International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(18s) 2026 pp. 1035-1058
https://doi.org/10.70917/ijcisim-2026-4940
Article

AI Driven Cyber Security Framework for Next Generation Optical Communication Networks

Pradyumn Warasee[1], Naveen Kumar[2]

[1] Amity Institute Of Information Technology, Amity University Patna, Amity Institute Of Information Technology, Amity University Patna. Email:-pradyumn707@gmail.com
[2]Amity Institute Of Information Technology, Amity University Patna. Email:-nav.bharti@gmail.com

Abstract: Next generation optical communication networks will be progressively more programmable, software based, elastic, multi-domain and highly-connected with AI-intensive cloud and edge workloads. This progress expands capabilities and automation, but also expands cyber-physical attack surface in terms of optical physical layer, management plane, control plane and AI analytics pipeline. This paper describes an Artificial Intelligence-Driven Cyber Security Framework (AI-DCF) for future optical communication networks. This framework integrates optical telemetry, machine-learning-based anomaly identification, zero-trust access restriction, secure network governance, cryptographic agility, post-quantum capability and quantum-key-distribution aware key services into a unified framework. This paper applies a theoretical framework based on existing literature to consolidate recent research on smart optical networking, optical-layer attack detection, SDN security, adversarial machine learning, and quantum-safe security standards. The solution presents a framework that includes five nested layers (governance and risk; AI security analytics; optical-control security; optical-physical security; crypto-agile resilience). Results are represented in a threat taxonomy, a framework architecture, evidence mapping, and capability assessment tables. The study found that AI-based optical security needs to be approached not as an intrusion detection solution per se, but as a closed-loop, validating and standards-aligned cyber-physical system. The work provides insight for researchers, network organizations, cloud providers, and telecom operators in creating hardened optical infrastructures for 6G, AI data centers, smart cities and other critical digital services. In response to implementation-focused review feedback, the revised version also adds a small synthetic simulation, a proof-of-concept deployment architecture, an ML workflow demo, and an illustrative attack-detection example to demonstrate how the framework may be operationalized before field-test validation.
Keywords: Artificial intelligence, optical communication network, cyber security, machine learning, anomaly detection, software-defined optical network, zero trust, quantum key distribution, post-quantum cryptography, optical network security. proof-of-concept, simulation, ML workflow, attack detection

1. Introduction

Optical communications networks are the high-capacity transport substrate for cloud computing, mobile backhaul, financial networks, government communications, research networks, submarine systems, and emerging 6G services. They are valuable due to huge bandwidth, low latency, protocol transparency, and the ability to transport data across long distances with high reliability. The same transparency and scale that make optical systems attractive also pose security challenges, however. Modern optical networks are no longer isolated point-to-point fiber routes; they consist of elastic optical networks like the aforementioned, reconfigurable optical add-drop multiplexers, wavelength-division multiplexing,passive optical networks, software-defined optical controllers, coherent transceivers, telemetry systems, and automated management platforms. A recent survey of optical security found physical-layer threats, control-plane manipulation, jamming, crosstalk exploitation, wavelength injection, fiber tapping, and SDN misuse need to be addressed simultaneously in a cross-layer security model [25].

Optical communications is thus a cyber-physical cyber security problem. Attackers can target bits, protocols, controllers, management accounts, software supply chains, AI models, or the optical medium itself. A completely IP-

layer firewall will fail to detect slight optical power alteration on a wavelength channel, an optical time-domain reflectometer will not detect compromised SDN credentials, machine-learning algorithms capable of providing robust, high-performing classification methods will become untrustworthy if trained on poisoned data and deployed without drift monitoring. This is what is leading the current guidance in cyber security and AI governance to adopt a risk-management perspective over a one-tool perspective. NIST CSF 2.0 is a general taxonomy to understand, assess, prioritise and communicate cyber risks [1], whereas NIST zero trust rejects implicit trust in network location and relies on continuous authentication and authorization for access to resources [2]. Next-generation networks are also transitioning towards quantum security. In 2024, NIST finalized its first post-quantum cryptography standards, which encouraged administrators to initiate transition planning [5].

Optical infrastructure can be particularly important for this transition as the transport with high capacities can carry encrypted data that may be stored today and decrypted later if cryptographic migration is delayed. Meanwhile, optical channels provide quantum key distribution (QKD) infrastructure to enable keys to be produced with quantum properties of optical signals. For organizations desiring infrastructure safeguarding beyond traditional public-key presumptions, ETSI considers QKD as a complementary quantum-safe method, which can be further augmented by post-quantum cryptography [4]. Artificial intelligence is increasingly employed in optical communications to support performance monitoring, failure prediction, resource management, digital signal processing, and autonomous network operation. Gu et al. described the need for intelligent optical networks so there is less manual interaction and more automation [6]. Wang and Zhang have already described techniques for deep learning (DL) in optical communication such as convolutional neural networks, recurrent neural networks, generative adversarial networks, and deep reinforcement learning for physical and network layer tasks [8].

These approaches lay a strong foundation for AI-driven security, since optical networks produce massive quantities of time-series, spectrum, power, quality-of-transmission, alarm, control, and flow data that are not easily processed manually during real time. The security challenge is that AI isn't trustworthy. Adversarial machine learning also introduces evasion, poisoning, privacy, Trojan, and abuse threats to the AI lifecycle [17]. AI-driven optical security must thereby encompass ML models for identification: model governance and integrity of training data; adversarial testing; explainability; human oversight; policy enforcement; and rollback mechanisms. ISO/IEC 42001:2023 provides a management-system approach for organizations that use or provide AI systems [19], while the NIST AI RMF fosters trustworthiness across AI design, development, use, and evaluation [16]. This article presents an AI-Driven Cyber Security Framework (AI-DCF) for future optical communication networks. The paper presents a standards-compatible, multi-layer framework that integrates recent optical security literature, AI risk guidance, zero-trust principles, optical management security, and quantum-safe migration. The paper seeks to establish a research-oriented conceptual framework rather than a lab experiment. Its goal is to assist network operators and researchers in organizing future implementations, evaluations, and deployment of prototypes.

2. Literature Review / Thematic Discussion

2.1 Intelligent Optical Networks and AI-Based Automation

The literature on intelligent optical network indicates that AI and ML are applied with the capability of organizing complexity in routing, resource allocation, quality-of-transmission estimation, fault management, optical performance monitoring and survivability. Gu et al. categorize ML deployments in optical networks as network control and resource management and monitoring and survivability [6]. This classification is crucial for cyber security because the same telemetry and prediction pipelines used for performance automation can also identify abnormal behavior. Such issues, defined in quality-of-transmission metrics that may differ from learned evolution patterns, could suggest failures, misconfiguration, or adversarial threats. Deep learning builds on this capability by learning spatial, spectral, and temporal patterns from optical data. The work by Wang and Zhang discusses CNNs for image-like optical data, RNNs and LSTMs for sequential data, GANs for data augmentation, and deep reinforcement learning for adaptive network decisions [8]. Security frameworks might employ these techniques on different scales: CNNs to solve for constellation diagrams or spectral images; recurrent models to model the evolution of alarms and traffic; autoencoders

to uncover newness; GANs to enrich unusual attack datasets, and reinforcement learning to facilitate adaptive mitigations once an attack is detected. But these methods have to be limited by policy and safety checks since autonomous mitigation in carrier-grade transport networks would impact important services. The shift toward AI-driven traffic and data-center interconnection elevates the need for optical readiness. The industry report of 2025 stated that AI workloads in metro and long-haul optical networks are already affecting this, while optical capacity, latency, and automation needs at the communication service provider level are currently being evaluated in the light of the increase in AI traffic. This trend means optical cyber security is no longer confined to the realm of a back office transport issue, but rather it becomes integral to national digital resilience; cloud service provision, AI-factory scalability and crucial infrastructure security.

2.2 Optical-Layer Threats: Fiber Tapping, Jamming, Crosstalk, and Signal Injection

The physical-layer security in optical fibers is underestimated, because optical fiber links are perceived as challenging to tap. Recent work disputes that presumption. Spurny et al. examine physical-layer component risks in optical fiber infrastructures and classify splitter insertion, macrobending leakage, back-reflection, crosstalk, and multiplexer-related leakage as practical issues [12]. Their measurements demonstrate that in certain cases, an attack can produce only modest attenuation changes, thus manual inspection alone is not adequate for timely detection. Thus, the security mechanism required must combine optical monitoring, baseline learning, alarms, and physical-route protection. Gazani et al. 2026 survey takes this viewpoint across WDM, TDM, PON, EON, and IP-over-WDM architectures and identifies physical, control, and cross-layer threats such as fiber tapping, optical jamming, crosstalk exploitation, signal injection, malicious signaling, and SDN manipulation [25]. This indicates that optical cyber security should be cross-layer. The same attack might start out in physical form, propagate through the optical layer and eventually result in service degradation, influence routing decisions, and eventual manifestation as a higher-layer availability event. An IP-only observation of the traffic only will misclassify an event as congestion instead of optical manipulation. Machine learning techniques are already being considered for optical-layer attack detection and localization. Furdek et al. present cognitive security diagnostics including supervised, semi-supervised, and unsupervised learning for optical-layer attack detection and localization, and discuss integration with network management systems (NMS) [7]. Abdelli et al. integrate an autoencoder and an attention-based bidirectional GRU to identify and localize anomalous fiber activity, such as fiber cuts and optical eavesdropping attacks, yielding an F1 score of 96.86% for anomaly detection and 98.2% average identification accuracy [10]. Behera et al. suggest real-time anomaly detection based on encoder-decoder prediction of QoT evolution along with statistical hypothesis testing for less dependency on labeled anomalies [9].

2.3 Software-Defined Optical Networks and Control-Plane Security

Software-defined networking provides centralized view, programmability, and fast reconfiguration. These features in optical networks enable elastic bandwidth, dynamic lightpath provisioning, and cross-layer optimization. That said, the security risks of centralization and programmability are also a concern. SDN security surveys have pointed out that controller compromise, application vulnerabilities, malicious northbound or southbound interfaces, and denial-of-service attacks on the controller are possible impacts to a wide scale [14]. In the scope of optical system, a hacked controller can misroute lightpaths, inappropriately allocate spectrum, suppress alarms, or author policies which lead to crosstalk or resource starvation. ETSI TS 103 961 is very specific and directly applicable since it provides a framework for the security and management of optical network equipment and services. It describes functions including real-time topology discovery, device and service configuration,

FCAPS management and network resilience assurance capability [3]. It also highlights the relationship of trust between the system entities and optical network devices. This is all a part of a zero-trust model in which any management action, device identity, controller service, telemetry collector, and policy actuator must be authenticated, authorized, logged, and continuously validated. The control-plane element of the proposed framework therefore treats SDN and management systems as security-critical assets rather than mere automation tools. Intent verification, least-privilege controller roles, signed policies, secure boot, attestation, configuration backups, anomaly scoring, and

rollback are necessary. Similarly, the control plane needs model outputs from AI analytics but should not blindly execute them. Finally, a policy decision point should consider model confidence, service criticality, possible collateral damage and human approval before enforcing any changes.

2.4 AI Security, Adversarial Machine Learning, and Trustworthy Analytics

The AI lifecycle itself must secure an AI-driven security framework. NIST articulates adversarial machine learning in attacker objectives, knowledge, capabilities, lifecycle stages, and mitigations [17]. Some examples of the adversarial ML threats in optical networks are poisoning telemetry while training, generating optical perturbations that avoid anomaly detection, manipulating labels in incident databases, stealing detection models, or denial of service attacks against inference pipelines. These attacks are particularly insidious since the network operator may give a lot of operational trust to an AI system that seems statistically true in normal circumstances. NIST AI RMF and ISO/IEC 42001 both serve as the governance frameworks for AIs. The AI RMF seeks to ensure the degree of trustworthiness of AI design, development, use, and assessment [16]. ISO/IEC 42001 sets forth the requirements for the implementation and periodic improvements of an AI system in organizations that supply or operate an AI product or service [19]. In optical cyber security, these requirements mean that models should include written guidance on purpose, data lineage, risk assessment, measures of performance, bias or coverage determination, fallback process, and change governance. Generative AI and self-governing security agents also pose another level of risk. OWASP finds specific LLM risks like prompt injection, insecure output handling, training data poisoning, model denial of service, supply chain vulnerabilities, sensitive information disclosure, excessive agency, overreliance, and model theft [18]. To a certain extent, if LLMs are indeed used as security copilots for optical network operations, the approach needs to secure them away from direct command execution, validate all generated configurations, restrict access to topology secrets, log tool calls, and require deterministic policy checks before operations can begin.

2.5 Quantum-Safe Security and Crypto-Agility

Post-quantum transition is a central requirement for long-lived optical infrastructures. NIST finalized the first three post-quantum encryption standards in 2024 and encouraged system administrators from all sectors (application designers and service providers) to begin transitioning to the new standards

[5]. The harvest-now-decrypt-later threat applies where high-value optical networks transport data over significant distances and where confidentiality must remain intact for years. This paper treats crypto-agility as a core framework layer rather than an optional upgrade. QKD is relevant because it can generate shared secret keys using optical signals. ETSI describes QKD as a complementary quantum-safe technique to post-quantum cryptography and notes key work areas such as optical characterization, implementation security, security proofs, interoperability, key management services, network architectures, and QKD in SDN [4]. QKD does not remove the need for authentication, device security, implementation assurance, or PQC; instead, it provides an additional key-establishment option for selected high-value routes. The proposed framework therefore uses QKD as part of a layered and risk-based strategy rather than as a universal replacement for cryptography. Recent deployments and research on quantum-safe optical transport highlight the practical direction of the field. NTT announced a post-quantum secure transport system based on an open optical transponder and rapid transition mechanisms [22]. Other work on quantum-safe metro-optimized optical transport networks integrates QKD with transport infrastructure to strengthen the optical physical layer against future quantum threats [24]. The implication for AI-DCF is that security analytics must not only detect attacks but also track cryptographic inventory, key lifetimes, algorithm status, route sensitivity, and migration readiness.

Table 1: Literature themes informing the AI-DCF design.

Theme	Representative recent evidence	Security implication	Framework layer
AI for optical networks	ML supports optical-network monitoring, control, resource	Telemetry can support security detection and predictive defense.	AI analytics

	management, and survivability [6], [8].		
Optical-layer attacks	Fiber tapping, splitter insertion, crosstalk, jamming, and signal injection are documented risks [12], [25].	Physical-layer telemetry must be part of cyber security.	Optical-physical
ML attack detection	Cognitive diagnostics and autoencoder/BiGRU approaches detect and localize optical anomalies [7], [10].	Models can identify faults and attacks faster than manual review.	AI analytics
SDN/control security	SDN introduces programmable control but also controller and API attack surfaces [14].	Controllers require zero trust, policy signing, and runtime verification.	Optical-control
Optical management security	ETSI TS 103 961 defines optical management functions and trust associations [3].	NMS/EMS and device management must be secured as first-class assets.	Governance/control
AI model security	NIST AML taxonomy addresses evasion, poisoning, privacy, and related attacks [17].	The detector itself needs monitoring and adversarial testing.	Governance/AI
Quantum-safe transition	NIST PQC standards and ETSI QKD guidance support quantum-safe migration [4], [5].	Optical transport must become crypto-agile.	Crypto-resilience

3. Methodology

This paper uses a literature-based conceptual design methodology. The purpose is not to report a new laboratory experiment; instead, it synthesizes recent peer-reviewed literature, standards, and industry evidence into a deployable cyber security framework for next-generation optical networks.

• Scope definition: The study focuses on terrestrial next-generation optical communication networks, including WDM, PON, EON, IP-over-WDM, SDN-enabled optical transport, data-center interconnect, and quantum-safe optical transport. The security scope covers physical, control, management, cryptographic, and AI-model layers [3], [25].

• Literature selection: Sources were selected from 2020-2026, with emphasis on optical network security, intelligent optical networks, ML-based anomaly detection, SDN security, zero trust, AI risk management, adversarial ML, post-quantum cryptography, and QKD [1]-[25].

• Threat modeling: Threats were grouped into five categories: optical-physical attacks, control-plane attacks, management-plane compromise, AI-pipeline attacks, and quantum-era cryptographic risk. This structure follows recent cross-layer optical-security literature [25] and NIST risk-management principles [1].

• Framework synthesis: The proposed AI-DCF was derived by mapping each threat category to preventive, detective, corrective, and governance controls. Zero trust was used as the access-control model [2], ETSI TS 103 961 was used for optical management functions [3], and NIST AI/AML guidance was used for model-risk controls [16], [17].

• Evidence mapping: Reported results from ML-based optical anomaly studies were used to justify AI-assisted detection. Abdelli et al. report 96.86% F1 for anomaly detection and 98.2% identification accuracy in fiber-monitoring experiments [10]. Behera et al. show how multi-step QoT prediction supports real-time anomaly detection without labeled anomalies [9].

• Evaluation logic: Because the framework is conceptual, the results section uses capability mapping, design tables, and literature-derived performance indicators rather than a fabricated dataset. Maturity scores are used as a planning tool, not as empirical proof.

• Validation criteria: A framework component is considered relevant if it addresses an identified threat, has support in recent literature or standards, and can be integrated into optical network operations without violating reliability requirements.

• Limitations: The paper does not measure performance on a new physical testbed. Future work should implement AI-DCF in a real SDN-enabled optical testbed with attack emulation, adversarial model testing, and operational cost evaluation.

3.1 Proposed AI-DCF Architecture

Figure 1 illustrates the five-layer AI-Driven Cyber Security Framework. The layers are intentionally ordered from governance to cryptographic resilience, but implementation is closed-loop rather than strictly sequential. Telemetry flows upward from optical devices, ROADMs, amplifiers, transceivers, controllers, NMS/EMS, flow collectors, identity systems, and key-management services. Decisions flow downward through policy decision points, orchestration systems, and controlled response actions.

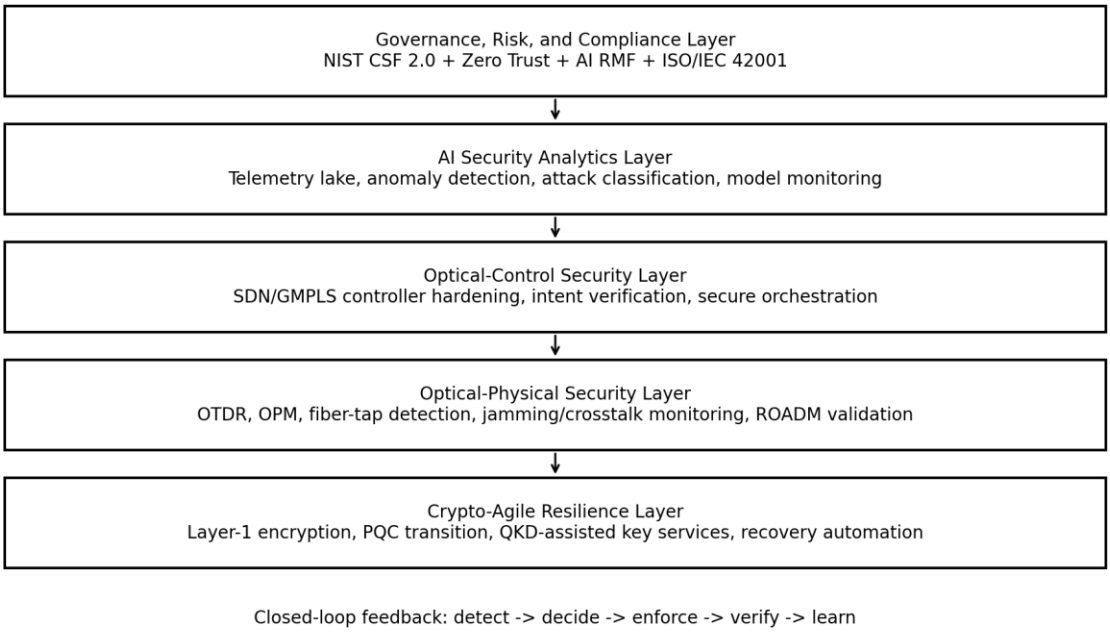

Figure 1: AI-DCF layered architecture for next-generation optical communication networks.

The governance layer sets out risk appetite, responsibility, policies, compliance, service criticality, and assurance measures. It conforms security results with NIST CSF 2.0 functions and zero-trust principles. It also defines model governance, data governance, change control, audit logging, and incident accountability. This layer avoids a common failure in AI-security projects where models are deployed without clear policy boundaries. The AI security

analytics layer takes in telemetry and establishes the situational awareness necessary for closed-loop security. Its roles encompass data normalization, feature extraction, feature stores, supervised detection, semi-supervised detection,

unsupervised anomaly detection, sequence prediction, graph-based topology analytics, explainability, and model drift monitoring. The objective is not just to call attention to attacks, but to gauge confidence, compromised routes, potential root causes, and mitigation suggestions. The optical-control security layer secures the SDN controller, orchestration APIs, GMPLS/PCEP/NETCONF or other management interfaces, controller applications, and policy deployment. It leverages least privilege, signed intents, configuration validation, attestation, secure controller clustering, and automated rollback.

This layer is critical because a high-confidence detection event is of little usable functionality when the response path is neither secure nor reliable. In this section, the optical-physical security layer is used to guarantee fiber and photonic component security. OTDR traces, optical performance monitoring, power levels, bit-error-rate trends, spectrum occupancy, ROADM states, amplifier telemetry, laser parameters, and physical access events are integrated. The goal is to identify fiber tapping, splitter insertion, macrobending leakage, jamming, crosstalk, signal injection, and unexpected changes of transmission behavior. The crypto-agile resilience layer will protect both confidentiality and long-term security. This encompasses Layer-1 encryption and MACsec or OTN encryption wherever applicable, secure key management, PQC migration inventory, QKD-assisted key services for sensitive routes, cryptographic algorithm agility, and recovery automation. This layer realizes that detection is limited if sensitive data is transmitted over links that will be unable to weather future cryptanalytic threats.

3.2 Threat Model

The threat model assumes adversaries with heterogeneous capabilities: external remote attackers, malicious insiders, compromised vendors, physically proximate attackers near fiber routes or cabinets, adversarial tenants in shared infrastructure, and advanced adversaries preparing for future quantum decryption. Table 2 summarizes the threat model.

Table 2: Threat model for AI-DCF.

Threat class	Attack examples	Likely indicators	AI-DCF controls	Key references
Optical-physical	Fiber tapping, splitter insertion, macrobending, jamming, crosstalk, signal injection	Subtle attenuation change, spectral anomaly, route-specific BER rise, unexpected OTDR trace	OPM/OTDR baselines, anomaly models, route risk scoring, physical access correlation	[10], [12], [25]
Control plane	Controller spoofing, malicious app, API abuse, false intent, route manipulation	Unexpected lightpath changes, policy conflicts, repeated controller errors	Zero trust, signed policies, controller attestation, least privilege, rollback	[2], [14]
Management plane	Credential theft, NMS/EMS manipulation, insecure SNMP/NETCONF, config tampering	Unauthorized configuration, topology mismatch, missing logs	Trust associations, MFA, secure management protocols, immutable logging	[3]
AI pipeline	Data poisoning, evasion, model theft, drift, excessive	Model confidence instability, false negatives,	AML testing, data lineage, explainability, drift	[16]-[19]

	agency of AI copilot	unusual data distribution	detection, human approval	
Cryptographic	Harvest-now-decrypt-later, weak algorithms, key compromise, poor migration inventory	Legacy algorithms, long-lived secrets, untracked high-value routes	PQC migration, QKD-assisted keys, crypto-agile inventory, key rotation	[4], [5], [22], [24]

3.3 Proof-of-Concept Architecture

To convert the conceptual AI-DCF into an implementable prototype, a proof-of-concept architecture can be organized around six minimum components: optical telemetry sources, secure telemetry collectors, a normalized telemetry lake, an ML analytics service, a zero-trust policy decision point, and an SDN/NMS response orchestrator. This prototype is intentionally small enough for a university or operator lab, while still preserving the closed-loop cyber-physical structure required for production optical networks. The design supports shadow-mode detection first, then policy-validated response once the model and playbooks are trusted.

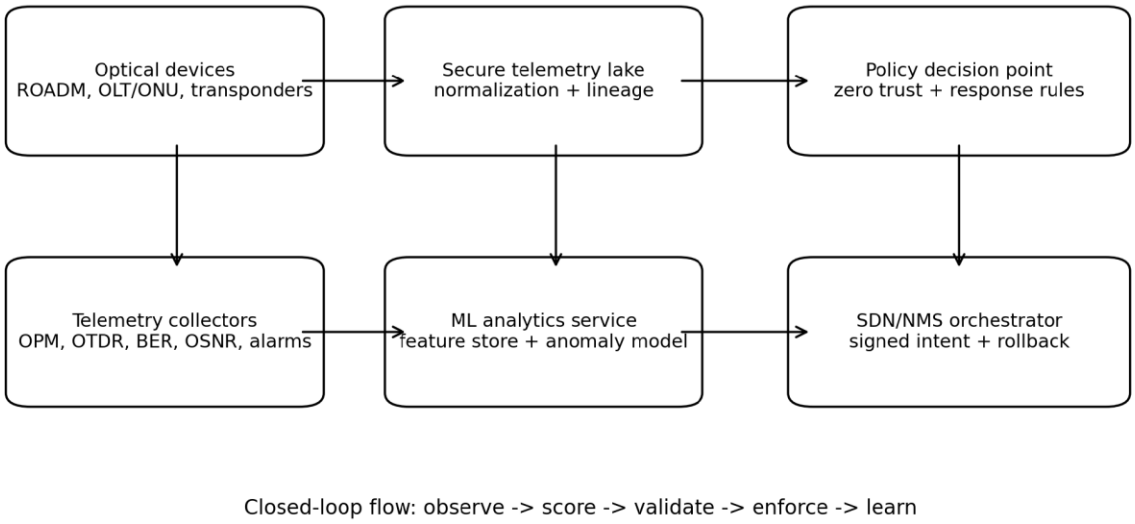

Figure 2: Proof-of-concept architecture for implementing AI-DCF in a lab or pilot environment.

Table 3: Minimum proof-of-concept components for AI-DCF implementation.

PoC component	Input	Processing role	Output
Telemetry collectors	OPM/OTDR traces, BER, OSNR, power, alarms, controller logs	Authenticate devices, timestamp readings, validate integrity and forward telemetry	Trusted streaming telemetry events
Telemetry lake and feature store	Raw and normalized telemetry	Clean missing values, align route/service identifiers, compute features and retain lineage	Training and inference-ready feature vectors

ML analytics service	Feature vectors and model registry	Run anomaly scoring, drift checks, explainability and route-risk estimation	Anomaly score, likely class and confidence
Policy decision point	Model output, service criticality, identity context	Apply zero-trust rules and response authority levels	Allow, require approval, or block response

3.4 ML Workflow Demo

A compact ML workflow demo for AI-DCF can be implemented with synthetic or lab-collected telemetry. The objective is not to replace field validation, but to demonstrate how optical-security telemetry becomes a repeatable detection pipeline. The demo begins with baseline normal telemetry, injects controlled anomalies, trains an unsupervised or semi-supervised detector, and routes the output through a policy layer before any operational action is taken.

Data ingestion: collect optical power, OSNR, BER/Q-factor, alarm rate, route identifier, wavelength/channel, maintenance flag, controller API events and route criticality score.

Feature engineering: compute short-window deltas, slopes, rolling means, alarm bursts, controller-change count, and service-impact indicators.

Model training: train an unsupervised detector, such as an isolation forest or autoencoder, only on clean baseline data; reserve attack and fault scenarios for evaluation.

Inference and explanation: output anomaly score, top contributing features, affected route and recommended response level.

Policy validation: allow only low-impact automation by default, such as increasing sampling rate, opening an incident ticket or freezing suspicious noncritical changes.

MLOps control: log model version, training window, feature schema, drift status, false-positive review and rollback condition.

Pseudo-code demo: collect telemetry -> clean/align timestamps -> compute features -> train baseline anomaly detector -> score live telemetry -> explain anomaly -> validate with policy -> open ticket or trigger controlled response.

4. Results

The results are presented as design outputs and literature-based assessment artifacts. Since this is a conceptual framework paper, results do not claim a new empirical benchmark. Instead, they show how the proposed framework translates recent evidence into deployable security functions and measurable capability targets.

4.1 Framework Output: Security Functions and Controls

Table 4: Security functions produced by the proposed AI-DCF.

AI-DCF layer	Core controls	Operational outputs	Expected benefit
Governance and risk	Asset inventory, route criticality, risk appetite, compliance mapping, cyber-physical incident ownership	Risk register, policy catalogue, service criticality map	Improves accountability and auditability [1], [16]

Identity and zero trust	Device identity, controller identity, operator MFA, service-to-service authentication, continuous authorization	Identity graph, policy decision logs, access denials	Reduces implicit trust in location or device ownership [2]
Telemetry lake	OPM, OTDR, alarms, logs, topology, flow data, controller events, cryptographic inventory	Normalized feature store, lineage metadata, retention policy	Enables cross-layer detection [6], [8]
AI detection	Autoencoders, RNN/GRU/LSTM, graph analytics, hypothesis testing, federated learning where needed	Anomaly score, confidence interval, affected path, root-cause probability	Supports rapid anomaly detection and localization [7], [9], [10]
Model assurance	Data quality checks, adversarial testing, drift detection, explainability, model rollback	Model card, drift alert, AML test report	Protects the AI lifecycle [17], [19]
Control enforcement	Intent verification, signed policy, controller segmentation, API security, safe automation	Validated change plan, rollback token, human approval flag	Prevents unsafe autonomous response [3], [14]
Physical-layer defense	Route hardening, fiber access controls, OTDR baselining, OPM thresholds, sensor fusion	Physical risk score, tap suspicion alert, location estimate	Addresses fiber and component risks [12], [25]
Crypto-agility	PQC inventory, Layer-1 encryption, QKD-assisted key service, key rotation, algorithm lifecycle tracking	Algorithm bill of materials, key status, protected route list	Improves quantum-era resilience [4], [5]

4.2 Literature-Derived Performance Evidence

The strongest quantitative evidence in the reviewed literature concerns ML-assisted detection of optical anomalies. Abdelli et al. report that an autoencoder detected fiber faults or anomalies with an F1 score of 96.86%, while an attention-based bidirectional GRU identified anomalies with 98.2% average accuracy and localized faults with 0.19 m RMSE [10]. Figure 2 visualizes the two reported classification-related metrics. These results do not prove that every operator will achieve the same performance; rather, they demonstrate that optical-layer telemetry can contain enough information for high-performing data-driven detection under controlled conditions.

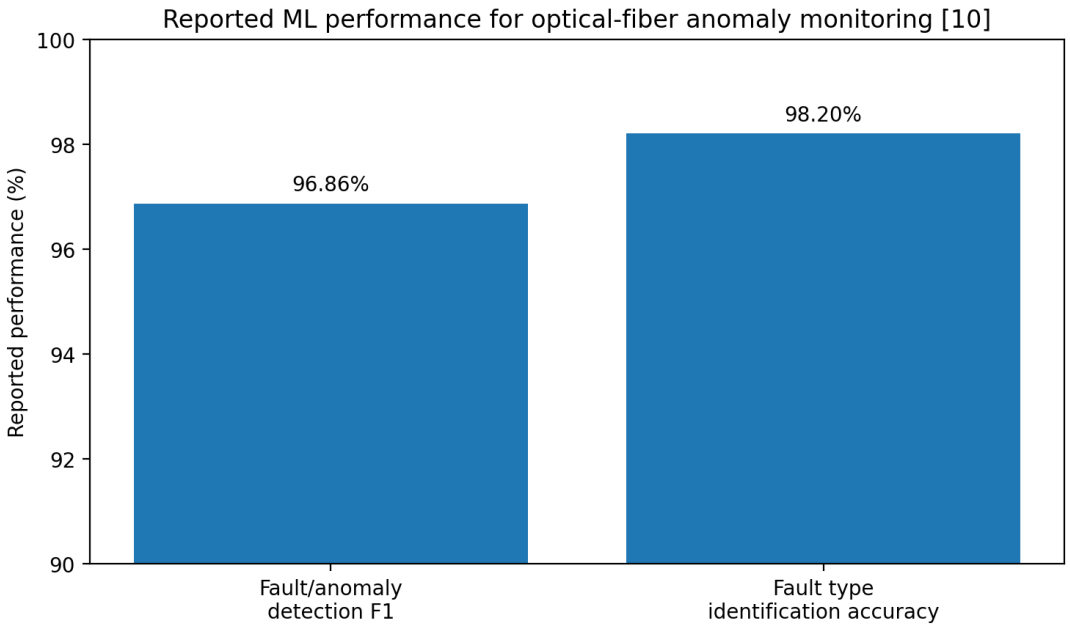

Figure 3: Reported ML performance for fiber anomaly monitoring from Abdelli et al. [10].

Behera et al. provide complementary evidence by using encoder-decoder prediction of QoT evolution and hypothesis testing for real-time anomaly detection [9]. This is useful for practical optical networks because labeled attack data are limited, expensive, and sometimes unavailable. Furdek et al. also emphasize supervised, semi-supervised, and unsupervised approaches and discuss practical integration with network management systems [7]. Therefore, the AI-DCF analytics layer deliberately supports multiple learning paradigms rather than assuming that one universal detector can cover all attacks.

Table 5: Evidence mapping from recent optical security and AI studies.

Study	Security focus	Method / contribution	Relevance to AI-DCF
Furdek et al. [7]	Optical-layer attack detection and localization	Supervised, semi-supervised, unsupervised learning; WAD; NMS integration	Shows practical framework for cognitive optical security diagnostics
Abdelli et al. [10]	Fiber cuts and optical eavesdropping/tapping anomaly detection	Autoencoder + attention-based BiGRU	Reports 96.86% F1 detection and 98.2% identification accuracy
Behera et al. [9]	Real-time anomaly detection using QoT evolution	Encoder-decoder prediction + statistical hypothesis testing	Reduces dependence on labeled anomaly datasets
Wang and Zhang [8]	AI in optical communications	CNN, RNN, GAN, DRL, end-to-end learning	Provides algorithmic basis for optical security analytics
Spurny et al. [12]	Physical-layer component security risks	Measured risk analysis of splitter, macrobending, crosstalk, and leakage	Shows why physical-layer telemetry is security-relevant

4.3 Capability Maturity Assessment

Figure 3 presents a qualitative capability mapping. The baseline represents a conventional posture where optical transport security relies mainly on perimeter controls, basic monitoring, and conventional encryption. The AI-DCF target posture represents the proposed framework with AI analytics, zero trust, optical-physical telemetry, model assurance, and crypto-agility. The scores are a planning tool on a 1-5 maturity scale and are not empirical measurements.

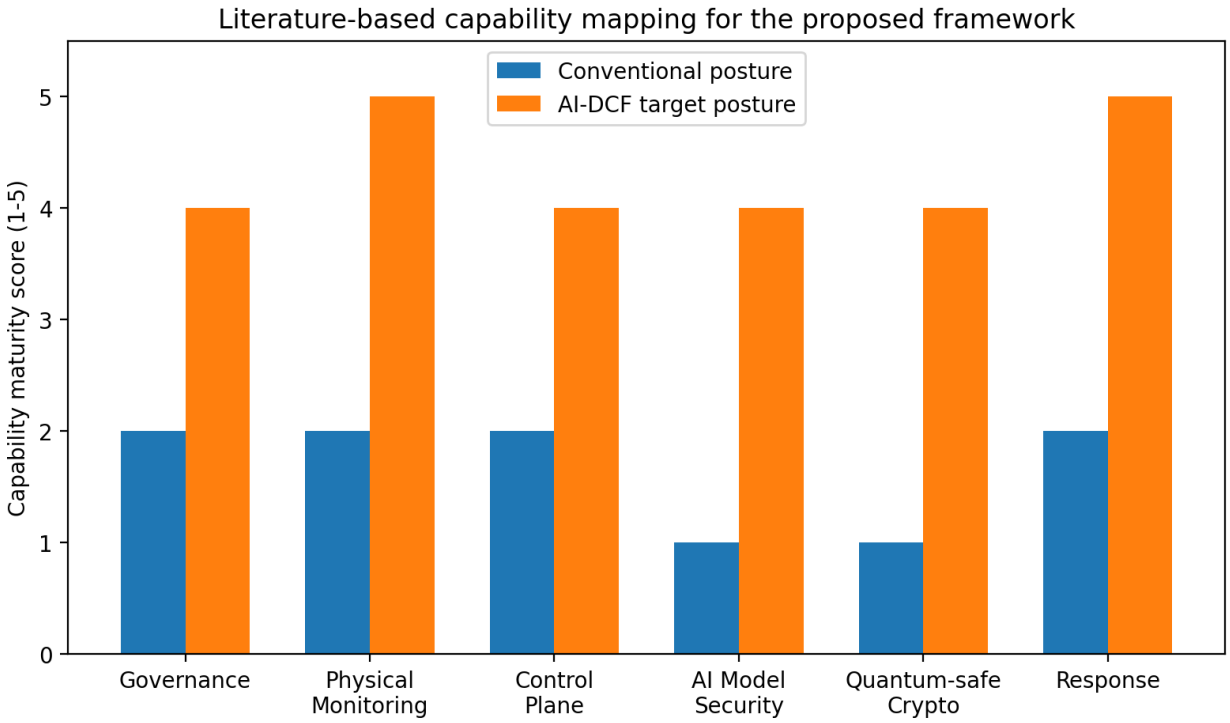

Figure 4: Qualitative capability mapping for conventional posture and AI-DCF target posture.

Table 6: AI-DCF maturity model.

Maturity level	Characteristics	Recommended transition action
1 - Initial	Ad hoc security monitoring; limited asset inventory; manual response	Inventory optical assets and management interfaces; define critical routes
2 - Managed	Basic alarms, access control, encryption on selected services	Normalize telemetry; implement MFA and centralized logging
3 - Defined	Documented processes; anomaly detection pilots; SDN/API controls	Integrate optical telemetry with SIEM/SOAR; deploy test ML models
4 - Quantitatively managed	Model performance monitoring; risk-based automated response; crypto inventory	Add drift monitoring, signed policies, and controlled automation
5 - Optimizing	Closed-loop cyber-physical defense; adversarial testing; crypto-agile and QKD-ready operations	Run continuous red teaming, PQC transition, QKD-assisted service for high-value routes

4.4 Deployment Blueprint

A practical deployment can be organized into phases. The first phase builds trusted visibility. The second deploys AI analytics in shadow mode. The third uses controlled enforcement through the SDN/NMS layer. The fourth hardens crypto-agility and model assurance. The fifth optimizes response through continuous validation. Table 6 gives a roadmap suitable for a carrier, cloud provider, research network, or critical infrastructure operator.

Table 7: Deployment roadmap for AI-DCF.

Deployment phase	Key actions	Success metrics	Main risks
Phase 1: Discover and baseline	Asset inventory, topology discovery, telemetry collection, identity mapping	Complete inventory, baseline optical telemetry, define service criticality	Incomplete legacy data; vendor interface diversity
Phase 2: Detect in shadow mode	Train anomaly models, compare against expert labels, add explainability	Precision/recall targets, false positive review, model-risk register	Rare attack data; model drift; domain transfer
Phase 3: Enforce safely	Connect analytics to policy decision point; signed intents; human approval for high-impact action	Reduced mean time to detect; validated rollback	Automation fear; change-management resistance
Phase 4: Harden AI and crypto	AML testing, model rollback, PQC inventory, QKD pilot for high-value routes	Model assurance evidence, algorithm bill of materials	Cost, key-management complexity, interoperability
Phase 5: Optimize and red-team	Continuous attack simulation, chaos testing, federated learning, multi-domain sharing	Resilience score, recovery time, validated playbooks	Data sharing, legal constraints, operator workload

4.5 Small Simulation: Synthetic Optical Attack Detection

To make the framework more demonstrable, a small proof-of-concept simulation was added using synthetic optical telemetry. The simulated data should be interpreted only as a workflow demonstration, not as a real operator benchmark. A baseline set of normal observations was generated with low optical-power variation, stable OSNR, low BER change and few controller events. Three anomaly classes were then injected: fiber tapping or splitter insertion, optical jamming and controller/API abuse. An unsupervised isolation-forest detector was trained only on normal telemetry and tested on mixed normal and attack observations.

Table 8: Synthetic simulation setup for a small AI-DCF workflow demonstration.

Simulation item	Value used in demo	Reason	AI-DCF mapping
Normal baseline	900 observations; 700 for training and 200 for testing	Represents steady-state route behavior	Telemetry lake and feature store
Attack scenarios	100 fiber-tap, 100 jamming, 100 controller/API-abuse observations	Shows physical and control-plane attack coverage	Optical-physical and optical-control layers
Features	Power drop, OSNR drop, BER change, API-change	Small feature schema suitable for a lab demo	AI analytics layer

	count, alarm rate, route-risk score		
Detector	Isolation forest trained on normal-only data	Useful where labelled optical attacks are rare	Unsupervised anomaly detection

In this deterministic demonstration, the detector achieved precision = 0.84, recall = 1.00, and F1 score = 0.91 on the synthetic mixed test set. The result is included to show the end-to-end workflow from telemetry to detection output; it does not replace the literature-derived evidence or a physical optical testbed.

Table 9: Synthetic detection output from the small AI-DCF simulation.

Metric / class	Value	Interpretation	Operational action	Caution
Overall precision	0.84	Share of generated alerts that were true anomalies	Review alert quality and tune thresholds	Synthetic distribution may be cleaner than field telemetry
Overall recall	1.00	Share of attacks correctly detected	Estimate missed-attack risk	Real stealth attacks may reduce recall
Overall F1 score	0.91	Balance between precision and recall	Compare models during pilot testing	Not a production benchmark
Confusion matrix	TN=144, FP=56, FN=0, TP=300	Compact summary of detector behavior	Use to support threshold decisions	Depends on simulated attack mix

4.6 Attack Detection Example

The following illustrative example shows how the proposed system would handle a suspected low-loss fiber-tapping attack. In the scenario, the optical layer detects a small but persistent drop in received power on one route, while BER and alarm rate increase mildly. No authorized maintenance ticket is active, and physical access logs show an unexpected cabinet-door event near the route segment. The ML analytics layer classifies the event as suspicious fiber access rather than ordinary congestion because the anomaly is localized, physically plausible, and correlated with route-risk and access-control context.

Table 10: Example AI-DCF detection and response flow for suspected fiber tapping.

Step	Telemetry / evidence	Model interpretation	Policy decision	Response level
1. Observe	Power drop of about 0.7 dB; minor BER slope; route-risk score high	Possible splitter insertion or macrobending	Increase sampling and preserve evidence	Level 1
2. Correlate	No planned maintenance; cabinet access event near route; controller config unchanged	Physical-layer event more likely than control-plane change	Open SOC/NOC incident and notify route owner	Level 1
3. Classify	Anomaly score exceeds threshold; top features are power drop, BER	Suspected fiber tapping or unauthorized access	Require human approval for service-affecting actions	Level 2

	slope and physical access			
4. Contain	Critical service on affected wavelength; alternate path available	Rerouting is possible but may affect SLA	Validate signed intent and rollback plan	Level 2/3
5. Recover and learn	Post-incident OTDR comparison and field inspection results	Confirmed attack or benign physical fault updates labels	Update model data, playbook and risk score	Closed-loop learning

This example highlights the most important design principle of AI-DCF: the model does not independently disable routes or reconfigure the optical network. It produces evidence, confidence and recommended response level. The policy decision point then checks service criticality, change authority, signed intent, rollback availability and human-approval requirements before enforcement. This separation makes the demonstration suitable for safety-sensitive optical transport environments.

5. Discussion

The proposed framework repositions optical cyber security as a holistic cyber-physical control challenge. Traditional security architectures commonly divide responsibility among optical transport engineers, IP security teams, NOC operators, SOC analysts, cloud teams, and compliance officers. This separation can lead to blind spots. For example, a SOC may notice traffic degradation (but not the optical telemetry); a NOC might see a power anomaly (but is not able to see threat intelligence); and a transport engineer might think this change in configuration is operational but a security analyst may see it as a possible compromise. AI-DCF solves this by applying telemetry, identity, policy, and response to work in coordination across layers. The trade-off between automation and operating safety is a crucial area of discussion. AI can identify anomalies faster than manual review but optical transport networks carry services that cannot be interrupted by aggressive automated responses. Hence, the framework differentiates between detection automation and enforcement automation. Detection can be highly automated, while enforcement should be risk-aware. Low-impact actions, including increasing monitoring frequency or creating an incident ticket, might be automated. If the use of high-impact actions (like rerouting critical services or disabling controller apps) need to get policy validation and potentially a human intervention. In addition, another critical question is on data availability.

Supervised learning works effectively with labeled data but labeled optical attack data are rare. This makes it applicable for unsupervised, semi-supervised learning, prediction-based anomaly detection, and synthetic data augmentation. But synthetic data can add bias if they do not model real optical attack scenarios. The framework therefore advises employing a hybrid model: physics-based thresholds for known failure modes, unsupervised anomaly detection for non-established trends, supervised classifiers where labeled as known, or expert-guided evaluation for ambiguous high-impact events. A cross-layer correlation is required. Optical-layer attacks may appear in the form of microphysical changes that lead to degrading QoT, localized BER changes, alarm bursts, changes in control-plane routes, or changes in packet levels at the higher-layer. Another telemetry source may not give you enough evidence. Hence, AI-DCF focuses on feature fusion among optical performance monitoring, OTDR traces, controller logs, NMS/EMS events, identity logs, cryptographic inventory and service impact. Graph-based analysis is particularly attractive due to the fact that optical networks inherently possess graph structure: nodes, links, wavelengths, spectrum slots, routes, services, controllers and trust domains.

The framework also highlights the growing need on telecom operations for AI model security. An optical security model with a false negative can tap in secret, while a false positive could disrupt service without reasonable justification. Model assurance must include some training-data lineage, drift detection, adversarial testing, confidence thresholds, explainability, rollback and continuous evaluation. The NIST adversarial ML taxonomy is a helpful tool

in that it provides operators the language necessary for discussing poisoning, evasion, privacy and abuse [17]. The OWASP LLM Top 10 also comes in handy when generative AI copilots are integrated into operational workflows [18]. Quantum-secure migration also calls for pre-emptive planning among operators of optical networks. PQC migration isn't just a cryptographic engineering problem but also an inventory and orchestration problem. Operators need to figure out which optical service has long-lived confidential data, which devices can upgrade algorithms, which management channels use vulnerable public-key systems, and which routes are able to accommodate QKD-assisted access to key services. So, a crypto-agile optical network must be able to change algorithms and key management approaches without completely re-engineering the whole transport layer. There are also cost and complexity issues. AI-DCF is typically associated with telemetry normalization, storage, compute, model operation, on-staff training and integration with legacy NMS systems. Smaller providers will not deploy the entire framework for the time being.

The maturity model of Table 5 deals with this by enabling staged deployment. Even a simple implementation can increase security by establishing asset inventory, implementing zero-trust management access, having cryptographic inventory, and tying optical alarms back into the security events. More advanced deployments can also include live AI, adversarial testing, and QKD support applications for key routes. Standards play a big role. NIST CSF 2.0 favors general cyber-risk governance [1]; NIST SP 800-207 promotes zero-trust design [2]; ETSI TS 103 961 covers optical management security [3]; ETSI QKD work endorses quantum-safe key distribution [4]; and NIST PQC standards offer guidance during cryptographic transition [5]. These guidelines should not be supplanted by AI-DCF. It operationalizes them for the next generation of optical communication networks by mapping standards into controls, telemetry, AI models, and response processes. AI-DCF provides 3 improvements over traditional optical security. One, it treats optical telemetry not only as performance data but as security evidence. Second, it adopts AI as a part of a governed closed-loop system as opposed to a black-box detector. Third, it expands security beyond threats we face now to the adversarial ML and quantum-era cryptographic risk. This combination is vital, since the next-generation optical networks will enable critical AI workloads, multi-domain services, and high-value data flows for which downtime, eavesdropping, & delayed cryptographic migration can have severe impacts. The framework also has some drawbacks.

The capability scores in Figure 3 are qualitative, and should not be interpreted as observed mitigation of risk. Although the literature gives promising ML results, many of them are based on specific testbeds, datasets, and types of attack. Real networks range among vendor equipment, telemetry resolution, route design, maintenance practice, service mix, and staff maturity. In the coming research, AI-DCF needs to conduct multi-vendor optical testbeds, detect it under both stealthy and adaptive attack, understand false-positive cost measurement, and test model robustness to adversarial manipulation. Finally, a last gap is the adversarial analysis of AI-powered optical defense. A model capable of working with traditional attack datasets can be effectively attacked but cannot detect an adaptive attacker with the same ability to understand the model. Future research should investigate mechanisms for evading, poisoning, stealing models from others and falsification of telemetry systems. Evaluations should cover the cyber and physical attack channels because optical networks are cyber-physical systems. Not just accuracy, but resilience and robustness under intelligent opposition will be crucial for the future of AI-driven optical security. A fourth gap is the integrated orchestration of PQC and QKD. PQC migration and QKD deployment are investigated separately, but actually real optical networks may employ both approaches. AI-DCF believes that cryptographic modes should be selected for future systems based on route risk, data sensitivity, device capability, algorithm status, & key availability.

AI can help by keeping cryptographic inventory, suggesting migration prioritizations, but the final cryptographic policy must also continue to be governed and auditable. Therefore, research is needed to establish interoperable control models, and perform cost-benefit analysis and failure handling for mixed PQC-QKD transport. A third gap is explainable AI in optical security. Many effective models show up as hard to interpret, while network operators need clear justification before they alter transport paths that matter. Exploratory approaches suitable for optical engineers can be pursued in the research ahead in terms of explainability: wavelength power, OSNR trend, BER slope, route segment, controller action or alarm sequence. Explanations should be used to support decisions, not just meet a machine-learning transparency prerequisite. One gap is integration between optical digital twins and

security analytics. Digital twins can simulate optical behavior that is expected for the normal, failure and, under the assumption of an attack, attack actions that you would anticipate. A digital twin when integrated with AI-DCF might check whether an observed anomaly is physically possible, to analyze the location of an attack, to project the impact on service performance, and to assess response before enforcement. This would minimize risk from automated mitigation.

Digital twins are essential with accurate models, current topology, and calibration against live telemetry. Currently, one area for research is that open optical-security datasets are not frequently present. Most operators can't provide real telemetry, as it exposes topology and customer details. This makes researchers almost always use small testbeds or synthetic data. Thus, reproducibility and model comparison is limited. One potential solution is anonymized benchmark datasets with topology abstraction, controlled attack labels, and privacy-preserving feature transformation. One additional approach is federated benchmarking, in which models are tested across operator environments without moving data to the next. More broadly speaking, AI-DCF points to an indication of increasing integration of optical-network security into a strategic approach to digital sovereignty and economic resilience. Because AI workloads, cloud services and critical infrastructure rely more on optical transportation, attacks that target the optical layer have ramifications that go beyond telecom operations. A compromised route can impact financial transactions, public services, research data, AI model training, or emergency communications. Accordingly, optical security should be integrated into national cyber exercises, critical infrastructure standards, and enterprise risk management programs.

5.1 Broader Implications and Research Gaps

The importance of multi-vendor evaluation is particularly relevant. Optical networks are heterogeneous, and the semantics of telemetry can vary between devices. A model trained on one vendor or topology may not be generalizable to another. Transfer learning, domain adaptation, and federated learning may assist but need to be validated carefully. AI-DCF thus suggests that future testbeds include multiple vendors, different link lengths, varied modulation formats, SDN controller diversity, and realistic operational noise. This validation would make the framework more real and believable for application outside laboratory conditions. A useful benchmarking design would consist of benign failures, maintenance events, and malicious events. This is based on two sources of failure to be considered: benign events such as fiber bends, amplifier drift, transceiver degradation, and planned rerouting. Malicious events may include low-loss tap insertion, optical jamming, wavelength injection, controller abuse, and telemetry manipulation. The benchmark should not only indicate if the system detects the events, but if it also categorizes the event classifier correctly and advises proper response. This distinction is important since a framework that identifies anomalies but does not discriminate attack, failure, and maintenance will create excessive operational burden.

Cost metrics must also be assessed. AI-DCF may involve new telemetry collectors, storage, compute, staff training, and vendor integration. Cost per protected route, incremental telemetry load, inference latency, storage growth, as well as operational efficiencies from early detection are the parameters that a strong empirical study should measure. It should contrast centralized analytics to edge or in-network analytics. For high-speed packet-optical nodes, some security functions might have to run near the data plane while others might be able to run in a centralized analytics platform. This trade-off impacts scalability and response latency. We recommend a more empirical perspective for assessing AI-DCF in the future using a multi-dimensional metric set. Detection metrics are precision, recall, F1 score, receiver operating characteristic area, detection delay, and localization accuracy. Operational metrics are mean time to detect, mean time to respond, mean time to recover, service downtime, number of affected customers, and operator workload. Security metrics: attack coverage, adversarial robustness, access-control violations blocked, unauthorized configuration attempts detected, cryptographic migration progress. The model-governance metrics are drift frequency, retraining intervals, explainability coverage, and number of unsafe recommendations blocked by policy.

5.2 Evaluation Metrics for Future Empirical Validation

Security automation should first be tested in exercises before production. Tabletop simulations, digital twins, lab testbeds, and controlled field trials can show whether playbooks are safe. You should include ambiguous incidents, maintenance overlap, false-positive bursts, and multiple concurrent failures. These scenarios are realistic because operational events of optical networks often match attacks. Hence, the framework advises gradual activation: monitoring only, recommending only, automating low-impact responses, and finally automating selected high-confidence responses. There is still a big element of human supervision. AI helps support operators by minimizing noise and identifying correlations; it does not eliminate accountability. Operators require explanations that link the output from each of these models to optical facts: affecting the route, changing a metric, potentially suggesting a threat, confidence, supporting telemetry, and taking the appropriate action. This is critically important in rules-driven settings where all decisions must be auditable. AI-DCF recommends role-specific dashboards: NOC engineers need optical details, SOC analysts need threat context, executives need risk and service impact, auditors need evidence trails. Recovery needs verified rollback. As with any SDN-enabled optical network, this can change states fast, but this is only helpful if changes are reversed safely.

The framework recommends configuration snapshots, signed known-good policies, route-state checkpoints, controller rollback tokens. When things are resolved with an incident over, post-incident review should revise detection rules, model data, playbooks, and risk scores. The resulting closed-loop learning cycle, as shown in Figure 1, is the product of this. Without learning from incidents, AI-DCF would become static monitoring rather than adaptive cyber defense. Reliability expectations of optical networks have relevance to response design. An erroneous mitigation might be worse than the original anomaly if it disrupts vital traffic. AI-DCF consequently defines the levels of the response. Level 1 responses are observational, such as increasing sampling rate, opening a ticket, requesting further telemetry etc. Level 2 responses contain relatively minor service impact, like isolating a suspicious management session or freezing noncritical configuration changes. Level 3 responses are service-affecting actions like rerouting traffic, disabling an optical channel, or rolling back a controller application. Level 3 actions ought to require policy validation plus often human authorization for these actions as well.

5.3 Response, Recovery, and Human Oversight

A model operations should involve version control, reproducible training, rollback, and model cards. All deployed models should state their training data period, intended use, known limitations, evaluation results, feature dependencies, and approved response actions. Drift monitoring should compare live feature distributions with training distributions. If drift exceeds a threshold, the model should move to advisory mode or require review before enforcement. This matters because optical networks evolve: you can have new transceivers, firmware upgrades, route changes, seasonal temperature effects, traffic shifts, all altering normal behavior. Adversarial testing is key. Attackers can analyze the detector's behavior and then develop optical or telemetry patterns used to bypass detection. They even can poison training data by introducing repeated low-level anomalies that become accepted as normal.

The NIST adversarial ML taxonomy also specifies a vocabulary for these threats such as evasion, poisoning, privacy, Trojan, and abuse threats [17]. AI-DCF suggests red-team exercises in which controlled optical perturbations, telemetry manipulation, controller misuse, and model-extraction attempts are simulated. The output should be a model-risk report and a set of hardening actions. The model validation has to include both statistical and operational criteria. So, statistical indicators (precision, recall, F1 score, false-positive rate, false-negative rate, detection latency, and localization error) are important but inadequate. Operational metrics include: service impact, alert fatigue, mitigation safety, explainability to operators, and compatibility with change-management procedures. As an example, a model that is very sensitive will find more attacks but overload the NOC with false alarms.

Conversely, a conservative model might safeguard availability but miss more subtle tapping. AI-DCF advises using risk-based thresholds, where the crucial services need more sensitivity in combination with increased human review. The lifecycle of the AI model is initiated by problem definition. It is a temptation to train a general anomaly detector and then deploy it widely in optical cyber security. Yet every model should have a defined operational purpose — fiber-tap suspicion, jamming detection, soft-failure separation, controller-abuse detection, route-risk scoring, or

cryptographic-inventory prioritization. Purpose definition supports feature selection, label selection, evaluation metrics, and response thresholds. It also helps operators determine when model output is advisory and when it is allowed to trigger automated action. Such approach prevents overreliance, one of the AI risks identified in AI security guidance [18].

5.4 AI Model Lifecycle Controls

Data retention requires balance. Retention for long time helps for model training, trends analysis, and forensic evaluation which can make some privacy, confidentiality, and storage issues arising. AI-DCF suggests tiered retention, high resolution raw telemetry for incident analysis of recent events and compressed feature histories for long term trend learning and anonymized or aggregated datasets for research and model building. Where data is shared among domain or organization users privacy-preserving techniques and contractual principles must be employed. So federated learning could be a place where operators require an improvement to their detection models without revealing sensitive topology or customer data. The framework also needs to facilitate secure telemetry transport. Should these attackers manipulate telemetry, then they can blind the AI detector. Telemetry collectors should validate the devices and protect the integrity of the data, use only encrypted transport, and log collection gaps. High value telemetry should be considered sensitive because it details the topology, route design, protection schemes, vendor information, and the traffic behavior of the device. Telemetry access must be monitored and follow least privilege. This is in line with zero trust, where trust does not originate when a collector or an analyst sits on the network perimeter [2]. Quality of telemetry influences quality of the model.

Missing values, inconsistent vendor semantics, mismatch between sampling and rate rates, alarm storms and configuration noise all contribute to a lowered detection accuracy. For instance a power drop from planned maintenance should not be considered malicious, whereas a low-loss splitter insertion may have sensitive detection thresholds. AI-DCF recommends maintaining data lineage, maintenance annotations, service-impact labels, known benign-change records. These records help to prevent false positives and contribute to ongoing learning efforts. They can also serve to discern cyber security anomalies from what would be expected with optical engineering like fiber repair, programmed route change, amplifier balancing or transceiver replacement. Most importantly, data engineering is one of the key practical aspects of the application.

Optical security analytics rely on telemetry that is synchronized, normalized, and reliable in nature. Data You'll need information like optical power, OSNR, BER, Q-factor, spectrum occupancy, amplifier gain, ROADM switching states, OTDR traces, alarms, configuration changes, controller API calls, device identity events, maintenance tickets, physical access logs and cryptographic status. The AI system may not correlate an optical anomaly with a suspicious controller action or physical cabinet access event without unified time stamps and consistent identifiers. Thus, AI-DCF regards the telemetry lake as a security control and not merely as an operations database.

5.5 Data Engineering and Telemetry Requirements

The fourth use case is multi-tenant optical infrastructure. Cloud providers, wholesale carriers, and neutral-host networks can deliver traffic for multiple customers over shared optical platforms. It increases the risk of accidental exposure, deliberate side-channel exploitation, and complex responsibility boundaries. AI-DCF suggests tenant-aware segmentation, per-service encryption policies, route isolation for regulated workloads, and explainable security events that can be shared with customers without exposing other tenants. AI analytics is beneficial in determining whether an anomaly is localized to a single wavelength, tenant service, ROADM node, or broader physical route. The third use case is smart-city and critical infrastructure transport. Intelligent transport systems, healthcare networks, surveillance platforms, government services, and emergency response systems all rely on fiber infrastructure.

If multiple public services share a fiber route, a physical attack can generate cascading effects. AI-DCF encourages asset classification coupled with route criticality scoring to encourage greater monitoring and faster response for high-value routes. The framework also outlines service-aware incident playbooks: a suspected fiber tap on a financial data route, a jamming incident on a public-safety route, or a controller anomaly affecting a smart-grid

connection should not be treated the same way. The second use case is AI data-center interconnection. AI training and inference must also be done under very large east-west and inter-data-center traffic flows. Optical interconnects and co-packaged optics are strategic because electrical interconnects are limited in distance, power, and density. According to industry reports, AI workloads are reshaping optical transport planning, and co-packaged optics are being developed to improve both power efficiency and scalability [20], [21].

In such an environment, a compromised optical route might affect the availability of model training, data confidentiality, and synchronization between distributed AI clusters. AI-DCF allows this use case to work by monitoring high-bandwidth links, verifying route integrity, protecting management interfaces, and maintaining cryptographic readiness for data that may stay sensitive for extended retention periods. The first use case of the system is secure 6G transport and mobile fronthaul/backhaul. In future, advanced radio must have high-capacity and low-latency optical transport between radio units, distributed units, centralized units, edge clouds, and core functions. As the pathways could have applications such as emergency services, autonomous systems, industrial control, and public safety applications, the optical layer must be monitored in the security chain. AI-DCF can correlate radio-slice identifiers, transport paths, wavelength assignments, QoT trends, and controller actions. If an optical-power variation can be detected along a route serving a critical slice, the framework can raise a security-priority incident instead of assuming the event is a regular degradation. This connects optical telemetry to service criticality, where security response becomes more business-aware [1], [25].

5.6 Operational Use Cases for AI-DCF

The fourth research recommendation is to incorporate crypto-agility in optical-network simulators and testbeds. Most of the optical-network work focuses on routing, modulation, spectrum, and failure recovery, but cryptographic migration is treated separately. In applications, route choice, key availability, service sensitivity, and algorithm status will impact each other. A future optical controller must choose not only the shortest feasible path but also the one with the appropriate encryption mode, key freshness, PQC support, or QKD-assisted key service. A rich area of research is opened from here at the intersection of optical networking, cyber security, and cryptographic engineering. The third suggestion for research is to investigate adversarial robustness.

AI-based optical security will inevitably encounter adaptive attackers. A realistic evaluation may also contain attackers who have an understanding of the detection window themselves, would make incremental changes, manipulate telemetry, exploit maintenance periods, or could have access to the model pipeline itself. Testing of this kind should be ethical and controlled but it is necessary. In the absence of adversarial testing, these reported high performances might give false confidence. The most useful future systems will be the ones that continue to be relevant if the attacker attempts to evade, poison, or confuse the security analytics layer. The second research recommendation is to integrate physics-informed models with data-driven models. It can detect unusual patterns with pure statistical AI but optical engineering knowledge limits the search space and makes them more interpretable. Physical relationships characterize attenuation, OSNR, BER, dispersion, nonlinear impairments, and amplifier behavior. Using these relationships, a hybrid model can differentiate physically plausible attacks from measurement noise or data-quality issues.

Similarly, this method can also reduce training-data requirements as the model does not have to be trained to learn all of the optical constraints from example data. According to researchers, the first recommendation is to perform experiments that account for operational cost. A lot of papers report classification accuracy, but operators still need to know false-positive cost, response safety, operator workload, and recovery time. A model leading to a one percent increase in F1 score may not be as valuable as a model that reduces explanation time by fifty percent. Future work should show detection performance together with operational metrics. They also need to report characteristics of the dataset, available features, labeling process, attack realism, and how much the tool generalizes across topology or equipment changes.

The fourth recommendation is to test the framework in conditions of maintenance, not only attack. Optical networks suffer fiber repair, planned rerouting, amplifier replacement, transceiver firmware changes, link balancing,

and emergency restoration. Events like this can be similar to attacks on telemetry. If the AI system doesn't learn from clean normal data and clear malicious data, it could create too many alerts during the maintenance window with alerts for the fault that are generated during maintenance windows. Hence, AI-DCF would suggest that maintenance calendars, work orders, and change tickets be added to the analytics pipeline with automation. This leads the model to know the difference between authorized operational change versus unscheduled change. The third recommendation is to develop a security data model on optical telemetry.

Operators should define common identifiers for links, services, wavelengths, spectrum slots, devices, cards, transceivers, ports, controllers, routes, and customer services. They should also coordinate time stamps with optical devices and security systems. Often, AI projects fail because they can be technically provided but semantically inconsistent data. Normalized data model allows detection, incident reconstruction, auditability, and better alignment between NOC and SOC teams. It also minimizes vendor lock-in because security analytics can work not based on proprietary alarm formats but on common features. The second recommendation is to be clear about analytics confidence versus response authority. While a model can yield a high anomaly score, the authority to change a route, disable a wavelength, or revoke a controller session should take place in the context of policy, service criticality, and operational considerations. This difference is important for trust. Thereby, operators are more likely to accept AI assistance when they are aware that actions with high impact are subject to deterministic policy checks.

A useful design pattern is for AI to generate evidence and recommended actions and for a policy engine to decide whether the action is allowed, has to be approved, or must be blocked. For network operators, the initial recommendation is to start with visibility rather than sophisticated AI. An organization does not protect what it does not detect or measure. Operators should build an optical cyber asset inventory, which will include fibers, ROADMs, amplifiers, transponders, controllers, NMS/EMS platforms, management APIs, telemetry collectors, key-management services, and privileged operator accounts. And mapping each asset to services and business criticality. This inventory should specify which links contain regulated data, long-lived confidential data, or safety-critical services. If this foundation is set up, AI models may be trained on meaningful baselines instead of fragmented operational points.

5.7 Practical Recommendations for Researchers and Network Operators

In educational and research settings, AI-DCF can act as a teaching model for cyber-physical network security. This is a bridge for optical engineering topics such as OSNR, BER, OTDR, ROADMs, and WDM to cyber security topics such as zero trust, anomaly detection, adversarial ML, and post-quantum migration. This interdisciplinary perspective is necessary because next-generation optical security will necessarily rely on teams that understand photonic behavior in conjunction with cyber risk. A model that connects these fields may enhance curriculum development, lab programs, and industry-academic cooperation. Policy makers and standardization organizations can aid this endeavor by promoting optical-security reporting, interoperable telemetry models, secure management interfaces, and quantum-safe transition planning.

As optical networks serve as a cornerstone for AI factories, public cloud, 6G, and critical infrastructure, security needs to be described so that the security requirement statements are measurable and testable. This should have future standards for minimum telemetry, identity, logging, model-assurance and crypto-agility requirements for optical devices and management systems. All of which means that operators should view AI-DCF as a governance and engineering program, instead of a software purchase. Purchasing a monitoring platform does not eliminate route criticality, model accountability, cryptographic migration, or response authority in and of itself. The best deployments will be people, process, telemetry, secure architecture, and proven automation combined. This is why the framework opens with governance and closes with resilience: It is this balance that makes AI valuable in operational environments where reliability and accountability are as integral as detection accuracy.

6. Conclusions

This paper proposes an AI-Driven Cyber Security Framework for next-gen optical communication networks. The framework fills a key gap in modern optical security: high-capacity optical infrastructures are increasingly

programmable, software-defined, AI-driven, and quantum-era sensitive, but most security initiatives are still fragmented in the physical, control, management, cryptographic, and AI-model domains. The AI-DCF consolidates these domains into a layered and closed-loop architecture. The literature review demonstrates that AI and ML are already critical for intelligent optical networks, optical performance monitoring, fault prediction, and anomaly detection [6], [8]. In terms of security-focused studies, ML-based optical attack detection and localization can be practically applied, with demonstrated high performance for both autoencoder and BiGRU-based fiber anomaly detection [10]. The revised paper further strengthens implementation readiness by adding a proof-of-concept architecture, a compact ML workflow demo, a synthetic simulation, and an attack-detection example that clarify how AI-DCF can be piloted before a full physical testbed deployment.

Physical-layer risk studies and recent optical-security surveys show that fiber tapping, splitter insertion, jamming, crosstalk, signal injection, and SDN manipulation still constitute serious threats [12], [25]. The introduced model comprises five holistic perspectives: governance and risk, AI security analytics, optical-control security, optical-physical security, and crypto-agile resilience. It employs NIST CSF 2.0 for risk governance, NIST zero trust for access decisions, ETSI optical network security provisions for management-plane controls, NIST AI/AML guidance for AI model assurance, and NIST/ETSI quantum-safe guidance for crypto-agile design [1]-[5], [16], [17]. The broad finding is that AI-driven optical cyber security should not be considered an individual intrusion detection model. It must be a continuously governed, continuously verified, and continuously improving system. Detection models should demand secure telemetry, trustworthy data, explainability, drift monitoring, adversarial testing, and safe routes of enforcement. Network responses require both policy validation and rollback. Migration planning and quantum-safe readiness are required by cryptography. Human operators must have clear responsibility and confidence in the automation. Some further work should leverage tools such as the SDN-enabled control, multi-source telemetry, attack emulation, and crypto-agile key management in an eventual deployment in a field-scale optical-network testbed and verify AI-DCF to be effectively implemented on an actual optical network. Evaluation should provide insight into detection performance, localization accuracy, false-positive cost, response latency, operator usability, adversarial ML resilience, and PQC/QKD interoperability. This work would transition the framework from conceptual synthesis to actual, quantifiable deployment guidance for secure 6G backhaul, AI data centers, cloud interconnects, smart-city infrastructure, and mission-critical optical transport.

REFERENCES

1. National Institute of Standards and Technology, “The NIST Cybersecurity Framework (CSF) 2.0,” NIST CSWP 29, Feb. 2024, doi: 10.6028/NIST.CSWP.29.
2. S. Rose, O. Borchert, S. Mitchell, and S. Connelly, “Zero Trust Architecture,” NIST Special Publication 800-207, Aug. 2020, doi: 10.6028/NIST.SP.800-207.
3. ETSI, “CYBER; Optical Network and Device Security; Security provisions for the management of Optical Network devices and services,” ETSI TS 103 961 V1.1.1, Dec. 2023.
4. ETSI, “Quantum Key Distribution (QKD); Common Criteria Protection Profile - Pair of Prepare and Measure Quantum Key Distribution Modules,” ETSI GS QKD 016 V2.1.1, Jan. 2024.
5. National Institute of Standards and Technology, “NIST Releases First 3 Finalized Post-Quantum Encryption Standards,” Aug. 2024.
6. R. Gu, Z. Yang, and Y. Ji, “Machine Learning for Intelligent Optical Networks: A Comprehensive Survey,” Journal of Network and Computer Applications, vol. 157, 2020, Art. no. 102576, doi: 10.1016/j.jnca.2020.102576.
7. M. Furdek, C. Natalino, F. Lipp, D. Hock, A. Di Giglio, and M. Schiano, “Machine Learning for Optical Network Security Monitoring: A Practical Perspective,” Journal of Lightwave Technology, vol. 38, no. 11, pp. 2860-2871, 2020, doi: 10.1109/JLT.2020.2987032.
8. D. Wang and M. Zhang, “Artificial Intelligence in Optical Communications: From Machine Learning to Deep Learning,” Frontiers in Communications and Networks, vol. 2, 2021, Art. no. 656786, doi: 10.3389/frcmn.2021.656786.
9. S. Behera, T. Panayiotou, and G. Ellinas, “Machine Learning for Real-Time Anomaly Detection in Optical Networks,” in Proc. 23rd International Conference on Transparent Optical Networks (ICTON), 2023, doi: 10.1109/ICTON59386.2023.10207370.
10. K. Abdelli, J. Y. Cho, F. Azendorf, H. Griesser, C. Tropschug, and S. Pachnicke, “Machine Learning-based Anomaly Detection in Optical Fiber Monitoring,” arXiv:2204.07059, 2022, doi: 10.48550/arXiv.2204.07059.

11. J. Lun, M. Fu, X. Liu, Y. Wu, L. Yi, W. Hu, and Q. Zhuge, “Soft Failure Identification for Long-haul Optical Communication Systems Based on One-dimensional Convolutional Neural Network,” Journal of Lightwave Technology, vol. 38, no. 11, pp. 2992-2999, 2020.
12. V. Spurny, P. Munster, A. Tomasov, T. Horvath, and E. Skaljo, “Physical Layer Components Security Risks in Optical Fiber Infrastructures,” Sensors, vol. 22, no. 2, Art. no. 588, 2022, doi: 10.3390/s22020588.
13. S. Rothe, N. Koukourakis, H. Radner, A. Lonnstrom, E. Jorswieck, and J. Czarske, “Physical Layer Security in Multimode Fiber Optical Networks,” Scientific Reports, vol. 10, Art. no. 2740, 2020, doi: 10.1038/s41598-020-59625-9.
14. N. Ahmed et al., “Network Threat Detection Using Machine/Deep Learning in SDN-Based Platforms: A Comprehensive Analysis of State-of-the-Art Solutions, Discussion, Challenges, and Future Research Direction,” Sensors, vol. 22, no. 20, Art. no. 7896, 2022, doi: 10.3390/s22207896.
15. N. N. Dao, M. Park, J. Kim, and S. Cho, “SDN-Based Cyber Defense: A Survey,” Future Generation Computer Systems, vol. 115, pp. 652-675, 2021.
16. National Institute of Standards and Technology, “Artificial Intelligence Risk Management Framework (AI RMF 1.0),” NIST AI 100-1, Jan. 2023, doi: 10.6028/NIST.AI.100-1.
17. A. Vassilev, A. Oprea, A. Fordyce, and H. Anderson, “Adversarial Machine Learning: A Taxonomy and Terminology of Attacks and Mitigations,” NIST AI 100-2e2023, Jan. 2024, doi: 10.6028/NIST.AI.100-2e2023.
18. OWASP Foundation, “OWASP Top 10 for Large Language Model Applications,” Version 2025, 2025.
19. ISO/IEC, “ISO/IEC 42001:2023 Information technology — Artificial intelligence — Management system,” International Organization for Standardization, 2023.
20. Ciena and Heavy Reading, “Optical Transport Networks for AI: 2025 Heavy Reading Survey Analysis,” 2025.
21. NVIDIA, “Scaling AI Factories with Co-Packaged Optics for Better Power Efficiency,” 2025.
22. NTT, “World’s first post-quantum secure transport system capable of switching cryptographic technologies,” News Release, Oct. 2024.
23. JPMorgan Chase, “JPMorgan Chase establishes quantum-secured crypto-agile network,” 2024.
24. Y. Zhang et al., “Research on Key Technologies of Quantum-Safe Metro-Optimized Optical Transport Networks,” Applied Sciences, vol. 15, no. 5, Art. no. 2809, 2025.
25. A. Gazani, A. Mantzavinos, P. Tsompanoglou, K. Kantelis, S. Petridou, P. Nicopolitidis, and G. Papadimitriou, “Optical Network Security: Threats, Techniques, and Future Directions,” Electronics, vol. 15, no. 4, Art. no. 878, 2026, doi: 10.3390/electronics15040878.
26. X. Liu, Y. Zhang, Q. Qiu, Y. Cheng, W. Hu, and Q. Zhuge, “Field trial of an LLM-powered AI agent for autonomous optical networks: A full-lifecycle demonstration,” Journal of Optical Communications and Networking, vol. 18, pp. A179–A190, 2026, doi: 10.1364/JOCN.575402.
27. Y. Zhang, Q. Qiu, X. Liu, X. Yu, D. Fu, X. Liu, Z. Wang, H. Lin, Y. Chen, L. Yi, W. Hu, and Q. Zhuge, “AI agent for autonomous optical networks: Architectures, technologies, and prospects [Invited Tutorial],” Journal of Optical Communications and Networking, vol. 18, pp. A159–A178, 2026, doi: 10.1364/JOCN.576017.
28. M. Ruiz, J. Comellas, and L. Velasco, “Network digital twinning solutions to increase the security of multi-domain optical transport networks,” Journal of Optical Communications and Networking, vol. 18, no. 2, pp. 150–164, 2026, doi: 10.1364/JOCN.581853.
29. A. Yeganehfallah, A. Sgambelluri, E. Paolini, K. S. Mayer, M. F. Silva, D. A. A. Mello, and L. Valcarenghi, “Confidentiality-preserving real-time localization of soft failures in optical networks based on PCA and MLaaS,” Journal of Optical Communications and Networking, vol. 17, pp. D137–D143, 2025.
30. F. Musumeci and M. Tornatore, “Failure management in optical networks with ML: A tutorial on applications, challenges, and pitfalls [Invited],” Journal of Optical Communications and Networking, vol. 17, no. 8, pp. C144–C155, 2025, doi: 10.1364/JOCN.551910.
31. O. González de Dios, P. Armingol Robles, L. Roelens, A. Muñiz-Da-Costa, I. de Miguel, R. J. Durán Barroso, and J. P. Fernández-Palacios, “Automation of multi-layer multi-domain transport networks and the role of AI [Invited],” Journal of Optical Communications and Networking, vol. 17, no. 2, pp. A124–A133, 2025.
32. C. Sun, X. Yang, N. Di Cicco, R. Ayassi, V. V. Garbhapu, P. A. Stavrou, M. Tornatore, G. Charlet, and Y. Pointurier, “Experimental demonstration of local AI-agents for lifecycle management and control automation of optical networks,” Journal of Optical Communications and Networking, vol. 17, no. 8, pp. C82–C92, 2025, doi: 10.1364/JOCN.550286.
33. A. Abishek, D. Adanza, P. Alemany, L. Gifre, R. Casellas, R. Martínez, R. Muñoz, and R. Vilalta, “End-to-end transport network digital twins with cloud-native SDN controllers and generative AI [Invited],” Journal of Optical Communications and Networking, vol. 17, no. 7, pp. C70–C81, 2025, doi: 10.1364/JOCN.550864.
34. C. Natalino, “AI-driven optical network automation: Opportunities and challenges,” in Advanced Photonics Congress 2024, Technical Digest Series. Optica Publishing Group, 2024, paper NeW2C.1, doi: 10.1364/NETWORKS.2024.NeW2C.1.
35. C. Natalino, A. Panahi, N. Mohammadiha, and P. Monti, “AI/ML-as-a-Service for optical network automation: Use cases and challenges [Invited],” Journal of Optical Communications and Networking, vol. 16, pp. A169–A179, 2024.
36. A. Pagano, R. Mercinelli, M. Valvo, and A. Manzalini, “Quantum key distribution in optical fibers: A comprehensive design view of the overall quantum layer beyond transmission,” Journal of Optical Communications and Networking, vol. 16, pp. D111–D118, 2024.

37. N. Aquina, S. Rommel, and I. T. Monroy, “Quantum secure communication using hybrid post-quantum cryptography and quantum key distribution,” in 2024 24th International Conference on Transparent Optical Networks (ICTON), Bari, Italy, 2024, Art. no. 10648124, doi: 10.1109/ICTON62926.2024.10648124.
38. D. Sequeira, M. Ruiz, N. Costa, A. Napoli, J. Pedro, and L. Velasco, “OCATA: A deep-learning-based digital twin for the optical time domain,” Journal of Optical Communications and Networking, vol. 15, pp. 87–97, 2023.
39. C. Natalino, L. Gifre, F.-J. Moreno-Muro, S. Gonzalez-Diaz, R. Vilalta, R. Muñoz, P. Monti, and M. Furdek, “Flexible and scalable ML-based diagnosis module for optical networks: A security use case [Invited],” Journal of Optical Communications and Networking, vol. 15, no. 8, pp. C155–C165, 2023, doi: 10.1364/JOCN.482932.
40. M. F. Mello da Silva, A. Sgambelluri, A. Pacini, F. Paolucci, A. W. Green, D. D. L. Mascarenas, and L. Valcarenghi, “Confidentiality-preserving machine learning algorithms for soft-failure detection in optical communication networks,” Journal of Optical Communications and Networking, vol. 15, no. 8, pp. C212–C222, 2023, doi: 10.1364/JOCN.481690.