection appears both rational and urgent. Yet, the idea of "absolute security" is conceptually flawed and politically consequential. This chapter argues that the pursuit of total protection gradually transforms protective mechanisms into pervasive systems of surveillance, thereby altering the relationship between the state, technology, and the individual.

The illusion lies in the belief that risk can be eliminated. Security, by its nature, is probabilistic. Technological systems are vulnerable, human actors are unpredictable, and threats continually evolve. Nonetheless, political discourse and corporate marketing often portray advanced surveillance technologies as capable of delivering near-total safety. This narrative fosters public consent for intrusive measures by framing privacy as an obstacle to security rather than as its complement.

This chapter critically examines how the shift from protection to surveillance occurs in three interconnected stages. First, it explores the securitization of fear, where exceptional threats are used to justify extraordinary measures. Second, it analyses the technological mediation of security, highlighting how data-driven governance prioritizes predictive control over reactive protection. Third, it interrogates the social consequences of normalized surveillance, including chilling effects on expression, behavioural conformity, and the marginalization of vulnerable groups through algorithmic bias.

Importantly, the chapter does not deny the legitimacy of security concerns. Rather, it challenges the assumption that more surveillance necessarily equates to greater safety. By tracing the historical and theoretical foundations of security discourse, it reveals how the pursuit of absolute protection can undermine democratic values, erode civil liberties, and concentrate power within opaque institutional and corporate structures. The paradox is that measures introduced to preserve freedom may gradually constrain it. The chapter ultimately advocates for a reframing of security as a balanced, accountable, and rights-respecting practice. Instead of striving for unattainable absolutes, policymakers must acknowledge the inevitability of risk and focus on proportionality, transparency, and oversight.

Keywords: Absolute security, surveillance society, cyber security, data governance, digital privacy, securitization, state power, algorithmic profiling, biometric identification, risk management, civil liberties, technological governance, mass data collection; predictive policing; democratic accountability.

Introduction

The contemporary world is increasingly shaped by narratives of security. Governments across the globe justify expansive technological infrastructures and legal frameworks in the name of protecting citizens from threats such as terrorism, cybercrime, and transnational insecurity. While these measures promise enhanced safety, they simultaneously generate complex concerns regarding privacy, civil liberties, and democratic accountability. We would examine the concept of the "illusion of absolute security," arguing that the pursuit of total safety often results in the normalization of pervasive surveillance. This chapter draws an interdisciplinary perspective from law, political theory, and surveillance studies, it explores how security discourses legitimize monitoring practices that extend beyond their original purposes. It further highlights the role of digital technologies, biometric identification systems, algorithmic governance, and public-private partnerships in expanding the scope of surveillance. Despite an extensive body of literature on surveillance and privacy, a critical problem persists in understanding how contemporary societies internalize and normalize surveillance in everyday governance structures while maintaining the belief that such systems guarantee complete protection.[223] This chapter addresses that problem by analysing the structural shift from protective security to surveillance-based governance. It concludes by proposing a rights-based framework that reconciles security objectives with democratic values through transparency, proportionality, and institutional accountability.

Security has historically been a central objective of the modern state. It has long been regarded as the fundamental responsibility of modern state. The legitimacy of political authority has often been justified by the promise of protecting citizens from violence, disorder, and external threats. In the contemporary era, however, the concept of security has expanded beyond traditional notions of military defence and policing. Governments now confront complex threats such as cyber warfare, terrorism, digital espionage, and organized transnational crime. These developments have encouraged the adoption of technologically advanced surveillance systems that monitor individuals, spaces, and communications.

In many contemporary societies, surveillance infrastructures are introduced under the promise of ensuring public safety. Digital monitoring tools, biometric databases, facial recognition systems, and algorithmic risk assessments are presented as necessary instruments to prevent harm. However, the mechanisms used to guarantee such protection as in mass data collection, biometric identification, predictive policing, and algorithmic monitoring have simultaneously expanded state surveillance. Yet these mechanisms also reshape the relationship between citizens and the state. The line between legitimate security practices and intrusive surveillance becomes increasingly blurred.

The notion of absolute security plays a significant role in this transformation. Political narratives frequently suggest that comprehensive monitoring and data analysis can eliminate risks and guarantee safety. However, such promises often obscure the social and legal consequences of surveillance expansion. The pursuit of total security may gradually normalize practices that undermine privacy, autonomy, and democratic oversight.

[223] Julie E Cohen, 'What Privacy Is For' (2013) 126 *Harvard Law Review* 1904.

This chapter examines the illusion of absolute security by analysing how protective frameworks evolve into surveillance regimes.[224] It also examines the paradox between the promise of absolute security and the gradual normalization of surveillance practices. It argues that the pursuit of complete security is inherently illusory because it requires intrusive monitoring that can erode civil liberties, privacy, and democratic accountability.[225] It argues that modern security policies rely heavily on technological infrastructures that expand state power while creating new vulnerabilities and rights concerns. The chapter contributes to ongoing debates regarding the balance between security and liberty in the digital age. This chapter explores the illusion of absolute security by examining how contemporary security policies increasingly blur the line between protection and surveillance. It argues that while surveillance technologies can enhance security capabilities, their unchecked expansion risks transforming democratic societies into environments of continuous monitoring. The chapter concludes that democratic societies must critically reassess the narrative of absolute security and instead pursue a balanced approach that safeguards both safety and fundamental rights. Only by acknowledging the limits of security and reinforcing transparency, accountability, and rights-based regulation can societies prevent protection from gradually becoming pervasive surveillance.

Conceptual Foundations: Security, Surveillance and Governance

The relationship between security and governance has long been a subject of political and legal scholarship. Classical theories of the state emphasize that individuals consent to political authority in exchange for protection. According to this perspective, the state functions as a guarantor of security, ensuring order and stability within society.

Modern security governance, however, operates within a far more complex technological and institutional environment. Surveillance technologies have transformed the way states manage risks and monitor populations. Unlike traditional policing methods, digital surveillance allows authorities to collect and analyse vast amounts of information in real time.

Scholars of surveillance studies have observed that modern societies increasingly operate through systems of continuous observation and data collection. These systems are embedded not only within state institutions but also within everyday technologies such as smartphones, online platforms, and digital payment systems. As a result, surveillance has become an ordinary feature of social life rather than an exceptional practice.

The transformation of security governance reflects a broader shift from reactive to preventive strategies. Instead of responding to threats after they occur, modern security systems aim to predict and prevent potential risks. Predictive analytics, behavioural monitoring, and algorithmic profiling are increasingly used to identify patterns that may indicate future criminal or terrorist activity.

The expansion of surveillance is often justified by the argument that modern threats require sophisticated technological responses. Yet this reasoning also raises concerns about proportionality and

[224] Bruce Schneier, *Beyond Fear: Thinking Sensibly About Security in an Uncertain World* (Springer 2003) 38.
[225] Glenn Greenwald, *No Place to Hide: Edward Snowden, the NSA and the US Surveillance State* (Metropolitan Books 2014) 95.

accountability. When surveillance becomes pervasive, it alters the relationship between citizens and the state, shifting from trust-based governance to suspicion-based monitoring.

While preventive security offers certain advantages, it also raises important normative questions.[226] The reliance on predictive technologies can result in extensive monitoring of individuals who have not committed any wrongdoing. This shift alters the traditional presumption of innocence and raises concerns about fairness, accountability, and discrimination.

The Expansion of Surveillance Technologies: Rise of Surveillance Society

Technological innovation has dramatically increased the capacity of both governments and private actors to monitor social behaviour. Advances in digital technology have significantly enhanced the capacity of governments and institutions to collect, store, and analyse personal data. Surveillance today is no longer limited to physical observation; it encompasses a vast array of digital processes that operate invisibly within everyday technologies. Contemporary surveillance infrastructures combine multiple forms of data collection, including biometric identification, location tracking, financial records, and digital communications.

Biometric technologies represent one of the most prominent developments in modern surveillance systems. Fingerprint databases, facial recognition tools, and iris-scanning technologies allow authorities to verify identities and track individuals across different contexts. While these technologies are often introduced to improve efficiency and security, they also create centralized repositories of sensitive personal information.

Another significant development is the rise of algorithmic governance. Security agencies increasingly rely on data analytics and artificial intelligence to identify suspicious patterns and predict potential risks. Predictive policing models analyse historical crime data to determine areas that require increased surveillance or policing presence.

The integration of surveillance technologies into everyday infrastructure further complicates the security landscape. Public spaces are now monitored through extensive networks of cameras, sensors, and digital tracking systems. At the same time, private technology companies collect enormous volumes of data through online platforms and digital services.

Moreover, surveillance is not solely conducted by governments. Private technology companies play an increasingly significant role in collecting and managing data. Their platforms generate vast datasets that can be accessed by state authorities through legal mandates or cooperative arrangements. This public-private collaboration further complicates questions of accountability and transparency.

The collaboration between public institutions and private corporations creates new forms of surveillance governance. Governments may access privately collected data for security purposes, while corporations develop technologies that shape surveillance practices. This relationship raises important questions regarding transparency, accountability, and regulatory oversight.

[226] David Lyon, *Surveillance Society: Monitoring Everyday Life* (Open University Press 2001) 2.

Security Narratives and the Normalization of Surveillance

Public acceptance of surveillance policy measures often depends on the narratives used to justify them and emphasise fear and risk. Governments frequently invoke the language of emergency, threat, and protection to legitimize expanded monitoring powers. In times of crisis, citizens may be more willing to accept intrusive measures if they believe such actions are necessary to ensure safety.

This dynamic can lead to the gradual normalization of surveillance, as described as security narrative, where extraordinary measures become normalized through the language of protection. Temporary measures introduced during periods of crisis may become permanent features of governance. Over time, societies adapt to these practices, and surveillance becomes embedded within routine administrative functions and temporary emergency powers may become permanent features of governance.

The politics of fear can also influence legislative processes. Laws enacted in response to security crises often prioritize rapid action over careful consideration of rights implications. As a result, surveillance authorities may be granted broad powers with limited oversight.

The normalization process is reinforced by the perception that surveillance technologies are neutral and objective. However, these systems are designed and implemented within specific political and institutional contexts. Decisions regarding data collection, algorithmic criteria, and risk assessment often reflect broader social priorities and biases.

While such measures may initially target specific threats, their scope frequently expands. Surveillance technologies introduced for counter-terrorism purposes, for instance, may later be used for routine policing or administrative monitoring.[227]

Moreover, the widespread use of digital technologies encourages individuals to voluntarily share personal information through social media platforms, mobile applications, and online services.[228] This voluntary participation contributes to a culture in which surveillance becomes both institutional and self-imposed.

Data, Algorithms, and Predictive Security

One of the most significant developments in modern surveillance is the integration of algorithmic analysis into security systems. Predictive policing tools analyse historical crime data and behavioural patterns to anticipate potential incidents. Similarly, border security systems use risk-assessment algorithms to identify individuals considered suspicious.

Although these technologies promise greater efficiency, they also raise serious ethical and legal concerns. Algorithms rely on datasets that may reflect existing social biases. When such data are used to predict risk, they can reinforce discriminatory patterns in policing and law enforcement.

Furthermore, algorithmic decision-making often lacks transparency. The criteria used to determine risk scores or threat assessments may not be publicly accessible. This opacity makes it difficult for individuals to challenge decisions that affect their rights.

[227] Daniel J Solove, *Nothing to Hide: The False Trade-Off Between Privacy and Security* (Yale University Press 2011) 3.
[228] Mireille Hildebrandt, *Smart Technologies and the End(s) of Law* (Edward Elgar 2015) 67.

Another concern relates to data centralization. Large-scale security databases combine information from multiple sources, including biometric records, financial transactions, travel histories, and online communications. The concentration of such sensitive data increases the risk of misuse, unauthorized access, or data breaches.

Specifically, there is limited research addressing how security narratives construct the perception that surveillance technologies can guarantee complete safety, thereby legitimizing the expansion of monitoring infrastructures. This chapter addresses that gap by examining the relationship between security discourse, technological governance, and the social acceptance of surveillance.

By analysing the illusion of absolute security, the chapter highlights how political narratives shape public perceptions of risk and protection. It also demonstrates that surveillance expansion is not solely driven by technological capability but also by the cultural and institutional belief that more monitoring inevitably leads to greater safety.

Legal and Human Rights Implications

The expansion of surveillance raises significant legal and ethical concerns. Privacy is widely recognized as a fundamental human right, essential for protecting individual autonomy, dignity, and freedom of expression.[229] Excessive monitoring can undermine these rights by creating an environment in which individuals feel constantly observed.

Legal frameworks attempt to balance security interests with individual rights through principles such as necessity, proportionality, and legality. Surveillance measures should therefore be justified by legitimate objectives and implemented only to the extent required to achieve those objectives.

However, contemporary surveillance infrastructures often operate on a scale that challenges traditional legal safeguards. Mass data collection and algorithmic decision-making can occur without meaningful public oversight or transparency. A key challenge lies in ensuring that surveillance measures remain proportionate and necessary. Broad or indiscriminate data collection may conflict with constitutional principles and human rights standards. Courts in various jurisdictions have increasingly scrutinized surveillance laws to determine whether they adequately protect individual liberties.

Another important concern relates to data security. Centralized databases containing biometric or behavioural information present significant risks if they are misused, hacked, or accessed without authorization. Data breaches involving sensitive personal information can have long-term consequences for individuals and institutions alike. Another significant issue concerns the role of private corporations in surveillance infrastructures.[230] Technology companies that manage communication networks, cloud storage, and digital platforms possess vast quantities of user data. When governments seek access to such data, questions arise regarding jurisdiction, consent, and the limits of corporate cooperation with state authorities.

Courts in various jurisdictions have increasingly recognized the importance of privacy protections in the digital age. Judicial decisions have emphasized that technological advancements must not erode the fundamental rights that underpin democratic societies.

---

[229] *Universal Declaration of Human Rights* (adopted 10 December 1948 UNGA Res 217 A(III)) art 12.
[230] Shoshana Zuboff, *The Age of Surveillance Capitalism* (Profile Books 2019) 8.

The Illusion of Absolute Security

The pursuit of absolute security is inherently problematic because complete protection from all threats is unattainable. Societies are inherently dynamic and complex, and new forms of risk continuously emerge. Attempting to eliminate every potential threat would require an unrealistic level of control over social behaviour.

The belief in absolute security therefore functions as a political and technological myth. Surveillance systems can reduce certain risks, but they cannot eliminate uncertainty entirely.[231] In fact, excessive reliance on surveillance may generate new vulnerabilities, including data misuse, institutional overreach, and technological dependence.

Furthermore, large-scale data collection can create information overload for security agencies. When authorities must analyse enormous datasets, identifying genuine threats becomes more difficult rather than easier. The pursuit of total security may therefore reduce rather than enhance effectiveness.

Recognizing the limits of security is essential for maintaining democratic values. Instead of striving for unattainable certainty, societies must focus on building resilient institutions that can respond effectively to emerging threats while respecting fundamental rights.

The illusion of absolute security therefore lies in the belief that technological surveillance can fully eliminate uncertainty. In reality, security strategies must acknowledge the limits of prediction and control.

Towards Balanced Security Governance and Rights-Based Security Framework

A balanced approach to security governance requires rethinking the relationship between protection and surveillance. Several policy recommendations can help address the challenges identified in this chapter.

1. Strengthening Legal Safeguards

Surveillance practices must operate within clear legal frameworks that define the scope and limitations of monitoring powers. Laws should specify the purposes for which data may be collected and ensure that surveillance activities are subject to judicial authorization.

2. Enhancing Institutional Oversight

Independent oversight bodies should be empowered to review surveillance practices and investigate potential abuses. Parliamentary committees, data protection authorities, and judicial institutions can play crucial roles in ensuring accountability.

3. Promoting Transparency

Transparency is essential for maintaining public trust. Governments should provide clear information regarding the technologies used for surveillance, the types of data collected, and the safeguards implemented to protect individual rights.

[231] Michel Foucault, *Discipline and Punish: The Birth of the Prison* (Alan Sheridan tr, Vintage Books 1995) 195.

4. Implementing Privacy-by-Design

Technological systems should incorporate privacy protections at the design stage. Data minimization, encryption, and decentralized storage can reduce the risks associated with large-scale data collection.

Conclusion

The relationship between security and surveillance represents one of the defining challenges of the digital age. The pursuit of absolute security represents one of the most significant paradoxes of modern governance. While surveillance technologies promise protection against emerging threats, they also risk undermining the fundamental freedoms that democratic societies seek to protect.[232]

This chapter has argued that the belief in total security is largely illusory. Surveillance systems cannot eliminate uncertainty, and their unchecked expansion may create new vulnerabilities and rights concerns. By identifying a critical research gap in understanding how security narratives normalize surveillance, the chapter contributes to broader debates within legal and political scholarship. Ultimately, the challenge for contemporary societies lies in maintaining a delicate balance between safety and liberty. A sustainable approach to security requires recognizing that safety and liberty are not mutually exclusive but must be carefully balanced.

Security policies must be grounded in principles of transparency, accountability, and respect for fundamental rights. Only by acknowledging the limits of surveillance and strengthening democratic safeguards can societies prevent the transformation of protective governance into pervasive monitoring. By reinforcing transparency, accountability, and rights-based regulation, societies can ensure that the pursuit of protection does not gradually transform into a regime of pervasive surveillance.

[232] David Lyon, *Surveillance Studies: An Overview* (Polity Press 2007) 14.

Chapter - XII

Anticipating Tomorrow's Threats: Strategic Foresight in Cybersecurity

Dr. Chandrima Chakraborty, Assistant Professor, Department of Law, Swami Vivekananda University and Ankita Mukherjee, Assistant Professor, Department of Law, Swami Vivekananda University.

Abstract

In an increasingly interconnected digital landscape, cybersecurity is no longer a reactive discipline but a strategic imperative that demands anticipation and adaptability. This paper explores the role of strategic foresight in identifying and mitigating emerging cyber threats before they materialize into large-scale disruptions. By integrating methodologies such as horizon scanning, scenario planning, and trend analysis, organizations can move beyond traditional defense mechanisms and adopt a proactive security posture.

The study highlights how evolving technologies—including artificial intelligence, quantum computing, and the Internet of Things—are reshaping the threat landscape, introducing complex vulnerabilities alongside unprecedented opportunities. Cyber adversaries are leveraging automation, deepfakes, and sophisticated attack vectors, necessitating a forward-looking approach that accounts for uncertainty and rapid technological change. Strategic foresight enables decision-makers to envision multiple plausible futures, assess potential risks, and design resilient systems capable of withstanding unforeseen challenges.

Furthermore, the paper emphasizes the importance of cross-sector collaboration, intelligence sharing, and policy alignment in strengthening global cybersecurity readiness. It examines case studies where foresight-driven strategies have enhanced threat detection, improved incident response, and reduced organizational risk exposure. The integration of foresight into cybersecurity governance frameworks is presented as a critical factor in achieving long-term digital resilience.

Ultimately, this research argues that anticipating tomorrow's cyber threats requires a shift in mindset—from short-term risk management to continuous, future-oriented planning. Organizations that embed strategic foresight into their cybersecurity practices will be better equipped to navigate uncertainty, safeguard critical assets, and maintain trust in an increasingly volatile digital environment.

Keywords: Strategic Foresight, Cyber Threat Intelligence, Emerging Technologies, Proactive Cybersecurity, Digital Resilience

Introduction

Cybersecurity is often compared with arms race. Every new defence act as an inspiration for a new attack. Every patch tries to uncovers some vulnerability. Yet, in the persistent myopia, the organisation often overlooks a crucial truth: the most dangerous cyber threats are not the ones we see today, but the ones silently shaping beneath the surface. Strategic anticipation in cybersecurity is therefore not merely an academic idea; it is an indispensable requirement for governments, businesses, and societies that depend on digital configuration.

In classrooms, boardrooms, and policy discussions, cybersecurity is still too frequently treated as a technical problem in lieu of strategic one. Firewalls are installed, antivirus software is updated, and compliance checklists are completed. But adversaries are not. They innovate, cooperate, and adapt. Criminal groups use artificial intelligence, whereas state actors' experiment with quantum technologies, and hacktivists mobilise through social platforms. A purely passive approach cannot keep pace. Strategic foresight assists organisations envisage possible futures, identify emerging threats, and prepare before crises materialise.

This chapter demonstrate how strategic foresight can modify cybersecurity from reactive defence into proactive resilience. It scrutinises the changing threat landscape, methods of foresight, empirical implementation strategies, ethical challenges, and the role of education and governance. The objective is to humanise cybersecurity by exhibiting that behind every attack are persons, institutions, and choices—and foresight allow us to shape better outcomes.

Understanding the Changing Cyber Threat Landscape

Cyber threats have evolved from isolated acts of curiosity into complex universal operations. In the formative years of computing, hackers frequently sought reputation or intellectual challenge. Today, cybercrime is a multi-billion-dollar industry. Ransomware gangs run like big agencies, with customer service teams and profit-sharing models. Nation-states manage cyber espionage to influence elections, steal intellectual property, and disrupt infrastructure. Terrorist groups traverse digital sabotage. Even standard users unknowingly participate through botnets or data leaks.[233]

Several trends define today's threat landscape:

First, automation has accelerated attacks. Malware can now scan millions of devices in seconds. Artificial intelligence permits attackers to spawn realistic phishing emails, deep-fake voices, and 'automated vulnerability discovery tools'.[234]

Second, interconnectivity has enlarged risk. Smart homes, wearable devices, industrial control systems, and cloud computing create new gateways. A vulnerability in a small supplier can compromise a complete supply chain.

Third, geopolitical tensions progressively play out in cyberspace. Critical infrastructure such as power grids, hospitals, and financial networks has become targets during international disputes.

---

[233] Bruce Schneier, 'Click Here to Kill Everybody: Security and Survival in a Hyper-connected World', WW Norton 2018.

[234] Bruce Schneier, 'Click Here to Kill Everybody: Security and Survival in a Hyper-connected World', WW Norton 2018.

Fourth, data has become the new currency. Personal information, trade secrets, and biometric identifiers are valuable assets. Once stolen, they cannot easily be restored.

These changes portrays that cybersecurity is no longer about protecting computers only. It is about safeguarding trust in digital society.

What is Strategic Foresight?

Strategic foresight is a disciplined approach to thinking about the time ahead. It does not endeavour to predict exactly the future. Instead, it reconnoitres multiple plausible futures to assist decision-makers prepare. In cybersecurity, foresight means asking questions such as: What new technologies might generate vulnerabilities? How criminal groups might evolve? What social trends could enlarge exposure to digital harm?

Foresight differs from conventional hazards assessment in three ways. First, it considers beyond known threats. Second, it emphasises enduring thinking. Third, it amalgamates technological, political, economic, and social factors. A new encryption method, for instance, may affect privacy law, international relations, and consumer trust simultaneously.[235]

Common foresight tools include 'scenario planning, horizon scanning, trend analysis, Delphi surveys, and red teaming exercises.'[236] Each method encourages organisations to envisage alternative futures and test their preparedness.

Why Strategic Foresight Matters in Cybersecurity

Cybersecurity failures often appear from surprise. Organisations focus largely on present risks while neglecting growing ones. Strategic foresight reduces surprise by increasing awareness.

One benefit is resilience. When businesses anticipate interruptions, they may create flexible systems that adjust quickly. For example, preparing for supply chain assaults may encourage vendor variety and the use of strong monitoring technologies.

Another advantage is improved policy. Governments that forecast cyber dangers might create legislation to encourage safe software development, responsible data exchange, and critical infrastructure security.

Foresight is also useful in enhancing schooling. Universities can aim to integrate cybersecurity jobs, rather than outmoded technology, in future curriculum teaching capabilities.

Finally, foresight builds trust. When citizens notices institutions prepared for cyber crises, confidence in digital systems increases.[237]

[235] Paul Cornish and others, 'On Cyber Warfare', Chatham House Report 2010.

[236] Ibid.

[237] Joseph Nye Jr, 'Deterrence and Dissuasion in Cyberspace' (2017) 41 International Security 44.

Methods of Strategic Foresight in Cybersecurity

- **Horizon Scanning**

Horizon scanning recognises weak signals. Analysts monitor technological developments, academic research, social trends, and criminal activity. For example, early research into quantum computing divulges the possibility that future machines could shatter current encryption. Organisations that identified this risk began developing post-quantum cryptography.

- **Scenario Planning**

Scenario planning envisages detailed future worlds. A cybersecurity team might traverse scenarios such as 'global internet fragmentation, widespread AI-driven cybercrime, or collapse of trust in digital identity systems.' These scenarios help organisations test strategies under various conditions.

- **Red Teaming and War Gaming**

Red teaming entails conducting simulated assaults to uncover weaknesses. War gaming takes this concept to a strategic level, investigating how enemies may use political or economic factors. These exercises reveal blind spots and encourage creative thinking.

- **Technology Forecasting**

Technology forecasting examines emerging innovations such as artificial intelligence, biotechnology, and quantum computing. Each new technology generates opportunities and risks. Forecasting assists organisations prepare for both.

- **Stakeholder Engagement**

Cybersecurity is not just technical. Lawyers, sociologists, economists, psychologists, and educators all influence outcomes. Foresight exercises involving manifold stakeholders produce more realistic insights.[238]

- **Implementing Strategic Foresight in Organisations**

Strategic foresight should not endure an isolated workshop. It must flatter parts of organisational culture.

First, leadership commitment is essential. Executives must allot resources for foresight research and merge results into planning. Without leadership, foresight becomes symbolic rather than practical.

Second, interdisciplinary squad are necessary. Cybersecurity professionals should collaborate with specialist in public policy, and behavioural science. Human error remains a major reason of breaches, so understanding psychology matters.

Third, continuous learning is pivotal. Cyber threats develop quickly, therefore foresight must be ongoing. Scenarios, threat models, and training courses are regularly updated to ensure their relevance.

[238] Joseph Nye Jr, 'Deterrence and Dissuasion in Cyberspace' (2017) 41 International Security 44.

Fourth, communication is vital. Technical specialists must translate prescient judgement into understandable terms for decision-makers. Storytelling may make abstract dangers seem more logical.

Fifth, coordination between organisations strengthens security. Sharing threat intelligence across businesses and nations reduces redundancy and improves readiness.[239]

Ethical Challenges in Cybersecurity Foresight

Apprehending menace promote ethical questions. Surveillance system stewardship can threaten privacy. Predictive algorithms could conserve prejudice. Governments may substantiate surveillance in the name of national security.

Strategic foresight must therefore integrate 'holistic ethical analysis'. Decisions about 'data collection, artificial intelligence, and digital governance' should contemplate human rights. 'Transparency and accountability' boost legitimacy.[240]

Another challenge is inequality. Affluent institutions can fund next generation defense, while small businesses and public institutions suffer. Foresight shall eliminate the ambiguities by advocating for accessible, secure technology and training.

Strategic Foresight and Public Policy

Governments contribute a critical role in advancing the cyber security paradigm. 'National strategies, international agreements, and legislative frameworks' influence behaviour. Public policy can foster sturdy software development, demand breach disclosure, and protect critical infrastructure. It can also advance pioneering research such as 'quantum-resistant encryption'.

The Human Dimension of Cybersecurity

Behind every 'cyber breach' lies a personal story. Interacting with a malicious URL because they are tired. A student tends to downloading unauthorized software because it is affordable. A criminal group recruits emerging tech talent from marginalised communities. Strategic foresight must therefore combat social conditions that create vulnerability. Awareness campaigns, ethical technology design, and inclusive digital policies mitigate risk. When systems are user-friendly and transparent, people choose other options which are secure. When education is accessible, less people turn to cybercrime.[241]

Education policy matters are important concern. Schools and universities should educate on' digital literacy, ethical hacking, and critical thinking'. A society that recognises cybersecurity is more resilient.

[239] Herbert Lin, 'Cyber Conflict and International Humanitarian Law' (2012) 89 International Review of the Red Cross 515.
[240] Ibid.

[241] Michel Godet, 'Creating Futures: Scenario Planning as a Strategic Management Tool', Economica 2001.

Future Threats to Watch

Several emerging trends may shape cybersecurity in coming decades:

- Quantum computing could decipher existing cryptanalysis.
- Artificial intelligence may automatize cyberattacks and generate persuade misinformation.
- Biometric systems could be misused, creating Synthetic Identity Theft.
- Space-based infrastructure such as satellites may eventually become cyber targets.
- Internet-of-Things devices could allow large-scale disruption of cities.[242]

Strategic foresight helps various organisations systematize for these possibilities without panic. It spurs strategic thinking instead of crisis-driven reaction.

Looking ahead, cybersecurity experts must focus to an array of next-generation technologies and social trends that are likely to metamorphosis of digital vulnerabilities. These are not faint possibilities; many are already surreptitiously developing, and their ramifications may be intense if left unprepared for.

One of the most important future risks is 'quantum computing'. When practical, extensive quantum machines arrive, they may shatter widely used 'encryption systems' that protect banking, 'government communication', and personal data. This does not mean security will fall apart overnight, but organisations that causes hindrance in transitioning to quantum-resistant cryptography could face sudden exposure. Planning for this transition now is important because cryptographic change takes years.

Artificial intelligence represents a second wave of transformation. While AI tools reinforce defence through automatic detection and quick response, they also authorize the attackers. 'Deep fake audio and video' can be used for 'fraud, political manipulation, or corporate espionage'. AI-generated phishing emails are increasingly imperceptible as they imitate human writing techniques and local languages. In the future, malware might evolve on its own, learning from failed attacks and upgrade automatically.

A third concern lies in biometric security. 'Fingerprints, facial recognition, and voice identification' are incrementally used for authentication. Unlike 'passwords, biometric data' cannot be modified once stolen. If huge biometric databases are breached, individuals are likely to endure prolonged disruption of identity risks. Strategic foresight therefore demands combining biometric systems with enhanced 'confidentiality safeguards and backup verification methods'.

The sudden expansion of the 'Internet of Things'[243] is another critical area. 'Smart homes, connected cars, medical implants, and industrial sensors'[244] generate convenience but also widen the attack surface. Weak security in one device can enable the hackers to enter larger networks. Future cyber breaches may target complete cities by disrupting the whole network of the satellite systems, water treatment plants, or big payment systems of any institutes. Securing web and devices is therefore a

[242] Michel Godet, 'Creating Futures: Scenario Planning as a Strategic Management Tool', Economica 2001
[243] Lawrence Freedman, 'Strategy: A History', Oxford University Press 2013.
[244] Ibid.

concern of public safety, not merely a consumer choice. Space infrastructure is also venturing within cybersecurity planning. Satellites negotiate with communication, navigation, and weather forecasting. As more countries and private corporation launch satellites, the risk of cyber interference enhances. A synchronised attack on satellite systems could obstruct banking transactions, aviation routes, and emergency services co-occurrent.

Finally, social and psychological threats must be taken into account. 'Disinformation campaigns, online radicalisation, and data-driven political manipulation'[245] can lead to democratic backsliding without a single line of malicious code. Cybersecurity foresight must therefore embrace media literacy, ethical platform design, and strong legal frameworks.

None of these threats should amplify terror or technological pessimism. Instead, they remind us that cybersecurity is an unwavering allegiance. By identifying risks at an initial phase, investing in research, giving importance to education, and encouraging various organisations across sectors, societies can meet future challenges with utmost confidence rather than crisis or fear.

Creating a foresight-driven organisation requires patience. Employees must feel inspired to think creatively and further be open to ask question in various assumptions. Training programmes should include simulation drills and interdisciplinary dialogue. Organisations should focus on long-term thinking rather than short-term fixes.

Conclusion

In an era of an increasing ttechnological modernization, *Anticipating Tomorrow's Threats: Strategic Foresight in Cybersecurity* postulates a major truth: the future of cybersecurity cannot be affixed by counteractive measures alone. As advanced technologies such as 'artificial intelligence, quantum computing, sovereign systems, and hyperconnected infrastructures' reshape the international landscape, the threat environment advances in equidistant, comparatively faster than conventional defense mechanisms can adapt. Strategic foresight therefore appears not as an opulence, but as an essential requirement for organizations and nations looking for continued digital resilience.

The chapter emphasizes that anticipating tomorrow's threats entails more than technical proficiency or advanced security tools. Cybersecurity is not only a technical discipline; it is a social, political, and ethical challenge shaped by human decisions, power structures, and global inequalities. Every vulnerability outlines alternatives in system design, governance, and human behavior. Strategic foresight dispenses a structured way to envision possible futures, prophesy emerging risks, and prepare responsibly for uncertainty. It calls for the integration of technological expertise with fascination, empathy, and ethical reflection.

Through 'horizon scanning, scenario planning, and predictive analytics', executives can identify weak signals before they escalate into cataclysmic upheaval. By inserting foresight into organizational culture, institutions shift from a reactive posture of troubleshooting to a proactive stance of preparedness and adaptability. This kind of new initiatives encourages investment in resilient architectures, flexible policies, and professional development capable of responding to rapidly

[245] Herbert Lin, 'Cyber Conflict and International Humanitarian Law' (2012) 89 International Review of the Red Cross 515.

changing threats. Collaboration is midway to this approach. Governments, private enterprises, academia, and civil society must share intelligence and coordinate strategies to mitigate risk that transcend borders and sectors. Collective learning, transparent communication, and shared responsibility strengthen global cybersecurity ecosystems and reduce the likelihood of strategic startle.

Ultimately, the chapter delivers a simple but compelling message: the fate of cybersecurity is not predetermined. Strategic foresight repudiates to prognosticate the future with certainty; rather, it equips communities to anticipating probable developments with confidence and purpose. Through ‘deliberate planning, ethical awareness, and cooperative action’, we can convert ‘uncertainty into opportunity and construct resilient technological futures’. In doing so, strategic foresight becomes the cornerstone of ‘sustainable cybersecurity’—ensuring that tomorrow’s cyberspace is not merely defended, but circumspectly and responsibly shaped.