EST. 2021 EMC
EDITORIAL MAR CARIBE

Inteligencia artificial agéntica y criptografía poscuántica

Escrito por:
Juan Carlos Lázaro Guillermo
Edwin Edgard Torre Suarez
Dulio Oseda Gago
Jorge Luis Armijos Carrión
Gustavo Arredondo Castillo
William Martín Enríquez Maguiña
Jose Luis Cuya Camara

www.editorialmarcaribe.es

ISBN: 978-9915-698-90-8

9 789915 698908

Inteligencia artificial agéntica y criptografía poscuántica

Lázaro Guillermo, Juan Carlos; Torre Suarez, Edwin Edgard; Oseda Gago, Dulio; Armijos Carrión, Jorge Luis; Arredondo Castillo, Gustavo; Enríquez Maguiña, William Martín; Cuya Camara, Jose Luis

Primera edición (1.ª ed.): abril, 2026

Editado por:

Editorial Mar Caribe ®

www.editorialmarcaribe.es

Av. Gral. Flores 547, 70000 Col. del Sacramento, Departamento de Colonia, Uruguay.

Diseño de carátula e ilustraciones: *Luisa Fernanda Lugo Rojas*

Libro electrónico disponible en:

https://editorialmarcaribe.es/ark:/10951/isbn.9789915698908

Formato: Electrónico

ISBN: 978-9915-698-90-8

ARK:

ark:/10951/isbn.9789915698908

Editorial Mar Caribe (OASPA): Como miembro de la Open Access Scholarly Publishing Association, apoyamos el acceso abierto de acuerdo con el código de conducta, la transparencia y las mejores prácticas de OASPA para la publicación de libros académicos y de investigación. Estamos comprometidos con los más altos estándares editoriales en ética y deontología, bajo la premisa de «Ciencia Abierta en América Latina y el Caribe»

OASPA

Editorial Mar Caribe, firmante Nº 795 de 12.08.2024 de la Declaración de Berlín

"... Nos sentimos obligados a abordar los retos de Internet como un medio funcional emergente para la distribución del conocimiento. Obviamente, estos avances pueden modificar significativamente la naturaleza de la publicación científica, así como el actual sistema de garantía de calidad...." (Max Planck Society, ed. 2003, pp. 152-153).

Editorial Mar Caribe

Inteligencia artificial agéntica y criptografía poscuántica

Colonia, Uruguay

2026

Inteligencia artificial agéntica y criptografía poscuántica

Editorial Mar Caribe
Av. General Flores 547, Colonia, Colonia, Uruguay. CP 70000
www.editorialmarcaribe.es

Encuéntranos en:
Facebook
X

Contáctenos:

contacto@editorialmarcaribe.es

contacto@editorialmarcaribe.uy

admin@edicionesmarcaribe.com

ISBN *978-9915-698-90-8*

Título: Inteligencia artificial agéntica y criptografía poscuántica

El contenido de este libro es propiedad intelectual de:

Lázaro Guillermo, Juan Carlos, *ORCID:* 0000-0002-4785-9344, *Correo:* jlazarog@unia.edu.pe, *Universidad Nacional Intercultural de la Amazonía, Perú*
Torre Suarez, Edwin Edgard, *ORCID:* 0009-0005-3590-7950, *Correo:* edwtorre@gmail.com, *Universidad Nacional Hermilio Valdizán, Perú*
Oseda Gago, Dulio, *ORCID:* 0000-0002-3136-6094, *Correo:* doseda@undc.edu.pe, *Universidad Nacional de Cañete, Perú*
Armijos Carrión, Jorge Luis, *ORCID:* 0000-0003-0312-786X, *Correo:* jlarmijos@utmachala.edu.ec, *Universidad Técnica de Machala, Ecuador*
Arredondo Castillo, Gustavo, *ORCID:* 0000-0002-2199-4660, *Correo:* garredondoc@unmsm.edu.pe, *Universidad Nacional Mayor de San Marcos, Perú*
Enríquez Maguiña, William Martín, *ORCID:* 0000-0003-1819-191X, *Correo:* wenriquezm@unmsm.edu.pe, *Universidad Nacional Mayor de San Marcos, Perú*
Cuya Camara, Jose Luis, *ORCID:* 0000-0003-3412-6608, *Correo:* jose.cuya@upn.pe, *Universidad Privada del Norte, Perú*

2026

Índice

Prólogo

En la historia del desarrollo tecnológico existen momentos de inflexión donde dos corrientes disciplinares, hasta entonces paralelas o tangenciales, convergen de forma acelerada para redefinir los cimientos mismos de nuestra infraestructura digital. Nos encontramos en medio de uno de esos hitos históricos: la confluencia inevitable entre la Inteligencia Artificial Agéntica (IAA) y la Criptografía Poscuántica (PQC).

Del análisis histórico, la criptografía ha operado bajo la premisa de supuestos matemáticos de dureza computacional. Algoritmos como RSA o ECC han sostenido el comercio electrónico, la privacidad de las comunicaciones y la soberanía de los datos a nivel global. Sin embargo, la inminente madurez de la computación cuántica —capaz de resolver de manera eficiente la factorización de enteros y el problema del logaritmo discreto mediante el algoritmo de Shor— exige una migración sin precedentes hacia esquemas resistentes a ataques cuánticos.

Por otro lado, la Inteligencia Artificial ha trascendido la era de los modelos predictivos pasivos para adentrarse en el paradigma agéntico. Los agentes autónomos de IA procesan y analizan información; perciben su entorno, toman decisiones multinivel, ejecutan acciones complejas de manera independiente y coordinan redes de colaboración sintética. Esta convergencia plantea desafíos y oportunidades de una magnitud inédita que este libro aborda de manera pionera y exhaustiva:

- *Seguridad criptográfica en entornos autónomos:* ¿Cómo pueden los modelos criptográficos lattice-based o basados en isogenias garantizar la autenticidad, confidencialidad e integridad de las transacciones ejecutadas por agentes sin intervención humana?
- *Agentes como custodios y defensores:* la IA agéntica surge como una

herramienta crítica para supervisar la compleja transición hacia la criptografía poscuántica, facilitando la agilidad criptográfica, identificando vulnerabilidades en código heredado y automatizando la migración de protocolos a escala masiva.

- Nuevas superficies de ataque: la capacidad adaptativa y el razonamiento autónomo de los agentes artificiales abren el camino hacia esquemas de ciberataques impulsados por IA que podrían explotar las fases de transición e implementación del estándar PQC.

La presente investigación reúne un análisis riguroso que va desde las bases teóricas de la seguridad poscuántica hasta las arquitecturas de software de los sistemas agénticos avanzados. A través de sus capítulos, el profesional y académico lector encontrará un marco conceptual y práctico para comprender cómo la autonomía de la máquina y la resistencia cuántica deben coevolucionar para construir el ecosistema digital seguro del mañana. Esta obra es una hoja de ruta conceptual para investigadores, arquitectos de seguridad, desarrolladores de IA y responsables de políticas de ciberseguridad que buscan anticiparse a los desafíos del próximo decenio.

Invitamos a la comunidad científica a adentrarse en esta lectura para adquirir conocimientos y reflexionar sobre la responsabilidad de diseñar sistemas autónomos capaces de proteger el bien más valioso de nuestra era: la confianza en la información.

Juan Carlos Lázaro Guillermo

Introducción

El panorama tecnológico actual está a punto de experimentar dos grandes cambios interrelacionados: el desarrollo de la Inteligencia Artificial (IA) Agéntica y el avance hacia la Criptografía Poscuántica (PQC). La IA agéntica representa una mejora significativa sobre los modelos de lenguaje tradicionales, mostrando sistemas que no solo generan contenido de forma reactiva, sino que funcionan como agentes autónomos capaces de realizar razonamientos complejos, tomar decisiones en escenarios complicados, emplear herramientas digitales y llevar a cabo tareas estratégicas sin necesidad de supervisión constante.

Por otro lado, el avance inminente de la computación cuántica representa una amenaza existencial para la infraestructura global de seguridad cibernética. Algoritmos como el de Shor podrían desestabilizar los fundamentos del cifrado asimétrico actual (RSA, ECC, Diffie-Hellman), lo que obliga a adoptar con urgencia esquemas criptográficos resistentes a ataques cuánticos (como los basados en retículas, códigos, isogenias y multivariables). La interacción de estas áreas genera un nuevo campo de estudio, sumamente complejo.

i. *La IA agéntica como vector de amenaza y herramienta defensiva:* los agentes de IA pueden ser entrenados para automatizar la criptoanálisis, descubrir vulnerabilidades en implementaciones poscuánticas a velocidad sobrehumana o coordinar ciberataques dinámicos. En contraposición, estos mismos agentes resultan indispensables para orquestar la compleja migración hacia estándares PQC, gestionar claves dinámicas y detectar anomalías en tiempo real.
ii. *Resiliencia de la IA agéntica:* los propios sistemas agénticos —sus comunicaciones en red, bases de conocimiento distribuido y protocolos de consenso entre agentes— requieren arquitecturas de cifrado

poscuántico para garantizar su integridad y prevenir la suplantación o manipulación maliciosa de sus decisiones.

Este libro examina rigurosamente el empalme entre la autonomía de la inteligencia artificial (IA) y los nuevos estándares de cifrado poscuántico, analizando tanto los riesgos emergentes como las oportunidades estratégicas para la seguridad de la información en la era cuántica. ¿De qué manera la IA agéntica transforma la implementación, evaluación de vulnerabilidades y gestión de la criptografía poscuántica (PQC), y cuáles son los requerimientos criptográficos necesarios para garantizar la autonomía segura de los agentes de IA frente a la amenaza cuántica?

Con base en la pregunta de investigación, se plantea como objetivo analizar la relación bidireccional entre la IA agéntica y la criptografía poscuántica, evaluando cómo la autonomía de los agentes inteligentes optimiza o vulnera los sistemas PQC, y estableciendo un marco conceptual y técnico para la implementación de arquitecturas agénticas ciberresilientes en el escenario post-RSA.

Para delimitar con precisión el campo de análisis, la investigación abarca los siguientes frentes:

- *Alcance Teórico-Conceptual:*
    - Estudio de los principales estándares de PQC formalizados (NIST) como ML-KEM (Kyber), ML-DSA (Dilithium) y SLH-DSA (SPHINCS+).
    - Caracterización de las arquitecturas de IA agéntica (sistemas multiagente, memoria a largo plazo, uso de herramientas y marcos de orquestación).
- *Alcance Técnico y Operativo:*
    - Evaluación del rol de la IA agéntica en el proceso de agilidad criptográfica (*crypto-agility*) y la automatización de la migración tecnológica.

    - *Análisis de ciberseguridad ofensiva y defensiva:* simulación conceptual del uso de agentes autónomos en auditorías de código PQC y gestión de claves.
- *Alcance Temático:*
    - La investigación se centra en el ecosistema informático civil y corporativo de alta sensibilidad (financiero, crítico e industrial).

El estudio se justifica bajo el criterio de conveniencia y oportunidad, pues, la transición a la PQC no es inmediata; requiere años de reacondicionamiento de infraestructura. Estudiar la integración de agentes de IA en este momento permite automatizar un proceso que de otro modo sería lento, costoso y propenso al error humano. A su vez, en cuanto a relevancia social y económica, garantizar la transición segura a PQC es vital para proteger la privacidad ciudadana, el comercio electrónico, las infraestructuras críticas y la seguridad nacional frente a ataques del tipo "*Harvest Now, Decrypt Later*" (recolectar ahora, descifrar después).

Se propone un enfoque de investigación mixta secuencial exploratoria y explicativa (QUAL-QUAN-Integración): primero, se elabora un análisis formal y una revisión sistemática; luego, el diseño e implementación de entornos de simulación; y, finalmente, la triangulación y redacción de hallazgos técnicos. Esto a través de técnicas de benchmarking de algoritmos de firma PQC en redes de agentes y tablas comparativas de rendimiento y propuesta de protocolo de identidad sintética.

En la investigación, se aportan pautas concretas para diseñadores de sistemas y directores de seguridad (CISO), ofreciendo estrategias para integrar PQC en redes de comunicación entre agentes autónomos, así como los ecosistemas que cierran la brecha en la literatura académica actual, que suele tratar la IA agéntica y la criptografía cuántica de forma aislada, sin abordar su convergencia analítica y operativa.

Por ello, los autores sugieren un análisis técnico-científico ante la carencia de un modelo integrado que combine las capacidades operativas de la IA agéntica con los estándares de la criptografía poscuántica. Este vacío deja a las organizaciones vulnerables: por un lado, ante una migración lenta e ineficiente a la PQC frente a la amenaza cuántica; y por otro, ante la proliferación de sistemas de IA autónomos sin protección contra intercepciones cuánticas y con riesgo de uso malicioso en ciberataques automatizados.

Capítulo 1
Identidad de máquinas, defensa de redes y gobernanza automatizada con inteligencia artificial

La transición global hacia la criptografía poscuántica (PQC, por sus siglas en inglés) representa la reestructuración más compleja y crítica de la infraestructura de seguridad digital en la historia de la informática. La convergencia de la computación cuántica a gran escala —capaz de ejecutar el algoritmo de Shor para factorizar enteros de gran tamaño y calcular logaritmos discretos en tiempo polinómico— invalidará de manera irreversible los pilares fundamentales de la criptografía asimétrica clásica, como RSA, Diffie-Hellman (DH), la criptografía de curva elíptica (ECC), ECDH y ECDSA (Riquelme y Pereyra, 2025).

Esta vulnerabilidad sistémica exige una evolución coordinada que trasciende el reemplazo aislado de algoritmos en aplicaciones individuales. Requiere la refundación de la identidad de máquinas en entornos *Zero Trust*, la adaptación de los protocolos de transporte y de la defensa de redes para gestionar sobrecargas de tamaño de clave significativas, y la adopción de herramientas impulsadas por inteligencia artificial (IA) para la gobernanza, el descubrimiento continuo de activos y la mitigación de ataques a canales laterales (*side-channel attacks*).

Estándares criptográficos poscuánticos y marco normativo global

El Instituto Nacional de Estándares y Tecnología de EE. UU. (NIST) Marcó

el hito operacional decisivo al publicar las primeras versiones finales de los Estándares Federales de Procesamiento de Información (FIPS) orientados a la resistencia cuántica. Estas especificaciones constituyen el nuevo marco de referencia técnico para la infraestructura de clave pública (PKI) a nivel global (Dziechciarz y Niemiec, 2025).

Especificaciones de los estándares NIST finalizados

Los estándares promulgados se dividen en mecanismos de encapsulamiento de claves (KEM) y esquemas de firma digital, sustentados en problemas matemáticos duros sobre mallas estructurales (*structured lattices*) y funciones hash:

1. **FIPS 203 — ML-KEM (Module-Lattice-Based Key-Encapsulation Mechanism Standard):** Basado en el algoritmo CRYSTALS-Kyber, está diseñado para el intercambio de claves y el transporte cifrado en protocolos como TLS y IPsec, reemplazando a RSA KEM y ECDH. Su seguridad se fundamenta en la dureza del problema de aprendizaje con errores sobre módulos (*Module Learning With Errors*, M-LWE).
2. **FIPS 204 — ML-DSA (Module-Lattice-Based Digital Signature Standard):** Derivado de CRYSTALS-Dilithium, constituye el estándar primario recomendado para firmas digitales en certificados X.509, autenticación de protocolos, firma de código y documentos. También se basa en el problema M-LWE y en el problema de aprendizaje con redondeo sobre módulos.
3. **FIPS 205 — SLH-DSA (Stateless Hash-Based Digital Signature Standard):** Basado en SPHINCS+, ofrece un esquema de firma digital sin estado cuya seguridad descansa exclusivamente en las propiedades de resistencia a colisiones y preimagen de las funciones hash criptográficas. Funciona como una alternativa conservadora de respaldo frente a eventuales debilidades teóricas descubiertas en las mallas algebraicas.

Próximos estándares en desarrollo técnico incluyen FIPS 206 (FN-DSA, basado en FALCON) y la estandarización de HQC (*Hamming Quasi-Cyclic*) como mecanismo de encapsulamiento de claves alternativo no basado en mallas (véase la tabla 1).

Tabla 1: Especificaciones de los estándares NIST finalizados

Estándar	Criptosistema de Origen	Tipo de Criptografía	Primitivas Clásicas Reemplazadas	Niveles de Seguridad (Parámetros)	Tamaño de Clave Pública (Bytes)	Tamaño de Texto Cifrado / Firma (Bytes)	Caso de Uso Primario
FIPS 203	CRYSTALS-Kyber	Encapsulamiento de Claves (KEM)	RSA key exchange, ECDH, DH	ML-KEM-512 ML-KEM-768 ML-KEM-1024	800 1,184 1,568	768 1,088 1,568	Intercambio de claves TLS 1.3, IPsec, túneles VPN
FIPS 204	CRYSTALS-Dilithium	Firma Digital (Mallas)	RSA-2048/4096, ECDSA, Ed25519	ML-DSA-44 ML-DSA-65 ML-DSA-87	1,312 1,952 2,592	2,420 3,293 4,627	PKI, firmas de certificados, firma de código
FIPS 205	SPHINCS+	Firma Digital (Hash)	RSA, ECDSA	12 conjuntos de parámetros (p. ej., SLH-DSA-SHA2-	32 - 64	7,856 - 49,856	Root CAs, firmas de larga duración, respaldo

		128f)	conservador

Mandatos gubernamentales y cronogramas de cumplimiento normativo

Las directivas de seguridad nacional han establecido límites operacionales estrictos que transforman la adopción de PQC de una recomendación teórica en una obligación regulatoria vinculante.

La NSA, mediante el conjunto de algoritmos de seguridad nacional comercial (*Commercial National Security Algorithm Suite 2.0*, CNSA 2.0), exige la transición exclusiva a algoritmos resistentes a ataques cuánticos en todos los Sistemas de Seguridad Nacional (NSS). La actualización del FAQ V2.1 de CNSA 2.0 especifica que los entornos de alta seguridad no permiten implementaciones en borrador ni variantes intermedias; únicamente se aceptan las configuraciones de parámetros de mayor fortaleza: ML-KEM-1024 para establecimiento de claves y ML-DSA-87 para firmas digitales (véase la tabla 2) (Płoszaj et al., 2024).

Tabla 2: Conjunto de algoritmos de seguridad nacional comercial

Organismo / Marco	Hito / Fecha Límite	Requisito Técnico o Mandato Operativo	Alcance del Sistema
NSA CNSA 2.0	2025	Soportar y preferir opciones PQC (ML-KEM-1024 / ML-DSA-87)	Navegadores, servidores web, servicios en la nube
NSA CNSA 2.0	1 Enero 2027	Cumplimiento obligatorio de CNSA 2.0 en todas las nuevas adquisiciones de NSS	Redes y equipamiento de Seguridad Nacional

NSA CNSA 2.0	2030 - 2033	Uso exclusivo de PQC; prohibición total de RSA, ECDH, ECDSA y Diffie-Hellman	Equipamiento de red, sistemas operativos, servidores
NIST IR 8547 / OMB	2030	Depreciación formal de algoritmos vulnerables (RSA/ECC) e incremento del nivel de seguridad mínimo a 128 bits	Sistemas de información federales civiles
NIST IR 8547	2035	Prohibición absoluta de criptografía clásica vulnerable en todos los sistemas federales	Infraestructura gubernamental completa
EE. UU. Orden Ejecutiva 14412	22 Junio 2026	Directiva ejecutiva que formaliza la exigencia de un CBOM (*Cryptographic Bill of Materials*)	CISA/NIST publican especificaciones en 270 días
EE. UU. Orden Ejecutiva 14412	200 (Claves) / 2031 (Firmas)	Migración completa de Activos de Alto Valor (*High-Value Assets*) a PQC	Infraestructura crítica federal

El marco temporal impone un desafío estructural crítico: el fenómeno *Harvest Now, Decrypt Later* (HNDL) —donde adversarios estatales capturan hoy tráfico cifrado clásico para desencriptarlo retroactivamente cuando dispongan de una computadora cuántica criptográficamente relevante (CRQC)— implica que la información confidencial con un ciclo de vida o vigencia que supere el año 2030 ya se encuentra comprometida si transita sobre algoritmos tradicionales.

Identidad de máquinas y arquitectura zero trust poscuántica

En arquitecturas distribuidas modernas, la identidad de máquinas ha reemplazado al perímetro de red como el plano de control primario de la seguridad. La autenticación mutua sobre *Transport Layer Security* (mTLS) sustentada en microservicios, contenedores, cargas de trabajo en la nube y agentes de IA requiere que las credenciales criptográficas se emitan, verifiquen y roten con frecuencias elevadas y latencias mínimas (Ngema et al., 2026).

Integración de SPIFFE/SPIRE en la infraestructura de servicio Mesh

El estándar abierto *Secure Production Identity Framework for Everyone* (SPIFFE) y su entorno de ejecución de referencia *SPIFFE Runtime Environment* (SPIRE) definen un mecanismo multiplataforma para atestar e identificar cargas de trabajo de forma criptográfica.

SPIRE emite un *SPIFFE Verifiable Identity Document* (SVID), generalmente estructurado como un certificado X.509 de corta duración, el cual codifica el *SPIFFE ID* de la carga de trabajo en la extensión *Subject Alternative Name* (SAN). En un ecosistema *Zero Trust* orquestado por Service Meshes como Istio o mallas basadas en Envoy, el servidor de SPIRE actúa como Autoridad de Certificación (CA) central y proveedor de identidades mediante la interfaz *Secret Discovery Service* (SDS).

El ciclo de vida de la atestación e intercambio de credenciales poscuánticas dentro del entorno SPIFFE/SPIRE se articula mediante el siguiente flujo operativo secuencial:

- Atestación del Nodo y la Carga de Trabajo: Cuando una carga de trabajo se inicializa, el agente SPIRE local ubicado en el nodo realiza una verificación

fuera de banda inspeccionando las propiedades del proceso del sistema (como los hashes del contenedor, el espacio de nombres de Kubernetes o los identificadores del kernel) para certificar su legitimidad antes de conceder cualquier credencial.

- Solicitud de Firma de Certificado (CSR): Una vez validada la identidad del proceso, la carga de trabajo consulta la API local de SPIFFE expuesta por el agente para solicitar su identidad. El agente SPIRE genera una solicitud CSR que transmite al servidor SPIRE central.
- Emisión del SVID Poscuántico por la CA: El servidor SPIRE, actuando como la CA raíz o intermedia del dominio de confianza, firma la solicitud utilizando una clave privada resistente a ataques cuánticos (como ML-DSA-65 o ML-DSA-87) y devuelve al agente un SVID X.509 junto con el paquete de confianza (*Trust Bundle*) correspondiente.
- Establecimiento del Handshake mTLS: El proxy Sidecar (como Envoy) obtiene el SVID X.509 PQC mediante la interfaz SDS y ejecuta el apretón de manos mTLS con el servicio destino, empleando ML-KEM para el intercambio de claves efímeras y ML-DSA para la autenticación de la identidad de la máquina.

Estandarización IETF de certificados híbridos y compuestos

Durante la fase de transición, la incompatibilidad entre pilas de red poscuánticas y clientes legados impide la migración inmediata a certificados PQC puros. Para resolver esta bifurcación, el Grupo de Trabajo LAMPS del IETF ha desarrollado estándares para estructurar credenciales X.509 híbridas.

Existen dos filosofías arquitectónicas principales dentro de los borradores del IETF para abordar la convivencia de primitivas criptográficas:

1. Firma Compuesta Poscuántica/Tradicional (draft-ietf-lamps-pq-composite-sigs)

Define combinaciones atómicas de algoritmos donde una clave pública PQC (p. ej., ML-DSA) y una clave pública clásica (p. ej., ECDSA o RSA) se fusionan en un único elemento criptográfico identificado por un solo Identificador de Objeto (OID).

La firma final se produce combinando las dos firmas componentes mediante una función combinadora (*signature combiner*) que incluye pre-hashing y etiquetas de contexto. Un certificado compuesto requiere que **ambas** firmas se verifiquen con éxito para dar por válida la identidad. Esta aproximación proporciona compatibilidad a nivel de protocolo (*protocol backward compatibility*), ya que el sistema no requiere ser plenamente "consciente del modo híbrido" en las capas superiores; la estructura es tratada como un algoritmo único e indivisible.

2. Mecanismo de Extensiones Alternativas (draft-truskovsky-lamps-pq-hybrid-x509)

A diferencia del enfoque compuesto atómico, esta especificación introduce extensiones v3 en la estructura ASN.1 del certificado X.509 para transportar material criptográfico secundario sin alterar los campos principales del certificado:

- id-subjectAltPublicKeyInfo: Transporta la clave pública alternativa (poscuántica).
- id-altSignatureAlgorithm: Identifica el algoritmo de firma alternativo utilizado por la CA.
- id-altSignatureValue: Contiene el valor de la firma digital calculada por la CA sobre la estructura preTBSCertificate usando el algoritmo alternativo.

Esta arquitectura permite que un cliente legado procese la clave clásica

y la firma convencional alojadas en los campos estándar X.509 (SubjectPublicKeyInfo y signatureValue), mientras que un cliente actualizado extrae y verifica las credenciales poscuánticas ubicadas en las extensiones, logrando una interoperabilidad dual transparente.

Prevención de Inyecciones de Degradación (*Downgrade Attacks*): La Extensión PQCHC

La coetaneidad de credenciales tradicionales y poscuánticas expone a la infraestructura a ataques de degradación (*downgrade attacks*). Un atacante activo con acceso a una computadora cuántica podría interceptar la negociación de seguridad, suprimir las credenciales PQC o compuestas presentadas por el servidor y forzar la degradación a un certificado clásico RSA/ECDSA, permitiéndole realizar un ataque *Man-in-the-Middle* (MitM) sin levantar alertas en el validador.

Para contrarrestar esta vulnerabilidad, la especificación draft-reddy-lamps-x509-pq-commit introduce la extensión X.509 denominada *Post-Quantum or Composite Hosting Continuity* (PQCHC), identificada formalmente por el OID id-pe-pqchc. Mediante esta extensión, el sujeto del certificado y la CA firman una declaración de intención explícita en la que se comprometen a continuar sirviendo credenciales PQC o compuestas durante el período de validez del certificado y durante un intervalo extendido denominado continuityPeriod, aplicable tras la fecha de expiración notAfter. Si una entidad de verificación observa que un servidor con la extensión PQCHC omite repentinamente el soporte PQC en el apretón de manos, identifica la inconsistencia como una manipulación maliciosa de degradación y aborta la conexión inmediatamente (véase la tabla 3).

Tabla 3: Credenciales tradicionales y poscuánticas

Parámetro / Propiedad	Esquema Compuesto Atómico (draft-ietf-lamps-pq-composite-sigs)	Esquema de Extensiones Alternativas (draft-truskovsky-lamps-pq-hybrid-x509)	Certificados PQC Puros (FIPS 204 / 205)
Representación ASN.1	Un solo OID en campos de clave y firma X.509 estándar.	Clave y firma clásicas en campos estándar; PQC en extensiones v3.	Clave y firma ML-DSA / SLH-DSA en campos estándar X.509.
Lógica de Verificación	Conjunción estricta ("AND"): requiere la validación exitosa de ambas primitivas.	Semántica condicional: el cliente elige procesar la clave principal o la extensión.	Validación única de la primitiva poscuántica.
Compatibilidad con Clientes Legados	Baja. Los analizadores ASN.1 antiguos rechazan OIDs compuestos no reconocidos.	Alta. Los clientes antiguos ignoran extensiones no críticas y procesan el bloque RSA/ECC.	Nula. Incompatible con clientes que no soporten algoritmos FIPS 203/204.
Resistencia a Fallos Algorítmicos	Excelente. Se mantiene seguro mientras al menos uno de los componentes permanezca intacto.	Moderada. Depende de si la pila del cliente fuerza el uso de la extensión PQC.	Sensible a vulnerabilidades matemáticas imprevistas en la teoría de mallas.
Impacto en Tamaño de Certificado	Suma lineal de clave pública y firmas de ambas primitivas.	Incremento superior debido al encapsulamiento redundante ASN.1	Amplia sobrecarga en firmas (ML-DSA: ~2.4–4.6 KB).

Impacto en la defensa de redes y protocolos de transporte

La integración de la criptografía poscuántica en la capa de transporte (TLS 1.3, DTLS 1.3, QUIC) introduce cambios sustanciales en la dinámica de tráfico de red. Las propiedades algebraicas de las mallas y los arreglos hash requieren un intercambio de datos considerablemente superior al de los grupos Diffie-Hellman sobre curvas elípticas clásicas (p. ej., X25519) (Cherkaoui et al., 2024).

Análisis de sobrecarga en la capa de transporte y fragmentación de paquetes

En el apretón de manos TLS 1.3 clásico, un intercambio de claves efímero basado en ECDH requiere enviar un vector *KeyShare* de aproximadamente 32 bytes dentro del mensaje inicial ClientHello. En contraste, al implementar ML-KEM-768 (FIPS 203), la clave pública transmitida abarca 1,184 bytes, mientras que la respuesta del servidor (*ciphertext*) transporta 1,088 bytes.

Esta variación en las dimensiones del mensaje altera la relación con la Unidad Máxima de Transmisión (*Maximum Transmission Unit*, MTU) estándar en redes Ethernet (habitualmente 1.500 bytes).

Cuando la carga útil de la extensión KeyShare en el ClientHello sobrepasa el límite de la MTU física, el mensaje se fragmenta a nivel de la capa de red (IP fragmentation) o dentro del protocolo de transporte en escenarios como QUIC o DTLS. La trayectoria del tráfico durante la negociación PQC atraviesa un punto de falla crítico cuando ocurre la fragmentación de paquetes:

- **Emisión del ClientHello PQC:** El cliente genera el mensaje inicial de apretón de manos conteniendo la clave pública ML-KEM-768 (o la combinación híbrida X-Wing), con un volumen que ronda de 1,8 a 2,2 kilobytes.

- **Fragmentación en la Capa IP:** Al superar la MTU de la interfaz Ethernet de salida (1,500 bytes), la pila de red divide el datograma en múltiples fragmentos IP independientes (p. ej., un primer paquete de 1,500 bytes y un segundo de aproximadamente 400 a 700 bytes).
- **Intercepción por Dispositivos Intermedios (*Middleboxes*):** Los paquetes fragmentados transitan por la infraestructura de red donde son inspeccionados por cortafuegos, sistemas de detección de intromisiones (IDS) o balanceadores de carga con inspección de estado.
- **Descarte de Conexión (*Drop*):** Debido a políticas estrictas de seguridad diseñadas para prevenir ataques DoS o evasión de inspección, los dispositivos intermedios descartan los fragmentos IP secundarios o datagramas UDP incompletos, provocando la caída silenciosa de la conexión (*black-holing*) y retransmisiones fallidas.

Los efectos colaterales de este fenómeno sobre el rendimiento de red incluyen:

- **Latencia Exponencial en Redes de Acceso Radio (RAN):** En entornos móviles e inalámbricos, la adición de cada kilobyte extra a los mensajes del apretón de manos incrementa la latencia percibida en aproximadamente un 1,5% a 3% por cada *Round-Trip Time* (RTT), exacerbando los retrasos de cola en redes celulares congestionadas.
- **Incompatibilidad de Tampones en Hardware Antiguo:** Tarjetas de red con descarga de procesamiento criptográfico (*TLS offload*) y aceleradores de hardware legados sufren desbordamientos de buffer o excepciones al intentar procesar certificados cuya cadena excede los límites históricos, como ocurre con las firmas ML-DSA (de 2.4 KB a 4.6 KB) o SLH-DSA (hasta 49.8 KB).

Estrategias de optimización y mitigación mecánica

Para mitigar los efectos desestabilizadores de las sobrecargas de tamaño

PQC en la capa de red, las arquitecturas de red modernas implementan optimizaciones estructurales:

- **Mecanismos Híbridos KEM Eficientes (X-Wing):** Un esquema combinado que une ML-KEM-768 con X25519. Al optimizar la serialización del paquete *KeyShare*, se minimizan los bytes residuales para mantenerse en el límite de un único marco Ethernet cuando se combina con la reducción de extensiones TLS opcionales.
- **Compresión de Certificados TLS (RFC 8879):** Las entidades de red aplican algoritmos de compresión como zlib, Brotli o zstd a la cadena de certificados devuelta en el mensaje Certificate de TLS 1.3, reduciendo el volumen de transmisión de firmas ML-DSA en más de un 30-40%.
- **Despliegue Cautivo de PKI y Caching de CAs Intermedias:** En redes *Zero Trust* controladas (como microservicios interconectados), los nodos intermedios almacenan en caché las cadenas de certificados raíz e intermedios. De este modo, los servicios intercambian únicamente certificados de hoja (*leaf SVIDs*) o referencias hash directas, aliviando la saturación de los canales mTLS (Lee et al., 2025).
- **Procesamiento de Fragmentación en QUIC y DTLS 1.3:** Uso de extensiones específicas de PMTUD (*Path MTU Discovery*) para calcular dinámicamente el tamaño de datagrama soportado antes de emitir los mensajes ClientHello con claves poscuánticas.

Gestión criptográfica asistida por IA e inventario automatizado (CBOM)

La escala exponencial de activos criptográficos distribuidos en infraestructuras multinube, pipelines de CI/CD, arquitecturas serverless e IoT imposibilita el inventario manual. La visibilidad y la agilidad criptográfica (*crypto-agility*) requieren la implementación de herramientas automatizadas de

descubrimiento continuo que generen de forma estandarizada una Lista de Materiales Criptográficos (*Cryptographic Bill of Materials*, CBOM) (Ploszaj et al., 2024).

La lista de materiales criptográficos (CBOM):

Un CBOM es un registro estructurado, legible por máquinas y continuamente actualizado que extiende el concepto del *Software Bill of Materials* (SBOM) para proporcionar un mapa contextualizado e integral del ecosistema criptográfico de una organización.

Bajo esquemas abiertos como la especificación CycloneDX (desarrollada por OWASP) y los mandatos derivados de la Orden Ejecutiva 14412 firmada en junio de 2026, un CBOM funcional debe catalogar los siguientes componentes esenciales:

1. **Identificación de Primitivas y Algoritmos:** Nombre exacto del algoritmo, longitud de clave, modo de operación (p. ej., AES-256-GCM, RSA-2048, ML-KEM-768) e identificadores OID.
2. **Procedencia de Módulos y Librerías:** Trazabilidad de las librerías criptográficas activas en tiempo de compilación y ejecución (OpenSSL, BoringSSL, AWS-LC, Bouncy Castle), incluyendo versiones específicas y estado de validación FIPS 140-3.
3. **Gestión de Claves y Certificados:** Metadatos de identificadores de claves, tipo de almacenamiento (HSM, TPM, KMS en la nube, keystores de software), fechas de expiración, esquemas de rotación y datos de la cadena de certificación.
4. **Contexto Operacional e Identificación de Dependencias:** Vinculación directa entre el activo criptográfico y las aplicaciones, componentes de software (purl), servicios, endpoints de API o contenedores que dependen de él.
5. **Evaluación de Riesgo Poscuántico:** Clasificación explícita del nivel de

vulnerabilidad cuántica del activo (p. ej., Crítico/Vulnerable para RSA/ECC; Poscuántico/Seguro para ML-DSA)[3].

Descubrimiento continuo impulsado por inteligencia artificial

Las soluciones estáticas de inventario basadas en análisis de código fuente resultan insuficientes porque ignoran la configuración en tiempo de ejecución, la generación dinámica de claves y las dependencias criptográficas no documentadas ("criptografía en la sombra" o "*shadow crypto*"). La IA transforma la gobernanza de la seguridad mediante capacidades analíticas avanzadas.

El ecosistema de descubrimiento automatizado con IA opera mediante la convergencia de tres capas de ingesta y procesamiento de datos:

- Capa de Ingesta y Extracción Multifuente: Motores especializados recopilan información de múltiples puntos de la infraestructura. Ejecutan análisis estático de código fuente (AST), plantillas de infraestructura como código (IaC) y binarios compilados; inspeccionan en tiempo de ejecución el tráfico de red y las llamadas al sistema operativo mediante sondas eBPF (*Berkeley Packet Filter Extended*); y consultan las APIs de los almacenes de claves (KMS, HSM y keystores en la nube).
- Capa de Procesamiento e Inferencia con IA: Los datos crudos se canalizan hacia modelos de aprendizaje automático especializados. Un clasificador de patrones criptográficos identifica primitivas sin documentar o algoritmos implementados de forma opaca; un motor de grafos de dependencia mapea la relación entre las librerías criptográficas y los procesos de negocio; y un módulo de evaluación de riesgos analiza la exposición frente al fenómeno *Harvest Now, Decrypt Later* (HNDL).
- Capa de Estructuración y Orquestación: El resultado se sintetiza automáticamente en un documento CBOM estandarizado (en formato CycloneDX o JSON), el cual alimenta directamente las herramientas de

gobernanza, los sistemas de gestión del ciclo de vida de certificados (CLM) y las políticas de cumplimiento en los pipelines de CI/CD.

IA en la evaluación de vulnerabilidades de implementación: ataques de canal lateral (SCA)

Más allá del descubrimiento, la IA desempeña un papel dual en el análisis de seguridad de las implementaciones PQC. Algoritmos matemáticamente sólidos como ML-KEM y ML-DSA pueden ser vulnerables durante su ejecución física en hardware o procesadores embebidos si filtran información no intencionada (Alahmadi et al., 2022).

Los modelos de *Deep Learning* (como redes neuronales convolucionales y transformadores) son utilizados por auditores y adversarios para procesar mediciones de consumo de energía, variaciones electromagnéticas (EM) y delaciones de tiempo (*timing attacks*) emitidas por módulos criptográficos mientras procesan operaciones sobre mallas algebraicas (Wang et al., 2026).

La IA permite filtrar el ruido electromagnético ambiental para reconstruir las claves secretas sk en esquemas ML-KEM a partir de trazas físicas individuales, lo que exige que los módulos validados por el programa CMVP (FIPS 140-3) incorporen enmascaramientos aritméticos (*masking*) y desincronizaciones aleatorias verificadas por herramientas de análisis asistidas por IA.

Ejecución de la agilidad criptográfica (*Crypto-Agility*)

El valor operativo del CBOM y la IA se materializa en la agilidad criptográfica: la capacidad de una arquitectura para alternar o reemplazar algoritmos criptográficos, parámetros y claves sin alterar la infraestructura subyacente ni requerir la reescritura masiva de software. Esta agilidad se operacionaliza mediante:

1. Integración de *Policy-as-Code* en CI/CD: Herramientas como Open Policy Agent (OPA) evalúan el CBOM generado en cada *build*. Si un desarrollador introduce una librería que utiliza RSA-2048 o un cifrado no conforme con los parámetros CNSA 2.0, el pipeline de CI/CD bloquea automáticamente el despliegue.
2. Orquestación del Ciclo de Vida de Certificados (CLM): Soluciones CLM automatizadas conectan el inventario CBOM con las CA de la empresa (incluyendo SPIRE). Ante la revocación de un algoritmo o la emisión de un nuevo parámetro por parte del NIST, el sistema de CLM sustituye dinámicamente los certificados de los endpoints sin intervención humana ni interrupción del servicio.

La transición a la criptografía poscuántica excede la dimensión de una actualización tecnológica rutinaria: constituye un proyecto de reingeniería estructural que abarca la criptografía teórica, el rendimiento de las capas de red, la autenticación en arquitecturas *Zero Trust* y la gobernanza continua de datos (Dziechciarz y Niemiec, 2025).

Principales conclusiones analíticas

- El Rendimiento de Red como Factor de Bloqueo Criptográfico: La sobrecarga de tamaño de las primitivas FIPS 203 (ML-KEM) y FIPS 204 (ML-DSA) degrada la eficiencia del apretón de manos TLS 1.3 al cruzar el umbral de la MTU estándar de 1.500 bytes. Las organizaciones deben planificar la reconfiguración de sus infraestructuras de red y la adopción de esquemas de compresión (RFC 8879) para evitar caídas masivas de paquetes por fragmentación IP en dispositivos intermedios.
- Ineludibilidad del Estándar Compuesto en la Identidad de Máquinas: En ecosistemas *Zero Trust* orquestados por SPIFFE/SPIRE y service meshes, los certificados híbridos o compuestos (draft-ietf-lamps-pq-composite-sigs) son indispensables para garantizar la continuidad operativa. La

adopción complementaria de la extensión PQCHC (draft-reddy-lamps-x509-pq-commit) resulta obligatoria para blindar las negociaciones contra ataques de degradación ejecutados mediante computadoras cuánticas (Ngema et al., 2026).

- El CBOM Dinámico Asistido por IA como Requisito de Seguridad y Cumplimiento: Frente a las exigencias normativas impuestas por CNSA 2.0 y la Orden Ejecutiva 14412, la generación manual de inventarios es inviable. La IA debe integrarse de forma nativa en el descubrimiento dinámico de activos y en los pipelines de CI/CD para generar registros CBOM legibles por máquina, clasificando la exposición a ataques *Harvest Now, Decrypt Later*.

Hoja de ruta operativa para la transición poscuántica

Para garantizar una migración efectiva y ordenada, se establece un plan de acción estructurado en cuatro fases temporales definidas por hitos técnicos y regulatorios específicos:

- Fase 1: Descubrimiento Automatizado y Emisión de CBOM (Meses 1 – 12): Desplegar motores de análisis impulsados por IA y agentes eBPF para escanear el código fuente, contenedores, redes y keystores. Generar el CBOM estandarizado bajo el esquema CycloneDX y mapear los sistemas que custodian datos críticos con requerimientos de protección más allá de 2030, identificando el riesgo HNDL de la organización.
- Fase 2: Habilitación de Crypto-Agilidad y Pruebas Piloto de Red (Meses 12 – 24): Implementar el intercambio de claves híbrido (como X-Wing o ML-KEM-768 con X25519) en los balanceadores de carga de borde y proxies TLS. Configurar la compresión de certificados X.509 (RFC 8879) y verificar que las métricas de latencia y la tasa de fragmentación IP se mantengan dentro de los parámetros de tolerancia operacional en entornos de desarrollo y staging.

- Fase 3: Transición de Identidades de Máquina y Mallas de Servicio (Meses 24 – 48): Actualizar el plano de control SPIFFE/SPIRE para dar soporte a la firma de SVIDs mediante FIPS 204 (ML-DSA) o formatos compuestos (draft-ietf-lamps-pq-composite-sigs). Desplegar la extensión PQCHC (draft-reddy-lamps-x509-pq-commit) en la PKI interna de la empresa e integrar reglas de *Policy-as-Code* en los pipelines de integración continua.
- Fase 4: Erradicación Criptográfica Clásica y Cumplimiento Poscuántico Puro (Alineado con Hitos 2030–2035): Retirar progresivamente el soporte de algoritmos asimétricos clásicos (RSA, ECDH, ECDSA) de todos los endpoints, pasando de configuraciones híbridas a cifrado PQC puro conforme a los mandatos CNSA 2.0 y directivas del NIST. Sustituir los módulos HSM legados por plataformas validadas FIPS 140-3 equipadas con contramedidas contra ataques de canal lateral verificadas por IA.

Capítulo 2
La ciberdefensa como quinto dominio de la soberanía estatal: análisis bibliométrico y estratégico

La conceptualización del ciberespacio como el quinto dominio de la guerra —al lado de los dominios físicos tradicionales de tierra, mar, aire y espacio— representa una reconfiguración estructural en la teoría de la seguridad nacional y la doctrina militar contemporánea. A diferencia de los dominios naturales, el ciberespacio es un entorno sintético, no intrínsecamente físico y de naturaleza híbrida, compuesto por infraestructura tangible, datos lógicos, software y redes interconectadas que trascienden las fronteras geográficas tradicionales. Esta condición desdibuja la noción westfaliana de soberanía territorial, pues las operaciones cibernéticas permiten a los actores estatales y no estatales proyectar poder, realizar espionaje, causar sabotajes e interferir en asuntos internos sin necesidad de movilización física ni proximidad geográfica.

El reconocimiento oficial del ciberespacio como dominio operativo formal fue impulsado inicialmente por el Departamento de Defensa de los Estados Unidos y posteriormente consolidado a nivel multilateral durante la Cumbre de la OTAN en Varsovia en julio de 2016. En dicha cumbre, los países aliados declararon que el ciberespacio es un dominio operativo en el que la OTAN debe defenderse y disuadir de manera tan eficaz como en los dominios terrestres o marítimos.

Estrategias nacionales de defensa, como la Estrategia Nacional de Ciberseguridad de la República Checa y marcos normativos similares en otras

naciones, formalizaron la diferenciación teórica y operativa entre la ciberseguridad general —orientada a la protección global de la confidencialidad, integridad y disponibilidad de los datos— y la ciberdefensa, concebida como un componente autónomo y específico de la defensa nacional encargado de prevenir, mitigar y neutralizar ciberataques que amenacen la supervivencia del Estado o el funcionamiento de sus infraestructuras críticas.

La dinámica del quinto dominio altera la naturaleza de la confrontación estratégica debido a varias características estructurales distintivas:

- Conflicto persistente y no declarado: El ciberespacio opera en un estado de confrontación continua por debajo del umbral del conflicto armado formal, eliminando la distinción binaria tradicional entre paz y guerra.
- La infraestructura civil como campo de batalla primario: A diferencia de la guerra convencional, donde los activos civiles sufren daños colaterales, en la ciberdefensa las redes privadas de energía, agua, telecomunicaciones, transporte y cadenas de suministro constituyen el objetivo estratégico directo y primario para desestabilizar la capacidad estatal.
- Posicionamiento previo y capacidad de primer ataque no cinético: Los adversarios infiltran e instalan accesos persistentes en redes críticas durante tiempos de paz, permitiendo la activación de capacidades destructivas o disruptivas en el momento estratégico que el atacante elija.
- Ambigüedad en la atribución y fricción jurídica: La capacidad de ocultar el origen de un ataque mediante servidores intermediarios, proxies criminales y herramientas de acceso abierto retrasa la toma de decisiones estatales y debilita los mecanismos de respuesta defensiva o retaliación bajo el derecho internacional.

El análisis bibliométrico de la literatura científica publicada entre los años 2000 y 2025 evidencia la rápida institucionalización académica de la ciberdefensa y la soberanía estatal como subdisciplinas transversales a las

ciencias de la computación, las relaciones internacionales y los estudios de defensa. La producción científica acumulada refleja una tasa de crecimiento promedio anual del 41,91%, lo que señala un marcado incremento en el interés académico y normativo, impulsado por eventos catalizadores globales como los ciberataques masivos a Estonia en 2007, las revelaciones de vigilancia global, las ciberoperaciones en el conflicto ruso-ucraniano y la acelerada adopción de tecnologías de inteligencia artificial (Delgado y Romero, 2025).

Utilizando la metodología estandarizada del protocolo PRISMA para revisiones sistemáticas y análisis cientométricos, el corpus académico se estructura conceptualmente en torno a cuatro grandes clústeres o corrientes analíticas:

1. Gobernanza y Arquitectura Institucional: Estudios enfocados en la convergencia de marcos de trabajo como TOGAF (arquitectura de sistemas), DoDAF (operaciones militares) y COBIT (gobernanza y gestión de riesgos) para diseñar arquitecturas de ciberdefensa estatal adaptables.
2. Capacidades Operativas y Aumento Tecnológico: Investigaciones sobre detección avanzada de amenazas, *threat hunting*, inteligencia de ciberamenazas (*Cyber Threat Intelligence*), aprendizaje automático (*machine learning*) y modelos de engaño cognitivo para contrarrestar ciberataques asimétricos.
3. Derecho Internacional y Normas de Comportamiento Estatal: Literatura volcada en el análisis del *jus ad bellum* y *jus in bello* aplicados al ciberespacio, articulada predominantemente en torno a los trabajos del Centro de Excelencia de Ciberdefensa Cooperativa de la OTAN (NATO CCDCOE) y el Manual de Tallinn.
4. Soberanía Digital y Geopolítica de la Información: Análisis sobre los modelos conceptuales de control estatal sobre las capas física, lógica y de contenido de la red, confrontando las visiones multipartitas occidentales

frente a las doctrinas de soberanía ciberespacial estatales (véase la tabla 4).

Tabla 4: Indicador bibliométrico y enfoque metodológico

Indicador Bibliométrico / Dominio	Tasa de Crecimiento / Prevalencia	Principales Temas Motores y Conceptos Clave	Enfoque Metodológico Dominante
Producción Académica Global	~41.91% Anual (2017–2023/25)	*Machine Learning*, Ciberseguridad, *Threat Hunting*, *Threat Intelligence*	Cuantitativo y Modelado Computacional
Marcos de Arquitectura de Defensa	73% Dominio de COBIT / 67% Sector Público	TOGAF (Arquitectura), DoDAF (Operaciones), COBIT (Gobernanza y Riesgo)	Cualitativo y Revisiones Sistemáticas (53%)
Gobernanza y Soberanía Ciberespacial	Expansión en la literatura de RI desde 2016	Profundidad Estratégica Cibernética, Resiliencia Nacional, Soberanía Digital	Análisis Documental y Legal-Normativo
Derecho Internacional e Infraestructura	Alta tasa de citación en el Manual de Tallinn 2.0	Principio de Soberanía, Diligencia Debida, Infraestructuras Críticas, *Lex Lata*	Doctrinario, Jurídico y Teórico

La evolución de la literatura denota un desplazamiento analítico fundamental: el debate académico ha dejado de tratar la seguridad cibernética como un mero problema técnico o de ingeniería de software para reinterpretarlo como un problema de gobernanza del riesgo sistémico, arquitectura institucional y proyección del poder nacional. Las investigaciones enfatizan que la efectividad de la ciberdefensa estatal no depende de la instalación de controles perimetrales tradicionales, sino de la integración

estratégica de marcos normativos, operacionales y tecnológicos estructurados bajo un mando unificado.

Taxonomía conceptual de la soberanía en la era digital

Uno de los principales aportes de la literatura cientométrica sobre la materia es la delimitación analítica entre términos que con frecuencia se emplean de forma imprecisa o intercambiable en el discurso político: soberanía de datos, soberanía digital, soberanía tecnológica y cibersoberanía. La falta de precisión terminológica genera fricciones operativas y doctrinales en el diseño de políticas públicas de defensa nacional, por lo que la bibliografía especializada ha establecido fronteras teóricas claras entre estos ámbitos.

La soberanía cibernética y tecnológica se articula en un esquema escalar que abarca desde la microgestión de activos de información hasta la macrogestión de la autoridad geopolítica del Estado. En la cúspide de este esquema se sitúa la Soberanía Tecnológica, orientada a la independencia industrial, el control de las cadenas de suministro de hardware crítico y el desarrollo de componentes esenciales como semiconductores e infraestructura pesada de telecomunicaciones. En un nivel intermedio opera la Soberanía Digital, la cual busca garantizar la interoperabilidad, el desarrollo de capacidades tecnológicas locales y la autodeterminación normativa sobre el entorno lógico y de servicios informáticos sin caer en el aislamiento (Santos, 2020).

Paralelamente, la ciberautoridad ejerce la autoridad jurisdiccional y la defensa militar sobre las fronteras lógicas y el espacio de información nacional, asegurando la protección de las infraestructuras críticas estatales contra intrusiones extranjeras. Finalmente, en la base operativa se encuentra la Soberanía de Datos, enfocada en la gobernanza, propiedad y custodia de los

activos de información frente a accesos transfronterizos no autorizados (véase la tabla 5).

Tabla 5: Taxonomía conceptual de la soberanía de datos

Concepto de Soberanía	Ámbito de Aplicación Primario	Agentes e Instituciones Clave	Prioridades Estratégicas y Medidas Operativas
Soberanía de Datos	Capa de datos y activos de información a nivel micro y organizacional.	Individuos, empresas, entidades públicas, custodiadores de datos.	Propiedad legítima, localización de datos, protección contra accesos transfronterizos no autorizados (p. ej. *US CLOUD Act*).
Soberanía Digital	Capa lógica y de servicios a nivel meso; interoperabilidad y competencias tecnológicas.	Estados, bloques regionales (p. ej. Unión Europea), industrias de software.	Autodeterminación digital, capacidad de moldear la transformación digital sin aislamiento, marcos normativos, reducción de dependencias.
Soberanía Tecnológica	Capa física e industrial a nivel macro; independencia de hardware y cadenas de suministro.	Conglomerados industriales, ministerios de defensa/comercio, potencias mundiales.	Producción nacional de semiconductores, autonomía de cadenas de valor tecnológicas, regulación de transferencias internacionales.
Ciber Soberanía	Capa de autoridad política y jurisdicción estatal sobre el espacio virtual y la información.	Gobiernos estatales, fuerzas armadas, agencias de inteligencia cibernética.	Control del espacio de información nacional, protección de infraestructura crítica, defensa cibernética de fronteras lógicas.

A nivel internacional, la literatura documenta una bifurcación geopolítica profunda respecto al modelo dominante de cibersoberanía. Las naciones occidentales priorizan un modelo de gobernanza abierto y de múltiples partes interesadas (*multistakeholder*), que defiende la neutralidad de la red, el flujo libre de información y la vigencia del derecho internacional de los derechos humanos en el entorno digital. Bajo este enfoque, las medidas defensivas se orientan a proteger la resiliencia de la infraestructura sin menoscabar las libertades civiles.

Por otro lado, potencias como China y Rusia impulsan la doctrina de la cibersoberanía de corte westfaliano o estatal centralizado. Este paradigma afirma el derecho inalienable del Estado a ejercer control absoluto sobre el segmento nacional del ciberespacio, legitimando la censura, el filtrado de contenidos transfronterizos, el control centralizado de los puntos de intercambio de tráfico (*Internet Exchange Points*) y proyectos de "Internet Soberana" destinados a desconectar el espacio digital nacional de la red global en situaciones de crisis estratégica.

Mando unificado y desafíos jurídico-institucionales

La institucionalización de la ciberdefensa como función esencial de la soberanía estatal exige estructurar arquitecturas organizacionales complejas que superen las rigideces burocráticas tradicionales. Las investigaciones enfocadas en la implementación de sistemas de información de defensa demuestran que la integración coherente de los marcos COBIT (gobernanza), DoDAF (operaciones militares) y TOGAF (arquitectura empresarial de TI) permite establecer capacidades defensivas adaptativas y estratificadas. Sin embargo, la efectividad operativa a menudo tropieza con tensiones legales e institucionales

internas.

Un hallazgo recurrente en los estudios del sector público es la presencia de una antinomia legal o de una fragmentación normativa interinstitucional. En múltiples jurisdicciones estatales, las competencias sobre el entorno cibernético se encuentran superpuestas y repartidas entre tres esferas administrativas sin una demarcación clara de umbrales operativos: el dominio administrativo-civil (agencias de ciberseguridad o CERT/CSIRT nacionales encargados de la resiliencia regulatoria), el dominio de seguridad interior (fuerzas policiales dedicadas a la investigación de ciberdelitos bajo procedimiento penal) y el dominio de la defensa nacional (fuerzas armadas e inteligencia militar capacitadas para ejecutar operaciones defensivas y ofensivas de intensidad estratégica).

La falta de mecanismos legislativos de escalamiento oportuno crea situaciones de parálisis en la toma de decisiones durante incidentes en zonas grises. Para resolver este vacío, la literatura estratégica postula la urgente reconstrucción normativa mediante leyes de ciberseguridad y ciberdefensa redactadas como *lex specialis*, institucionalizando un Sistema de Mando Unificado que regule con precisión los umbrales de transferencia de control operativo de las autoridades civiles a la estructura militar en situaciones de crisis nacional. Un ejemplo notable de estructuración operativa se observa en casos como el de la Inteligencia Militar en la República Checa o la asignación de unidades especializadas de cibercomando bajo control ministerial, garantizando la continuidad defensiva y la capacidad de detección temprana de anomalías lógicas en las redes del Estado.

En el plano internacional, la delimitación de la soberanía estatal en el ciberespacio enfrenta complejos vacíos normativos. El corpus doctrinal más consolidado de la *lex lata* sobre la materia es el Manual de Tallin (versiones 1.0 y 2.0), elaborado por expertos independientes bajo el patrocinio del NATO

CCDCOE. El *M*anual de Tallinn 2.0 reafirma que los principios generales del derecho internacional —incluidos la Carta de las Naciones Unidas, la soberanía territorial, la responsabilidad estatal, la prohibición de la intervención y el Derecho Internacional Humanitario (*jus in bello*)— se aplican íntegramente a las ciberoperaciones efectuadas por o contra los Estados (Castilloo et al., 2026).

El principio de soberanía opera en una doble dimensión: como derecho del Estado a ejercer control exclusivo sobre las infraestructuras ciberespaciales en su territorio, y como la obligación de diligencia debida (*due diligence*), según la cual un Estado no debe permitir a sabiendas que su territorio cibernético sea utilizado para realizar actos ilícitos que causen daño a terceros Estados. La respuesta legal estatal ante una ciberoperación adversa escala progresivamente en función de la severidad y los efectos del ataque:

- Infracciones menores o espionaje no destructivo: Permiten el despliegue de actos de retorsión, entendidos como medidas impopulares o inamistosas pero intrínsecamente legales bajo el derecho internacional, como la imposición de sanciones diplomáticas o restricciones comerciales.
- Violaciones del principio de soberanía o incumplimiento de diligencia debida: legitiman el uso de contramedidas no cinéticas por parte del Estado afectado, suspendiendo temporalmente el cumplimiento de obligaciones internacionales respecto al Estado infractor para forzar el cese de la actividad ilícita.
- Situaciones de peligro grave e inminente para intereses esenciales: Habilitan el recurso al estado de necesidad como causa de exclusión de la ilicitud para neutralizar la amenaza a infraestructuras críticas.
- Ciberoperaciones con efectos destructivos equiparables a un ataque armado: Superan el umbral del uso de la fuerza del Artículo 2(4) de la Carta de la ONU y activan el derecho inalienable a la legítima defensa individual o colectiva consagrado en el Artículo 51, permitiendo respuestas

proporcionales que pueden incluir medios cinéticos o militares (véase la tabla 6).

Tabla 6: la soberanía territorial y la responsabilidad estatal

Mecanismo de Respuesta Legal	Umbral de Intrusión / Gravedad de la Amenaza	Base Normativa en el Derecho Internacional	Medidas Operativas Permitidas
Retorsión	Ciberespionaje, interferencia de baja intensidad, intrusiones no destructivas.	Actos inamistosos pero lícitos bajo la soberanía estatal.	Sanciones diplomáticas, restricciones tecnológicas, expulsión de personal.
Contramedidas	Violación directa de la soberanía o incumplimiento de la diligencia debida.	Artículos sobre la responsabilidad del Estado por hechos internacionalmente ilícitos.	Interrupción temporal de servicios, ciberoperaciones defensivas proporcionales.
Estado de Necesidad	Peligro grave e inminente para la supervivencia de infraestructuras esenciales.	Causa de exclusión de ilicitud en el Derecho Internacional Consuetudinario.	Intervención o neutralización de activos de red no estatales o transfronterizos.
Legítima Defensa	Ciberataque con efectos destructivos o letales equivalentes a un ataque armado.	Artículo 51 de la Carta de las Naciones Unidas (*Jus ad Bellum*).	Uso de la fuerza militar de respuesta, operaciones cinéticas y ciberofensivas.

Vectores emergentes: inteligencia artificial, autonomía y conflictos de nueva generación

La velocidad a la que se desarrollan los ciberataques modernos ha superado la capacidad de respuesta humana en tiempo real, impulsando la integración masiva de la Inteligencia Artificial (IA) y de las Capacidades Cibernéticas Autónomas (ACC, por sus siglas en inglés) en la ciberdefensa. El estudio bibliométrico de temas motores señala una fuerte convergencia entre la ciberdefensa militar y el procesamiento masivo de datos mediante modelos de IA para la caza proactiva de amenazas (*threat hunting*), el desarrollo de modelos de engaño cognitivo (CDM) y la automatización de respuestas ante incidentes.

Las capacidades cibernéticas autónomas son sistemas informáticos capaces de evaluar el entorno lógico, detectar intrusiones y desplegar medidas de contención o contraataque cibernético (*hack-back*) sin intervención humana directa. Si bien estas herramientas reducen drásticamente el tiempo de mitigación de ataques, su empleo plantea dilemas jurídicos sustanciales bajo el marco del derecho de la responsabilidad estatal y la legítima defensa (Rojek et al., 2026). Un sistema autónomo defensivo no cuenta actualmente con la capacidad analítica para verificar con certeza la identidad del atacante ni para determinar si la operación atribuida responde a una acción dirigida por un Estado o a la actividad de actores criminales no estatales.

Además, la ejecución automática de contramedidas o respuestas intrusivas ante falsos positivos puede ser interpretada por un Estado adversario como un primer ataque cibernético de carácter estatal, detonando una espiral de escalada que alcance el conflicto cinético convencional. Por ello, la doctrina legal advierte que el uso defensivo de ACC en contextos de contramedidas o legítima defensa presenta graves limitaciones, siendo

utilizable de manera prudente casi de forma exclusiva para actos de retorsión o medidas defensivas puramente internas en la red propia.

La transición hacia la era de la IA ha redefinido el concepto de soberanía digital, transformando el acceso a modelos avanzados de lenguaje y arquitecturas de aprendizaje profundo en un factor directo de poder geopolítico. Eventos recientes en el ámbito estratégico demuestran la disposición de las grandes potencias para aplicar controles de exportación severos y restricciones operativas sobre modelos fundacionales de IA dirigidos a usuarios extranjeros. La imposición de directivas de seguridad nacional sobre desarrolladores de IA de frontera —exigiendo la revocación del acceso a modelos avanzados para usuarios e investigadores internacionales bajo amenaza de sanciones penales y civiles— refleja cómo la tecnología de inteligencia artificial ha sido redefinida como un insumo estratégico crítico (Tanase, 2026). Los Estados reconocen que la concentración del desarrollo de la IA en unas pocas empresas privadas transnacionales genera una vulnerabilidad sistémica de soberanía, forzando a los gobiernos a implementar iniciativas de "IA Soberana" para garantizar la autonomía de sus sistemas cibernéticos de defensa e inteligencia.

La invasión a gran escala de Ucrania iniciada en 2022 constituye el primer estudio de caso empírico de guerra híbrida total en la que el quinto dominio se integró plenamente con las operaciones terrestres, marítimas, aéreas y espaciales. A lo largo del conflicto, la combinación de ciberataques destructivos contra redes eléctricas, infraestructura satelital de telecomunicaciones y portales gubernamentales con bombardeos cinéticos convencionales confirmó que el ciberespacio es un componente estructural e indivisible de la guerra moderna.

La resiliencia ucraniana demostró que la soberanía en el quinto dominio no descansa exclusivamente en los activos de defensa militar, sino en la

capacidad de articular una respuesta colectiva ágil mediante alianzas con empresas multinacionales de la nube, la migración remota de infraestructuras críticas de datos a jurisdicciones aliadas y la distribución de inteligencia sobre amenazas en tiempo real.

El análisis bibliométrico y doctrinal ratifica que la ciberdefensa ha trascendido la esfera de la tecnología operativa para consolidarse como un pilar indivisible de la soberanía del Estado en el siglo XXI. La desterritorialización de las amenazas cibernéticas no invalida la noción de soberanía, sino que obliga a redefinirla en términos de resiliencia sistémica, capacidad de atribución, autonomía tecnológica y autoridad normativa sobre la infraestructura digital nacional.

Para responder de forma eficaz a los retos del quinto dominio, la literatura especializada sugiere que los Estados adopten una serie de medidas prioritarias en su arquitectura de seguridad nacional:

1. Reforma Legal e Institucional Unificada: Promulgar marcos regulatorios especiales de ciberseguridad y ciberdefensa (*lex specialis*) que eliminen la fragmentación entre agencias civiles, policiales y militares, estableciendo un Mando Unificado con umbrales normativos claros para la gestión de crisis.
2. Adopción de Arquitecturas Integradas de Gobernanza: Estructurar los sistemas defensivos del Estado combinando COBIT para la gestión de riesgos y cumplimiento, DoDAF para la articulación de operaciones de comando militar y TOGAF para el diseño de arquitecturas tecnológicas adaptativas.
3. Fomento de la Autonomía Tecnológica e IA Soberana: Implementar políticas de soberanía industrial que aseguren el control estatal sobre componentes críticos de hardware, algoritmos de detección activa de amenazas y modelos propios de inteligencia artificial para ciberdefensa,

mitigando la dependencia respecto a proveedores extranjeros.

4. **Participación Activa en la Diplomacia Ciberespacial:** Impulsar la codificación de normas internacionales de comportamiento estatal responsable, la fijación de estándares de diligencia debida y el fortalecimiento de alianzas estratégicas para el intercambio proactivo de inteligencia sobre ciberamenazas.

Capítulo 3
Estándares criptográficos poscuánticos y marco normativo global

El avance vertiginoso en el desarrollo de ordenadores cuánticos con capacidad criptográficamente relevante (CRQC, por sus siglas en inglés) representa una amenaza estructural para la infraestructura global de seguridad de la información (Agencia de Seguridad Nacional, 2022). Los esquemas de criptografía asimétrica tradicionales, basados en la dificultad computacional de la factorización de enteros (RSA) y el logaritmo discreto sobre campos finitos o curvas elípticas (DH, ECDH, ECDSA), son vulnerables al algoritmo cuántico de Shor, el cual puede resolver estos problemas matemáticos en tiempo polinómico.

Como respuesta a este desafío, los organismos internacionales de estandarización, las agencias gubernamentales de defensa y los entes reguladores del sector privado han establecido una arquitectura unificada de criptografía poscuántica (PQC, por sus siglas en inglés). Esta transición no constituye una simple actualización de software, sino una reingeniería profunda del ecosistema criptográfico global (Cherkaoui et al., 2024).

Finalización de estándares criptográficos por el NIST

El 13 de agosto de 2024, el Instituto Nacional de Estándares y Tecnología de los Estados Unidos (NIST) marcó el hito operativo más relevante en la historia de la criptografía moderna al publicar las versiones finales de sus tres primeros

estándares *Federal Information Processing Standards* (FIPS) poscuánticos. Estos estándares culminaron un proceso de evaluación pública internacional de ocho años que inició en diciembre de 2016 con 69 propuestas iniciales.

FIPS 203 especifica el mecanismo de encapsulamiento de claves basado en módulos de reticulados (ML-KEM), estandarizado a partir del algoritmo CRYSTALS-Kyber. ML-KEM es el estándar primario recomendado por el NIST para el establecimiento de claves simétricas compartidas a través de canales no confiables sin transmitir la clave en sí misma. Su seguridad matemática se basa en la dureza del problema algebraico *Module Learning with Errors* (MLWE) sobre reticulados algebraicos. El estándar define tres juegos de parámetros alineados con los niveles de seguridad del NIST: ML-KEM-512 (Nivel 1, equivalente a la resistencia de clave de AES-128), ML-KEM-768 (Nivel 3, equivalente a AES-192 y recomendado como configuración predeterminada) y ML-KEM-1024 (Nivel 5, equivalente a AES-256). Adicionalmente, el NIST seleccionó en marzo de 2025 el algoritmo HQC (*Hamming Quasi-Cyclic*), basado en códigos correctores de errores, como un mecanismo KEM de respaldo fundado en supuestos matemáticos no reticulares (NIST, 2024).

FIPS 204 establece el Algoritmo de Firma Digital Basado en Módulos de Reticulados (ML-DSA), derivado de la propuesta CRYSTALS-Dilithium. Designado como el estándar principal de firma digital para uso general en infraestructuras de clave pública (PKI), firmas de certificados, autenticación de protocolos y firma de documentos, ML-DSA se estructura en los parámetros ML-DSA-44 (Nivel 2), ML-DSA-65 (Nivel 3) y ML-DSA-87 (Nivel 5).

FIPS 205 define el Algoritmo de Firma Digital Sin Estado Basado en Funciones Hash (SLH-DSA), estandarizado a partir de SPHINCS+. A diferencia de los esquemas basados en reticulados, SLH-DSA cimenta su seguridad exclusivamente en la resistencia a colisiones y preimágenes de funciones hash criptográficas tradicionales (SHA-256, SHA-512, SHAKE128 y SHAKE256).

Aunque SLH-DSA produce firmas significativamente más grandes y requiere un mayor tiempo de procesamiento que ML-DSA, el NIST aprobó su estandarización paralela para garantizar la diversidad algorítmica. En el supuesto teórico de que se descubra una debilidad algebraica inesperada en el problema MLWE, las implementaciones de ML-DSA se verían comprometidas, mientras que las redes desplegadas con SLH-DSA mantendrían su integridad.

De forma complementaria, el estándar FIPS 206 (en fase de borrador) estandariza el algoritmo FN-DSA (derivado de FALCON), el cual emplea problemas de reticulados NTRU y transformadas rápidas de Fourier (FFT). FN-DSA genera firmas altamente compactas en comparación con ML-DSA, pero requiere implementaciones complejas con operaciones de punto flotante de alta precisión y muestreo gaussiano propenso a ataques de canal lateral. Para aplicaciones de firma de firmware y código, el marco del NIST se complementa con la publicación especial SP 800-208, que rige los esquemas basados en hash con estado LMS (*Leighton-Micali Signature*) y XMSS (*Extended Merkle Signature Scheme*).

Tabla 7: Estándares criptográficos

Estándar	Algoritmo	Función Criptográfica	Base Matemática	Parámetros Principales	Tamaño de Clave Pública / Firma	Nivel de Seguridad NIST
FIPS 203	ML-KEM	Encapsulamiento de Claves (KEM)	Reticulados (MLWE)	ML-KEM-512 ML-KEM-768 ML-KEM-1024	Clave: 800 bytes (512) Cifrado: 768 bytes	Nivel 1 (128-bit) Nivel 3 (192-bit) Nivel 5 (256-bit)
FIPS 204	ML-DSA	Firma Digital	Reticulados	ML-DSA-44	Clave: 1,952	Nivel 2 (128-bit)

		General	(MLWE)	ML-DSA-65 ML-DSA-87	bytes (65) Firma: 3,309 bytes	Nivel 3 (192-bit) Nivel 5 (256-bit)
FIPS 205	SLH-DSA	Firma Digital Conservadora	Funciones Hash (Sin Estado)	Parámetros con SHA-2 y SHAKE	Firma: > 7,000 bytes (múltiples KB)	Niveles 1, 3 y 5
FIPS 206 (Draft)	FN-DSA	Firma Digital Compacta	Reticulados NTRU	FN-DSA-512 FN-DSA-1024	Firma: ~666 bytes (512) Clave: ~897 bytes	Nivel 1 (128-bit) Nivel 5 (256-bit)
SP 800-208	LMS / XMSS	Firma de Código y Firmware	Funciones Hash (Con Estado)	Árboles Merkle específicos	Firma variable según altura del árbol	Equivalente a 128/256 bits

Marcos regulatorios y cronogramas de migración obligatoria

El proceso de estandarización técnica ha desencadenado una respuesta regulatoria global donde las directivas gubernamentales han pasado de ser meras recomendaciones teóricas a mandatos ejecutivos con fechas de cumplimiento estrictas.

Mandatos en Estados Unidos: CNSA 2.0 y directivas federales

La Agencia de Seguridad Nacional de los Estados Unidos (NSA) promulgó la *Commercial National Security Algorithm Suite 2.0* (CNSA 2.0), estableciendo el calendario de migración más agresivo y detallado del mundo para los Sistemas de Seguridad Nacional (NSS). CNSA 2.0 exige la adopción exclusiva de los conjuntos de parámetros de mayor fortaleza: ML-KEM-1024 para intercambio de claves, ML-DSA-87 o SLH-DSA para firmas digitales, LMS/XMSS para firma de firmware, AES-256 para cifrado simétrico y SHA-384/512 para hashing (National Security Agency, 2024).

El marco temporal de la NSA clasifica el cumplimiento por categorías tecnológicas con dos hitos operativos: una fecha para soportar y preferir algoritmos poscuánticos, y una fecha límite de uso exclusivo. La firma de software y firmware lidera el calendario debido a su rol como raíz de confianza del sistema, exigiendo la preferencia en 2025 y el uso exclusivo obligatorio para 2030.

El equipamiento de red tradicional (VPN, routers) debe soportar los algoritmos en 2026 y ser exclusivo en 2030. Los navegadores web, servidores, servicios en la nube y sistemas operativos deben alcanzar el uso exclusivo poscuántico para 2033. Además, a partir del 1 de enero de 2027, todas las nuevas adquisiciones de equipos para NSS deben ser conformes con CNSA 2.0 por defecto (véase la tabla 8).

Tabla 8: Algoritmo especificado por categoría tecnológica

Categoría Tecnológica	Soporte y Configuración Preferida	Uso Exclusivo Obligatorio	Algoritmo Especificado (CNSA 2.0)
Firma de Software y Firmware	2025	2030	LMS / XMSS, ML-DSA-87
Equipamiento de Red Tradicional (VPNs, Routers)	2026	2030	ML-KEM-1024, ML-DSA-87
Navegadores Web, Servidores y Nube	2025	2033	ML-KEM-1024, ML-DSA-87
Sistemas Operativos	2027	2033	ML-KEM-1024, ML-DSA-87
Adquisiciones Nuevas de Equipamiento NSS	Cumplimiento por defecto: 1 enero 2027	2033 (Límite General)	Suite Completa CNSA 2.0

A nivel del gobierno federal civil, el Memorando de Seguridad Nacional (NSM-10) y la Ley de Preparación Cibersegura ante la Computación Cuántica coordinan la transición. La Oficina de Gerencia y Presupuesto (OMB), mediante la circular M-23-02, obliga a las agencias federales a presentar inventarios anuales de sus activos criptográficos vulnerables prioritarios. Por su parte, el documento NIST IR 8547 fija el calendario de depreciación algorítmica para el sector civil, declarando obsoletos los algoritmos RSA-2048, ECDH y ECDSA para nuevos sistemas a partir de 2030, y prohibiendo totalmente su uso en cualquier infraestructura federal para 2035.

Directivas en la Unión Europea y el Reino Unido

La Comisión Europea emitió la Recomendación (UE) C(2024) 2393, impulsando una Hoja de Ruta de Implementación Coordinada para la Transición Criptográfica Poscuántica. Esta recomendación insta a los Estados Miembros a sincronizar sus planes nacionales de migración para evitar la fragmentación del Mercado Único Digital y garantizar la interoperabilidad transfronteriza. La PQC se está integrando directamente en la regulación operativa comunitaria:

El Reglamento de Resiliencia Operativa Digital del Sector Financiero (DORA) exige a las entidades financieras que evalúen de forma continua los riesgos de dependencia algorítmica y diseñen planes de reemplazo criptográfico dentro de sus marcos de gestión de riesgos TIC. Paralelamente, la Directiva NIS2 considera la actualización de los mecanismos de cifrado de red como un requisito de higiene de ciberseguridad obligatorio para los operadores de servicios esenciales y proveedores digitales. Por su parte, el *Cyber Resilience Act* (CRA), aplicable a partir de 2027, impone que los productos de hardware y software comercializados en la UE incorporen identidades criptográficas únicas y soporte para actualizaciones que permitan la agilidad poscuántica.

En la esfera de los organismos nacionales de ciberseguridad, la agencia francesa ANSSI estableció una postura pionera al prescribir el uso obligatorio de arquitecturas híbridas (combinando un algoritmo clásico probado con uno poscuántico) como requisito previo indispensable para obtener visados de seguridad en productos criptográficos. En el Reino Unido, el *National Cyber Security Centre* (NCSC) estructuró un plan sectorial articulado en tres fases: Fase 1 de descubrimiento y planificación (hasta 2028), Fase 2 de actualización en sistemas de alta prioridad (2028–2031) y Fase 3 de completitud de la migración poscuántica para las infraestructuras críticas (2031–2035).

Normalización internacional y esfuerzos multilaterales

El Subcomité ISO/IEC JTC 1/SC 27 (Grupo de Trabajo WG 2 y JWG 7) lidera la armonización de estándares internacionales. La norma ISO/IEC 14888-4:2024 fue actualizada formalmente para incluir mecanismos de firma basados en hash con estado (XMSS y LMS). Paralelamente, el estándar ISO/IEC 23837 establece los criterios de evaluación de seguridad y métodos de prueba bajo el marco de Criterios Comunes (ISO/IEC 15408) para dispositivos de Distribución de Claves Cuánticas (QKD), reconociendo a QKD como una solución física complementaria en entornos restringidos de red óptica.

En Asia, el Instituto de Inteligencia Nacional de Corea del Sur gestiona el programa KpqC, el cual seleccionó algoritmos nacionales derivados que convergerán en un plan de adopción completa para 2035. China ha avanzado mediante la iniciativa ICCS para evaluar algoritmos poscuánticos propietarios que se promulgarán como estándares nacionales obligatorios dentro de sus fronteras.

Impacto operativo e integración sectorial

La transición hacia los estándares poscuánticos impone demandas técnicas dispares según la naturaleza operativa de cada industria. El sector de las telecomunicaciones afronta retos severos de volumen y latencia. El *Post-Quantum Telco Network Task Force* de la GSMA —que agrupa a más de 60 operadores globales, reguladores y fabricantes— ha publicado la serie de directivas PQ.01 a PQ.07. Estas guías establecen metodologías específicas para integrar PQC en enlaces MACsec, túneles IPsec, TLS en planos de señalización 5G y el perfil IoT SAFE.

Las especificaciones de la GSMA imponen identidades de dispositivos arraigadas en hardware (raíz de confianza) para la autenticación de millones de conexiones M2M y redes no terrestres. La migración del empaquetado de red

en 5G requiere actualizar los mecanismos de red cifrados sin degradar los tiempos de procesamiento en milisegundos requeridos por el *network slicing*.

Sector financiero y medios de pago

En las instituciones bancarias y pasarelas de pago, la urgencia regulatoria está impulsada por el cumplimiento del estándar PCI DSS v4.0. Específicamente, el Requisito 12.3.3 de PCI DSS obliga a las organizaciones a documentar, inventariar y revisar anualmente todos los conjuntos de cifrado, claves y protocolos en uso, lo cual actúa como el marco legal que exige la creación implícita de un inventario de activos criptográficos.

El impacto más crítico en la infraestructura financiera recae sobre los Módulos de Seguridad de Hardware (HSM). Los HSM bancarios de alta velocidad (como las plataformas Thales payShield o Utimaco) que autentican transacciones con tarjeta, transferencias interbancarias y emiten certificados PKI deben sufrir un ciclo de reemplazo de plataforma. Los HSM tradicionales carecen de la memoria fija en chip (*SRAM/Flash*) y de la capacidad de procesamiento vectorial necesarias para ejecutar las operaciones de reticulados de ML-KEM-1024 o ML-DSA-87. Como resultado, el mercado de HSM está migrando aceleradamente hacia dispositivos certificados bajo FIPS 140-3 Nivel 3 y arquitecturas HSM basadas en la nube (*HSM-as-a-Service*), capaces de reprogramar sus microcódigos mediante capas de abstracción como la API PKCS#11 adaptada a PQC.

Vectores de Amenaza: Harvest Now, Decrypt Later (HNDL) frente a Targeted Network Forge / Tamper Later (TNFL)

Las estrategias de priorización técnica se basan en dos modelos de ataque diferenciados por su dinámica temporal:

El modelo *Harvest Now, Decrypt Later* (HNDL) describe la práctica de adversarios estatales y grupos cibernéticos que interceptan y almacenan

actualmente grandes volúmenes de tráfico cifrado confidencial que transita por canales públicos. Aunque no pueden descifrarlo hoy, lo conservarán hasta la llegada de un ordenador cuántico operativo. Los datos con una vida útil de confidencialidad de 10 a 30 años (secretos comerciales, datos médicos, inteligencia militar, registros financieros) están comprometidos en este instante. HNDL impulsa la sustitución inmediata de los esquemas de intercambio de claves (RSA/ECDH) por ML-KEM.

Por otro lado, la amenaza *Targeted Network Forge / Tamper Later* (TNFL) afecta directamente la integridad y autenticidad de los activos digitales. Si un atacante cuántico puede falsificar firmas digitales clásicas, podría emitir actualizaciones de firmware maliciosas firmadas válidamente, alterar certificados raíz de PKI o falsificar identidades de software. Esto explica por qué las agencias de defensa exigen la adopción inmediata de PQC en la firma de código y firmware: las raíces de confianza implantadas hoy deben seguir garantizando la integridad de los dispositivos durante todo su ciclo de vida útil de más de una década.

Cripto-agilidad y gestión de inventarios (CBOM)

La lección operativa fundamental derivada del proceso de estandarización es que la sustitución directa y estática de un algoritmo por otro es inviable a escala empresarial. La agilidad criptográfica —la capacidad técnica de un sistema para alternar entre primitivas, longitudes de clave y algoritmos mediante cambios de configuración sin alterar el código fuente subyacente— se posiciona como el paradigma arquitectónico indispensable.

El Inventario de activos criptográficos (CBOM) y el modelo de madurez

Cualquier programa de transición estructural debe iniciar con la

construcción automatizada de un Inventario de Activos Criptográficos (*Cryptographic Bill of Materials* o CBOM). Un CBOM formalizado (estructurado en formatos estándar como CycloneDX CBOM) debe catalogar de forma exhaustiva los atributos de cada componente del entorno corporativo. Específicamente, debe registrar la identificación del criptosistema (algoritmo, tamaño de clave, modo de operación y versión de la librería implementada), el contexto de protocolo y certificado (uso en protocolos TLS, SSH o IPsec y su cadena de Autoridades de Certificación), la ubicación y propiedad del activo (dueño del sistema, dependencias de código de terceros y canal de actualización) y la clasificación de riesgo en función de la sensibilidad de los datos y el nivel de exposición a ataques HNDL o TNFL.

La ejecución técnica del ciclo de migración se despliega a través de un modelo de madurez estructurado secuencialmente. Inicialmente, la fase de descubrimiento y CBOM cataloga la totalidad de las primitivas criptográficas existentes. A continuación, la evaluación de riesgo califica la exposición de los sistemas ponderando la longevidad de los datos frente a las ventanas temporales regulatorias. Posteriormente, la fase de diseño e hibridación define las arquitecturas de abstracción necesarias y selecciona los algoritmos de reemplazo. Le sigue el despliegue agilizado, donde se prueban e integran las capas de software actualizables en entornos de producción. Finalmente, la fase de mantenimiento continuo monitorea la aparición de criptoanálisis sobre los nuevos estándares y gestiona la obsolescencia algorítmica de forma automatizada.

Implementación de esquemas híbridos y desafíos de rendimiento en red

Durante la fase intermedia de transición (prevista entre 2025 y 2030), la recomendación estándar del IETF, la ENISA y la ANSSI exige el despliegue de

soluciones híbridas. Un intercambio de claves híbrido combina un algoritmo clásico probado (como X25519 o ECDH P-256) con un mecanismo poscuántico (como ML-KEM-768) en una sola operación de protocolo. La clave simétrica resultante se deriva combinando matemáticamente ambos secretos mediante una función de derivación de claves (KDF). Este enfoque garantiza que la seguridad del sistema no se vea reducida si se descubriera un fallo no detectado en la nueva matemática de reticulados, manteniendo la protección clásica existente mientras se añade la resistencia cuántica.

No obstante, la integración de las primitivas FIPS 203 y FIPS 204 introduce severas restricciones de rendimiento e infraestructura que los arquitectos de sistemas deben gestionar. En primer lugar, se produce una inflación sustancial del tamaño de claves y firmas; mientras que una firma digital ECDSA P-256 ocupa únicamente 64 bytes, una firma ML-DSA-65 requiere 3.309 bytes (aproximadamente 50 veces mayor), y su clave pública asciende a 1.952 bytes.

Esta disparidad provoca problemas severos de fragmentación IP a nivel de red. El incremento en los tamaños de las firmas de certificados en las cadenas de autoridades de certificación (CA) causa que los mensajes ServerHello en apretones de manos TLS 1.3 excedan la unidad de transmisión máxima (MTU) de la red. Esto desencadena la fragmentación de paquetes TCP/UDP que los routers intermedios a menudo descartan por políticas de seguridad, resultando en fallos de conexión o latencias severas. Finalmente, en el entorno de la Internet de las Cosas (IoT), los microcontroladores con recursos estrictamente limitados en memoria RAM no pueden albergar las matrices requeridas para procesar las operaciones polinomiales de ML-KEM-1024 o almacenar los grandes buffers de firma de SLH-DSA, obligando al rediseño del hardware perimetral.

La publicación de los estándares FIPS 203, 204 y 205 por el NIST en agosto de 2024 clausura definitivamente la etapa de debate teórico sobre la

criptografía poscuántica e inaugura la era de cumplimiento obligatorio e ingeniería de despliegue. Las fechas de caducidad fijadas por la NSA en la directiva CNSA 2.0 y por el NIST en la publicación IR 8547 establecen un marco donde el uso de algoritmos asimétricos clásicos quedará erradicado de las infraestructuras clave entre 2030 y 2035.

Para mitigar los riesgos operativos, legales y de cumplimiento, las organizaciones corporativas y gubernamentales deben abordar la transición mediante una serie de acciones estratégicas continuas:

En primer lugar, es indispensable establecer una gobernanza centralizada de criptoagilidad mediante la creación de una oficina de programa poscuántico que asegure presupuesto plurianual y coordine las áreas de ciberseguridad, arquitectura de TI, cumplimiento normativo y gestión de la cadena de suministro.

En segundo lugar, se debe automatizar la generación e integración continua del CBOM desplegando herramientas de escaneo pasivo de red y análisis estático de código fuente para inventariar todas las primitivas criptográficas, catalogándolas bajo estándares interoperables que permitan identificar dependencias críticas.

En tercer lugar, la prioridad técnica inmediata debe centrarse en la firma de código e integridad de los sistemas, migrando los pipelines de CI/CD, las firmas de software y las raíces de confianza de firmware hacia estándares SP 800-208 (LMS/XMSS) o ML-DSA, asegurando que la infraestructura de actualización distribuida sea inmune a vectores TNFL.

En cuarto lugar, se deben desplegar de forma gradual protocolos híbridos en los canales de comunicación expuestos a la red pública (TLS 1.3, VPNs IPsec) combinando ECDH con ML-KEM-768/1024 para neutralizar la amenaza inminente de captura e interceptación de datos sensibles vulnerables a ataques HNDL.

Por último, resulta crítico auditar la cadena de suministro tecnológica exigiendo a los proveedores de software, equipos de red y servicios en la nube declaraciones juradas de cumplimiento y hojas de ruta de agilidad poscuántica, condicionando las licitaciones a la compatibilidad nativa con las especificaciones FIPS del NIST.

Capítulo 4

Migración empresarial a la criptografía postcuántica: análisis arquitectónico del marco de migración PQC cuántica aplicado

La inminente transición de la criptografía clásica a la poscuántica (PQC) representa la renovación estructural más fundamental de la arquitectura de confianza digital en la historia moderna de la computación. A medida que las capacidades de procesamiento cuántico avanzan hacia los ordenadores cuánticos criptográficamente relevantes (CRQC), la criptografía estándar de clave pública —específicamente familias de algoritmos ancladas en la factorización entera (RSA), logaritmos discretos (DSA) y logaritmos discretos de curva elíptica (ECDSA, ECDH)— quedará completamente insegura mediante la ejecución del algoritmo de Shor.

Para conceptualizar la urgencia temporal de esta exposición, los arquitectos de seguridad se basan en el Teorema de Mosca, que evalúa el riesgo estructural utilizando la desigualdad:

$$D + M > T$$

Donde:

- D representa la vida útil requerida de confidencialidad o la vida operativa de datos y sistemas sensibles.
- M representa la duración operativa necesaria para ejecutar una

migración poscuántica a nivel empresarial.

- T representa el tiempo transcurrido antes de que un adversario realice una CRQC.

Cuando la suma combinada de los requisitos de retención de datos (D) y la duración de ejecución de migración M () supera el tiempo restante hasta la realización de capacidades cuánticas (T), el sistema entra en un estado de precompromiso cuántico. En esta condición, los datos protegidos por primitivas clásicas son activamente vulnerables a la interceptación y a futuras descifraciones antes de que se pueda completar la migración operativa (véase la tabla 9).

Tabla 9: Vectores de amenaza distintos: Cosecha-Ahora-Descifrar-Después (HNDL) y Confiar-Ahora-Forja-Después (TNFL)

Dimensión de amenaza	Cosecha-Ahora-Descifrado-Después (HNDL)	Confianza-Ahora-Forja-Después (TNFL)
Objetivo principal	Establecimiento e intercambio de claves (por ejemplo, ECDH, transporte de claves RSA).	Firmas digitales e infraestructura de clave pública (por ejemplo, RSA, ECDSA).
Objetivo de seguridad	Confidencialidad de cargas útiles sensibles de larga duración.	Autenticidad, integridad y no repudio.
Mecanismo de compromiso	Descifrado retroactivo de textos cifrados interceptados pasivamente.	Falsificación de firmas en tiempo real, suplantación de identidad y actualizaciones de rebeldes.
Vectores primarios	Registros de tráfico TLS, flujos transfronterizos, datos almacenados en reposo.	Firma de código, certificados raíz, verificación de firmware, registros de transacciones.

El riesgo operativo que requiere la migración se divide en dos vectores de amenaza distintos: Cosecha-Ahora-Descifrar-Después (HNDL) y Confiar-Ahora-Forja-Después (TNFL). HNDL representa un vector de amenaza de confidencialidad en el que las entidades hostiles interceptan y almacenan hoy tráfico de datos cifrados, dirigiéndose a los mecanismos de establecimiento e intercambio de claves. La intención del adversario es descifrar retroactivamente estas cargas útiles de datos una vez que un CRQC entre en funcionamiento (POSTQUANTUM, 2026). La HNDL afecta a datos sensibles de larga duración, incluidos secretos de seguridad nacional, propiedad intelectual, acuerdos financieros transfronterizos e historiales médicos.

Por el contrario, TNFL es un vector de amenaza de integridad y autenticación que amenaza las firmas digitales y la infraestructura de clave pública (PKI). Una vez que surge una CRQC, los adversarios pueden romper las suposiciones de dureza criptográfica que sustentan las firmas clásicas. Esto permite a los atacantes falsificar certificados de identidad digitales, subvertir mecanismos de firma de códigos, comprometer actualizaciones de firmware en infraestructuras nacionales críticas y repudiar transacciones históricas o contratos legales.

Un punto crítico de fallo en la estrategia empresarial es tratar la migración PQC como un ciclo de parches o intercambio de algoritmos convencional de TI. Los datos empíricos de programas complejos de migración demuestran que las grandes implementaciones empresariales generan más de 120.000 tareas discretas, de las cuales menos del 25% implican reemplazos directos de código algorítmico. El 75% restante abarca gobernanza de la cadena de suministro del proveedor, actualizaciones de módulos de seguridad de hardware (HSM), reconfiguración de middlebox y pila de red, adaptaciones de protocolos, revisiones del ciclo de vida de certificados y mapeo continuo de cumplimiento. Los programas que abordan la migración PQC estrictamente

como una tarea de ingeniería de software localizada suelen quedarse atascados en un plazo de 6 a 12 meses debido a dependencias multisistema no gestionadas y fricciones operativas.

Para abordar este desafío sistémico, el Marco de Migración Applied Quantum PQC —creado por Marin Ivezic y publicado bajo la licencia Creative Commons Attribution 4.0 International (CC BY 4.0)— ofrece una metodología de acceso abierto basada en profesionales. Refinado durante dos años de despliegue operativo, establece un modelo holístico de ejecución en 8 fases diseñado para llevar a las organizaciones desde asegurar un mandato ejecutivo inicial hasta mantener la criptoagilidad operativa.

Desglose del ciclo de vida de ejecución en 8 fases

El Marco de Migración Applied Quantum PQC organiza la transición empresarial en ocho fases secuenciales y superpuestas. Mientras que las fases iniciales se convierten en hitos posteriores, las fases operativas (Fases 5 y 6) se ejecutan como bucles de retroalimentación iterativos y paralelos, apoyados continuamente por la gobernanza de proveedores de la Fase 7 desde el inicio del programa (véase la tabla 10) (Campbell, 2026).

Tabla 10: Desglose del ciclo de vida de ejecución en 8 fases

Número de fase	Nombre de la fase	Entregables Principales y Alcance Operativo	Dependencias interfases
Fase 0	Mandato ejecutivo y caso de negocio	Patrocinio ejecutivo, carta de programa, presupuesto plurianual dedicado, mapeo de obligaciones regulatorias, identificación de	Desbloquea toda la financiación y aplicación de políticas posteriores en las unidades de negocio empresariales.

		activos críticos entre los 20 principales.	
Fase 1	Descubrimiento e inventario	Escaneo automatizado, análisis de protocolos de red, análisis estático de código, identificación de primitivas criptográficas, longitudes de clave y cadenas de certificados.	Introduce directamente datos criptográficos en bruto de dependencias en la documentación CBOM de la Fase.
Fase 2	CBOM y Documentación	Lista de Materiales Viables Mínimos Consultables (MV-CBOM) a través de cuatro capas arquitectónicas.	Establece la verdad básica requerida para la puntuación algorítmica de riesgo de la Fase 3.
Fase 3	Puntuación y priorización de riesgos	Priorizó el backlog de migración utilizando ponderación de vulnerabilidades de algoritmos, vida útil de los datos y exposición ambiental.	Transforma inventarios de activos en secuencias de ejecución accionables y clasificadas por riesgo para la Fase 4.
Fase 4	Hoja de ruta y gobernanza	Calendario maestro plurianual, configuración de la Oficina de Gestión de Programas (PMO), indicadores clave de riesgo (KRIs), puertas de decisión.	Define asignaciones de recursos, calendarios de oleadas y compuertas de preparación para la ejecución de las Fases 5 y 6.

Fase 5	Pilotos y migración	Despliegue de protocolos híbridos, refactorización de código, emisión de certificados de doble pila, reingeniería de la capa de aplicación.	Opera en iteración estrecha y paralela con la Fase 6; descubre limitaciones de infraestructura que informan las actualizaciones.
Fase 6	Infraestructura y Rendimiento	Actualizaciones HSM, implementación de doble pila PKI, autorización de midbox de red, evaluación de overhead de rendimiento.	Desbloquea directamente las migraciones de objetivos de la Fase 5; actualiza las entradas de la Fase 2 del CBOM a medida que la infraestructura se moderniza.
Fase 7	Gobernanza de proveedores y cadena de suministro	Evaluaciones de preparación de PQC por terceros, cláusulas contractuales de contratación de PQC, aplicación de SLA, seguimiento de la hoja de ruta de proveedores.	Comienza en el mes 1 y se desarrolla de forma continua en todas las fases para mitigar la deuda en la cadena de suministro.

Nota: Interpretación de los autores desde Campbell (2026)

El impulso del programa se establece mediante una secuencia operativa estructurada de ejecución de 90 días diseñada para evitar la parálisis analítica durante la configuración inicial. El calendario de ejecución estructura la madurez temprana del programa en tres hitos operativos mensuales principales:

Durante el Mes 1 (Movilización), el programa se centra en establecer la autoridad ejecutiva, asegurar un patrocinador ejecutivo, redactar el caso de

negocio principal, mapear las obligaciones regulatorias (como DORA, CNSA 2.0 y PCI DSS v4.0) e identificar los 20 sistemas críticos más importantes de la organización junto con las 10 principales dependencias de proveedores externos.

En el Mes 2 (Descubrir), el equipo despliega herramientas de descubrimiento criptográfico dirigidas a los 3 a 5 sistemas de mayor prioridad, inicia la comunicación con proveedores de software e infraestructura de nivel 1, comienza a ensamblar el MV-CBOM inicial y lanza programas de formación basados en roles básicos.

Para el Mes 3 (Plan), el programa completa las evaluaciones iniciales de alcance, presenta el caso de negocio y la puntuación de riesgos ante la dirección ejecutiva y la junta directiva, asegura compromisos presupuestarios plurianuales, formaliza el Comité Directivo del Programa (SteerCo) y publica la hoja de ruta migratoria del Año.

El modelo de migración de dos vías y los cuellos de botella en el ecosistema

La versión 2.0 y la versión 2.1 del marco Applied Quantum introducen el Modelo de Migración de Dos Vías, que separa la transformación criptográfica empresarial en dos pistas de ejecución paralelas basadas en primitivas matemáticas distintas, impulsores de amenazas y dependencias del ecosistema (véase la tabla 11) (Ivezic, 2026).

Tabla 11: Versión 2.0/2.1 del marco Applied Quantum

Parámetro	Vía A: Confidencialidad / Intercambio de claves	Vía B: Integridad y Autenticación
Factor de urgencia primario	Exposición Harvest-Now-Decrypt-Later (HNDL).	Confianza en Ahora-Forja-Después (TNFL) de

		exposición.
Funciones criptográficas	Intercambio de claves TLS, túneles VPN, sesiones SSH, establecimiento de claves de sesión, claves de cifrado de datos en reposo.	Certificados web PKI, jerarquías internas de CA, firma de código, actualizaciones de firmware, tokens de identidad, logs de no repudiación.
Estándar / Algoritmo Principal	ML-KEM (FIPS 203) desplegado en modo híbrido con intercambio clásico de claves (por ejemplo, X25519MLKEM768).	ML-DSA (FIPS 204), SLH-DSA (FIPS 205), firmas basadas en hash con estado (NIST SP 800-208 / LMS / XMSS).
Viabilidad actual del despliegue	Alto: Inmediatamente desplegable en producción en navegadores modernos, pilas TLS y servicios de borde.	Restringido: La alta sobrecarga de carga útil y los atrasos de validación FIPS restringen el despliegue rápido a nivel empresarial.
Cuellos de botella arquitectónicos	Caídas en el middlebox de red, fragmentación de paquetes Client-Hello, incompatibilidades con balanceadores de carga heredados.	Bifurcaciones arquitectónicas PKI, sobrecarga de la cadena de certificados, requisitos de reemplazo de hardware HSM, disponibilidad de módulos FIPS 140-33.

Nota: Interpretación de los autores desde Ivezic (2026).

La Pista A aborda mecanismos clave de establecimiento para proteger los datos contra la interceptación pasiva y el descifrado a largo plazo. Los despliegues utilizan construcciones híbridas de intercambio de claves, combinando un algoritmo clásico de intercambio de claves (como X25519 o ECDH) con un mecanismo de encapsulación de claves post-cuántico (ML-KEM-768). Este enfoque de doble capa garantiza que se mantengan las garantías de seguridad incluso si la implementación poscuántica o la construcción clásica

sufre un fallo de día cero.

La telemetría empresarial indica una rápida preparación del lado del cliente; los principales navegadores web y las redes de borde han habilitado X25519MLKEM768, con Cloudflare informando de que más del 70% del tráfico humano entrante soporta el intercambio híbrido de clave post-cuántico. Sin embargo, el soporte de origen en el lado del servidor se retrasa significativamente, aproximadamente en un 10%. La restricción principal de ingeniería de la Vía A implica que los middleboxes de red heredados (por ejemplo, firewalls de inspección profunda de paquetes, firewalls de aplicaciones web) eliminen las solicitudes de conexión del cliente porque el mensaje ampliado ClientHello —ampliado con la inclusión de cargas útiles de intercambio de claves PQC— supera los tamaños tradicionales de MTU de red o rompe las suposiciones del protocolo.

La Pista B se centra en firmas digitales y anclajes de confianza. A diferencia del intercambio de claves, donde los parámetros híbridos pueden negociarse dentro de los handshakes de sesión, los esquemas de firma post-cuántica introducen aumentos dramáticos en los tamaños de clave y firma. El cambio del ECDSA clásico (que utiliza firmas de ~64 bytes) a ML-DSA-65 (~3.309 bytes de firmas) incrementa el tráfico de red de handshake TLS hasta en diez veces, causando latencias inaceptables y tasas de fallo de conexión en entornos web PKI públicos.

Esta limitación técnica ha forzado una bifurcación estructural entre las arquitecturas de fideicomisos públicos y privados. En entornos privados de PKI —como redes corporativas internas, mallas de servicio a servicio y sistemas empresariales controlados— la infraestructura de red puede tolerar cargas útiles de certificados mayores, permitiendo a las organizaciones desplegar certificados ML-DSA directos o certificados de doble pila compuesta.

Por el contrario, el entorno público de la Web PKI no puede acomodar esta

sobrecarga de carga útil. En consecuencia, los proveedores de navegadores y las Autoridades Certificadoras (incluyendo Let's Encrypt y Google Chrome) se han comprometido a adoptar los Certificados Merkle Tree (MTC) como solución a largo plazo para la autenticación pública post-cuántica. Los MTC combinan certificados de corta duración con registros de pruebas criptográficas derivados de la infraestructura de Transparencia de Certificados, eliminando la necesidad de transmitir grandes cadenas de firmas durante cada handshake TLS individual.

Un cuello de botella sistémico secundario en la ejecución de la Pista B es la Brecha de Validación FIPS 140-3. En entornos regulados, los módulos criptográficos deben obtener la certificación NIST FIPS 140-3 antes del despliegue en producción. Con la transición de los certificados FIPS 140-2 del NIST a la lista histórica en septiembre de 2026, las empresas reguladas se enfrentan a un grave cuello de botella en la cadena de suministro: los proveedores de software y hardware están enfrentando retrasos de varios años para obtener la certificación formal FIPS 140-3 para las primitivas PQC recién implementadas (Ivezic, 2026).

Para evitar un estancamiento total en el despliegue, la versión 2.1 del framework designa las firmas basadas en hash con estado NIST SP 800-208 (LMS y XMSS) como el componente Track B de despliegue ahora para firmware, sistemas de firma de código y autenticación de bootloader. Debido a que LMS y XMSS dependen de funciones hash bien entendidas (SHA-256) en lugar de matemáticas de red, han madurado a través de evaluaciones de seguridad y ofrecen usabilidad inmediata en producción (NIST, 2020). Para gobernar dónde y cómo los algoritmos postcuánticos pueden entrar en producción sin infringir las políticas legales u operativas, el marco categoriza los activos empresariales en cuatro niveles de despliegue (véase la tabla 12):

Tabla 12: Principales restricciones regulatorias según el nivel de entorno

Nivel de Entorno	Alcance del Activo Objetivo	Permisibilidad PQC de producción	Principales restricciones regulatorias
Nivel 1: Sin restricciones	Aplicaciones web públicas, servicios de borde no regulados.	Despliegue inmediato permitido para intercambio híbrido de clave.	Monitorización operativa estándar del SLA.
Nivel 2: Consciente de FIPS	Sistemas corporativos internos, aplicaciones empresariales no reguladas.	Producción permitida usando bibliotecas de proveedores preestandarizadas.	Supervisión de políticas de seguridad interna.
Nivel 3: FIPS-Requisito	Sistemas financieros, sanitarios y CNI regulados.	Restringido hasta la validación del módulo FIPS 140-3 o el uso del SP 800-208.	Mandatos de cumplimiento estatutarios (por ejemplo, PCI, DORA).
Nivel 4: CNSA 2.0	Sistemas de seguridad nacional, Base Industrial de Defensa.	Cumplimiento obligatorio de los estrictos horarios de algoritmos de la NSA.	Directivas federales (NSM-10, puerta de enero de 2027).

Nota: Interpretación de los autores desde NIST (2020)

Innovaciones metodológicas fundamentales

Los proyectos convencionales de inventario suelen caer en una trampa de completitud: un modo de fallo operativo en el que los equipos de proyecto intentan catalogar cada llamada a una función criptográfica en todas las aplicaciones empresariales antes de iniciar la remediación, lo que resulta en retrasos infinitos (Lam, 2026). El modelo de CBOM VIABLE MÍNIMO (MV-CBOM) reemplaza los escaneos exhaustivos y sin prioridad por una arquitectura

estructurada de cuatro capas.

La Capa 1 encapsula la Capa de Infraestructura, capturando proxies de borde, endpoints de terminación TLS externos, VPNs de red y certificados internos de CA raíz o intermedios.

La Capa 2 representa la Capa de Plataforma, cubriendo sistemas operativos de servidor, middleware, tiempos de ejecución de contenedores y cifrado de almacenamiento en el motor de base de datos.

La Capa 3 aborda la Capa de Aplicación, abarcando código fuente personalizado de primera mano, bibliotecas de lógica de negocio, marcos de autorización de API y componentes de procesamiento de pagos.

La Capa 4 comprende la Capa Embebida y la Capa de Terceros, que capturan paquetes de software comerciales listos para usar (COTS), dependencias SaaS, binarios de firmware y equipos de Tecnología Operativa (OT).

Al priorizar la recogida de inventario desde la Capa 1 hasta la Capa 4, las organizaciones obtienen visibilidad accionable sobre superficies externas de alta exposición en cuestión de semanas, permitiendo mitigaciones de la Vía A mientras inventarian dependencias complejas de software aguas abajo. Para garantizar la usabilidad en motores de escaneo automatizados y plataformas de gobernanza, el *MV-CBOM* se asigna directamente al estándar CycloneDX, capturando campos de esquema estructurados (véase la tabla 13).

Tabla 13: Campos de esquema estructurados y propósito operativo

Nombre del campo	Descripción y Tipo de Datos	Propósito operativo
ID de activo del componente	Identificador único de cadena (por ejemplo, UUID / etiqueta de activo).	Vincula el uso criptográfico con entradas específicas de inventario del sistema.
Función criptográfica	Categórico (intercambio de claves, firma, cifrado).	Determina la asignación de la vía de migración (Vía A vs

		Vía B).
Primitivo y algoritmo	Cadena específica de algoritmos (por ejemplo, RSA-2048, ECDH-P256).	Evalúa la exposición matemática en función del algoritmo de Shor.
Longitud / Fuerza de la llave	Valor entero que representa bits de seguridad (por ejemplo, 112, 128, 256).	Alimenta modelos de puntuación de riesgo para determinar la urgencia de mitigación.
Biblioteca de implementación	Nombre y versión de la biblioteca de software (por ejemplo, OpenSSL 1.1.1).	Identifica dependencias en la cadena de suministro y esfuerzos de refactorización.
Caducidad del certificado	Marca de tiempo ISO de los límites del ciclo de vida del certificado.	Guías de automatización del ciclo de vida de PKI y renovación de doble pila.
Nivel de despliegue	Categórico (Sin restricciones, FIPS-Required, CNSA 2.0).	Regula la elegibilidad para la producción de primitivas postcuánticas.

El objetivo final de la migración poscuántica no es simplemente realizar una sustitución estática de algoritmos clásicos por primitivas PQC; está estableciendo una criptoagilidad permanente. La criptoagilidad representa la capacidad organizativa para intercambiar algoritmos, ajustar tamaños de clave y modificar arquitecturas de certificados mediante parámetros de configuración en lugar de refactorización de código. El marco operacionaliza la criptoagilidad como una disciplina que abarca cinco dimensiones estructurales: Arquitectura, Operaciones, Gobernanza, Habilidades y Cadena de Suministro.

La versión 2.0 y la versión 2.1 establecen una integración operativa explícita entre los marcos de Gobernanza, Riesgo y Cumplimiento (GRC) y el Centro de Operaciones de Seguridad (SOC). Reconociendo que las implementaciones poscuánticas introducen nuevas superficies operativas de amenaza durante la transición de varios años, el marco exige reglas específicas

de detección de SOC.

Estas incluyen monitorizar los handshakes TLS para detectar respaldo anómalo desde conjuntos híbridos de cifrado post-cuántico (por ejemplo, X25519MLKEM768) hacia suites heredadas vulnerables (por ejemplo, ECDHE-RSA-AES128), lo que puede indicar manipulación de adversarios o explotación activa de middlebox.

Además, las reglas de deriva criptográfica activan alertas automáticas cuando primitivas clásicas no aprobadas o certificados caducados entran en producción mediante pipelines de despliegue no autorizados.

Finalmente, la monitorización de integridad TNFL impone la verificación continua de firmas de código, sumas de comprobación binarias y cadenas de actualizaciones de firmware para detectar intentos de falsificación de firma en fases tempranas contra anclas de confianza heredadas.

Para Rama (2019), mantener una supervisión ejecutiva continua, las operaciones de gobernanza se basan en una jerarquía en cascada de 17 indicadores clave de riesgo (KRIs) distribuidos en tres niveles organizativos. En el nivel de la Junta Ejecutiva, las métricas rastrean la Puntuación de Exposición Cuántica Residual de la Empresa y los riesgos de penalización por incumplimiento regulatorio proyectados. En el CISO y el Nivel de Gestión, los KRI miden el porcentaje de activos críticos no inventariados y las tasas de incumplimiento de SLA de proveedores de nivel. En el Nivel Operativo y Técnico, la telemetría monitoriza los conteos de eventos de degradación de cifrado híbrido y los cuellos de botella de la canalización de validación FIPS 140-36.

Diferentes sectores industriales operan bajo entornos operativos distintos, restricciones de infraestructura heredada y requisitos regulatorios. El marco Applied Quantum proporciona seis extensiones sectoriales específicas para abordar estos desafíos verticales (véase la tabla 14).

Tabla 14: Extensiones verticales de sectores

Extensión vertical	Líneas Base Regulatorias Reguladoras	Cuellos de botella técnicos clave y vectores de exposición	Estrategia de Adaptación Marco
Servicios financieros	DORA (UE), G7 CEG, MAS (Singapur), HKMA (Hong Kong), PCI DSS v4.01.	Exposición a HNDL en flujos interbancarios transfronterizos de larga duración; sistemas bancarios básicos heredados; límites de espacio del firmware HSM.	La Fase 1 prioriza los canales de mensajes transfronterizos; la Fase 6 exige hojas de ruta tempranas para la actualización de proveedores de HSM.
Pagos	Fase 2 del Proyecto BIS Leap, Normas de Seguridad PCI, Reglas de Red de Tarjetas.	Una huella terminal masiva (POS, cajeros automáticos); límites de ancho de banda en redes de mensajería de pago; algoritmos HSM con tarjeta codificada.	Establece encapsulación proxy gateway para terminales de punto de venta; separa las claves de emisión de tarjetas del intercambio de claves de transacción.
Activos digitales	Marcos financieros globales, gobernanza de protocolos descentralizados.	Derivaciones de direcciones de clave pública codificadas; contratos inteligentes inmutables; Armaduras de clave validadoras de prueba de estaca (PoS); Vulnerabilidad a prueba de ZK.	Planificación de coordinación de hard forks; migración en capas de verificación de conocimiento cero; despliegue de firmas hash con estado para monederos custodiales.
Telecomunicaciones	GSMA PQ.01–PQ.07, Estándares 3GPP, 5G-AKA, Estandarización 6G	Interfaces de intercambio de roaming (SEPP); provisión	Integra PQC en la planificación de lanzamientos 3GPP; exige un

	Windows.	criptográfica de tarjetas SIM; restricciones de latencia de la puerta de intercepción legal.	intercambio híbrido de claves a través de rutas de señalización entre operadores.
OT e Infraestructura Crítica	NERC CIP, NIST SP 800-82, IEC 62443, Directrices CISA.	Ciclos de vida físicos de equipos de 15 a 25 años; RTUs SCADA de bajo consumo; límites de tiempo de inactividad para la recertificación de la caja de seguridad.	Despliega pasarelas proxy de seguridad bump-in-the-wire; prioriza a los historiadores de procesos que poseen telemetría de alta confidencialidad.
Gobierno y Defensa	CNSA 2.0, NSM-10, FedRAMP/FIPS, Directivas de Adquisiciones del DoD.	Chips embebidos en sistemas de armas; Módulos criptográficos Tipo 1 / CSfC; Interoperabilidad de las Redes Tácticas de la Coalición (OTAN).	Hace cumplir el cumplimiento de las puertas de adquisición de enero de 2027; mapas los ciclos de sostenimiento de la plataforma directamente a las hojas maestras de la Fase 41.

En los servicios financieros, las autoridades reguladoras aplican estrictas normas de resiliencia operativa (como las directrices DORA y MAS de la UE) que exigen el mapeo inmediato de los riesgos de la cadena de suministro de terceros y los canales de transacciones transfronterizas de larga duración sujetos a la interceptación de HNDL.

En Pagos, los terminales de punto de venta heredados y los Módulos de Seguridad de Hardware (HSM) enfrentan límites de capacidad de hardware que prohíben el procesamiento en línea de grandes tamaños de clave PQC, lo que requiere gateways intermediarios de API proxy.

Los Activos Digitales se enfrentan a desafíos estructurales debido a

contratos inteligentes inmutables y esquemas de direcciones elípticas codificados de forma fija, lo que requiere bifurcaciones duras a nivel de protocolo y actualizaciones de prueba de conocimiento cero.

Para Chow y Ma (2022), la infraestructura de telecomunicaciones debe equilibrar las restricciones de latencia en tiempo real entre los protocolos de autenticación 5G-AKA y los intercambios de roaming interoperadores (SEPP), impulsando la integración de PQC directamente en la alineación de los estándares 3GPP. La Tecnología Operativa (OT) y la Infraestructura Crítica están especialmente limitadas por ciclos de vida de equipos de 15 a 25 años y estrictas normas de recertificación de casos de seguridad, que requieren proxies criptográficos no intrusivos de tipo "bump-in-the-wire" para proteger los sistemas de control industrial.

Por último, los entornos gubernamentales y de defensa operan bajo plazos legales estrictos, como la puerta federal de adquisiciones de enero de 2027 y los mandatos de CNSA 2.0, que requieren interoperabilidad multinacional entre coaliciones de redes clasificadas.

Para apoyar la adopción estandarizada y proporcionar cualificaciones profesionales verificables, la Quantum Academy ofrece el programa de credenciales profesionales Post-Quantum Certificate, derivado directamente del Applied Quantum PQC Migration Framework (véase la tabla 15).

Tabla 15: Certificación institucional y ecosistema de marco

Designación de credenciales	Nombre de la credencial	Enfoque profesional principal y alcance	Público objetivo
PQCS	Especialista post-cuántico certificado	Principios fundamentales de riesgo cuántico, inventarios criptográficos, planificación de	Analistas de seguridad, responsables de riesgos, responsables de TI.

		línea base.	
PQCM	Gestor Certificado Post-Quantum	Gobernanza de programas, presupuestos plurianuales, gestión de proveedores, integración GRC.	CISOs, responsables de PMO, responsables de cumplimiento.
PQCA	Arquitecto Certificado Post-Quantum	Selección de algoritmos PQC, diseño de protocolos híbridos, arquitectura de doble pila PKI.	Arquitectos empresariales, ingenieros cripto.
PQCV	Validador certificado post-cuántico	Pruebas de implementación, verificación de interoperabilidad, validación de reglas SOC.	Responsables de QA, auditores de seguridad, responsables de SOC.
PQCX	Experto Certificado Post-Cuántico	La acreditación Capstone se otorga automáticamente al obtener el PQCM, PQCA y PQCV.	Directores ejecutivos sénior de programas cripto.

Los candidatos comienzan con la credencial fundamental de Especialista Certificado Post-Cuántico (PQCS), que establece la competencia básica en evaluación de riesgos cuánticos y elaboración de inventarios. De PQCS, los profesionales seleccionan itinerarios específicos para cada puesto: PQCM para la gobernanza de la gestión y el cumplimiento, PQCA para el diseño técnico e híbrido arquitectónico, o PQCV para las pruebas y la validación de la implementación. Los candidatos que obtengan con éxito las tres credenciales especializadas (PQCM, PQCA y PQCV) recibirán automáticamente la acreditación capstone Post-Quantum Certified Expert (PQCX) (véase la tabla 16).

Tabla 16: Marco de Migración Applied Quantum PQC (v2.1)

Dimensión de característica / capacidad	Documentos de Orientación Estratégica (por ejemplo, ETSI TR 103 619, Manual Neerlandés)	Hyperscaler Playbooks (por ejemplo, Meta PQC Migration Level)	Marco de Migración Applied Quantum PQC (v2.1)
Alcance del ciclo de vida	Parcial: Recomendaciones de alto nivel y marcos de planificación.	Enfoque interno: Infraestructura y microservicios internos.	Completado: Metodología de ejecución de 8 fases de extremo a extremo.
Modelo de Migración Paralela	Recomendaciones implícitas o de secuencia única.	Parcial: Separación implícita de KEM y firmas.	Explícito: Modelo formal de dos vías (Pista A / Pista B).
Arquitectura CBOM	Listas de inventario de alto nivel.	Mapas internos de dependencias propietarias.	CBOM VIABLE MÍNIMO ESTRUCTURADO DE 4 CAPAS (MV-CBOM).
Guía de PKI en la web pública	Sugerencias genéricas para reemplazar certificados.	Céntrate en los anclajes internos de confianza.	Posición definitiva sobre los Certificados de Árbol Merkle (MTCs).
SOC e integración GRC	Especificaciones mínimas de seguridad operativa.	Medidas internas de protección y métricas de ingeniería.	17 KRIs en cascada, reglas y manuales explícitos de SOC.
Modelo de licencias	Distribución de cuerpos estándar.	Publicación del whitepaper corporativo.	Acceso abierto bajo Creative Commons CC BY 4.0.

La transición empresarial a la criptografía post-cuántica es una prioridad inmediata en la gestión de riesgos. Las organizaciones deben abandonar las estrategias pasivas que dependen completamente de los proveedores

comerciales de software para entregar actualizaciones oportunas. El retraso en la preparación introduce graves riesgos operativos, incluyendo deuda no gestionada en la cadena de suministro, penalizaciones por incumplimiento regulatorio, retrasos en la refactorización de aplicaciones y exposición a ataques Harvest-Now-Decrypt-Later (HNDL) y Trust-Now-Forge-Later (TNFL) (POSTQUANTUM, 2026). Para ejecutar una migración exitosa, el liderazgo en seguridad empresarial debe alinearse con tres directrices estratégicas principales:

- Primero, implementa inmediatamente la Estrategia de Ejecución de Dos Vías. Las organizaciones deben desvincular la migración de intercambios clave de la refactorización de firmas digitales. Desplegar hoy en día mecanismos híbridos de encapsulación de claves (Track A) en redes de borde de alta exposición y servicios web mitiga riesgos urgentes de HNDL, mientras se prepara sistemáticamente la infraestructura interna de PKI y se siguen los desarrollos del Certificado Merkle Tree para la migración a largo plazo de firmas (Track B).
- Segundo, evita la trampa de completitud del inventario usando el modelo MV-CBOM. Los equipos de seguridad deberían iniciar el descubrimiento criptográfico en la Capa 1 (Infraestructura) y la Capa 2 (Plataforma) para establecer un CBOM Mínimo Viable y accionable en las próximas semanas. La puntuación de riesgos, la implicación de proveedores y los despliegues piloto deben proceder en paralelo con una expansión más amplia del inventario, evitando la parálisis operativa (Płoszaj et al., 2024).
- Tercero, integrar la criptoagilidad operativa como una capacidad empresarial permanente. El liderazgo empresarial debe ver la migración posterior a la cantidad no como una actualización puntual del algoritmo, sino como una oportunidad para modernizar sistemas heredados, automatizar ciclos de vida de certificados, hacer cumplir la cadena de

suministro e implementar capas de abstracción criptográfica configurables.

Al anclar la transformación empresarial al Marco de Migración Applied Quantum PQC de acceso abierto, las organizaciones pueden reducir sistemáticamente la exposición cuántica, mantener el cumplimiento continuo y construir una resiliencia operativa permanente.

Capítulo 5
Seguridad, criptoanálisis autónomo y marcos de migración

El despliegue a escala de la Inteligencia Artificial Agéntica (AAI) introduce un paradigma donde los sistemas de software ya no se limitan a responder a peticiones reactivas, sino que ejecutan ciclos operativos autónomos caracterizados por la planificación, la invocación dinámica de herramientas, la coordinación entre múltiples agentes y la gestión de memoria persistente. Esta capacidad operativa descansa sobre una infraestructura oculta pero crítica: la criptografía de clave pública clásica. Cada decisión adaptativa, llamada a la interfaz de programación de aplicaciones (API) o al intercambio de estado inter-agente, depende de primitivas criptográficas tradicionales para garantizar la identidad no humana, el control de acceso, la integridad de los modelos y la no repudiación de las acciones ejecutadas.

La llegada inminente de un computador cuántico relevante criptográficamente (CRQC) destruye los supuestos de seguridad que sostienen este ciclo de funcionamiento agéntico. Los algoritmos cuánticos fundamentales imponen dos amenazas matemáticas directas (Telefónica, 2026):

- Algoritmo de Shor: Resuelve el problema del factoraje de enteros y el del logaritmo discreto en tiempo polinómico $\mathcal{O}((\log N)^3)$, lo que invalida de forma absoluta los criptosistemas asimétricos basados en RSA, la criptografía de curva elíptica (ECC), ECDSA y Diffie-Hellman.
- Algoritmo de Grover: Ofrece una aceleración cuadrática $\mathcal{O}(\sqrt{N})$ para la búsqueda no estructurada, reduciendo de manera efectiva a la mitad la

fuerza criptográfica de los cifrados simétricos. Bajo este modelo, un cifrado como AES-128 ofrece una seguridad efectiva de tan solo 2^{64} operaciones cuánticas, imponiendo la necesidad de migrar universalmente hacia tamaños de clave de al menos 256 bits (por ejemplo, AES-256) para mantener la resistencia.

A diferencia de las infraestructuras de TI empresariales estáticas, donde los activos criptográficos están controlados centralizadamente por operadores humanos, los ecosistemas agénticos presentan una densidad masiva de Identidades No Humanas (NHI) que se negocian en tiempo de ejecución (véase la tabla 17). Esta dinámica expone a los agentes autónomos a través de dos ejes temporales y operacionales diferenciados (POSTQUANTUM, 2026):

- Eje de Confidencialidad (*Harvest-Now-Decrypt-Later* / HNDL): Adversarios estatales y actores de amenaza interceptan y almacenan continuamente la mensajería interagente, los contextos de memoria persistente y las credenciales cifradas con algoritmos clásicos. Cuando se disponga de un CRQC, esta información interceptada será descifrada, revelando propiedad intelectual, sesgos estratégicos y secretos corporativos de largo plazo acumulados por la flota de agentes.
- Eje de Integridad (*Harvest-Now-Forge-Later* / HNFL): Los atacantes recopilan firmas digitales clásicas utilizadas para autenticar código de modelos, validar directivas de orquestación y autorizar la invocación de herramientas operativas. La capacidad futura de reconstruir las claves privadas asimétricas a partir de estas firmas permite a un atacante falsificar de forma retroactiva las identidades de los agentes, inyectar instrucciones maliciosas en la memoria persistente del sistema o asumir el control total de la infraestructura agéntica sin levantar alertas de autenticación.

Tabla 17: Dimensiones de riesgo de los ecosistemas agénticos

Dimensión de Riesgo	Mecanismo Criptográfico Afectado	Vector de Ataque Cuántico	Impacto en la IA Agéntica
Confidencialidad (HNDL)	Intercambio de Claves (RSA, ECDH)	Algoritmo de Shor / Intercepción de Tráfico	Descifrado retroactivo de datos de memoria, registros de chat y secretos transferidos entre agentes.
Integridad (HNFL)	Firmas Digitales (ECDSA, RSA)	Algoritmo de Shor / Reconstrucción de Clave Privada	Suplantación de identidades de agentes, falsificación de la proveniencia del modelo y ejecución no autorizada de herramientas.
Cifrado Simétrico	Cifrado de Datos (AES-128, ChaCha20)	Algoritmo de Grover / Reducción Cuadrática	Reducción de la seguridad efectiva a niveles vulnerables a la fuerza bruta cuántica.

El análisis de estos vectores revela que la IA agéntica representa una clase diferenciada dentro de la migración poscuántica. Si bien en las organizaciones tradicionales la migración es un problema finito de inventario y remediación, en la IA agéntica la adopción continua de nuevos agentes autónomos genera una acumulación constante de deuda criptográfica clásica. Cada nuevo agente desplegado con un certificado X.509 clásico o una clave ECC incrementa la superficie expuesta a ataques HNFL y HNDL, requiriendo un cambio conceptual desde el mantenimiento preventivo hacia la prevención en la fuente de despliegue.

Criptoanálisis autónomo mediado por IA de frontera y la vulnerabilidad de los algoritmos poscuánticos

La convergencia entre la inteligencia artificial y la criptografía poscuántica no se limita a la protección defensiva de los agentes. Paralelamente, los modelos de lenguaje de gran tamaño (LLM) y las arquitecturas agénticas avanzadas han comenzado a ser utilizados como herramientas autónomas de criptoanálisis, transformando radicalmente la evaluación de seguridad de los algoritmos de última generación. Un hito crítico en esta disciplina ocurrió con las investigaciones desarrolladas sobre el modelo *Claude Mythos Preview* de Anthropic. Sometido a bucles de razonamiento autónomo dirigidos a buscar vulnerabilidades matemáticas en primitivas criptográficas, el sistema identificó fallos de diseño estructurales en candidatos para la estandarización poscuántica, demostrando capacidades que anteriormente requerían años de análisis por parte de criptógrafos humanos.

El ataque autónomo sobre el esquema HAWK

El hallazgo más significativo afectó a HAWK, un esquema de firma digital poscuántica basado en reticulados (*lattices*) que competía en el proceso de evaluación adicional (*Signatures On-ramp*) del Instituto Nacional de Estándares y Tecnología (NIST). HAWK había ganado tracción en la industria debido a su alta velocidad de firma libre de operaciones en punto flotante y a sus claves compactas en comparación con otros esquemas de reticulados (Eum et al., 2024).

Sin embargo, para Tang et al. (2024), la seguridad de HAWK descansaba sobre un supuesto algebraico menos examinado: el Problema del Isomorfismo

de Reticulados en Módulos (*Module Lattice Isomorphism Problem*). El sistema de IA identificó una simetría latente (un automorfismo no explotado previamente) en la estructura del reticulado subyacente. Mediante la aplicación de este automorfismo, el modelo demostró que la recuperación de la clave privada asimétrica podía reducirse desde un problema de vector más corto (SVP) en una dimensión de rango a un problema SVP en una dimensión drásticamente menor$n/2+1$. Esta reducción de dimensión acelera exponencialmente la resolución del problema lattice subyacente.

Las consecuencias cuantitativas de este descubrimiento matemático redujeron los márgenes de seguridad teóricos del esquema a niveles inaceptables para la producción industrial:

- HAWK-256: El modelo logró ejecutar la extracción completa de la clave privada en un servidor estándar en cuestión de pocas horas.
- HAWK-512: El factor de trabajo computacional para la recuperación de la clave privada se redujo de un estimado de 2^{150} a 2^{108} operaciones.
- HAWK-1024: El nivel de seguridad cayó de 2^{288} a 2^{182} operaciones.

Aunque el ataque no degrada la seguridad del cifrado de reticulados en términos generales ni invalida esquemas consolidados como ML-DSA o Falcon, el hallazgo demostró que la reducción a la mitad de la dimensión del reticulado destruye la ventaja de eficiencia de HAWK. Para restaurar los niveles de seguridad requeridos, sería necesario duplicar el tamaño de sus parámetros, eliminando así la compacidad que motivó su selección.

En el ámbito de la criptografía simétrica, el sistema agéntico descubrió una optimización para los ataques sobre variantes reducidas del cifrado AES (específicamente AES-128 con 7 rondas de sustitución y permutación). Mediante el diseño autónomo de estructuras de almacenamiento en memoria para precomputar patrones de texto plano, la IA logró acelerar la tasa de

ejecución de los ataques conocidos en un factor de 200 a 800 veces. La capacidad de orquestar flotas de agentes impulsados por modelos de frontera democratiza la capacidad de auditoría criptográfica, pero también expone las arquitecturas de cifrado a ataques sofisticados a un costo comercialmente accesible para organizaciones cibercriminales bien financiadas.

Estándares criptográficos poscuánticos del NIST (FIPS 203, FIPS 204 y FIPS 205) y cripto-agilidad

Como respuesta a la amenaza cuántica, el NIST finalizó en agosto de 2024 el primer conjunto de estándares de criptografía poscuántica (PQC), diseñados para reemplazar de manera progresiva los algoritmos de clave pública clásicos. La integración de estos algoritmos en sistemas agénticos es indispensable para asegurar la identidad, la comunicación y el control de integridad de las decisiones autónomas. Los estándares abarcan dos funciones fundamentales: la encapsulación de claves (KEM) para garantizar la confidencialidad de la comunicación y las firmas digitales para asegurar la autenticidad e integridad (Cherkaoui et al., 2024):

1. **FIPS 203 (ML-KEM)**: Derivado del algoritmo CRYSTALS-Kyber, es el estándar primario para la encapsulación de claves asimétricas. Se fundamenta en la dureza matemática del problema del Aprendizaje con Errores en Módulos de Reticulados (*Module Learning with Errors* / M-LWE). Reemplaza directamente los intercambios de claves RSA, DH y ECDH en los protocolos TLS y en los canales de comunicación de memoria agéntica.
2. **FIPS 204 (ML-DSA)**: Basado en CRYSTALS-Dilithium, es el estándar primario para firmas digitales de uso general. También sustentado en reticulados estructurados (M-LWE y *Module Short Integer Solution* / M-SIS), reemplaza a RSA y ECDSA en la emisión de certificados de identidad

no humana, autenticación de invocación de herramientas y validación del código de los modelos.

3. **FIPS 205 (SLH-DSA)**: Derivado de SPHINCS+, es un esquema de firma digital basado exclusivamente en funciones hash sin memoria de estado (*stateless*). No depende de supuestos matemáticos de reticulados; su seguridad se reduce directamente a la resistencia a preimágenes de funciones hash estructuradas mediante árboles Merkle multinivel y esquemas FORS (*Forest of Random Subsets*). Aunque genera firmas considerablemente más grandes y exige mayor tiempo de procesamiento, actúa como la salvaguarda de respaldo (*conservative fallback*) en caso de que surjan ataques cuánticos o matemáticos imprevistos contra la criptografía de reticulados (véase la tabla 18).

Tabla 18: Claves (KEM) tipo para la autenticidad e integridad

Estándar FIPS	Algoritmo Criptográfico	Base Matemática	Reemplaza a	Tamaño Clave Pública (Bytes)	Tamaño Clave Privada / Firma (Bytes)	Aplicación Primaria en IA Agéntica
FIPS 203	ML-KEM (Kyber)	Reticulados Módulos (M-LWE)	RSA, ECDH, Diffie-Hellman	800 – 1.568	768 – 1.568 *(Texto cifrado)*	Cifrado de canales interagente y persistencia de memoria.
FIPS 204	ML-DSA (Dilithium)	Reticulados Módulos (M-LWE / M-SIS)	RSA, ECDSA, Ed25519	1.312 – 2.592	2.420 – 4.595 *(Firma)*	Firmas de identidad (SPIFFE), autenticación de llamadas a herramie

						ntas.
FIPS 205	SLH-DSA (SPHINCS +)	Árboles Hash / Hypertree FORS	RSA, ECDSA *(Respaldo)*	32 – 64	7.856 – 49.856 *(Firma)*	Validación de integridad de modelos a largo plazo y firmas de auditoría.

La exigencia de cripto-agilidad en sistemas agénticos

El análisis del colapso del esquema HAWK subraya la fragilidad de confiar ciegamente en una única familia matemática. Para los entornos agénticos, donde la interoperabilidad se da en tiempo de ejecución, la cripto-agilidad deja de ser una recomendación teórica para convertirse en un requisito arquitectónico indispensable (Tang et al., 2024).

La criptoagilidad en sistemas autónomos requiere abstracciones donde las primitivas cifradas no estén acopladas rígidamente al código del software. Los agentes deben poseer capas de orquestación adaptativas capaces de alternar dinámicamente entre cifrados híbridos (combinando un algoritmo clásico como ECDH/Ed25519 con un poscuántico como ML-KEM/ML-DSA) o cambiar a primitivas de respaldo como SLH-DSA si un análisis de amenazas señala un compromiso inminente en la teoría de reticulados.

Identidad, infraestructura de confianza y las siete superficies de migración

La integración de la criptografía poscuántica en la IA agéntica requiere desglosar la totalidad del ciclo de funcionamiento del agente para cartografiar

dónde residen las dependencias criptográficas asimétricas y simétricas. Mediante el rastreo sistemático del ciclo operativo del agente, se derivan analíticamente siete superficies de migración diferenciadas, cada una caracterizada por dependencias criptográficas específicas y dinámicas operativas singulares (Acosta, 2025):

1. Emisión y Autenticación de Identidad (*Identity Issuance*): Los agentes requieren certificados criptográficos para probar su existencia frente a gateways de orquestación y otros agentes. Depende de firmas asimétricas que deben migrarse de ECDSA/RSA hacia ML-DSA.
2. Invocación de Herramientas y APIs (*Tool Invocation*): La ejecución de acciones externas exige la firma de tokens de autorización y solicitudes en tiempo de ejecución. La alteración de estas solicitudes mediante ataques HNFL permitiría la inyección de comandos maliciosos sin detección.
3. Mensajería e Intercambio Inter-Agente (*Inter-Agent Messaging*): La colaboración entre múltiples agentes utiliza túneles TLS o cifrado de capa de aplicación que dependen de mecanismos KEM para establecer secretos compartidos dinámicos, siendo altamente vulnerables a la intercepción HNDL.
4. Carga e Integridad del Modelo (*Model Loading and Integrity*): Garantiza que los pesos del LLM y los hiperparámetros no hayan sido modificados por un atacante. La firma digital del artefacto del modelo debe validar la proveniencia antes de la ejecución mediante esquemas resistentes a Shor.
5. Memoria y Estado Persistente (*Memory and Persistent State*): Las bases de datos vectoriales y las cachés de contexto retienen datos confidenciales a lo largo del tiempo. Requieren cifrado simétrico robusto (AES-256) derivado mediante KEMs poscuánticos para evitar descifrados retroactivos.
6. Orquestación y Coordinación (*Orchestration*): La asignación de tareas

dentro de grafos dirigidos acíclicos (DAG) de agentes debe estar vinculada criptográficamente para evitar la manipulación no autorizada de flujos de trabajo.

7. Registro, Auditoría y Proveniencia (*Logging and Provenance*): Los registros de decisiones tomadas por los agentes deben ser inmutables y no repudiables para cumplir con auditorías normativas, asegurándose mediante estructuras de firma basadas en hash o reticulados.

El desacoplamiento de identidad: capa de credenciales vs. capa de anclas de confianza

Una de las contribuciones conceptuales más relevantes en la securización agéntica es la separación de la identidad no humana en dos niveles operacionales con dinámicas de escalabilidad e impacto inversas (Rahayu et al., 2025; Mandarino et al., 2026):

- Capa de Credenciales (*Credential Layer*): Comprende los tokens efímeros, secretos simétricos y claves de sesión de corta duración que sostienen las interacciones diarias del agente. Es numéricamente densa (millones de instancias en ejecuciones masivas), pero representa un riesgo individual reducido. Su migración es de bajo costo y esfuerzo operativo.
- Capa de Anclas de Confianza (*Trust-Anchor Layer*): Constituye la raíz asimétrica de la infraestructura de clave pública (autoridades de certificación raíz, claves de firma de tokens de identidad, raíces de confianza de orquestación) que otorga validez a toda la flota de agentes. Posee una población pequeña, pero su alcance de impacto (*blast radius*) ante un ataque de falsificación tardía (HNFL) es universal: el compromiso de un ancla de confianza invalida la autenticidad de cada agente, modelo y herramienta subordinada.

Esta escala inversa impone la regla de secuenciación fundamental: las

anclas de confianza deben migrarse en primer lugar. Maximizar el apalancamiento defensivo requiere asegurar las raíces de firma asimétrica antes de intentar remediar la totalidad de las credenciales efímeras distribuidas en la flota agéntica.

Estándares de identidad no humana: SPIFFE/SPIRE y pruebas de cero conocimiento poscuánticas

Para gestionar la identidad de los agentes en entornos de producción, marcos como SPIFFE (Infrastructura Abierta de Identidad de Servicios para Todos) y su implementación de tiempo de ejecución SPIRE ofrecen una base sólida al emitir Documentos de Identidad Verificables de SPIFFE (SVID) en forma de certificados X.509 de corta duración. En esta arquitectura, la identidad unívoca del agente (SPIFFE ID) se codifica directamente en el campo URI del *Subject Alternative Name* (SAN) del certificado X.509.

En la era poscuántica, los servidores SPIRE deben configurarse para emitir SVIDs firmados nativamente con ML-DSA (FIPS 204), sustituyendo las firmas ECDSA o RSA. La rotación automática de estos certificados efímeros (por ejemplo, con ciclos de vida de aproximadamente 60 minutos) limita la ventana de exposición para ataques de intercepción en tiempo real (NIST, 2024).

Adicionalmente, marcos gubernamentales como la iniciativa de Estonia (que comenzó a asignar códigos de identidad digital oficial a agentes de IA autónomos de forma independiente de sus propietarios humanos) sientan las bases para el reconocimiento legal de los agentes. En ecosistemas descentralizados y multiagente, estas identidades se complementan con pruebas poscuánticas de cero conocimiento (PQ-ZKP) sustentadas en reticulados (por ejemplo, construcciones zk-SNARKs poscuánticas). Las PQ-ZKP permiten que un agente pruebe a otro agente o regulador que posee la autorización o solvencia necesaria para ejecutar una transacción económica sin

revelar los datos confidenciales o los valores subyacentes.

Marcos defensivos autónomos y métricas de riesgo ajustadas cuánticamente

Frente a la complejidad de auditar infraestructuras de IA en constante evolución, han emergido paradigmas arquitectónicos e instrumentales diseñados para integrar la seguridad cuántica desde las fases primarias del diseño y automatizar el análisis de riesgos.

Quantum-Secure-by-Construction (QSC)

El modelo *Quantum-Secure-by-Construction* (QSC) propone tratar la comunicación segura contra ataques cuánticos como una propiedad arquitectónica nativa e intrínseca de los agentes de IA, en lugar de una actualización parcheada a posteriori. QSC articula su defensa integrando tres componentes esenciales: criptografía poscuántica (FIPS 203/204) para autenticación y cifrado, Generación Cuántica de Números Aleatorios (QRNG) para garantizar entropía física en la creación de claves efímeras, y Distribución Cuántica de Claves (QKD) para proporcionar seguridad incondicional en enlaces de infraestructura críticos.

QSC opera mediante una capa de orquestación consciente de las políticas que ajusta el perfil de seguridad del agente dinámicamente a lo largo de todo su ciclo de vida (sesión, coordinación inter-agente y memoria) según la disponibilidad de la infraestructura, los requisitos normativos y la latencia del entorno.

El Marco quantigence y la puntuación de riesgo ajustada cuánticamente (QARS)

Para contrarrestar la saturación de información en la gestión de

vulnerabilidades cuánticas, se ha desarrollado el marco Quantigence, un sistema multiagente diseñado para auditar autónomamente la postura criptográfica de sistemas complejos. Quantigence utiliza un principio denominado paralelismo cognitivo: descomponer los objetivos de análisis de seguridad en roles de razonamiento especializados (por ejemplo, agentes expertos en dureza de reticulados, agentes monitores de ataques de canal lateral y agentes de seguimiento de políticas de estandarización) (Alquwayfili, 2025). Cada rol razona de forma independiente para mantener la pureza del contexto y evitar la degradación de la atención del modelo. Los agentes se coordinan a través de un supervisor jerárquico e integran inteligencia externa mediante el Protocolo de Contexto de Modelo (*Model Context Protocol* / MCP) (véase la tabla 19).

Para priorizar las vulnerabilidades identificadas, el marco introduce la Puntuación de Riesgo Ajustada Cuánticamente (QARS), una extensión formal del Teorema de Mosca. El Teorema de Mosca clásico establece que si el tiempo de vida requerido para mantener la confidencialidad de los datos (x) sumado al tiempo necesario para migrar el sistema criptográfico (y) es mayor que el tiempo estimado para la aparición de un CRQC (z), expresado formalmente como $x + y > z$, el sistema entra en un estado de vulnerabilidad crítica e irrecuperable.

QARS extiende esta formulación estática transformándola en una métrica de urgencia continua y dinámica, incorporando variables operacionales específicas de los ecosistemas agénticos:

$$\mathrm{QARS} = f\left(\frac{x_{\text{persistencia}} + y_{\text{migración}}}{z_{\text{CRQC}}}, D_{\text{NHI}}, B_{\text{HNFL}}\right)$$

Donde:

- $x_{\text{persistencia}}$ representa el tiempo durante el cual los datos de contexto o memoria agéntica deben permanecer confidenciales.
- $y_{\text{migración}}$ mide el tiempo estimado para actualizar las anclas de confianza de la flota agéntica.
- z_{CRQC} Es la distribución de probabilidad temporal hasta la llegada de un ordenador cuántico criptográficamente relevante.
- D_{NHI} Es la densidad de identidades no humanas dinámicas que negocian certs efímeros en tiempo de ejecución.
- B_{HNFL} representa el alcance de impacto (*blast radius*) de las anclas asimétricas de la flota.

Las evaluaciones empíricas demuestran que el uso de arquitecturas multiagente dedicadas a la auditoría de riesgos poscuánticos reduce el tiempo de análisis en un 67% en comparación con los flujos de trabajo manuales desarrollados por expertos humanos, garantizando un seguimiento continuo del estado de caducidad criptográfica de los sistemas.

La intersección de la inteligencia artificial agéntica y la criptografía poscuántica marca una transformación inevitable en la ingeniería de la seguridad de la información. La vulnerabilidad de las primitivas asimétricas clásicas frente a los algoritmos cuánticos de Shor y Grover no es un problema futuro distante, sino una amenaza operativa inmediata debido a las dinámicas de intercepción *Harvest-Now-Decrypt-Later* y a la capacidad defensiva/ofensiva demostrada por los modelos de lenguaje de frontera mediante criptoanálisis autónomo.

Dado que los sistemas de IA agéntica acumulan deuda criptográfica de forma continua a medida que escalan, las organizaciones deben abandonar los

enfoques de inventario estáticos y adoptar infraestructuras cripto-ágiles gobernadas por políticas (Zia et al., 2026):

Tabla 19: Acciones de Mitigación Específicas en IA Agéntica (Fases)

Fase de Migración	Ventana Temporal Recomendada	Objetivos Prioritarios	Acciones de Mitigación Específicas en IA Agéntica
Fase 1: Preparación e Inventario de Anclas	Inmediata (2024 – 2026)	Mapeo de las 7 superficies y mitigación HNDL.	Implementar ML-KEM-768 en modo híbrido para TLS. Inventariar y aislar la Capa de Anclas de Confianza (raíces asimétricas) de la flota agéntica. Auditar aplicaciones con algoritmos de IA tipo Quantigence.
Fase 2: Transición de Identidades No Humanas	Corto Plazo (2026 – 2027)	Aseguramiento de la integridad y firmas (HNFL).	Desplegar certificados de identidad interna basados en FIPS 204 (ML-DSA) mediante servidores SPIRE adaptados. Exigir la firma poscuántica de los artefactos de modelos LLM.
Fase 3: Agilidad Externa y Cero Conocimiento	Mediano Plazo (2027 – 2028)	Interoperabilidad externa y gobernanza avanzada.	Migrar certificados y firmas de invocación de herramientas públicas a ML-DSA/SLH-DSA. Integrar pruebas poscuánticas de cero conocimiento (PQ-ZKP) para la verificación de

			permisos sin revelación de datos.
Fase 4: Migración Total y Erradicación	Largo Plazo (2028 – 2030)	Depreciación total del PKI clásico.	Eliminar completamente el soporte para RSA, ECDSA y ECDH. Establecer arquitecturas *Quantum-Secure-by-Construction* (QSC) automatizadas para prevenir la reinfección con criptografía obsoleta.

El establecimiento prioritario de la protección en las anclas de confianza, la automatización de la identidad de los agentes mediante protocolos como SPIFFE rediseñados para PQC, y el monitoreo continuo apoyado por sistemas de razonamiento agéntico constituyen la única estrategia sostenible para garantizar la resiliencia de la inteligencia artificial autónoma en la era cuántica.

Capítulo 6
Tríada estratégica digital: gobernanza, soberanía tecnológica y resiliencia operativa en el marco regulatorio global y nacional

La acelerada interconexión de los sistemas de información, combinada con la creciente sofisticación de las amenazas cibernéticas a nivel global, ha transformado radicalmente la gestión de las Tecnologías de la Información y la Comunicación (TIC). La seguridad digital ha dejado de ser una disciplina predominantemente técnica para convertirse en un pilar fundamental de la continuidad del Estado, la estabilidad financiera y la viabilidad corporativa. En este contexto, la confianza digital, la soberanía tecnológica y la resiliencia operativa constituyen una tríada estratégica interdependiente, indispensable para garantizar la supervivencia y funcionalidad de las organizaciones modernas ante escenarios de crisis o disrupción sistémica.

La confianza digital actúa como el cimiento que asegura que las interacciones en el ecosistema digital se desarrollen bajo parámetros de seguridad, privacidad, transparencia y certidumbre jurídica. No obstante, la confianza es irrealizable sin el ejercicio efectivo de la soberanía tecnológica, entendida como la capacidad de un Estado o institución de mantener el control estratégico y operativo sobre sus activos informacionales, infraestructuras críticas y capacidades de ciberdefensa, reduciendo la vulnerabilidad derivada de dependencias externas o sesgos jurisdiccionales (Kahn, 2020). Por su parte, la resiliencia operativa representa la capacidad pragmática del sistema para resistir, adaptarse y mantener sus operaciones esenciales aun en condiciones

adversas o tras la materialización de un ciberataque.

La confluencia de estas tres dimensiones responde a una profunda reconfiguración regulatoria internacional y nacional. Este nuevo paradigma exige trascender los esquemas tradicionales de protección perimetral estática para adoptar modelos dinámicos de gobernanza continua, gestión integral de riesgos de terceros y arquitecturas defensivas orientadas a la continuidad del servicio.

Gobernanza y confianza digital: reconfiguración estratégica del riesgo tecnológico

La confianza digital no representa un estado pasivo ni un atributo alcanzable mediante la mera adquisición de herramientas tecnológicas. Por el contrario, se define como la capacidad demostrable de una organización para diseñar, operar y gobernar sistemas digitales de manera segura, confiable y ética bajo condiciones reales y cambiantes. Cuando la gobernanza es deficiente, la implementación de controles técnicos resulta fragmentada e ineficaz, exponiendo a las entidades a fallas operativas severas y vulnerabilidades en la cadena de suministro.

El giro institucional de NIST CSF 2.0 y la función "Govern"

La evolución hacia una gobernanza integral alcanzó su punto de inflexión con la publicación del Marco de Ciberseguridad del Instituto Nacional de Estándares y Tecnología de los Estados Unidos en su versión 2.0 (NIST CSF 2.0). Al ampliar su alcance histórico centrado en infraestructuras críticas hacia todo tipo de organizaciones, el marco introdujo la función de Gobernar (*Govern*) como el núcleo que articula las restantes cinco funciones tradicionales: Identificar, Proteger, Detectar, Responder y Recuperar (Burga et al., 2026).

La incorporación de la función *Govern* reconoce formalmente que la

gestión del riesgo cibernético no es un problema exclusivo del departamento de TI, sino una prioridad estratégica que atañe directamente al directorio y a la alta dirección corporativa. Esta dimensión abarca la definición explícita de políticas de ciberseguridad, la asignación de roles y responsabilidades de supervisión, la integración con la gestión del riesgo empresarial (*Enterprise Risk Management*) y la alineación de las capacidades técnicas con los objetivos del negocio y la normativa aplicable. De este modo, la gobernanza proporciona la dirección y el contexto operativo necesarios para que las salvaguardas técnicas funcionen de manera coherente.

El marco normativo peruano: la política nacional de transformación digital al 2030 y la regulación sectorial

En el Perú, el desarrollo normativo ha buscado institucionalizar la gobernanza digital como una política transversal del Estado. Mediante el Decreto Supremo N° 085-2023-PCM, se aprobó la Política Nacional de Transformación Digital al 2030 (PNTD), conducida por la Secretaría de Gobierno y Transformación Digital (SGTD) de la Presidencia del Consejo de Ministros (PCM) en su condición de ente rector del Sistema Nacional de Transformación Digital (creado mediante Decreto de Urgencia N° 006-2020) (Silva, 2025). La PNTD establece un instrumento de gestión pública obligatorio para las entidades de la administración pública y orientativo para el sector privado, la academia y la sociedad civil; la PNTD estructura la estrategia nacional en seis Objetivos Prioritarios interrelacionados (Hiperderecho, 2023):

- OP1 Conectividad Digital: Enfocado en garantizar el acceso inclusivo, seguro y de calidad al entorno digital a toda la población, reduciendo la brecha digital territorial.
- OP2 Economía Digital: Orientado a vincular las tecnologías digitales a los procesos productivos sostenibles, promoviendo la inclusión financiera y la

digitalización de las Mipymes.

- OP3 Gobierno Digital: Centrado en asegurar la disponibilidad de servicios públicos digitales predictivos, accesibles e interoperables para la ciudadanía.
- OP4 Talento Digital: Destinado a fortalecer las habilidades, competencias y capacidades digitales de la población, con énfasis en grupos en situación de vulnerabilidad.
- OP5 Confianza Digital: Prioridad que busca garantizar la seguridad y confianza en el entorno digital mediante el fortalecimiento de capacidades nacionales para identificar, gestionar y mitigar riesgos cibernéticos, con especial protección a la infancia.
- OP6 Innovación Digital: Orientado a promover el aprovechamiento ético de tecnologías emergentes y exponenciales, como la Inteligencia Artificial y el Internet de las Cosas.

Esta arquitectura regulatoria se fundamenta operacionalmente en el marco de confianza digital instituido por el Decreto de Urgencia N° 007-2020. Asimismo, en el ámbito financiero y asegurador, la Superintendencia de Banca, Seguros y AFP (SBS) aprobó la Resolución SBS N° 504-2021, que establece el *Reglamento para la Gestión de la Seguridad de la Información y la Ciberseguridad*. Esta norma sustituyó esquemas regulatorios desactualizados y vinculó la ciberseguridad con el Reglamento para la Gestión del Riesgo Operacional (Resolución SBS N° 6523-2013 y Resolución SBS N° 778-2022).

La regulación de la SBS exige a las entidades bajo su supervisión la designación formal de un Oficial de Seguridad de la Información (RSI), la constitución de comités especializados, la ejecución de pruebas de penetración y la implementación de controles estrictos sobre la ciberseguridad en proveedores de servicios.

Soberanía tecnológica y autonomía de la infraestructura crítica

La soberanía tecnológica se define como la capacidad estratégica de un Estado o de una entidad corporativa para tomar decisiones autónomas sobre su infraestructura digital, salvaguardando sus datos, código fuente y operaciones frente a restricciones de acceso, contingencias geopolíticas o fallas catastróficas registradas en proveedores externos. En un entorno dominado por la concentración del mercado de almacenamiento y procesamiento en la nube, la soberanía implica reducir la vulnerabilidad sistémica que representa la dependencia crítica de un número reducido de actores tecnológicos globales (Hernández et al., 2022).

La ciberdefensa como quinto dominio de la soberanía estatal

A nivel estatal, la soberanía tecnológica trasciende el ámbito económico para convertirse en una dimensión del poder nacional. En el Perú, el ciberespacio ha sido formalmente reconocido como el quinto dominio de la soberanía del Estado —sumándose a los espacios terrestre, marítimo, aéreo y espacial— mediante la Ley N° 30999 (Ley de Ciberdefensa) y su Reglamento aprobado por el Decreto Supremo N° 017-2024-PCM.

Este marco legal otorga al Presidente de la República y al Ministerio de Defensa la facultad de dirigir las operaciones de ciberdefensa orientadas a neutralizar amenazas que pongan en riesgo la Seguridad Nacional, la integridad de los Activos Críticos Nacionales (ACN) y los Recursos Claves (RC). Bajo este esquema, el Comando de Operaciones Cibernéticas de las Fuerzas Armadas ejerce el mando y control operacional para garantizar la defensa de las infraestructuras informáticas militares e institucionales contra agresiones perpetradas por actores estatales o no estatales en y mediante el ciberespacio.

Soberanía tecnológica, nubes soberanas y marcos reguladores internacionales (DORA y NIS2)

A nivel internacional, la Unión Europea ha liderado el establecimiento de marcos legales que vinculan directamente la resiliencia operativa con la soberanía tecnológica. La Directiva NIS2 y el Reglamento de Resiliencia Operativa Digital (DORA, aplicable desde enero de 2025) redefinen las obligaciones de las entidades financieras y de los operadores de servicios esenciales en materia de ciberseguridad (EIOPA, 2025).

Un aspecto central de estas normativas es la exigencia de una rigurosa gestión del riesgo derivado de proveedores TIC de terceros. La alta dependencia de plataformas propietarias y proveedores hiperescaladores con sede fuera de la jurisdicción comunitaria expone a las organizaciones a riesgos legales, cambios unibancarizados de términos de servicio o la interrupción forzada de operaciones. Como respuesta estratégica, ha cobrado fuerza el despliegue de soluciones de nube soberana (*Sovereign Cloud*), infraestructuras autoalojadas (*self-hosted*), modelos de código abierto auditables y proyectos federados como Gaia-X. Estas arquitecturas aseguran la plena titularidad de los datos, garantizan la auditabilidad permanente de la infraestructura y eliminan el atrapamiento tecnológico (*vendor lock-in*), fortaleciendo la resiliencia de la organización.

Resiliencia operativa y arquitectura sustentable de respuesta

La resiliencia operativa representa la evolución de la ciberseguridad hacia una disciplina centrada en la continuidad del negocio en condiciones de estrés extremo. Mientras que los paradigmas históricos de Recuperación ante Desastres (DR) y Continuidad de Negocio (BC) asumían una condición binaria

donde la organización operaba plenamente o se encontraba fuera de línea, la resiliencia operativa parte del principio de que los ciberataques e incidentes tecnológicos son inevitables y que los sistemas deben continuar funcionando en estado de degradación controlada.

Mantenimiento de niveles mínimos de servicio viable (MVSL) y cero confianza (ZTA)

El objetivo central de la resiliencia operativa es salvaguardar los Niveles Mínimos de Servicio Viable (MVSL) durante y después de una disrupción de las TIC, garantizando la continuidad de las funciones vitales mientras se ejecutan los trabajos de restauración. Para lograr esta capacidad, las organizaciones avanzadas implementan arquitecturas de Cero Confianza (*Zero Trust Architecture*, ZTA), alineadas con la norma NIST SP 800-207 (El-hajj et al., 2023).

El principio fundamental de ZTA —"nunca confiar, siempre verificar"— elimina la confianza implícita asociada a la ubicación de un usuario o dispositivo dentro del perímetro de la red corporativa. Mediante la verificación continua de la identidad, el control de acceso bajo el principio de mínimo privilegio, la microsegmentación de la red y el monitoreo dinámico del contexto, ZTA limita la capacidad de movimiento lateral de los atacantes dentro de la infraestructura. Esta contención técnica directa resulta crítica para proteger los servicios del MVSL, previniendo que el compromiso de un activo no esencial escale hacia la parálisis completa de la institución.

Notificación obligatoria de incidentes y evaluación de capacidades mediante simulacros nacionales (SIMAC)

Un componente vital de la resiliencia operativa radica en la velocidad de detección y la articulación para la respuesta colectiva ante incidentes. En el Perú, la legislación impone obligaciones rigurosas de notificación que configuran un régimen de reporte dual para incidentes de seguridad digital

(Otero, 2026):

- Reporte Técnico de Ciberseguridad: Los proveedores de servicios digitales y los administradores de infraestructuras críticas están obligados a notificar al Centro Nacional de Seguridad Digital (CNSD) cualquier incidente de impacto significativo dentro de un plazo máximo de 48 horas desde su identificación, en conformidad con el Decreto de Urgencia N° 007-2020 y el Decreto Supremo N° 126-2025-PCM.
- Reporte de Brechas de Protección de Datos: Si el incidente compromete datos personales, la organización debe reportarlo de forma simultánea a la Autoridad Nacional de Protección de Datos Personales (ANPDP) en el mismo plazo de 48 horas, conforme a lo estipulado en el Decreto Supremo N° 016-2024-JUS y la Ley N° 29733, bajo el riesgo de sanciones administrativas que pueden alcanzar las 100 Unidades Impositivas Tributarias (UIT).

El CNSD, operado por la SGTD de la PCM, actúa como el organismo rector de la ciberseguridad civil y administra el CSIRT Nacional, emitiendo alertas integradas, coordinando la red nacional de Oficiales de Seguridad y Confianza Digital y brindando soporte especializado a las entidades públicas y privadas.

Para validar operativamente la efectividad de estos esquemas de respuesta, la PCM organiza periódicamente el Simulacro de Ataques Cibernéticos (SIMAC). En estos ejercicios, los equipos de defensa (*Blue Teams*) de las instituciones participantes son evaluados frente a escenarios simulados de alta complejidad que recrean vectores reales de agresión, como campañas de *ransomware*, compromiso de credenciales, explotación de vulnerabilidades en aplicaciones web y movimientos laterales no autorizados. Al concluir el simulacro, el CNSD proporciona informes confidenciales que identifican las brechas de capacidad de cada participante y establecen un plan de mejora continua. Este modelo de entrenamiento técnico guarda correspondencia

funcional con las Pruebas de Penetración Basadas en Amenazas (TLPT) prescritas por el marco europeo DORA.

Para evaluar el impacto de los diversos instrumentos legales e internacionales en la gestión del riesgo tecnológico, la tabla 20 resume sus características operativas centrales en relación con las dimensiones de gobernanza, soberanía y resiliencia (Silva, 2025):

Tabla 20: Análisis comparativo de marcos regulatorios e instrumentos normativos

Marco Regulatorio e Instrumento Normativo	Ámbito y Alcance Institucional	Mecanismo de Gobernanza Digital	Enfoque en Soberanía Tecnológica	Exigencias de Resiliencia Operativa y Pruebas
Política Nacional de Transformación Digital al 2030 (DS 085-2023-PCM / DU 006-2020)	Perú / Sector Público, Empresas Estatales, Sector Privado, Academia.	Rectoría de la SGTD-PCM. Articulación mediante el Objetivo Prioritario 5 de Confianza Digital.	Fomento de la autonomía digital en servicios públicos y regulación ética de la inteligencia artificial.	Notificación obligatoria de incidentes en 48 h e integración en la red de CSIRTs liderada por el CNSD.
Ley y Reglamento de Ciberdefensa (Ley 30999 / DS 017-2024-PCM)	Perú / Fuerzas Armadas, Policía Nacional, Activos Críticos Nacionales (ACN).	Conducción a cargo del Presidente de la República, MINDEF y Comando Conjunto de las FF.AA.	Definición del ciberespacio como 5.0 dominio soberano. Mando del Comando de Operaciones Cibernéticas.	Ejecución de operaciones militares cibernéticas para proteger ACN y garantizar la Seguridad Nacional.
Reglamento SBS de Seguridad de la Información y Ciberseguridad (Res. SBS 504-	Perú / Sistema Financiero, Asegurador y Fondos de Pensiones supervisados.	Responsabilidad directa del Directorio y de los Comités. Designación formal del RSI.	Control del riesgo operativo derivado de proveedores TIC de terceros	Obligatoriedad de realizar pruebas de penetración, gestionar vulnerabilidad

2021)			y canales digitales.	es y planificar planes de respuesta.
Reglamento de Resiliencia Operativa Digital de la UE (DORA - Reglamento UE 2022/2554)	Unión Europea / Entidades Financieras y Proveedores Críticos de Servicios TIC.	Gobernanza y responsabilidad directa del órgano de administración sobre riesgos de TIC.	Mitigación activa del riesgo de concentración en hiperescaladores y fomento de software soberano.	Pruebas avanzadas de resiliencia mediante Threat-Led Penetration Testing (TLPT) periódicas.
NIST Cybersecurity Framework 2.0 (NIST CSF 2.0)	Global / Voluntario para todo tipo de organizaciones e infraestructuras.	Inclusión estratégica de la función *Govern* (GV) vinculada a Enterprise Risk Management.	Evaluación del riesgo en la cadena de suministro (*Supply Chain Risk Management*).	Adopción de arquitecturas Cero Confianza (SP 800-207) e integración de contención y respuesta.

Implicancias ecosistémicas, fricción de cumplimiento y perspectiva futura

El acelerado despliegue de este marco normativo genera transformaciones profundas en la operativa diaria de las organizaciones, planteando retos complejos de articulación y adaptabilidad. Uno de los principales desafíos para las organizaciones estriba en la gestión de la superposición de competencias normativas. Ante un ciberataque significativo que involucre la filtración de bases de datos, una entidad peruana del sector regulado debe gestionar de manera concurrente tres obligaciones normativas de notificación: informando al CNSD conforme al DS 126-2025-PCM, notificando a la ANPDP según el DS 016-2024-JUS para evitar multas severas,

e informando al regulador sectorial (como la SBS) en virtud de la Resolución SBS N° 504-2021 (Alcalá, 2025).

Esta disparidad en los formatos, requisitos técnicos y autoridades receptoras genera una intensa fricción operativa durante las primeras horas de un incidente, que es cuando los recursos técnicos de ciberseguridad deben concentrarse prioritariamente en la contención del ataque y el aislamiento de las redes afectadas. La falta de una ventanilla única de notificación nacional incrementa el riesgo de incumplimiento involuntario de los plazos de 48 horas, lo que expone a las organizaciones a contingencias legales paralelas a la crisis informática.

Gestión de riesgos en la cadena de suministro y estandarización técnica

Las regulaciones analizadas convergen en que la transferencia tradicional de riesgos tecnológicos mediante contratos de subcontratación carece de validez legal y operativa. En la actualidad, las instituciones son legalmente responsables de las brechas ocurridas en las plataformas de sus proveedores de TI.

Esta responsabilidad ha impulsado un proceso de estandarización técnica en todo el ecosistema empresarial. Las organizaciones contratantes exigen progresivamente a sus proveedores la certificación en la norma ISO/IEC 27001 o la implementación de controles alineados con ISO/IEC 27002. Asimismo, en la gestión documental y en los repositorios digitales, se exige el cumplimiento de la NTP 392.030-2:2015 sobre microformas digitales para garantizar la inalterabilidad jurídica y la protección de los datos custodiados. Para las pequeñas y medianas empresas que operan como proveedoras de grandes corporaciones o del Estado, la imposibilidad de certificar estos estándares se convierte en una barrera de entrada al mercado.

El impacto disruptivo de la inteligencia artificial en la ciberseguridad

El surgimiento y difusión de herramientas de Inteligencia Artificial (IA) añaden un vector de complejidad inédito a la resiliencia operativa. Por una parte, las soluciones defensivas utilizan algoritmos de aprendizaje automático para correlacionar grandes volúmenes de eventos, detectar anomalías de comportamiento y responder de forma automatizada ante intrusiones a velocidad de ejecución informacional.

Por otra parte, la IA es aprovechada por actores maliciosos para automatizar la ingeniería social, crear campañas de phishing altamente personalizadas y desarrollar malware adaptativo capaz de evadir firmas tradicionales de detección. En el Perú, la Ley N° 31814 busca promover el uso ético y seguro de la IA, pero la velocidad de evolución de los modelos algorítmicos sobrepasa los tiempos legislativos ordinarios. En consecuencia, la gobernanza de la IA exige que las organizaciones actualicen sus matrices de riesgo para garantizar que el entrenamiento de modelos no exponga información confidencial ni infrinja la Ley de Protección de Datos Personales (Ley N° 29733).

El análisis integral de la tríada compuesta por la confianza digital, la soberanía tecnológica y la resiliencia operativa evidencia que la ciberseguridad ha dejado de ser un centro de costos puramente técnico para transformarse en un imperativo estratégico de gobernanza y continuidad institucional. La alineación entre los marcos globales (NIST CSF 2.0 y DORA) y el ordenamiento jurídico peruano (Política Nacional de Transformación Digital al 2030, Ley de Ciberdefensa y regulaciones SBS) demuestra una convergencia inequívoca hacia la exigencia de responsabilidad directa de la alta dirección, transparencia en la gestión de riesgos y capacidades prácticas de respuesta operativa (Burga

et al., 2026).

Para operar eficazmente en este entorno normativo y de amenazas avanzadas, las entidades públicas y privadas deben estructurar sus capacidades operativas en torno a cuatro prioridades estratégicas:

i. Asimilación directiva de la función de gobernanza

Los órganos de administración y las gerencias ejecutivas deben integrar formalmente la gestión del riesgo tecnológico en sus esquemas de supervisión, conforme a los principios de NIST CSF 2.0 y la Resolución SBS 504-2021. La ciberseguridad no debe evaluarse aisladamente, sino como un elemento constitutivo de la continuidad de negocio y del valor institucional. Es imprescindible asegurar presupuestos sostenidos para la formación del talento digital interno, la contratación de oficiales de seguridad calificados y el mantenimiento de licencias y herramientas avanzadas de monitoreo (Burga et al, 2026).

ii. Rediseño hacia cero confianza y niveles mínimos de servicio viable

Las organizaciones deben abandonar definitivamente el modelo de seguridad perimetral para implementar arquitecturas de Cero Confianza (NIST SP 800-207) basadas en la microsegmentación y la autenticación continua. Paralelamente, es necesario definir de forma precisa los Niveles Mínimos de Servicio Viable (MVSL) para cada proceso crítico de la entidad. Esta definición técnica permite establecer prioridades de contención y aislamiento durante un incidente cibernético, evitando la paralización total de los servicios públicos o de las operaciones de negocio esenciales (El-hajj et al., 2023).

a. Fortalecimiento de la soberanía tecnológica y control de terceros

Las instituciones deben realizar un mapeo exhaustivo de sus dependencias tecnológicas, identificando los riesgos asociados a la concentración de servicios en la nube o soluciones propietarias de terceros. Se debe priorizar el despliegue de infraestructuras con garantías de soberanía de datos, esquemas autoalojados o híbridos, software abierto auditable y exigencias contractuales estrictas de cumplimiento de estándares de seguridad (ISO/IEC 27001, NTP 392.030-2) a todos los proveedores de la cadena de suministro (Nextware, s/f).

Las entidades deben articular protocolos claros de triaje de incidentes para cumplir de forma oportuna con los requerimientos de notificación dual al CNSD y a la ANPDP dentro del plazo legal de 48 horas. Asimismo, los procedimientos teóricos deben someterse a prueba mediante la participación continua en los Simulacros de Ataques Cibernéticos (SIMAC) convocados por la SGTD-PCM o la realización de ejercicios internos de tipo *Red Team / Blue Team* y pruebas TLPT. La capacidad de respuesta ante un ciberataque no se mide por la sofisticación de los manuales acumulados, sino por la velocidad de detección, la eficacia de la contención y la resiliencia demostrada por la organización en condiciones reales de disrupción.

Conclusión

La analogía de los modelos de IA agéntica con la criptografía resistente a algoritmos cuánticos como los algoritmos de Shor y Grover reconfigura estructuralmente la ciberseguridad avanzada. La IA agéntica trasciende la automatización estática mediante arquitecturas compuestas por percepción, memoria de largo plazo, herramientas de ejecución y ciclos de razonamiento dinámico. La transición desde algoritmos tradicionales de clave pública (RSA, ECC, Diffie-Hellman) hacia esquemas poscuánticos estandarizados por el NIST (tales como ML-KEM/Kyber, ML-DSA/Dilithium y SLH-DSA/Sphincs+) presenta desafíos severos en sobrecosto computacional, tamaño de claves y longitud de firmas. Los agentes de IA transforman esta implementación mediante:

- *Refactorización dinámica de código:* agentes especializados inspeccionan repositorios de código fuente legacy, identificando primitivas criptográficas vulnerables e inyectando wrappers poscuánticos adaptativos sin interrupción de servicio.
- *Optimización de parámetros de retículos:* agentes de optimización ajustan automáticamente los parámetros algebraicos en problemas como LWE (Learning With Errors) y Ring-LWE para entornos restringidos (IoT, dispositivos periféricos), equilibrando el margen de seguridad cuántica con el consumo energético y la latencia.
- *Ensamblado y compilación adaptativa:* generación autónoma de código en ensamblador con vectorización para acelerar la multiplicación polinomial en la Transformada de Number-Theoretic (NTT).

La IA agéntica rediseña la auditoría de seguridad criptográfica, pasando de escaneos pasivos a un criptoanálisis activo, coordinado y multiagente:

- *Fuzzing criptográfico dirigido por razonamiento:* agentes exploradores generan vectores de prueba heurísticos diseñados para explotar

esquemas de relleno, condiciones de borde y fallas matemáticas complejas en implementaciones de PQC.

- *Detección y explotación de canales laterales:* redes de agentes supervisan trazas de consumo de potencia y emisiones electromagnéticas en hardware que ejecuta PQC. Mediante modelos de aprendizaje profundo acoplados a agentes de toma de decisiones, detectan fugas de información en operaciones de re-encapsulamiento de claves en microsegundos.
- *Redes adversarias agénticas:* equipos de agentes en roles de ataque (Red Team) y defensa (Blue Team) simulan escenarios de criptoanálisis híbrido (combinando supercomputación clásica, estimaciones cuánticas y heurísticas agénticas) para evaluar la resiliencia temporal de los esquemas desplegados.

La criptoagilidad —la capacidad de sustituir algoritmos criptográficos sin alterar la infraestructura subyacente— deja de ser un marco teórico para convertirse en un proceso dinámico gobernado por agentes:

- *Inventario criptográfico dinámico (CBOM Agéntico):* agentes de descubrimiento mapean continuamente la Bill of Materials Criptográfica (CBOM) de sistemas distribuidos, catalogando claves, certificados y algoritmos en tiempo real.
- *Rotación autónomamente desencadenada:* ante el descubrimiento de una debilidad matemática o un avance analítico en un esquema poscuántico (e.g., abismos en algoritmos multivariados o basados en isogenias), agentes gobernadores ejecutan la migración instantánea e ininterrumpida a primitivas alternativas de reserva.

Se evidencia que los agentes de IA asumen autonomía operacional y se convierten en objetivos primarios para adversarios equipados con capacidades cuánticas. Para garantizar la seguridad del tejido agéntico frente a ataques

cuánticos presentes y futuros (incluyendo el paradigma "Harvest Now, Decrypt Later"), se establecen los siguientes requerimientos criptográficos esenciales:

- *Identidad criptográfica verificable (PQC-DID):* cada agente debe poseer un par de claves poscuánticas asociadas a Identificadores Descentralizados (DIDs). Para prevenir la suplantación por parte de adversarios cuánticos, el protocolo de autenticación interagente debe emplear esquemas de firma como ML-DSA o SLH-DSA.
- *Firmas compactas para protocolos de alta frecuencia:* dado que las firmas PQC poseen un tamaño sensiblemente mayor que RSA/ECC, se requiere la estandarización e integración de firmas post-cuánticas de alta velocidad y menor sobrepeso de ancho de banda para la comunicación inter-agente (Zero-Trust Inter-Agent Communication).
- *Cifrado PQC en reposo y en tránsito:* los canales de transmisión entre agentes y el almacenamiento de la memoria de largo plazo (vectores de embeddings, grafos de conocimiento, bases de datos de contexto) deben protegerse obligatoriamente mediante esquemas de encapsulamiento de claves poscuánticas (ML-KEM) combinados con cifrado simétrico AES-256 (resistente al algoritmo de Grover).
- *Computación confidencial y homomórfica poscuántica (PQC-FHE):* para permitir que agentes colaboradores procesen datos confidenciales sin desvelar su contenido, es imperativo integrar esquemas de Cifrado Totalmente Homomórfico (FHE) construidos sobre problemas de retículos (RLWE), los cuales son inherentemente resistentes a ataques cuánticos.
- *Entornos de ejecución confiables con raíz de confianza poscuántica:* los enclaves seguros donde residen los pesos del modelo y el motor de razonamiento agéntico deben basar su atestación remota en primitivas PQC fijadas en el hardware

- *Firmas de cadena de suministro de IA (PQC Model Provenance):* todos los artefactos de la IA agéntica —pesos del modelo, conjuntos de entrenamiento, prompts del sistema y herramientas/plugins— deben estar firmados criptográficamente mediante PQC en cada etapa del ciclo de vida de desarrollo e inferencia.
- *Registros de auditoría inmutables poscuánticos:* las cadenas de decisiones y la trazabilidad de acciones ejecutadas autónomamente por los agentes deben consignarse en registros distribuidos protegidos por funciones hash poscuánticas y esquemas de firma agregada PQC, impidiendo que un atacante cuántico reescriba la historia del comportamiento del agente.

La transición mundial hacia la criptografía poscuántica (PQC) supone el cambio más complicado y determinante que ha habido en la infraestructura de seguridad digital desde que existe la computación. La convergencia de la computación cuántica a gran escala, que puede llevar a cabo el algoritmo de Shor para factorizar números enteros grandes y calcular logaritmos discretos en un tiempo polinómico, anulará de forma permanente las bases esenciales de la criptografía asimétrica clásica, entre ellas RSA, Diffie-Hellman (DH), ECDH, ECDSA y la criptografía de curva elíptica (ECC).

En conclusión, la era poscuántica y la era de la inteligencia artificial autónoma no pueden analizarse como fenómenos aislados. La IA agéntica proporciona la inteligencia adaptativa necesaria para navegar la complejidad técnica de la migración poscuántica masiva. Al mismo tiempo, la adopción rigurosa de algoritmos de clave pública basados en retículos, hashes y multivariados (estandarizados por el NIST) resulta imprescindible para blindar la identidad, el estado de memoria y la capacidad de decisión de los agentes de IA frente a adversarios equipados con ordenadores cuánticos. La construcción de

un ecosistema digital soberano, seguro y autónomo depende de la armonización inmediata de ambas disciplinas.

Bibliografía

Acosta Villamil, C. (2025). *Introducción a la criptografía poscuántica* (Tesis de Maestría, Universidad Zaragoza). Repositorio Institucional – Universidad Zaragoza

Agencia de Seguridad Nacional (NSA). (7 de septiembre de 2022). *La NSA publica los requisitos futuros para algoritmos resistentes a la computación cuántica (QR) para sistemas de seguridad nacional*. Gobierno de los Estados Unidos. https://www.nsa.gov/Press-Room/News-Highlights/Article/Article/3148990/nsa-releases-future-quantum-resistant-qr-algorithm-requirements-for-national-se/

Alahmadi, AN, Rehman, SU, Alhazmi, HS, Glynn, DG, Shoaib, H., & Solé, P. (2022). Amenazas de ciberseguridad y ataques de canal lateral para la agricultura digital. *Sensors*, *22* (9), 3520. https://doi.org/10.3390/s22093520

Alcalá Molina, F.A. (2025). Cibercriminalidad: retos y desafíos para la administración de justicia peruana. *Revista Oficial del Poder Judicial*, *17*(23), 205-248. https://doi.org/10.35292/ropj.v17i23.906

Alquwayfili, A. (2025). Quantigence: A Multi-Agent AI Framework for Quantum Security Research. *arXiv:2512.12989 [cs.MA]*. https://doi.org/10.48550/arXiv.2512.12989

Burga Durango, D.W., Flores Cometivo, A.R., y Rivera Cochachi, K.M. (2026). *Recomendaciones basadas en NIST CSF 2.0 para determinar la capacidad de gestión y respuesta ante ciberriesgos en entidades públicas del Perú* (Tesis de Pregrado, Universidad Peruana de Ciencias Aplicadas). Repositorio Institucional – Universidad Peruana de Ciencias Aplicadas

Campbell, R. (2026). Enterprise Migration to Post-Quantum Cryptography: Timeline Analysis and Strategic Frameworks. *Computers*, *15*(1), 9. https://doi.org/10.3390/computers15010009

Castillo Serracín, J.R., Aranda, O., Farnum Castro, F., & Gómez Solís, J.M. (2026). Digital Frontiers: Cybersecurity, Sovereignty, and the New Paradigms of Geopolitical Power . *Revista Iberoamericana De educación*, *10*(2), 238–251. https://doi.org/10.31876/rie.v10i2.372

Cherkaoui Dekkaki, K., Tasic, I., & Cano, MD. (2024). Exploring Post-Quantum Cryptography: Review and Directions for the Transition Process. *Technologies*, *12*(12), 241. https://doi.org/10.3390/technologies12120241

Chow, M.C., & Ma, M. (2022). A Secure Blockchain-Based Authentication and Key Agreement Scheme for 3GPP 5G Networks. *Sensors*, *22*(12), 4525. https://doi.org/10.3390/s22124525

Delgado, L.J. y Romero, J.R. (2025). Inteligencia artificial y amenazas híbridas en defensa y seguridad: un análisis bibliométrico y de clústeres en la literatura científica (2000–2025), 23 (43), 125-145. https://doi.org/10.18041/1900-0642/criteriolibre.2025v23n43.13701

Dziechciarz, D., & Niemiec, M. (2025). Efficiency Analysis of NIST-Standardized Post-Quantum Cryptographic Algorithms for Digital Signatures in Various Environments. *Electronics*, *14*(1), 70. https://doi.org/10.3390/electronics14010070

EIOPA. (2025). Ley de Resiliencia Operativa Digital (DORA). *EIOPA.* https://www.eiopa.europa.eu/digital-operational-resilience-act-dora_en

El-hajj, M., Mousawi, H., & Fadlallah, A. (2023). Análisis de algoritmos criptográficos ligeros en plataformas de hardware IoT. *Future Internet* , *15* (2), 54. https://doi.org/10.3390/fi15020054

Eum, S., Lee, M., & Seo, H. (2024). Optimizing HAWK Signature Scheme Performance on ARMv8. *Applied Sciences*, *14*(19), 8647. https://doi.org/10.3390/app14198647

Hernández Fuentes, A.P. (2022). Cooperación digital y soberanía tecnológica

para cerrar la brecha digital en la cuarta revolución industrial. *OASIS*, (36), 77–94. https://doi.org/10.18601/16577558.n36.06

Hiperderecho. (8 de agosto de 2023). *Gobierno aprobó Política Nacional de Transformación Digital: ¿en qué consiste?*. Hiperderecho. https://hiperderecho.org/2023/08/gobierno-aprobo-politica-nacional-de-transformacion-digital-en-que-consiste/

Ivezic, M. (2026). *The Applied Quantum PQC Migration Framework & Methodology*. PQC Framework. https://pqcframework.com/wp-content/uploads/2026/06/applied-quantum-marin-ivezic-pqc-migration-framework-universal-2-1.pdf

Kahn, C. (27 de enero de 2020). *La nueva sensación: Confianza Digital para interacciones en el Entorno Digital.* IUS360 - El portal jurídico de IUS ET VERITAS. https://ius360.com/la-nueva-sensacion-confianza-digital-para-interacciones-en-el-entorno-digital/

Lamm, H. (28 de enero de 2026). Why Enterprises Should Start Establishing a Cryptography Bill of Materials (CBOM) Now. *Sectigo*. https://www.sectigo.com/blog/cryptographic-bill-of-materials-cbom

Lee, S., Huh, J.H., & Woo, H. (2025). Security System Design and Verification for Zero Trust Architecture. *Electronics*, *14*(4), 643. https://doi.org/10.3390/electronics14040643

Mandarino, V., Pappalardo, G., & Tramontana, E. (2026). dAuth: A Hybrid Smart Contract-Based Architecture for Decentralized Authentication with Institutional Attestation. *Computers*, *15*(6), 398. https://doi.org/10.3390/computers15060398

National Security Agency (2024). The Commercial National Security Algorithm Suite 2.0 and Quantum Computing FAQ. https://media.defense.gov/2022/Sep/07/2003071836/-1/-1/0/CSI_CNSA_2.0_FAQ_.PDF

Nextware. (s/f). Identificación y mitigación de riesgos en infraestructuras IT. *Nextware.* https://nextware.com.ar/blog/mitigacion-riesgos-it/

Ngema, N., Nleya, B., & Maswanganyi, R.C. (2026). A Review of Zero Trust Architecture: Principles, Applications, and Implementation Challenges in Communication, Navigation, and Surveillance (CNS) Systems. *Sensors*, *26*(12), 3813. https://doi.org/10.3390/s26123813

NIST. (13 de agosto de 2024). *Estándar de mecanismo de encapsulación de claves basado en retícula modular*. Gobierno de los Estados Unidos. https://csrc.nist.gov/pubs/fips/203/final

NIST. (29 de octubre de 2020). *Recomendación para esquemas de firma basados en hash con estado: NIST SP 800-208*. NIST. https://csrc.nist.gov/news/2020/stateful-hash-based-signature-schemes-sp-800-208

NIST. (3 de enero de 2017). *Criptografía Postcuántica (PQC)*. NIST. https://csrc.nist.gov/projects/post-quantum-cryptography/faqs

Otero, C. (9 de abril de 2026). Nuevo Reglamento de Protección de Datos (D.S. 016-2024-JUS): Todo lo que cambió. *KOM agencia digital.* https://kom.pe/reglamento-ds-016-2024-jus-datos-personales/

Płoszaj-Mazurek, M., & Ryńska, E. (2024). Artificial Intelligence and Digital Tools for Assisting Low-Carbon Architectural Design: Merging the Use of Machine Learning, Large Language Models, and Building Information Modeling for Life Cycle Assessment Tool Development. *Energies*, *17*(12), 2997. https://doi.org/10.3390/en17122997

POSTQUANTUM. (2026). 120,000 Tasks: Why Post-Quantum (PQC) Migration Is Enormous. *POSTQUANTUM.* https://postquantum.com/post-quantum/quantum-security-pqc-program-plan/

Rahayu, M., Hossain, M.B., Huda, S., & Nogami, Y. (2025). Integrated Authentication Server Design for Efficient Kerberos–Blockchain VANET

Authentication. *Sensors*, *25*(21), 6651. https://doi.org/10.3390/s25216651

Rama Lingeswara, S.T. (2019). Integrado KRIs and KPIs ara una gestión eficaz del riesgo tecnológico. *ISACA,* 4. https://www.isaca.org/es-es/resources/isaca-journal/issues/2018/volume-4/integrating-kris-and-kpis-for-effective-technology-risk-management

Riquelme, B.M.A., & Pereyra, J.A. (2025). Ciberseguridad Post-Cuántica en América Latina: Desafíos y Estrategias Regulatorias. *Perspectivas*, (12). https://revistas.ucalp.edu.ar/index.php/Perspectivas/article/view/425

Rojek, I., Kotlarz, P., & Mikołajewski, D. (2026). AI-Based Autonomous Security for Cyber–Physical Systems 2.0 in IoT Ecosystems—A Narrative Review. *Electronics*, *15*(15), 3339. https://doi.org/10.3390/electronics15153339

Santos Alonso, J.M. (2020). La soberanía tecnológica de la gestión de las tecnologías de lainformación y la Política Tecnológica. *Serie Científica de la Universidad de las Ciencias Informáticas*, 13(6), 15-24. https://dialnet.unirioja.es/descarga/articulo/8590276.pdf

Silva Casas, C.Y. (2025). La gobernanza digital en América Latina: Una mirada desde la experiencia europea. *Política Internacional*, (138), 143–159. https://doi.org/10.61249/pi.vi138.246

Tanase, O. (2026). Las naciones del mundo buscan asegurar su soberanía digital en la era de la IA. *Revista Innovación.* https://revistainnovacion.com/nota/12875/las_naciones_del_mundo_buscan_asegurar_su_soberania_digital_en_la_era_de_la_ia/

Telefónica. (13 de enero de 2026). *Ordenadores cuánticos criptográficamente relevantes: impacto y preparación para la seguridad digital global*. Telefónica Tech. https://telefonicatech.com/blog/ordenadores-cuanticos-criptograficamente-relevantes-impacto-preparacion

Wang, J., Ouyang, Q., and Huanyu, W. (2026). A Systematic Evaluation of Deep-

Learning Side-Channel Attacks: Performance and Cost. In Proceedings of the 2026 5th International Conference on Cryptography, Network Security and Communication Technology (CNSCT '26). *Association for Computing Machinery,* New York, NY, USA, 149–155. https://doi.org/10.1145/3802927.3802950

Zia, M., Iqbal, R., and Ramasubbu, N. (2026). Governing Technical Debt in Agentic AI Systems. *arXiv:2605.29129 [cs.AI].* https://doi.org/10.48550/arXiv.2605.29129

De esta edición de "*Inteligencia artificial agéntica y criptografía poscuántica*", se terminó de editar en la ciudad de Colonia del Sacramento en la República Oriental del Uruguay el 17 de abril de 2026

Editorial Mar Caribe

Inteligencia artificial agéntica y criptografía poscuántica

ISBN: 978-9915-698-90-8

Colonia, Uruguay
Abril, 2026