JOURNAL OF LAW, INFORMATION & SCIENCE
ISSN 0729-1485• Vol. 3, Issue 1 (2025) • https://jlisjournal.com/

REVIEW ARTICLE Article ID: JLIS/2025/3/1/006

Facial Recognition Technology and the Right to Privacy: Legal Challenges in Automated Identification

Dr. Nandini Iyer [1]

[1] *Faculty of Law and Technology, Indian Institute of Digital Governance, Bengaluru, India*

Received: 18 Jan 2025 **Revised:** 28 Jan 2025 **Accepted:** 28 March 2025 **Published:** 29 May 2025

ABSTRACT

Facial recognition technology has emerged as a significant application of artificial intelligence, enabling automated identification and verification of individuals through biometric characteristics. Its increasing deployment in law enforcement, airports, financial services, public spaces, and digital platforms has generated substantial benefits in security and convenience while simultaneously raising serious legal concerns. Unlike conventional passwords or identification documents, facial characteristics are intrinsic and difficult to replace once compromised. This article examines the relationship between facial recognition technology and the legal protection of privacy, focusing on biometric data, surveillance, consent, accuracy, algorithmic bias, data security, and governmental use. It argues that unrestricted deployment of facial recognition can create significant risks to individual autonomy and informational privacy, particularly where individuals are identified or monitored without meaningful knowledge or control. The article proposes a rights-based regulatory framework incorporating necessity, proportionality, purpose limitation, data minimization, transparency, human oversight, accuracy requirements, independent auditing, and effective remedies. The article concludes that facial recognition technology can provide legitimate security and identification benefits, but its deployment should remain subject to strong legal safeguards designed to protect privacy, dignity, equality, and individual autonomy.

Keywords: *Facial Recognition; Artificial Intelligence; Right to Privacy; Biometric Data; Surveillance; Data Protection; Cyber Law.*

1. Introduction

Cyberattacks have become increasingly sophisticated, automated, and difficult to detect through conventional security mechanisms. Organizations now face threats capable of operating at a speed that may exceed the capacity of human security teams.

Artificial intelligence offers a potential solution.

AI-based cybersecurity systems can continuously analyze network traffic, identify anomalies, detect suspicious behavior, classify malware, and recommend or initiate defensive measures.

The development of autonomous cybersecurity systems represents a further technological step.

Instead of merely informing a human operator that an attack has occurred, an autonomous system may automatically isolate a compromised device, block network traffic, terminate a suspicious process, or modify security configurations.

Such automation can substantially reduce response time.

However, it also creates an important legal question:

Who is responsible when an autonomous cybersecurity system makes the wrong decision?

If an AI system mistakenly identifies legitimate activity as malicious and blocks an essential service, the resulting harm may affect individuals, organizations, or third parties.

This article examines the legal implications of autonomous cyber defence and proposes a governance framework for balancing technological autonomy with human accountability.

2. Concept of Autonomous Cybersecurity

Autonomous cybersecurity refers to security technologies capable of detecting, analyzing, and responding to cyber threats with limited human intervention.

A conventional security system may operate as follows:

Threat Detection → Human Alert → Human Decision → Defensive Action

An autonomous system may instead operate as:

Threat Detection → AI Analysis → Automated Decision → Defensive Action

The second model can significantly reduce response time.

However, increased autonomy also increases the importance of system reliability, authorization boundaries, and accountability.

3. AI-Based Threat Detection

AI systems can process large volumes of cybersecurity information.

They may analyze:

- network traffic;
- login behavior;
- system logs;
- application activity;
- malware characteristics;
- communication patterns; and
- endpoint behavior.

Machine-learning models can identify patterns associated with potentially malicious activity.

However, AI systems may produce false positives.

A legitimate user may be incorrectly classified as suspicious.

Therefore, automated cybersecurity decisions should incorporate appropriate confidence thresholds and escalation mechanisms.

4. Automated Incident Response

An autonomous cybersecurity system may take defensive actions after detecting a suspected attack.

Possible actions include:

- isolating a device;
- blocking an IP address;
- disabling an account;
- terminating a process;
- restricting network access; or
- applying a security configuration.

These actions can be highly effective during rapidly developing attacks.
However, inappropriate automated responses may interrupt legitimate activities.
Organizations should therefore classify defensive actions according to their potential impact.

---

5. Levels of Autonomy

Autonomous cyber defence can be understood through different levels.
Level 1 – Human Decision
AI identifies a threat, but humans make the final decision.
Level 2 – AI Recommendation
AI recommends a defensive action that a human approves.
Level 3 – Conditional Automation
AI automatically responds to predefined low-risk threats.
Level 4 – High Autonomy
AI independently detects, evaluates, and responds to threats across multiple systems.
Higher autonomy should generally correspond with stronger safeguards and oversight.

---

6. Legal Responsibility

The use of autonomous cybersecurity systems raises complex questions concerning legal responsibility.
Potentially relevant actors include:

- system operators;
- organizations deploying the technology;
- software developers;
- cybersecurity vendors;
- cloud providers; and
- security administrators.

Responsibility should not simply be attributed to the AI system.
Artificial intelligence is a technological mechanism rather than an independently accountable legal actor.
The organization deploying the system should therefore maintain appropriate responsibility for its authorized operation.

---

7. Human Oversight

Human oversight is particularly important for high-impact automated actions.
Organizations should establish clear rules concerning when human approval is mandatory.
For example:
Low-Risk Threat → Automatic Response

Medium-Risk Threat → Automated Containment + Notification
High-Risk Threat → Human Authorization
This approach allows organizations to benefit from automation while maintaining meaningful human control over consequential decisions.

---

8. Proportionality
Cybersecurity responses should be proportionate to the threat.
Suppose an AI system detects suspicious network traffic.
Blocking one connection may be sufficient.
Completely shutting down an entire organizational network may be excessive.
Automated systems should therefore incorporate proportionality principles.
The response should be based upon factors such as:

- threat severity;
- confidence level;
- potential consequences;
- affected systems; and
- availability of less disruptive alternatives.

---

9. False Positives
False positives represent a significant risk.
An AI system may incorrectly identify legitimate behavior as malicious.
For example, unusual activity by an employee traveling internationally may resemble suspicious account behavior.
If the system automatically disables the employee's access, legitimate business activity could be disrupted.
High-impact automated actions should therefore incorporate mechanisms for review and rapid restoration.

---

10. Privacy Concerns
Autonomous cybersecurity systems may continuously monitor network activity.
This monitoring can involve:

- employee communications;
- login information;
- browsing activity;
- device identifiers;
- location-related information; and
- behavioral patterns.

Although monitoring may be necessary for cybersecurity, excessive surveillance can create privacy concerns.
Organizations should therefore limit monitoring to legitimate security purposes and implement appropriate access controls.

---

11. Automated Network Defence

An autonomous system may potentially block suspicious traffic or isolate compromised systems.
However, network actions may affect third parties.
For example, an incorrectly configured defensive system could block an external service that is actually legitimate.
Organizations should therefore carefully define the scope of automated defensive authority.
Automated systems should generally be prevented from taking actions beyond their explicitly authorized environment.

---

12. The Problem of Automated Retaliation

One particularly sensitive issue is automated retaliation against suspected attackers.
A defensive system might theoretically attempt to interfere with the attacker's infrastructure.
However, attribution in cyberspace can be extremely difficult.
The apparent source of an attack may be a compromised third-party computer rather than the actual attacker.
Automated retaliatory actions could therefore unintentionally affect innocent third parties.
For this reason, autonomous cyber defence should emphasize **containment and protection** rather than uncontrolled offensive action.

---

13. Cybersecurity Algorithms and Explainability

Organizations should be capable of explaining why an autonomous system took an important defensive action.
A useful audit record might include:

- detected event;
- relevant indicators;
- confidence score;
- system recommendation;
- action taken;
- time of action; and
- subsequent human review.

Explainability does not require exposing every technical detail of an AI model.
Rather, it requires sufficient information to reconstruct and evaluate important decisions.

---

14. Audit Trails

Auditability is essential for autonomous systems.
Every significant automated action should generate an appropriate record.
For example:
Detection → Analysis → Decision → Action → Outcome
Such records can assist with:

- incident investigation;
- legal disputes;
- regulatory compliance;
- system improvement; and
- responsibility assessment.

Without adequate records, it may be difficult to determine why an automated action occurred.

---

15. Cybersecurity Vendors

Organizations frequently obtain AI cybersecurity technologies from external vendors.
This creates additional accountability questions.
Contracts should clearly establish:

- security obligations;
- system performance requirements;
- update procedures;
- vulnerability disclosure;
- incident reporting;
- audit rights; and
- allocation of responsibility.

Organizations should also conduct appropriate vendor-risk assessments before deploying autonomous systems.

---

16. AI Model Security

The AI system itself may become a target.
Attackers could potentially attempt to manipulate the system's inputs or exploit weaknesses in the underlying model.
Possible threats include:

- adversarial inputs;
- data poisoning;
- model manipulation;
- compromised training data; and
- unauthorized system access.

Therefore, securing the AI model is itself part of cybersecurity.

---

17. Autonomous Systems and Critical Infrastructure

Autonomous cybersecurity becomes particularly important in critical infrastructure.
Energy, healthcare, transportation, telecommunications, and financial systems require rapid threat detection.
However, an incorrect automated decision may also produce serious consequences.
Consequently, high-impact environments should employ:

- strict authorization limits;
- redundant monitoring;
- human escalation;
- fail-safe mechanisms;
- independent auditing; and
- tested recovery procedures.

---

18. Proposed Governance Framework

A comprehensive governance model can be based on seven principles.
18.1 Defined Authority

The system's permitted actions must be clearly specified.
18.2 Risk Classification
Cybersecurity events should be classified according to potential impact.
18.3 Human Oversight
High-impact actions should require appropriate human involvement.
18.4 Proportionality
Defensive measures should correspond to the severity of the threat.
18.5 Auditability
Important automated actions should be recorded.
18.6 Security by Design
The autonomous system itself should be protected against manipulation.
18.7 Accountability
Organizations deploying autonomous systems should maintain clearly defined responsibility.

---

19. Proposed Technical Architecture

A secure autonomous defence architecture could operate through the following model:
Sensors → Threat Intelligence → AI Analysis → Risk Classification → Policy Engine → Automated Response → Human Oversight → Audit System
The **Policy Engine** is particularly important.
It can define which actions the AI system is authorized to perform automatically.
For example, the system may automatically isolate a single compromised endpoint but require human approval before shutting down a larger network segment.

---

20. Emergency Override

Autonomous cybersecurity systems should include mechanisms allowing authorized personnel to suspend automated operations.
An emergency override may be necessary when:

- the AI system behaves unexpectedly;
- false positives increase;
- a security rule is incorrectly configured;
- critical services are affected; or
- the system itself appears compromised.

Emergency controls should be tested periodically.

---

21. International Cybersecurity Challenges

Cyberattacks frequently cross national borders.
An autonomous defensive action in one country may affect infrastructure located elsewhere.
This creates international legal and jurisdictional questions.
Organizations should therefore design autonomous systems with geographically limited authority and clear restrictions on actions affecting external networks.
International cooperation will remain important as autonomous cyber defence develops.

---

22. Future Development

Future autonomous cybersecurity systems may become increasingly capable of:

- predicting attacks;
- identifying unknown malware;
- automatically restructuring network defenses;
- coordinating multiple security systems; and
- adapting to changing attack strategies.

Such systems could significantly improve cyber resilience.
However, greater technological autonomy will make governance increasingly important.
The legal framework must evolve alongside the technology.

---

23. Conclusion

Artificial intelligence is transforming cybersecurity from a primarily reactive discipline into an increasingly automated and predictive field.
Autonomous cybersecurity systems can detect threats rapidly, analyze complex network behavior, and initiate defensive actions at machine speed. These capabilities can significantly improve organizational resilience.
However, autonomy creates legal and technological risks.
Incorrect automated decisions can disrupt legitimate services, affect third parties, invade privacy, or cause financial and operational harm.
The central legal challenge is therefore not whether autonomous cybersecurity should be permitted, but **under what conditions its autonomy should be exercised**.
A responsible framework should establish clearly defined authority, proportionality, human oversight, auditability, security controls, vendor accountability, and effective emergency-override mechanisms.
High-risk automated actions should remain subject to meaningful human supervision, while routine low-risk responses can benefit from greater automation.
The future of autonomous cyber defence should therefore be based on a principle of **controlled autonomy**: artificial intelligence should be sufficiently autonomous to respond rapidly to cyber threats, but sufficiently constrained to ensure that its actions remain lawful, proportionate, reviewable, and accountable.
In this model, technology and law operate together to create a cybersecurity environment that is not only faster and more intelligent, but also responsible and trustworthy.

1. Introduction

Cyberattacks have become increasingly sophisticated, automated, and difficult to detect through conventional security mechanisms. Organizations now face threats capable of operating at a speed that may exceed the capacity of human security teams.
Artificial intelligence offers a potential solution.
AI-based cybersecurity systems can continuously analyze network traffic, identify anomalies, detect suspicious behavior, classify malware, and recommend or initiate defensive measures.
The development of autonomous cybersecurity systems represents a further technological step.
Instead of merely informing a human operator that an attack has occurred, an autonomous system may automatically isolate a compromised device, block network traffic, terminate a suspicious process, or modify security configurations.
Such automation can substantially reduce response time.

However, it also creates an important legal question:

Who is responsible when an autonomous cybersecurity system makes the wrong decision?

If an AI system mistakenly identifies legitimate activity as malicious and blocks an essential service, the resulting harm may affect individuals, organizations, or third parties.

This article examines the legal implications of autonomous cyber defence and proposes a governance framework for balancing technological autonomy with human accountability.

---

2. Concept of Autonomous Cybersecurity

Autonomous cybersecurity refers to security technologies capable of detecting, analyzing, and responding to cyber threats with limited human intervention.

A conventional security system may operate as follows:

Threat Detection → Human Alert → Human Decision → Defensive Action

An autonomous system may instead operate as:

Threat Detection → AI Analysis → Automated Decision → Defensive Action

The second model can significantly reduce response time.

However, increased autonomy also increases the importance of system reliability, authorization boundaries, and accountability.

---

3. AI-Based Threat Detection

AI systems can process large volumes of cybersecurity information.

They may analyze:

- network traffic;
- login behavior;
- system logs;
- application activity;
- malware characteristics;
- communication patterns; and
- endpoint behavior.

Machine-learning models can identify patterns associated with potentially malicious activity.

However, AI systems may produce false positives.

A legitimate user may be incorrectly classified as suspicious.

Therefore, automated cybersecurity decisions should incorporate appropriate confidence thresholds and escalation mechanisms.

---

4. Automated Incident Response

An autonomous cybersecurity system may take defensive actions after detecting a suspected attack.

Possible actions include:

- isolating a device;
- blocking an IP address;
- disabling an account;
- terminating a process;
- restricting network access; or

- applying a security configuration.

These actions can be highly effective during rapidly developing attacks.
However, inappropriate automated responses may interrupt legitimate activities.
Organizations should therefore classify defensive actions according to their potential impact.

---

5. Levels of Autonomy

Autonomous cyber defence can be understood through different levels.
Level 1 – Human Decision
AI identifies a threat, but humans make the final decision.
Level 2 – AI Recommendation
AI recommends a defensive action that a human approves.
Level 3 – Conditional Automation
AI automatically responds to predefined low-risk threats.
Level 4 – High Autonomy
AI independently detects, evaluates, and responds to threats across multiple systems.
Higher autonomy should generally correspond with stronger safeguards and oversight.

---

6. Legal Responsibility

The use of autonomous cybersecurity systems raises complex questions concerning legal responsibility.
Potentially relevant actors include:

- system operators;
- organizations deploying the technology;
- software developers;
- cybersecurity vendors;
- cloud providers; and
- security administrators.

Responsibility should not simply be attributed to the AI system.
Artificial intelligence is a technological mechanism rather than an independently accountable legal actor.
The organization deploying the system should therefore maintain appropriate responsibility for its authorized operation.

---

7. Human Oversight

Human oversight is particularly important for high-impact automated actions.
Organizations should establish clear rules concerning when human approval is mandatory.
For example:
Low-Risk Threat → Automatic Response
Medium-Risk Threat → Automated Containment + Notification
High-Risk Threat → Human Authorization
This approach allows organizations to benefit from automation while maintaining meaningful human control over consequential decisions.

---

8. Proportionality

Cybersecurity responses should be proportionate to the threat.
Suppose an AI system detects suspicious network traffic.
Blocking one connection may be sufficient.
Completely shutting down an entire organizational network may be excessive.
Automated systems should therefore incorporate proportionality principles.
The response should be based upon factors such as:

- threat severity;
- confidence level;
- potential consequences;
- affected systems; and
- availability of less disruptive alternatives.

---

9. False Positives

False positives represent a significant risk.
An AI system may incorrectly identify legitimate behavior as malicious.
For example, unusual activity by an employee traveling internationally may resemble suspicious account behavior.
If the system automatically disables the employee's access, legitimate business activity could be disrupted.
High-impact automated actions should therefore incorporate mechanisms for review and rapid restoration.

---

10. Privacy Concerns

Autonomous cybersecurity systems may continuously monitor network activity.
This monitoring can involve:

- employee communications;
- login information;
- browsing activity;
- device identifiers;
- location-related information; and
- behavioral patterns.

Although monitoring may be necessary for cybersecurity, excessive surveillance can create privacy concerns.
Organizations should therefore limit monitoring to legitimate security purposes and implement appropriate access controls.

---

11. Automated Network Defence

An autonomous system may potentially block suspicious traffic or isolate compromised systems.
However, network actions may affect third parties.
For example, an incorrectly configured defensive system could block an external service that is actually legitimate.
Organizations should therefore carefully define the scope of automated defensive authority.

Automated systems should generally be prevented from taking actions beyond their explicitly authorized environment.

---

12. The Problem of Automated Retaliation

One particularly sensitive issue is automated retaliation against suspected attackers.
A defensive system might theoretically attempt to interfere with the attacker's infrastructure.
However, attribution in cyberspace can be extremely difficult.
The apparent source of an attack may be a compromised third-party computer rather than the actual attacker.
Automated retaliatory actions could therefore unintentionally affect innocent third parties.
For this reason, autonomous cyber defence should emphasize **containment and protection** rather than uncontrolled offensive action.

---

13. Cybersecurity Algorithms and Explainability

Organizations should be capable of explaining why an autonomous system took an important defensive action.
A useful audit record might include:

- detected event;
- relevant indicators;
- confidence score;
- system recommendation;
- action taken;
- time of action; and
- subsequent human review.

Explainability does not require exposing every technical detail of an AI model.
Rather, it requires sufficient information to reconstruct and evaluate important decisions.

---

14. Audit Trails

Auditability is essential for autonomous systems.
Every significant automated action should generate an appropriate record.
For example:
Detection → Analysis → Decision → Action → Outcome
Such records can assist with:

- incident investigation;
- legal disputes;
- regulatory compliance;
- system improvement; and
- responsibility assessment.

Without adequate records, it may be difficult to determine why an automated action occurred.

---

15. Cybersecurity Vendors

Organizations frequently obtain AI cybersecurity technologies from external vendors.
This creates additional accountability questions.
Contracts should clearly establish:

- security obligations;
- system performance requirements;
- update procedures;
- vulnerability disclosure;
- incident reporting;
- audit rights; and
- allocation of responsibility.

Organizations should also conduct appropriate vendor-risk assessments before deploying autonomous systems.

---

16. AI Model Security

The AI system itself may become a target.

Attackers could potentially attempt to manipulate the system's inputs or exploit weaknesses in the underlying model.

Possible threats include:

- adversarial inputs;
- data poisoning;
- model manipulation;
- compromised training data; and
- unauthorized system access.

Therefore, securing the AI model is itself part of cybersecurity.

---

17. Autonomous Systems and Critical Infrastructure

Autonomous cybersecurity becomes particularly important in critical infrastructure.

Energy, healthcare, transportation, telecommunications, and financial systems require rapid threat detection.

However, an incorrect automated decision may also produce serious consequences.

Consequently, high-impact environments should employ:

- strict authorization limits;
- redundant monitoring;
- human escalation;
- fail-safe mechanisms;
- independent auditing; and
- tested recovery procedures.

---

18. Proposed Governance Framework

A comprehensive governance model can be based on seven principles.

18.1 Defined Authority

The system's permitted actions must be clearly specified.

18.2 Risk Classification

Cybersecurity events should be classified according to potential impact.

18.3 Human Oversight

High-impact actions should require appropriate human involvement.

18.4 Proportionality

Defensive measures should correspond to the severity of the threat.

18.5 Auditability

Important automated actions should be recorded.

18.6 Security by Design

The autonomous system itself should be protected against manipulation.

18.7 Accountability

Organizations deploying autonomous systems should maintain clearly defined responsibility.

---

19. Proposed Technical Architecture

A secure autonomous defence architecture could operate through the following model:

Sensors → Threat Intelligence → AI Analysis → Risk Classification → Policy Engine → Automated Response → Human Oversight → Audit System

The **Policy Engine** is particularly important.

It can define which actions the AI system is authorized to perform automatically.

For example, the system may automatically isolate a single compromised endpoint but require human approval before shutting down a larger network segment.

---

20. Emergency Override

Autonomous cybersecurity systems should include mechanisms allowing authorized personnel to suspend automated operations.

An emergency override may be necessary when:

- the AI system behaves unexpectedly;
- false positives increase;
- a security rule is incorrectly configured;
- critical services are affected; or
- the system itself appears compromised.

Emergency controls should be tested periodically.

---

21. International Cybersecurity Challenges

Cyberattacks frequently cross national borders.

An autonomous defensive action in one country may affect infrastructure located elsewhere.

This creates international legal and jurisdictional questions.

Organizations should therefore design autonomous systems with geographically limited authority and clear restrictions on actions affecting external networks.

International cooperation will remain important as autonomous cyber defence develops.

---

22. Future Development

Future autonomous cybersecurity systems may become increasingly capable of:

- predicting attacks;
- identifying unknown malware;
- automatically restructuring network defenses;
- coordinating multiple security systems; and
- adapting to changing attack strategies.

Such systems could significantly improve cyber resilience.

However, greater technological autonomy will make governance increasingly important.
The legal framework must evolve alongside the technology.

23. Conclusion

Artificial intelligence is transforming cybersecurity from a primarily reactive discipline into an increasingly automated and predictive field.
Autonomous cybersecurity systems can detect threats rapidly, analyze complex network behavior, and initiate defensive actions at machine speed. These capabilities can significantly improve organizational resilience.
However, autonomy creates legal and technological risks.
Incorrect automated decisions can disrupt legitimate services, affect third parties, invade privacy, or cause financial and operational harm.
The central legal challenge is therefore not whether autonomous cybersecurity should be permitted, but **under what conditions its autonomy should be exercised**.
A responsible framework should establish clearly defined authority, proportionality, human oversight, auditability, security controls, vendor accountability, and effective emergency-override mechanisms.
High-risk automated actions should remain subject to meaningful human supervision, while routine low-risk responses can benefit from greater automation.
The future of autonomous cyber defence should therefore be based on a principle of **controlled autonomy**: artificial intelligence should be sufficiently autonomous to respond rapidly to cyber threats, but sufficiently constrained to ensure that its actions remain lawful, proportionate, reviewable, and accountable.
In this model, technology and law operate together to create a cybersecurity environment that is not only faster and more intelligent, but also responsible and trustworthy.

Declarations

Conflict of interest: The authors declare that they have no conflict of interest.

Funding: This review received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Ethical approval: Not applicable. This article is a review of published literature and did not involve human participants or animal subjects.

Data availability: No new data were generated or analyzed in support of this review. All sources relied upon are cited in the References.

Use of AI tools: No generative artificial intelligence tool was used in the conception, drafting or analysis of this article. The authors take full responsibility for the content.

Author contributions: Author contributed to the conception of the review, the survey of the literature, and the drafting and revision of the manuscript. Author read and approved the final version.