PREPRINT • ILLUSTRATED EDITION OF A PUBLISHED REVIEW ARTICLE

# Never Give Your Agent Your Key:

## Delegation and Revocation for AI Agents in Post-Quantum Decentralised Identity

**Petar Radanliev** Department of Computer Science, University of Oxford, United Kingdom
ORCID 0000-0001-5629-6857 • petar.radanliev@cs.ox.ac.uk

**Published version, please cite:** Radanliev P. (2026). Never give your agent your key: delegation and revocation for AI agents in post-quantum decentralised identity. *Frontiers in Blockchain* 9:1905568. doi:10.3389/fbloc.2026.1905568

**About this edition.** The published article is open access under the Creative Commons Attribution 4.0 licence. This edition is the author's own adaptation of it: the text has been rewritten for a wider technical readership and every figure is new. The argument, the quantitative values and the reference list follow the published article. No new data, results or claims are introduced. References corrected at proof stage are given here in their corrected form.

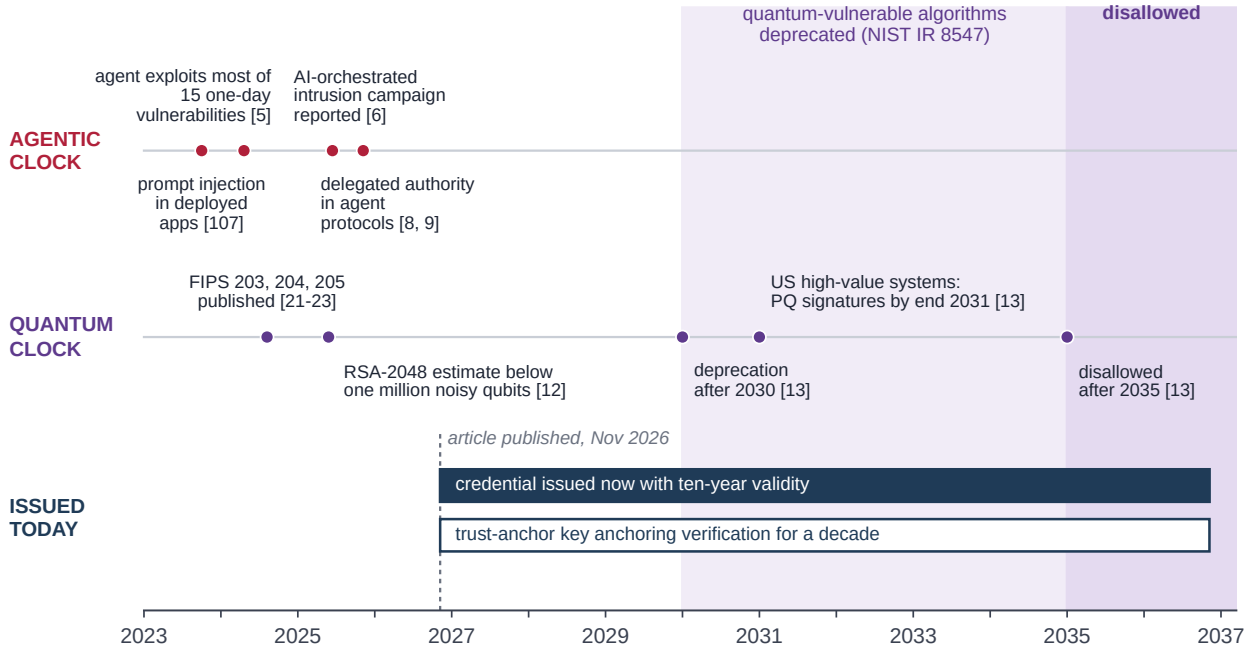



**Figure 1.** Two clocks run against decentralised identity. Agentic capabilities are documented now; the retirement dates for quantum-vulnerable public-key cryptography are fixed. A credential or trust-anchor key created today is still in use when those dates arrive. Dates and events as cited in the text.

**×61** larger identifier keys under ML-DSA-65

**×52** larger credential signatures

**2,953** bytes: the most one QR code holds, less than one ML-DSA-65 signature

**3.7 kB** added to a presentation by each delegation hop

**0** security properties that depend on an AI component

## Abstract

Autonomous AI agents now act on every side of digital identity infrastructure. They synthesise the faces and voices that defeat static biometric checks, run credential and Sybil attacks at machine speed, and increasingly hold and present credentials on behalf of the people who delegate to them. At the same time, the public-key cryptography beneath decentralised identity is exposed to a future quantum adversary, and credentials issued today will remain in use into the 2030 to 2035 deprecation window that NIST has now finalised. This review synthesises the literature and standards on decentralised identifiers, verifiable credentials, post-quantum cryptography, privacy-enhancing technologies and agentic AI security into a post-quantum deployment profile, and makes five contributions. First, it maps NIST FIPS 203 to 205 onto identity lifecycle operations at the level of parameter sets, wire formats and registered algorithm identifiers, and specifies how hash-based STARK predicate proofs bind to ML-DSA-signed credentials. Second, it quantifies the resulting overhead from normative constants and published benchmarks: identifier keys grow sixty-one-fold and credential signatures fifty-two-fold, which rules out single-QR presentation and on-chain signature anchoring. Third, it specifies cryptographic agility, a six-state key lifecycle and a three-phase hybrid transition aligned with current federal and European guidance. Fourth, it shows that accumulator revocation over RSA or pairing groups cannot be carried into a quantum-resilient design, and that the signature, not the status bitstring, dominates the cost of status lists. Fifth, it defines a bounded-autonomy model in which agents act as adversaries, delegates and operators: agents may observe, recommend and take quota-limited reversible actions, whereas revocation, key management and algorithm changes require keys held by accountable principals, so that no security property depends on an AI component. An STPA-Sec analysis supports the model. Delegation is carried by credentials issued to the agent's own identifier, never by sharing the person's key, and each delegation hop adds 3.7 kB of post-quantum signatures and keys. The result is an implementable blueprint, together with the validation programme that deployment still requires.

**Keywords:** agentic AI; decentralised identity; post-quantum cryptography; ML-DSA; verifiable credentials; delegated authority; credential revocation; zero-knowledge proofs

**HOW TO READ THIS EDITION**

One example runs through the whole document: a person asks a personal AI agent to rent a car, and the rental desk must confirm that a valid driving licence exists and that the driver is over 21, without learning anything else. Each section opens with the next step of that example, explains the mechanism behind it, and then states what the published article specifies. Grey boxes carry the example; blue boxes restate the point in plain terms; amber boxes say what is not claimed. Every number is either a constant taken from a published standard, a value reported in a cited benchmark, or arithmetic on those values that the text states. Section numbers differ from the journal version; the reference numbers are the same.

Published version: Frontiers in Blockchain 9:1905568, doi:10.3389/fbloc.2026.1905568

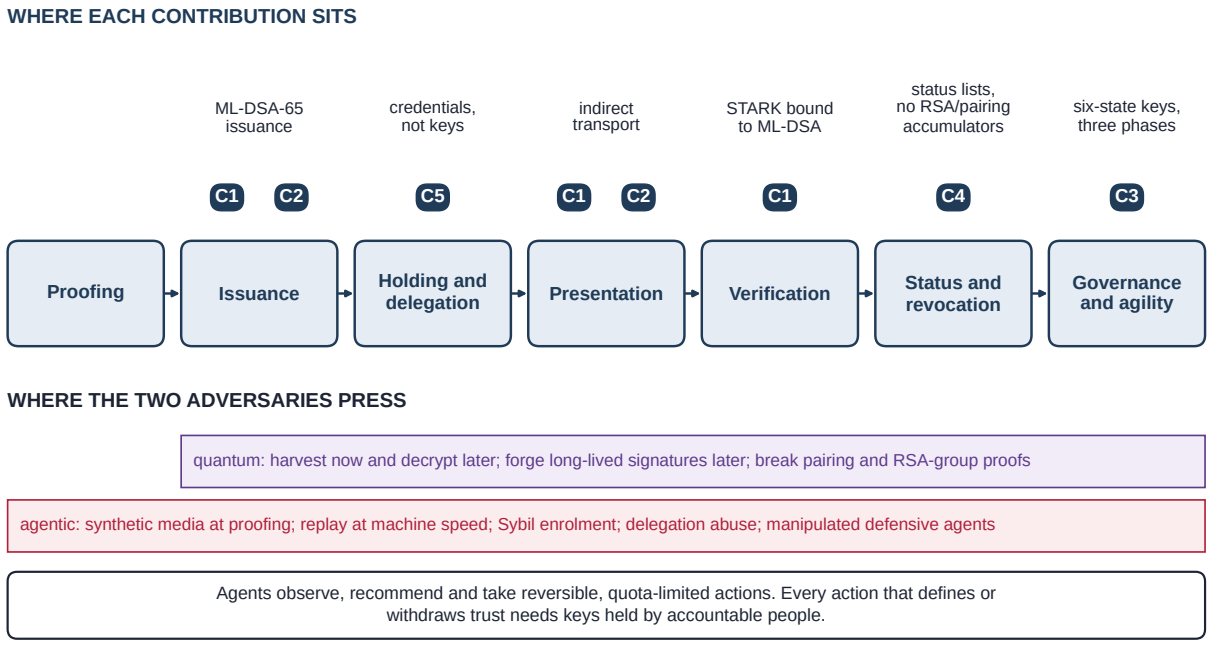



**Figure 2.** The paper on one page. The identity lifecycle runs left to right. Above it, the contribution that governs each stage (C1 to C5); below it, where each adversary presses. The bottom line is the design rule that the rest of the paper makes precise.

## 1 A car, a licence and an agent

**THE RUNNING EXAMPLE**

*A person asks a personal agent to rent a car for Saturday. The rental desk will not hand over keys until it knows two things: that a valid driving licence exists, and that the driver is over 21. It has no business knowing the driver's date of birth, address or licence number. The agent is software. It can be copied, redirected by text hidden in a web page, or impersonated by another agent. And the signatures that make the licence trustworthy today rest on mathematics that a large quantum computer would break.*

That small transaction contains the whole problem the paper addresses. Two developments bear on digital identity at once, and they are usually analysed apart. The first is agentic AI: systems that plan a sequence of actions, carry them out against live interfaces through tools, observe the outcome and adapt, without a person approving each step [1, 2]. Such systems already sit on every side of an identity transaction. As adversaries they synthesise faces and voices that defeat static biometric and liveness checks [3, 4] and exploit software vulnerabilities without human direction [5]; one developer has reported an intrusion campaign in which most tactical operations were carried out autonomously [6]. As delegates they are beginning to hold and present credentials for the people who instruct them [7–9]. As operators they are proposed for monitoring and triage of identity infrastructure itself [2, 10].

The second development is the prospect of a cryptanalytically relevant quantum computer. Shor's algorithm breaks the RSA and elliptic-curve schemes on which decentralised identifiers and verifiable credentials rest [11]; published estimates of the resources needed have fallen sharply [12]; and NIST has fixed a schedule under which quantum-vulnerable public-key algorithms are deprecated after 2030 and disallowed after 2035 [13].

Figure 3 follows the rental from request to verification. Each panel corresponds to a mechanism that the paper specifies, and the rest of this edition takes them in turn.

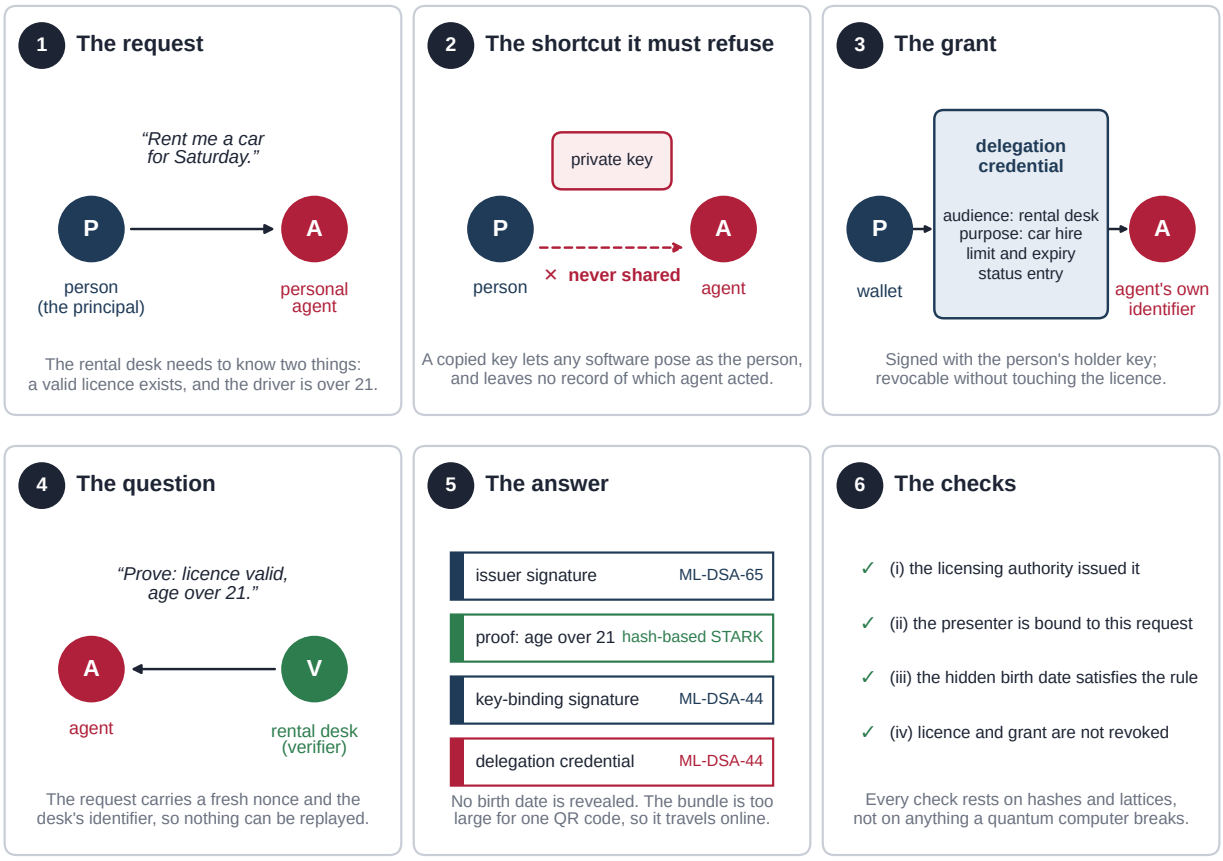



**Figure 3.** The running example in six panels. The agent never receives the person's key (2); it acts under a delegation credential issued to its own identifier (3); the verifier binds its request to a fresh nonce (4); the answer proves the age condition without revealing the birth date (5); and every check rests on hash-based or lattice-based assumptions (6).

The paper is a review-derived architectural study. From the literature and standards on decentralised identity, post-quantum cryptography, privacy-enhancing technologies and agentic AI security, it derives a deployment profile specified down to parameter sets, wire formats and registered algorithm identifiers. It makes five contributions:

**C1** A mapping of the NIST post-quantum standards onto every identity lifecycle operation, including how a hash-based predicate proof binds to an ML-DSA-signed credential (Sections 6 and 7).

**C2** A quantified account of what that mapping costs in identifier documents, credentials, presentations, delegation chains, revocation artefacts and ledger anchoring (Section 6).

**C3** Cryptographic agility as a mechanism, a six-state key lifecycle and a three-phase hybrid transition (Section 8).

**C4** An analysis of revocation under post-quantum constraints, which rules out a widely used mechanism and separates revocation from erasure (Section 9).

**C5** A bounded-autonomy model that separates the three roles agents occupy, specifies what an agentic operator may do, supports it with an STPA-Sec analysis and prices agent delegation (Section 10).

**Agents propose; keys dispose.**

That four-word principle governs the design. An agent never receives the key of the person it acts for, and never holds an issuer, root or governance key. Agents may observe, recommend and take bounded, reversible actions, but every action that defines or withdraws trust is taken under a cryptographic key held by an accountable principal. No security property of the architecture depends on an AI component.

## 2 Two clocks are running

### 2.1 Where decentralised identity stands today

Centralised identity systems concentrate sensitive data and decision authority in a few providers, which makes them attractive targets and single points of failure [25–27]. Identity fraud and theft carry large and persistent costs [28, 29], and concentrated biometric data adds harms that reissuing a credential cannot remedy [30, 31]. The European Digital Identity framework under eIDAS 2.0 [32] and India's Aadhaar [33] show two responses: a federated, wallet-centred model with strong legal assurance, and a centralised biometric registry whose privacy and security have been extensively questioned [25, 26]. Self-sovereign identity proposes instead that the subject holds credentials and discloses them selectively [34–37], and a large literature surveys its components, specifications and governance [38–45], including the needs of organisations [46].

Its technical core is standardised. Decentralised identifiers resolve to documents carrying verification methods [47]; verifiable credentials follow the W3C data model [48]; selective disclosure for JSON Web Tokens is defined in RFC 9901 [49]; and issuance and presentation are specified by OpenID for Verifiable Credentials [50]. Governance and layering are addressed by the Trust over IP stack and by mappings between eIDAS and self-sovereign trust frameworks [51, 52]; DIDComm provides messaging between the parties' agents and has been analysed formally [53]; and protocol work integrates these pieces with OpenID Connect and TLS [54, 55] and supports offline access control [56].

Deployed systems rest on classical cryptography throughout. Hyperledger Indy and Sovrin anchor identifiers on permissioned ledgers and use Camenisch–Lysyanskaya signatures over RSA groups for anonymous credentials [57–62], with a cryptographic toolbox built entirely on classical primitives [63]; earlier Ethereum-based designs such as uPort did the same [64]. Zero-knowledge identity schemes on blockchains and smart-contract verification of anonymous credentials rest on discrete-logarithm or pairing assumptions [65, 66], and performance studies of self-sovereign identity stacks measure only classical primitives [67, 68]. Security analyses find weaknesses in key management, recovery and wallets rather than in the primitives [69–71], because the primitives have so far been assumed sound. Three limitations recur: reliance on classical primitives, agility treated as an aspiration, and weak integration of privacy-preserving computation into verification [42, 72, 73].

Work on the quantum transition for blockchains and public-key infrastructures, which carry their own vulnerability classes [97], has identified the same problem at the level of ledgers and certificates [98–100], building on lattice signatures such as BLISS [82]. Work on post-quantum credentials themselves is recent: analyses of the quantum safety of current credential schemes [17], lattice-based zero-knowledge credential designs and a first implementation [18, 19], and a hybrid identifier and presentation implementation using composite verification methods [20]; biometric key derivation with post-quantum primitives addresses a neighbouring problem [103]. What the literature does not

yet provide is a mapping of the standardised primitives onto every operation of an identity lifecycle, with its costs.

### 2.2 The quantum clock: dates are now fixed

Shor's algorithm factors integers and computes discrete logarithms in polynomial time on a fault-tolerant quantum computer, which breaks RSA, finite-field Diffie–Hellman and elliptic-curve cryptography [11, 76]. NIST's response was a public standardisation process [77–79], whose first outputs were published in August 2024 [21–23]:

- **ML-KEM** (FIPS 203), a module-lattice key-encapsulation mechanism derived from CRYSTALS-Kyber [80], for agreeing keys;
- **ML-DSA** (FIPS 204), a module-lattice signature derived from CRYSTALS-Dilithium [81], for signing;
- **SLH-DSA** (FIPS 205), a stateless hash-based signature derived from SPHINCS+ [83].

Their security rests on the hardness of learning-with-errors and short-integer-solution problems over module lattices, and on properties of hash functions [84–88]. HQC was selected in March 2025 as a code-based backup key-encapsulation mechanism [89]; FN-DSA, derived from Falcon [90], remains in preparation [91]. The field has also narrowed by cryptanalysis: SIKE, an isogeny-based candidate that reached the fourth round, fell to a classical key-recovery attack [92], and no isogeny-based scheme is standardised.

The timing of a quantum computer is uncertain; the structure of the threat is not. Traffic recorded today can be decrypted once such a machine exists, so data whose confidentiality must outlast that date is already exposed. Mosca's inequality makes the point exact: if the time data must stay secret plus the time migration takes exceeds the time until a quantum computer arrives, the data is already at risk [14, 93, 94]. Published estimates for factoring 2048-bit RSA have fallen to fewer than a million noisy qubits running for under a week [12]. The final NIST IR 8547 deprecates quantum-vulnerable public-key algorithms after 2030 and disallows them after 2035; United States federal policy requires high-value assets and high-impact systems to adopt post-quantum key establishment by the end of 2030 and post-quantum signatures by the end of 2031 [13]; European and United Kingdom roadmaps set comparable milestones [95, 96].

> **IN PLAIN TERMS**
>
> Identity credentials live a long time. A credential issued in 2026 under an Ed25519 or ECDSA signature may still be presented after 2030, and a trust-anchor key that signs governance documents today may anchor verification for a decade [14]. Figure 1 on the cover shows the overlap. Migration cannot wait for the machine to appear.

### 2.3 The agentic clock: capabilities are already documented

Agentic systems combine a foundation model with a harness that assembles context, calls tools, executes actions and feeds the results back to the model [1,2,104]. The security literature on them has grown from surveys of their attack surface [2, 104] to taxonomies of agent-specific threats, in which identity and privilege abuse is ranked among the ten principal risks of agentic applications [105, 106], and to analyses of indirect prompt injection, in which instructions hidden in retrieved content redirect an agent [107].

Three capabilities bear directly on identity. First, generative models synthesise faces and voices that defeat static biometric verification, and injection of synthetic media into remote proofing is documented as an operational attack [3, 4, 108–110]; detection remains an arms race in which generators keep the initiative [111–113]. Second, a language-model agent exploited most of a set of fifteen real one-day vulnerabilities when given their public descriptions [5], and a 2025 report describes an intrusion campaign in which an AI system performed most tactical operations autonomously [6]. This changes the economics of attacks on identity rather than their nature [114, 115]: credential stuffing, replay and relay can run continuously, and the marginal cost of a fake (Sybil) identity falls. Sybil defences that infer honesty from social-graph structure assume that convincing identities are expensive to create [116–120], an assumption that automated fabrication weakens [121, 122]. Third, agents are becoming principals in their own right. Proposals for agent identifiers [123, 124], authenticated delegation [7] and agent identity management [125] converge on binding an agent to a principal and to a bounded authority, which tool-access and payment protocols now carry as OAuth tokens or signed mandates [8, 9, 126, 127], and which early designs express with decentralised identifiers and verifiable credentials [128–130].

AI is also widely proposed as a defence: machine learning for fraud and anomaly detection [10, 131], reinforcement learning for adaptive security [132–134], learned components for self-adaptive monitoring [135–137], with risk frameworks extended to AI itself [138, 139]. The evidence for such defences under attack is thin. Learned detectors are evaluated on data that rarely contains the adaptive attacks they are meant to resist [16, 140], and an AI component on the security path becomes a new attack surface with failure modes that are hard to bound [15, 107].

> **IN PLAIN TERMS**
>
> Answering an autonomous attacker with an autonomous defender whose behaviour is equally unvalidated offers no security argument [15, 16]. The paper therefore puts security in cryptography that neither kind of adversary can defeat, and gives agents only roles in which their failure cannot break a security property.

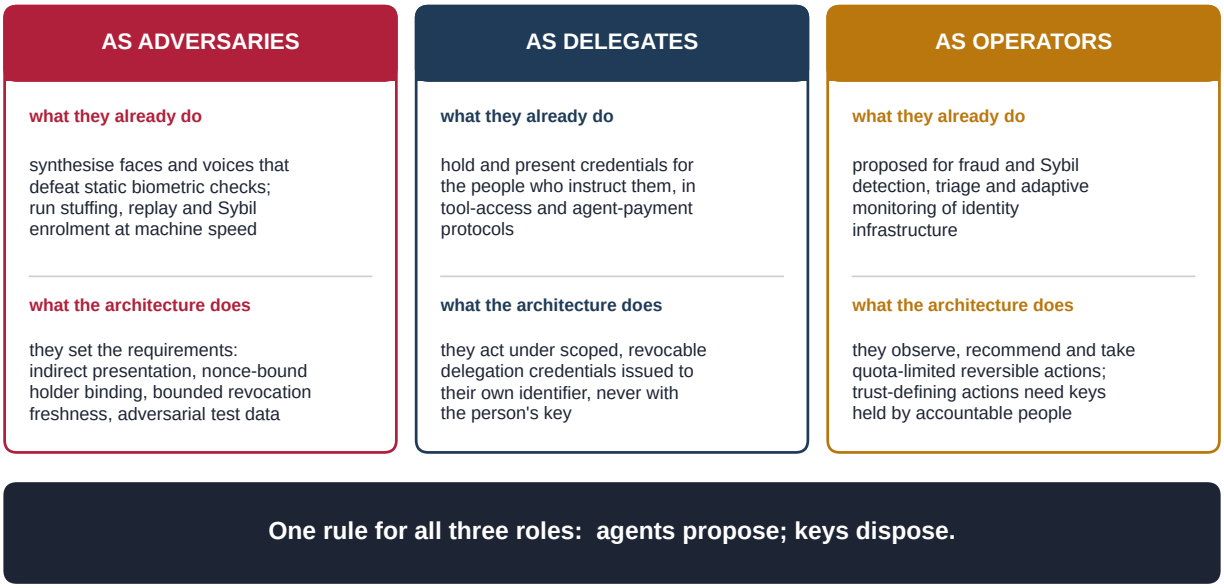



**Figure 4.** Agents occupy three roles in an identity system, and the architecture treats each differently. The asymmetry between well-documented agentic offence and weakly validated agentic defence is why AI is admitted only as a bounded operator.

## 3 How the evidence was gathered

The study is a structured narrative review informed by PRISMA principles [24]. Five research questions guided it: how the NIST post-quantum standards can be integrated into decentralised identity (RQ1); where existing systems fall short on agility and quantum resilience (RQ2); how privacy-enhancing technologies combine with post-quantum primitives (RQ3); which architectural patterns emerge for scalable, interoperable, regulation-compliant identity (RQ4); and where autonomous agents can act without any security property coming to depend on them (RQ5).

Literature was collected from IEEE Xplore, the ACM Digital Library, Scopus, Web of Science and arXiv, using two Boolean expressions: a post-quantum clause paired with an identity clause, and the same post-quantum clause paired with a privacy-enhancing clause. Each clause was also run without the post-quantum restriction, so that earlier identity and privacy work was not excluded by construction. Publications from 2010 to 2025 were searched, with foundational earlier works included where a mechanism originates. Standards were obtained directly from NIST, the IETF, the W3C, the OpenID Foundation and ISO. Because the literature on agent identity and delegation largely post-dates the main search, it was captured in 2026 by a targeted update search.

A source was included if it specified or evaluated a cryptographic primitive, protocol or architecture; reported measurements bearing on deployment cost; defined a standard or regulatory instrument; or was the primary source for a mechanism used in the architecture. It was excluded if it had no cryptographic or architectural content, restated a primitive without adding anything, or had been superseded by a later version of the same standard.

**WHAT THIS DOES NOT CLAIM**

Records were screened in three stages, but the number retained at each stage was not recorded prospectively. The study is therefore reported as a structured narrative review, not a PRISMA-conformant systematic review; no flow diagram is given and no claim of record-level reproducibility is made. What can be checked is the cited corpus, which is the reference list of this document.

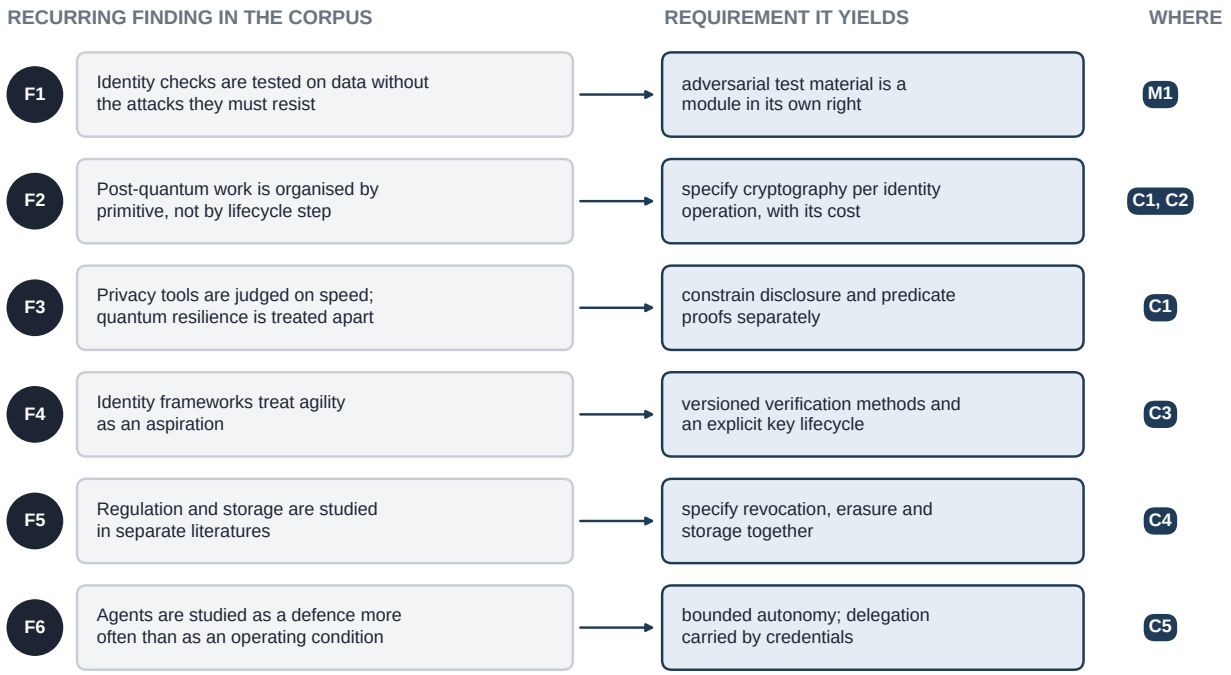



**Figure 5.** From findings to requirements. Each recurring finding in the corpus (F1 to F6) becomes a requirement and is answered by a module (M1) or a contribution (C1 to C5).

The step from review to architecture is made explicit, because a review that leaves it implicit cannot be told apart from a bibliography. Six findings recurred across the literature, and each yields a requirement (Figure 5).

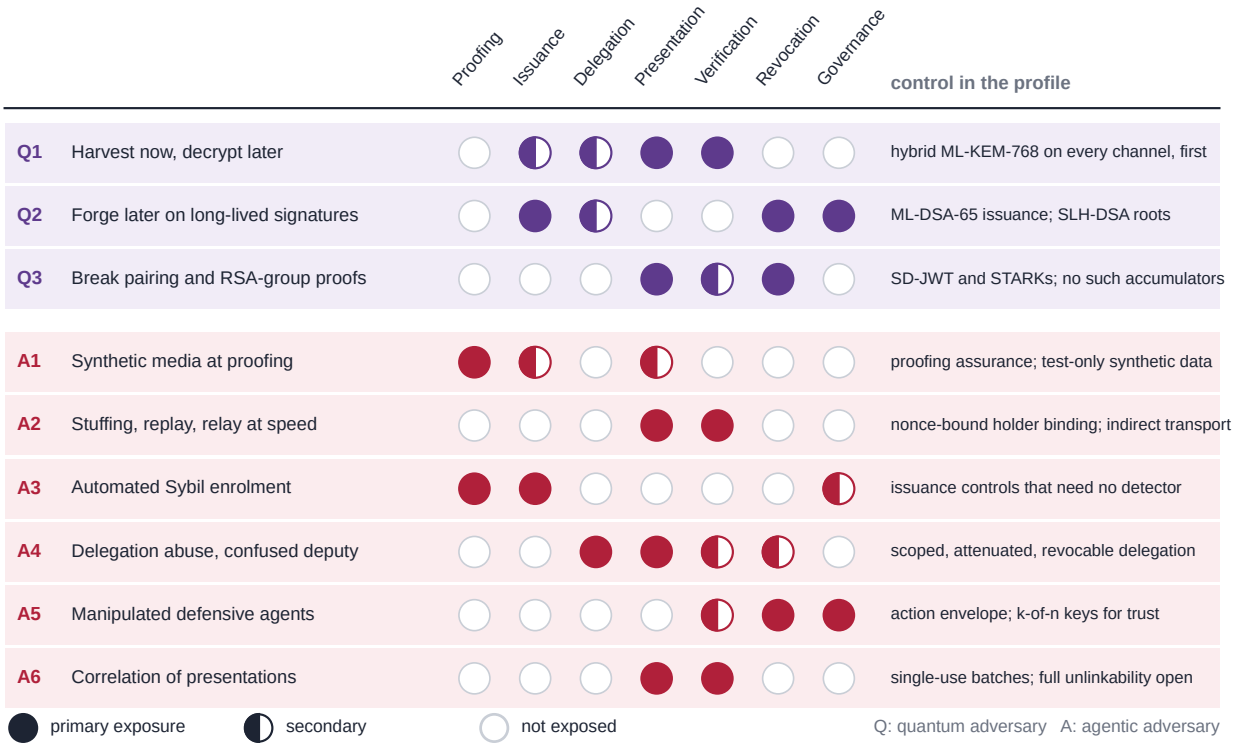



**Figure 6.** Where each threat lands on the identity lifecycle, and the control in the deployment profile that answers it. Q: quantum adversary; A: agentic adversary.

## 4 Who is attacking, and where

**THE RUNNING EXAMPLE**

*Two attackers watch the rental. One is patient: it records the encrypted exchange today and waits for a machine that can open it, or forge the licensing authority's signature years from now. The other is fast: it can generate a convincing face for a remote identity check, replay a stolen presentation thousands of times an hour, or plant an instruction in a web page that the person's agent will read.*

The threat model distinguishes two external adversaries and two internal failure modes [141] (Figure 6). The **quantum adversary** records traffic now and decrypts it later, forges signatures under quantum-vulnerable keys once it can, and breaks constructions whose security rests on factoring, discrete logarithms or bilinear pairings. For pairing-based proofs and credentials its effect falls on soundness and unforgeability: a system whose credentials or proofs can be forged is not quantum-resilient, even where the zero-knowledge property survives. The **agentic adversary** operates at machine speed and adapts: it injects synthetic media into proofing, drives presentation interfaces with stolen or replayed credentials, fabricates identities to enrol Sybils, and targets defensive agents through manipulated inputs [105, 107]. The two internal failure modes are a delegated agent that exceeds or abuses its authority, and an agentic operator whose judgement has been manipulated.

Five security properties are required: authenticity of issued credentials (P1); holder binding, so that only the subject or an agent with valid delegated authority can present (P2); soundness of predicate proofs (P3); integrity and bounded freshness of revocation status (P4); and minimal disclosure of attributes (P5). Unlinkability across presentations is a sixth property that the architecture identifies but does not deliver by default (Section 7). P1 to P5 must hold under arbitrary behaviour of any agent in the system. Issuer keys are assumed to sit in hardware security modules and holder keys in

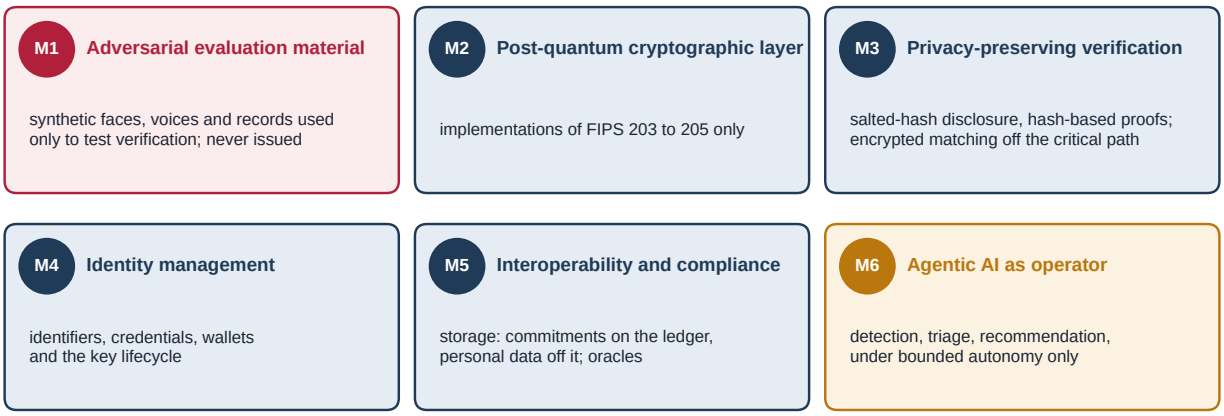



**Figure 7.** The six modules of the reference architecture. Module 2 admits only implementations of the finalised standards; Module 6 admits AI only under the bounded autonomy of Section 10.

a secure element; verifiers are honest but curious, and collusion between verifiers, or between an issuer and a verifier, is considered for linkability.

## 5 The architecture in one picture

The architecture has six functional modules (Figure 7); their composition, rather than their internal design, is the subject of the paper. Synthetic identities in Module 1 are test material only: they are never issued as credentials, bound to a subject or accepted as evidence of identity [108, 142, 143]. Module 5 keeps commitments on the ledger and personal data off it, with oracles where claims must be checked against external registers [144–148], on public or permissioned ledgers [149, 150].

**FOUR CONFUSIONS THE PAPER CLEARS UP**

1. **Homomorphic encryption is not post-quantum signing.** Libraries such as Microsoft SEAL, PALISADE and HElib implement the BFV, BGV and CKKS schemes [152–158]. They rest on ring learning-with-errors and are plausibly quantum-resistant, but they compute on ciphertexts; they provide neither key establishment nor signatures. Post-quantum primitives come from implementations of the FIPS standards, prototyped with libraries such as liboqs [159] and validated in production under NIST's module validation programme [13].
2. **Isogeny-based schemes are out.** SIKE was broken [92], and no other isogeny-based scheme is standardised.
3. **Running code on a quantum computer does not show quantum resilience.** Present devices cannot run Shor's algorithm at cryptographic scale, and no simulation measures the security of a lattice scheme against a quantum adversary. The argument rests on hardness assumptions and public cryptanalysis [13, 78, 79], with parameters assessed by lattice estimation [160], conformance tested against published vectors and side channels evaluated on target hardware [101].
4. **Indistinguishability obfuscation is not part of it.** No practical construction exists, and nothing in the profile depends on it.

**WHAT THIS DOES NOT CLAIM**

Modules 1 to 6 are a specification of proposed work. No component has been implemented or evaluated for this paper, no model trained, no benchmark executed and no prototype of the composed architecture

**Table 1.** Normative sizes of cryptographic material, in bytes. Post-quantum values from FIPS 203 Table 3, FIPS 204 Table 2 and FIPS 205 Table 2 [21–23]; classical values are the standard encodings.

| Scheme | Public or encapsulation key | Private key | Signature or ciphertext |
|---|---|---|---|
| Ed25519 | 32 | 32 | 64 |
| ECDSA P-256 | 65 (uncompressed) | 32 | 64 |
| RSA-2048 | 256 | — | 256 |
| ML-KEM-512 | 800 | 1,632 | 768 |
| ML-KEM-768 | 1,184 | 2,400 | 1,088 |
| ML-KEM-1024 | 1,568 | 3,168 | 1,568 |
| ML-DSA-44 | 1,312 | 2,560 | 2,420 |
| ML-DSA-65 | 1,952 | 4,032 | 3,309 |
| ML-DSA-87 | 2,592 | 4,896 | 4,627 |
| SLH-DSA-SHA2-128s | 32 | 64 | 7,856 |
| SLH-DSA-SHA2-128f | 32 | 64 | 17,088 |
| SLH-DSA-SHA2-192s | 48 | 96 | 16,224 |
| SLH-DSA-SHA2-256s | 64 | 128 | 29,792 |

deployed. The only measurements reported are published third-party benchmarks and normative constants from the standards, and neither validates the composed architecture.

## 6 Everything gets bigger

**THE RUNNING EXAMPLE**

*The agent assembles its answer for the rental desk and tries to put it into a QR code on the person's phone. It does not fit. It would not fit even with the licence details removed, because the signatures alone are larger than the largest QR code.*

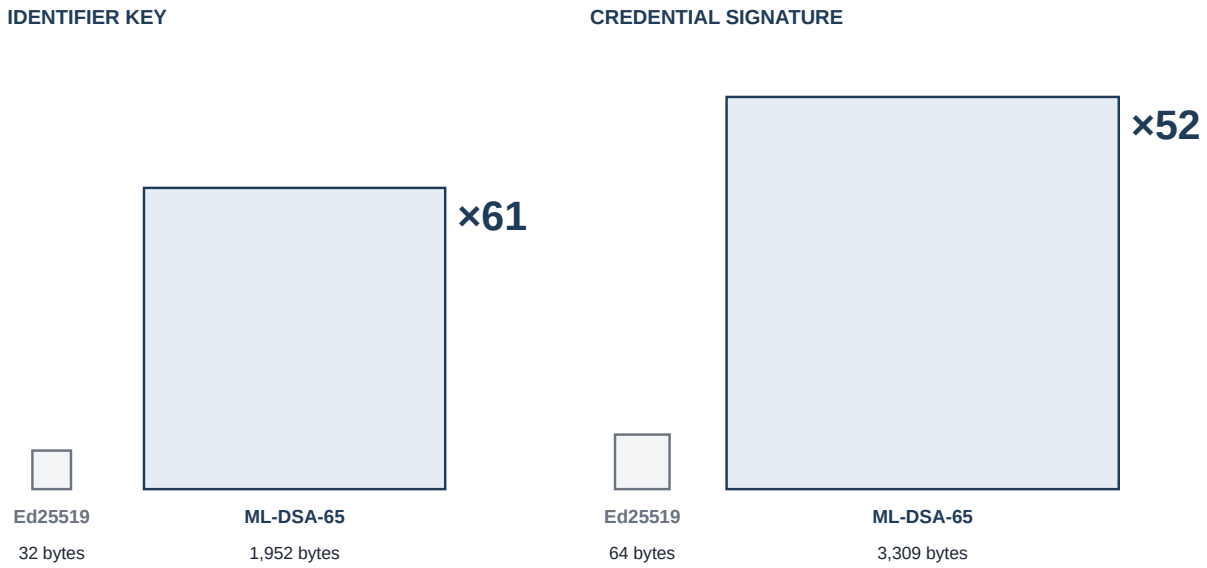
The dominant cost of post-quantum migration in identity systems is not computation but size. Table 1 gives the normative sizes from the standards [21–23] against the classical schemes in use today.

Four consequences follow by arithmetic, and each constrains the design.

**Keys and signatures.** A decentralised identifier document with one ML-DSA-65 verification method carries a 1,952-byte key instead of a 32-byte Ed25519 key, sixty-one times larger; base64url-encoded, the key grows from 43 to 2,603 characters. A credential signed with ML-DSA-65 under the JOSE and COSE profile carries a 3,309-byte signature instead of 64 bytes, fifty-two times larger, or 4,412 base64url characters (Figure 8).

**QR codes.** The largest QR code, version 40 with low error correction, holds 2,953 bytes in binary mode or 4,296 alphanumeric characters [161]. A single ML-DSA-65 signature already exceeds the binary limit, before any credential content is counted, and a minimal presentation carrying the issuer signature and the holder's key-binding signature needs 7,639 base64url characters for its signatures alone. The architecture therefore specifies *indirect presentation*: the QR code or NFC exchange carries only a request identifier, and the credential travels over an authenticated channel (Figure 9). Classical deployments never meet this constraint, and conceptual treatments of post-quantum identity do not surface it.

**Ledgers.** Writing post-quantum signatures to a blockchain is not economically viable. At sixteen gas




Square areas are proportional to size. Values from FIPS 204 [22]; Ed25519 standard encodings.

**Figure 8.** The size change at a glance. Square areas are proportional to size.

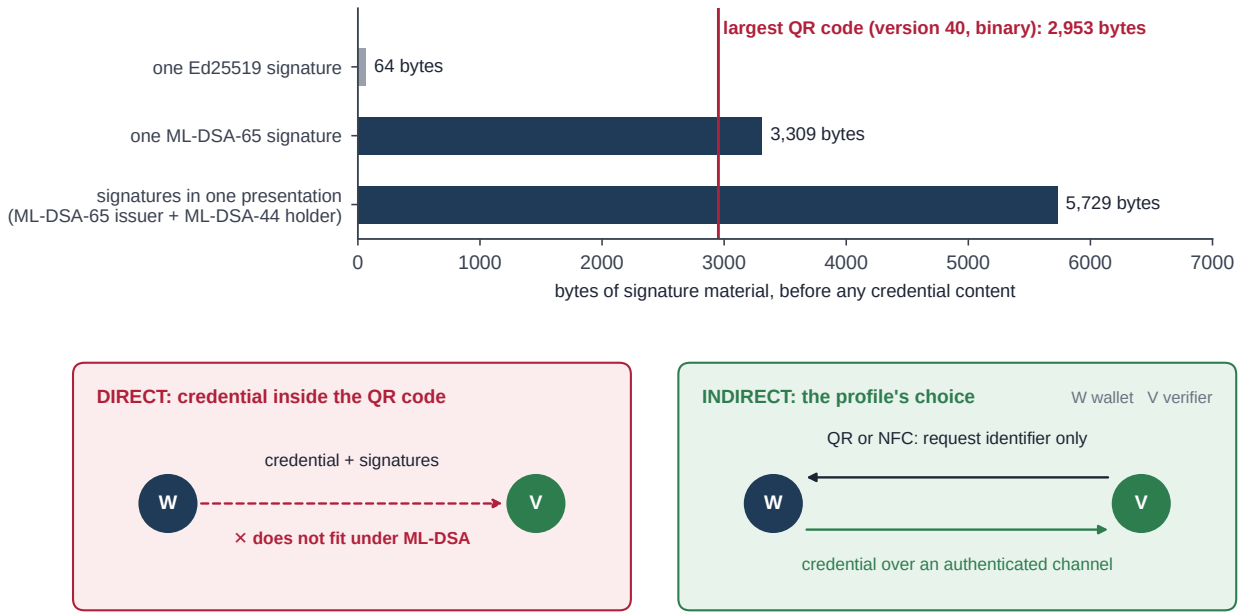



**Figure 9.** Why the credential cannot travel inside the QR code, and what the profile does instead. Sizes derived from Table 1; QR capacity from ISO/IEC 18004 [161].

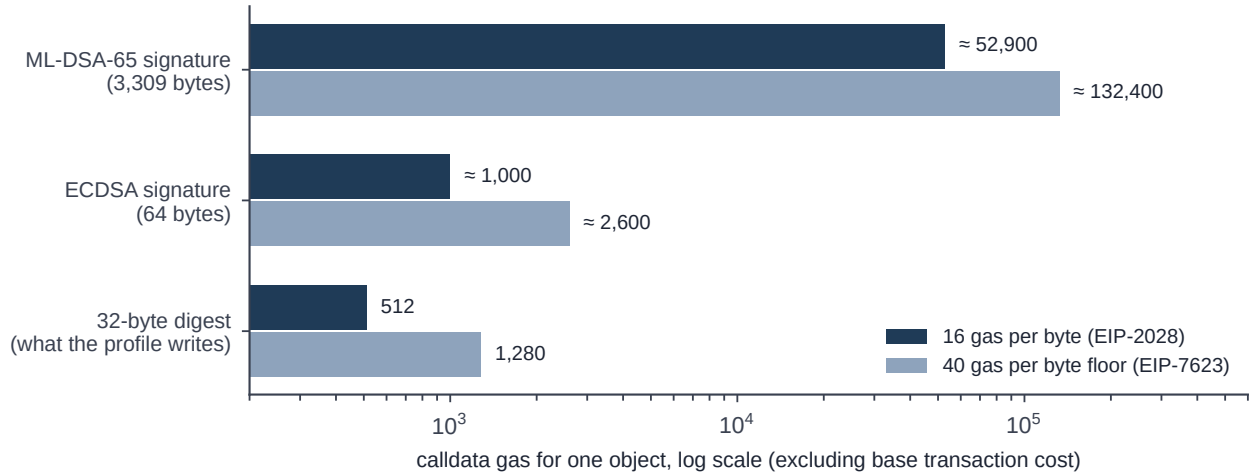



**Figure 10.** Calldata gas to write one object on-chain, excluding the base transaction cost. The digest values (512 and 1,280) are $32 \times 16$ and $32 \times 40$, assuming non-zero bytes.

per non-zero calldata byte under EIP-2028 [162], one ML-DSA-65 signature costs about 52,900 gas, against about 1,000 for a 64-byte ECDSA signature [149]. The floor introduced by EIP-7623 [163] applies to exactly the calldata-heavy transactions that anchoring produces; at forty gas per byte the figures become about 132,400 and 2,600. The profile writes no signature to the ledger: ledgers hold only 32-byte commitments (the digest of an identifier document version, a revocation artefact or a credential schema), and all post-quantum signature material travels off-chain in DIDComm or OpenID for Verifiable Presentations payloads [50, 53] (Figure 10).

**Small devices.** Computation is secondary on servers but decisive on the constrained hardware that holds wallets and credentials. On an ARM Cortex-M0+ running implementations of the final standards [164], ML-KEM-768 key establishment takes 56.62 ms with 17.7 KB peak RAM; ML-DSA-65 verification takes 72.21 ms; ML-DSA-65 signing averages 256.62 ms, with a 99th percentile of 953.6 ms and peak RAM of 84.8 KB (Figure 11). Signing latency varies widely (coefficients of variation of 65.6, 67.5 and 73.1 per cent for ML-DSA-44, -65 and -87) because of rejection sampling. Systems with a fixed presentation timeout must be sized against the tail, not the mean, and that variability must be checked for side-channel exposure [101]. The 84.8 KB requirement exceeds the working memory of many smart cards and secure elements in national identity programmes, an obstacle the architecture does not resolve.

**WHAT THIS DOES NOT CLAIM**

The transport-layer and certificate figures in the published article (for example, a TLS 1.3 handshake with Dilithium2 being 68.12 per cent faster than with RSA-4096) were measured by Raavi and colleagues on the round-three candidates Dilithium, Falcon and SPHINCS+ [165], not on the final standards. ML-DSA changes Dilithium's message hashing and domain separation, and FIPS 205 keeps neither the Haraka instantiation nor the robust parameter sets of SPHINCS+. Those figures show the order of magnitude only, and the deployment profile does not rest on them.

Table 2 states the resulting deployment profile. Each layer is bound to a specific standard, parameter set and registered identifier or wire format, so that the architecture can be implemented rather than only illustrated.

**IN PLAIN TERMS**

Post-quantum key establishment is cheap enough to adopt immediately everywhere. Post-quantum

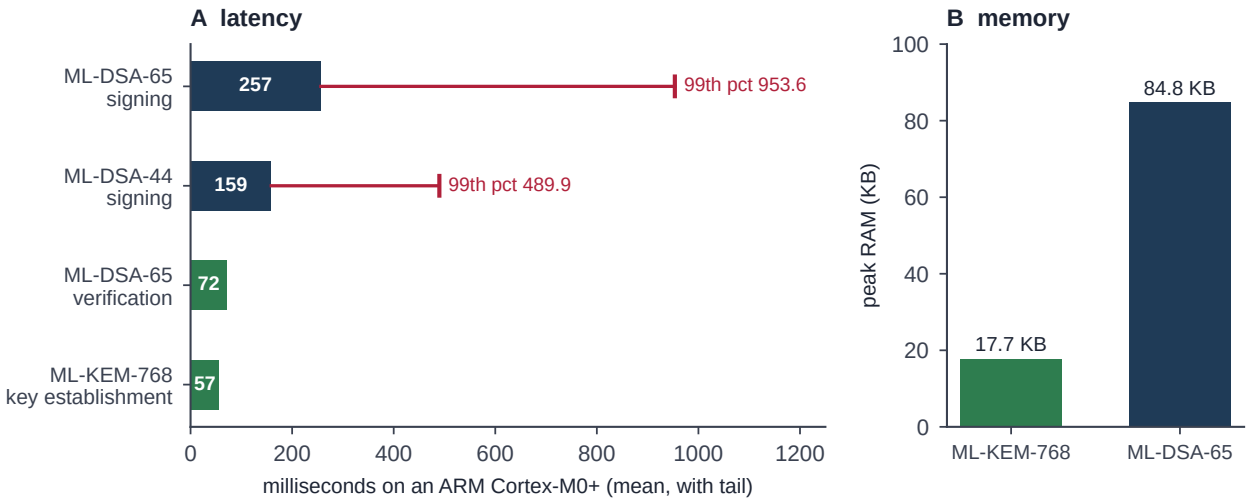



**Figure 11.** Latency and memory on an ARM Cortex-M0+ (RP2040), as reported by Chhetri [164]. Bars are means; red lines extend to the reported 99th percentile. No value was measured for this paper.

**Table 2.** Reference deployment profile.

| Function | Standard | Parameter set | Identifier or wire format |
|---|---|---|---|
| Key establishment | FIPS 203 [21] | ML-KEM-768 | hybrid group X25519MLKEM768, IANA 0x11EC [151] |
| Credential issuance | FIPS 204 [22] | ML-DSA-65 | JOSE ML-DSA-65; COSE $-49$; AKP key type [178] |
| Holder binding | FIPS 204 | ML-DSA-44 | JOSE ML-DSA-44; COSE $-48$ |
| Delegation credential | FIPS 204 | ML-DSA-44 | signed by the delegating holder; scope, expiry, status entry |
| Governance and trust-anchor root | FIPS 205 [23] | SLH-DSA-SHA2-128s | infrequent signing; 32-byte key, 7,856-byte signature |
| Credential data model | W3C VCDM 2.0 [48] | — | secured with JOSE and COSE |
| Selective disclosure | SD-JWT [49] | SHA-256 digests | salted-hash disclosures inside an ML-DSA envelope |
| Revocation | Bitstring Status List [179] | $\geq$131,072 bits | status list credential signed with ML-DSA-65 |
| Ledger anchoring | — | 32-byte digest | no signature material on-chain |
| Predicate proofs | — | hash-based, STARK-style | bound to the ML-DSA-signed digest; pairing-based zk-SNARK excluded |
| Encrypted matching | — | BFV or CKKS | off-chain, outside the verification path |

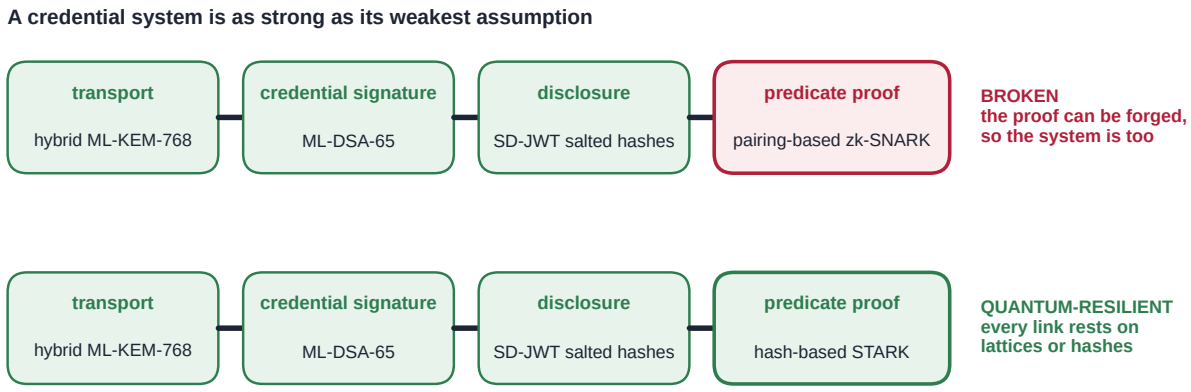



**Figure 12.** The weakest-link problem. Replacing the signature is not enough if the proof system still rests on an assumption that a quantum computer breaks.

signatures are acceptable on servers and marginal on small devices, and their size, not their speed, decides how credentials travel and what goes on a ledger. Hash-based signatures are viable only where signing is rare and bandwidth is not tight, which confines SLH-DSA to trust-anchor and governance roots.

## 7 Proving you are over 21 without saying when you were born

**THE RUNNING EXAMPLE**

*The rental desk asks whether the driver is over 21. The honest answer reveals a date of birth. A better answer proves the condition and nothing more. But the proof is only worth something if it is about a birth date that the licensing authority actually signed, and if a quantum computer cannot forge it.*

### 7.1 The weakest-link problem

Benchmarks of proof systems report that zk-STARK generates proofs fastest and Bulletproofs slowest, that zk-SNARK produces the smallest proofs and zk-STARK the largest, and that zk-SNARK verification is marginally faster [166, 167]. Efficiency is not the governing criterion in a quantum-resilient design. Pairing-based zk-SNARKs, including those used in anonymous payment systems [168–170], rest on discrete-logarithm and pairing assumptions, and Bulletproofs on the discrete logarithm alone [171]; the same quantum adversary breaks them all. A system that signs credentials with ML-DSA but proves predicates with a pairing-based zk-SNARK is not quantum-resilient (Figure 12).

The architecture takes a two-tier position. Selective disclosure uses salted-hash disclosure under SD-JWT [49], which relies only on hash collision resistance and is therefore quantum-safe inside an ML-DSA signature. Zero-knowledge predicate proofs, where needed, use transparent hash-based proof systems of the STARK family [172], whose soundness against quantum adversaries has been established in the quantum random-oracle model [173]; lattice-based proof systems are the second candidate family. Homomorphic encryption is confined to off-chain attribute matching.

### 7.2 Binding the proof to the signed credential

Without a binding, a holder could prove a true statement about a value that no issuer ever attested. The construction below achieves the binding from standard components and adds no assumption beyond those of FIPS 204 and the hash function (Figure 13).

1. **Issuance.** Each selectively disclosable claim $k$ with value $v_k$ is committed as a salted digest $d_k = H(s_k \,\|\, k \,\|\, v_k)$ with a fresh random salt $s_k$; this is the disclosure digest of RFC 9901 [49]. The issuer signs, with ML-DSA-65, the payload containing all digests, the holder's public key in the confirmation claim and the status-list reference [22]. The holder keeps the openings $(s_k, k, v_k)$ and never reveals them for claims that are to be proved rather than disclosed.
2. **Request.** The verifier names a predicate $P$ over claim $j$ (here, age over 21), a nonce $n$ and its own identifier $aud$.
3. **Proof and binding.** The holder produces a STARK $\pi$ for the relation "$H(s_j \,\|\, j \,\|\, v_j) = d_j$ and $P(v_j)$ holds", with public input $(d_j, P, n, aud)$ and witness $(s_j, v_j)$. Placing $n$ and $aud$ in the public input, and so in the Fiat–Shamir transcript, stops the proof being replayed to another verifier or session. The holder then signs a key-binding token with its ML-DSA-44 key over $n$, $aud$, the digest of the issuer-signed payload and the digest of $\pi$, so that holder binding covers the proof as well as the disclosures. This last element extends RFC 9901 by one claim.
4. **Verification.** The verifier checks (i) the ML-DSA-65 signature, which establishes that $d_j$ was issued; (ii) the ML-DSA-44 key-binding signature, which establishes holder binding and freshness and fixes $\pi$; (iii) the STARK against $(d_j, P, n, aud)$; and (iv) the status list.

> **IN PLAIN TERMS**
>
> Three properties combine. ML-DSA cannot be forged [181], so the digest really came from the issuer. The hash has no feasible collisions, so the digest opens to only one birth date. The STARK is knowledge-sound [172, 173], so anyone who convinces the verifier knows a birth date that opens the digest and satisfies the rule. None of the three depends on factoring, discrete logarithms or pairings.

STARK proofs are the largest of the benchmarked systems [166], so they travel over the authenticated channel of indirect presentation. The choice of hash $H$ is fixed at issuance and carries a trade-off: SHA-256 stays within the standardised family but is expensive to express inside a proof, whereas an arithmetisation-friendly hash such as Poseidon [182] reduces proving cost at the price of a younger primitive with less cryptanalytic history. An issuer that expects predicate proofs over particular claims can commit to those claims with the second and record the choice in the credential schema. Either way, the predicate circuits need formal verification, because an under-constrained circuit silently voids soundness [183].

### 7.3 Selective disclosure is not unlinkability

Selective disclosure and unlinkability are different properties, and SD-JWT provides only the first. The holder can withhold attributes, which limits what any single verifier learns. But two verifiers that compare records, or an issuer and a verifier acting together, can recognise repeated presentations, because the issuer-signed part and the disclosure digests are the same every time; the SD-JWT specification says so in its security considerations [49]. The binding above inherits the same limit: the issuer signature and $d_j$ appear unchanged in every presentation.

Where unlinkability is required, the profile issues batches of single-use credentials, each with distinct

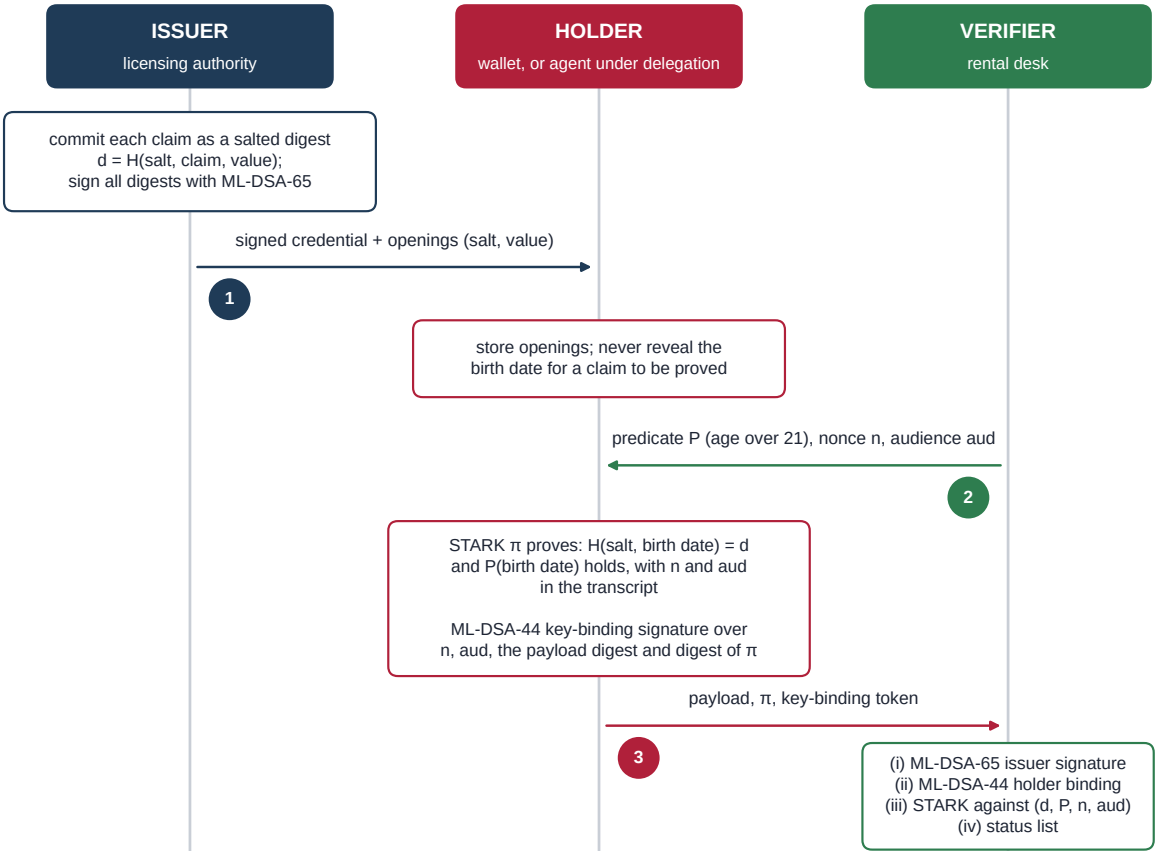



**Figure 13.** Binding a hash-based predicate proof to an ML-DSA-signed credential, shown as a message sequence for the running example.

salts and a distinct issuer signature. Under post-quantum signatures the cost is severe: a batch of twenty commits about 66 kB of signature material per credential type, against about 1.3 kB for Ed25519 (Figure 14). Pairwise pseudonymous identifiers, descended from Chaum's digital pseudonyms [174–176], remove one correlation handle but not this one; anonymous credentials from blind signatures onwards [60, 177] remove it, but only under classical assumptions.

**WHAT THIS DOES NOT CLAIM**

Full unlinkability against a colluding issuer and verifier needs a zero-knowledge presentation protocol, and the pairing-based and BBS-style constructions that provide it today are excluded on quantum-resilience grounds. A quantum-safe equivalent would mean proving knowledge of a valid ML-DSA signature inside the proof, or adopting a lattice-based anonymous credential framework [18, 19, 184]. Neither has a standardised construction, and the profile records unlinkability against colluding parties as an open dependency, not as a property it delivers.

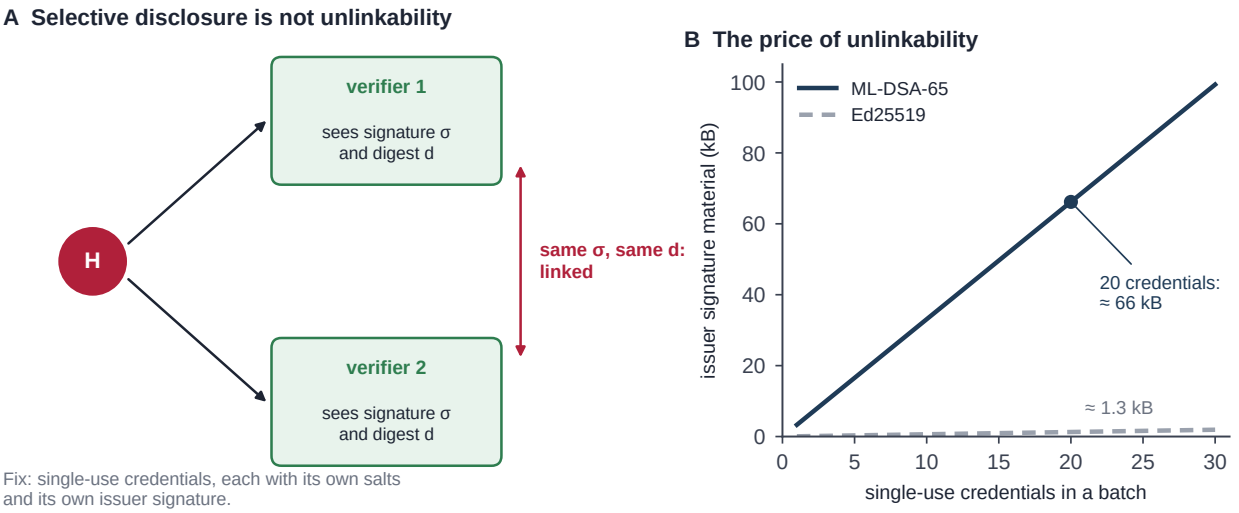



**Figure 14.** (A) Why two verifiers can link presentations of the same credential. (B) Issuer signature material committed when unlinkability is obtained by single-use credentials, derived from Table 1.

## 8 Keys that age well

**THE RUNNING EXAMPLE**

*The licensing authority rotates its signing key every two years, yet the licence it issued is valid for ten. In eight years' time the rental desk must still verify a signature made with a key that has long been retired, and must be able to tell a retired key from a stolen one.*

Much of the identity literature treats cryptographic agility as an aspiration. Here it is a mechanism. A decentralised identifier document carries an ordered set of verification methods, each expressed under the Algorithm Key Pair key type introduced for post-quantum algorithms in JOSE and COSE [178] and labelled with its registered algorithm identifier. Algorithm choice becomes a negotiation between holder, issuer and verifier over declared capabilities at the presentation layer, rather than a property fixed when the identifier is created. Because the document is versioned and its digest anchored, a method can be added and an obsolete one retired without changing the identifier, so agility never forces the subject to re-enrol.

Key lifecycle management is specified as six states (Figure 15). Archival is the state most often omitted, and the one that migration makes indispensable: a ten-year credential signed under a key rotated after two years must remain verifiable for eight more, which requires the retired key to stay resolvable and distinguishable from a revoked one. Superseded methods are therefore kept in the identifier document with an explicit expiry and a signed timestamp for the interval in which the key was active. Where an issuer changes algorithm family rather than key, credentials are re-issued rather than re-signed, because the security assumption behind the original attestation has changed. In Mosca's terms, credential validity and trust-anchor lifetime are the shelf-life term added to migration time [14], so long-validity classical credentials should be scheduled for re-issuance before the deprecation date.

The transition rests on an asymmetry that the identity literature seldom states. Confidentiality can be attacked after the fact: ciphertext captured today can be decrypted later, so key establishment must migrate first. Authenticity mostly cannot: forging a signature after a transaction was completed and verified does not undo it, so signature migration may follow, except for long-lived credentials

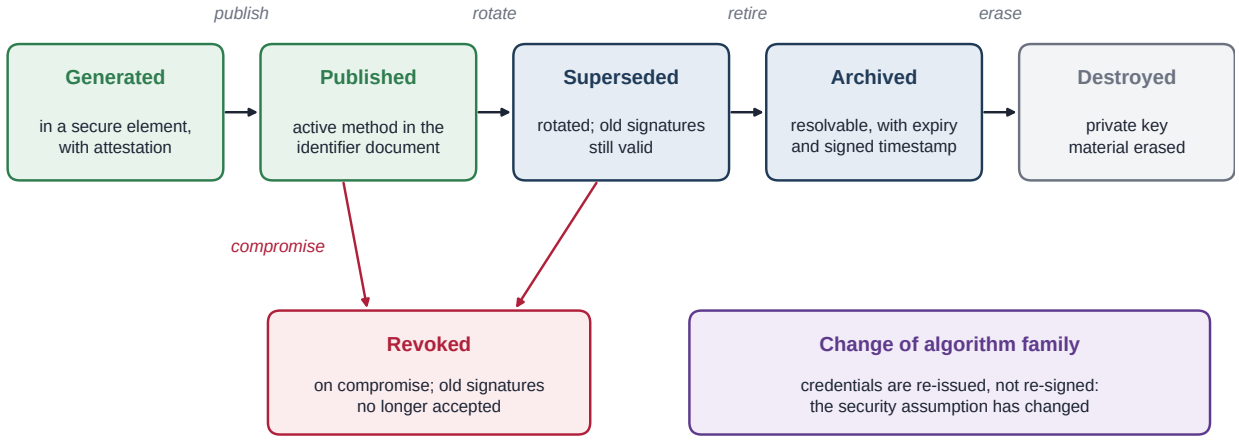



**Figure 15.** The six-state key lifecycle. Rotation preserves the validity of historical signatures; revocation does not; a change of algorithm family requires re-issuance.

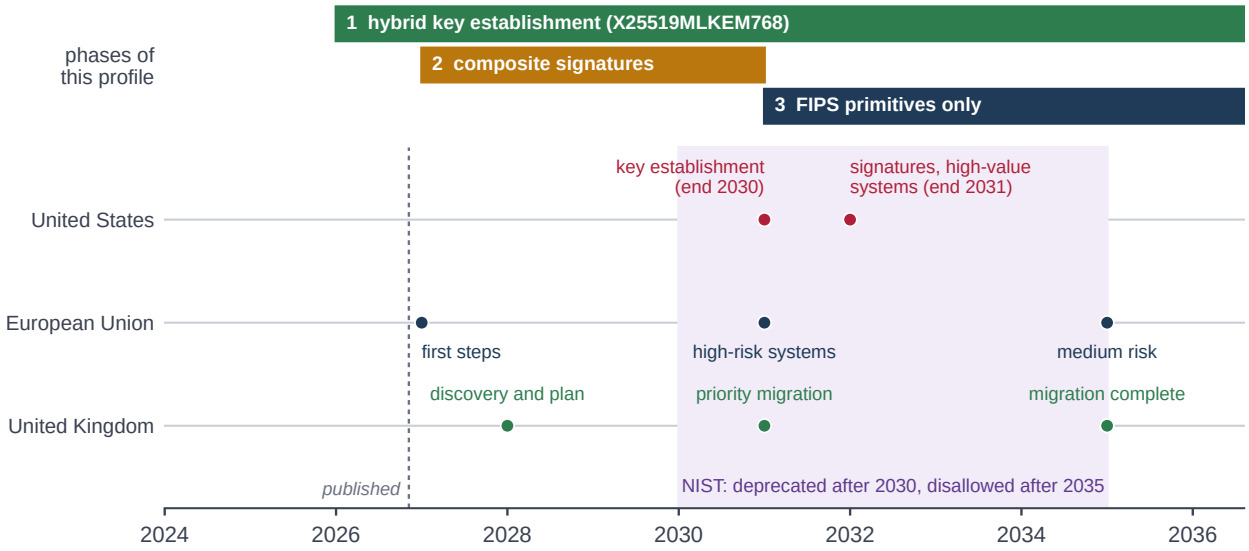



**Figure 16.** The three-phase transition against published milestones: NIST deprecation and disallowance dates and United States federal deadlines [13], and the European and United Kingdom roadmaps [95, 96].

and trust anchors, which are still verified later. Hence three phases (Figure 16):

1. **Hybrid key establishment now,** using the X25519MLKEM768 group specified for TLS 1.3 in RFC 10024 [151], which stays secure if either component holds. Signatures remain classical, but identifier documents are provisioned with post-quantum verification methods so that no re-enrolment is needed later.
2. **Composite signatures,** in which one signature object carries both a classical and an ML-DSA signature and verification requires both [185]. A verifier must implement the composite scheme, so composite signatures do not by themselves keep unmigrated verifiers working; an issuer that must serve them issues a classical credential in parallel during the transition window and withdraws it at the phase boundary. An Ed25519 and ML-DSA-65 composite is 3,373 bytes.
3. **FIPS primitives only,** once classical components are withdrawn.

FN-DSA and HQC are treated as optional profiles rather than dependencies. FN-DSA produces

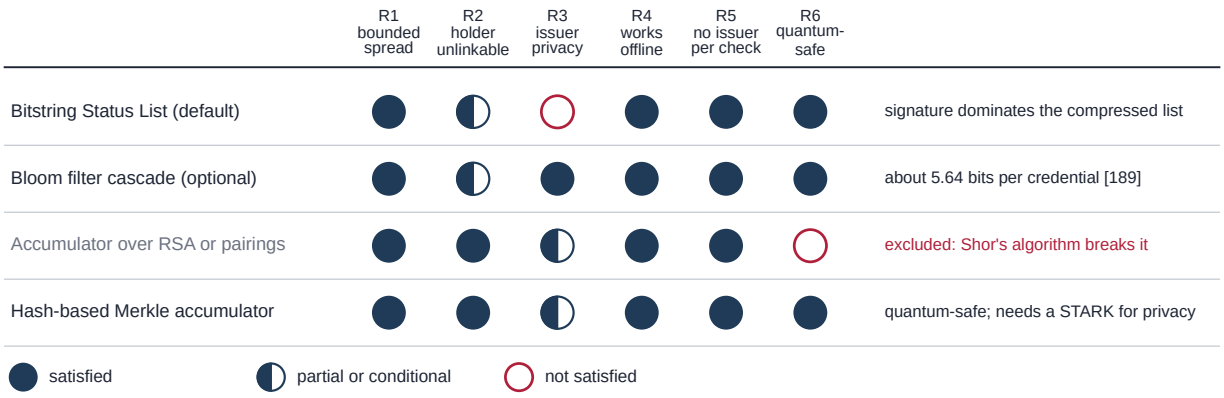



**Figure 17.** Four revocation mechanisms against the six requirements. Cascade figures from Hoops and colleagues [189].

markedly smaller signatures, 2,358 against 5,340 bytes at certificate level for the corresponding round-three candidates Falcon-512 and Dilithium2 [165], which would help the transport-constrained path; its draft was submitted for approval in August 2025 [91]. HQC offers key encapsulation under a different hardness assumption from ML-KEM [89, 186]. Because algorithm selection runs entirely through registered identifiers, both can be added by registration alone, without structural change. That is the practical content of agility.

**UPDATE SINCE ACCEPTANCE**

The composite ML-DSA specification, an Internet-Draft in the RFC Editor queue when the article was written, has since been published as RFC 10073 [185].

## 9 Taking it back: revocation under post-quantum constraints

**THE RUNNING EXAMPLE**

*On Friday evening the person's licence is suspended. On Saturday morning the agent presents it anyway. Whether the rental desk notices depends on how revocation works, how often the status list is published, and how long the desk caches it.*

Revocation is where the cryptographic, privacy and regulatory requirements of an identity system collide most directly, and it is the operation the decentralised identity literature treats least adequately [187]. Six requirements govern the design: bounded propagation (R1); no correlation of presentations (R2); no disclosure of the volume or timing of an issuer's revocations (R3); offline verification (R4); no issuer participation in each verification (R5); and quantum-safe assumptions throughout (R6).

R6 disqualifies a well-established mechanism. Accumulator-based revocation, as in AnonCreds-style systems, gives strong holder privacy by proving non-membership in zero knowledge [57, 188]. Built over RSA groups or bilinear pairings, it rests on assumptions that Shor's algorithm defeats. A system can sign credentials with ML-DSA and still be quantum-vulnerable through its revocation layer. Where accumulator semantics are needed, they must be built over hash-based Merkle structures, with the non-membership proof itself hash-based (Figure 17).

Status lists are the default. Bitstring Status List, a W3C Recommendation of May 2025, encodes

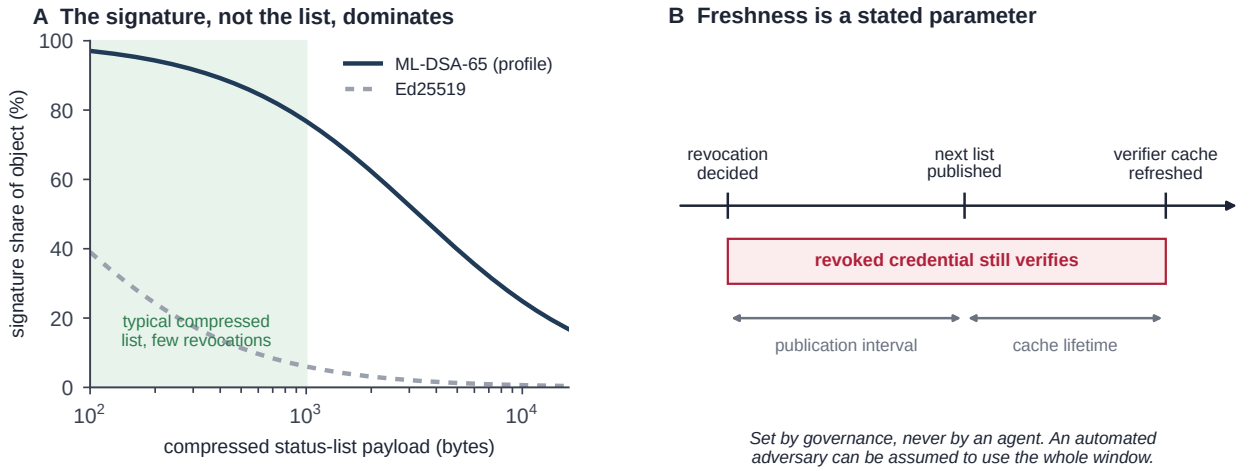



**Figure 18.** (A) Share of a signed status list taken by its signature, as a function of compressed payload size, derived from Table 1. (B) The window in which a revoked credential still verifies.

status as a bit position in a list of at least 131,072 bits (16 KB uncompressed); the minimum length gives herd privacy, since a query reveals only that some credential in a large set was checked [179]. Because most credentials are never revoked, the list compresses to a few hundred bytes. But the list is itself a credential and must be signed, and an ML-DSA-65 signature adds 3,309 bytes to that compressed payload. The signature becomes most of the object, and the compression advantage disappears (Figure 18A). The profile responds with longer refresh intervals, aggressive caching at the verifier and aggregation of many credential types into one signed list, so that one signature serves many status entries. Herd privacy protects the holder against the list publisher, not against verifiers: the status index in the credential is stable, so cross-verifier unlinkability again needs single-use credentials with distinct indices. The specification also separates revocation, which is permanent, from suspension, which is reversible [179]; the agent rules in Section 10 rely on that distinction.

Where revocation volume is itself sensitive (professional licensing, immigration-adjacent credentials), status lists fail R3, and the profile specifies a Bloom filter cascade. Hoops, Gebele and Matthes report that a cascade with fixed-size padding and a regular schedule hides both the number and frequency of revocations, that one Ethereum blob holds revocation data for about 170,000 credentials, and that it grows at about 5.64 bits per credential, at the cost of a lookup across up to about forty levels [189]. It is non-interactive and hash-based, so it works offline and is quantum-safe.

**IN PLAIN TERMS**

No non-interactive mechanism can make revocation instantaneous. A verifier checks against a signed artefact whose freshness is bounded by the publication interval plus its own cache lifetime. That bound is the security parameter a deploying organisation must set: short for high assurance, long for offline settings. An automated adversary can be assumed to use the whole window, which is why the bound is set by governance and never by an agent.

### 9.1 Revocation is not erasure

Revoking a credential withdraws its verifiability; it deletes no personal data and does not by itself satisfy a request under Article 17 of the General Data Protection Regulation [72, 74, 190]. Erasure rests on three further controls: no personal data or cleartext credential identifier is written to the ledger; the issuer deletes the credential record with its disclosure salts and the holder deletes the credential from the wallet, so the on-ledger commitment cannot be re-associated with the subject; and where attributes were encrypted for off-chain processing, the decryption keys are destroyed. Two limitations remain. A salted-hash commitment on an immutable ledger cannot be deleted, and whether it is still personal data depends on whether re-identification remains reasonably possible, a property of key management in the deployment [75]. Copies already disclosed to verifiers are outside the issuer's control.

## 10 Never give your agent your key

> **THE RUNNING EXAMPLE**
>
> *The person could simply give the agent a copy of the licence and the private key. It would work, until the agent is compromised, or until a second agent acquires the same secret. From then on, nobody can tell the person from any piece of software that holds the key, and nobody can tell which agent did what.*

Agents occupy three roles, and one rule unifies them: every boundary at which trust is created, transferred or withdrawn is enforced by cryptography and by deterministic policy that sit outside any agent.

### 10.1 Agents as delegates

An agent acting for a person must not receive the person's private key, or a copy of the person's credential to present as its own. Either defeats holder binding and leaves no record of which agent acted. Delegation is carried by a credential instead. The person's wallet issues a delegation credential to the agent's own decentralised identifier, signed with the holder key, carrying the scope of the grant (audience, purpose, value limit, validity interval) and a status-list entry under the person's control, so that the grant can be suspended or revoked without touching the underlying credential. This follows a lineage from delegation in federated assertions and OAuth [126, 191] through the actor chains of OAuth token exchange [127] to authenticated delegation for AI agents [7], and is consistent with agent identity designs based on decentralised identifiers [123, 128, 129] and with the signed mandates of agent payment protocols [9]. An agent that sub-delegates issues a further credential whose scope must be a subset of its own, and the verifier enforces that narrowing along the chain. Every hop names an agent, so every presentation is attributable to an identifiable agent acting under an identifiable grant [124, 130].

Delegation has a post-quantum cost the classical literature never meets. A presentation by the person carries the issuer's ML-DSA-65 signature, the holder's ML-DSA-44 public key and the holder's key-binding signature: $3{,}309 + 1{,}312 + 2{,}420 = 7{,}041$ bytes, against 160 under Ed25519. Each delegation hop adds one ML-DSA-44 signature and the delegate's ML-DSA-44 key, 3,732 bytes, against 96 classically. A person who instructs an agent that instructs a sub-agent therefore sends 14.5 kB of signatures and keys before any claim or proof, or about 19.3 kB base64url-encoded, against 352 bytes under Ed25519 (Figure 19).

Three rules follow: delegation depth is bounded by policy, with a depth of two covering the common

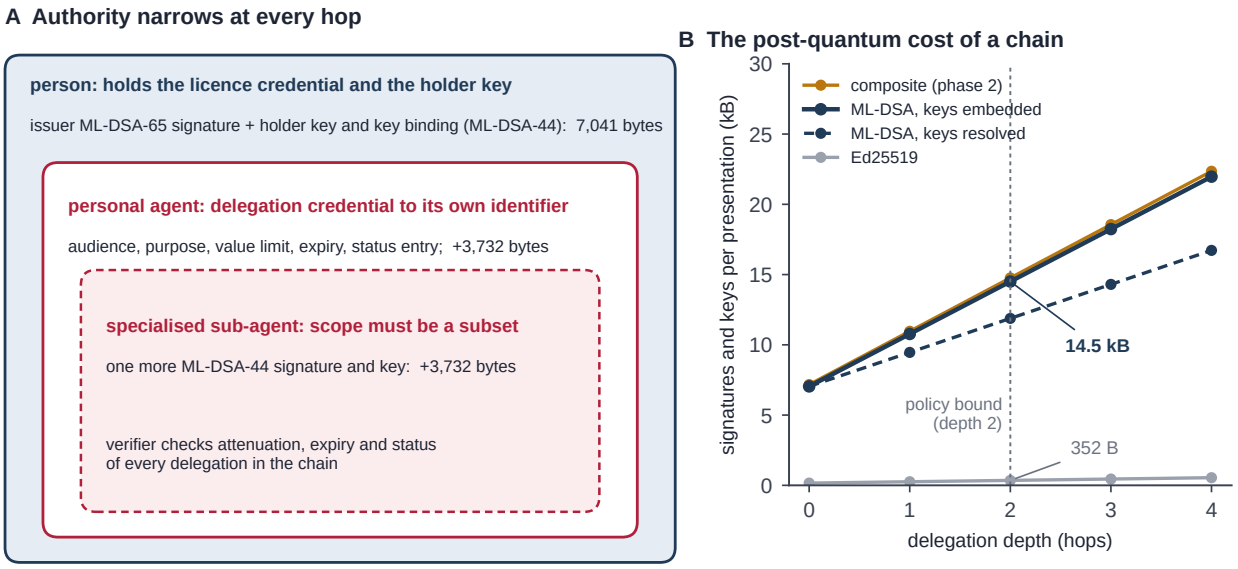



**Figure 19.** (A) A delegation chain from the person to a sub-agent, with the material each hop adds. (B) Signatures and keys per presentation against delegation depth, derived from Table 1: delegate keys embedded or resolved from identifier documents (saving 1,312 bytes per hop), composite signatures (64 bytes more per signature) and Ed25519.

case of a personal agent invoking a specialised one; delegation credentials, which change far less often than presentations, are cached by verifiers; and SLH-DSA, at 7,856 bytes per signature, is never used for delegation.

## 10.2 Agents as operators: the action envelope

An agentic operator is any AI component that observes the identity system and influences what it does: a fraud or Sybil detector that also acts, a triage agent that drafts revocation requests, a learned policy that adjusts rate limits. The difficulty is not that such components err, since every detector errs, but that their errors are hard to bound under attack [15, 16, 140] and that their inputs can carry instructions [107]. The architecture therefore bounds what an operator can do rather than how well it reasons, in line with recommended practice of constraining an agent's action space and requiring approval for consequential actions [1, 138]. Actions fall into four tiers, and a deterministic policy enforcement point outside the agent admits each by tier (Figure 20).

Two tensions in earlier proposals for AI-managed identity, including the reinforcement-learning operators that motivated this design [132, 192], are resolved. Cryptographic parameters are not tuned by learning: they are fixed by the profile and changed only by governance through registered identifiers. And revocation is not an agent action: an agent may suspend within quota and may propose revocation, but only the issuer revokes. Both map onto status purposes the W3C specification already defines [179], so the envelope needs no new mechanism.

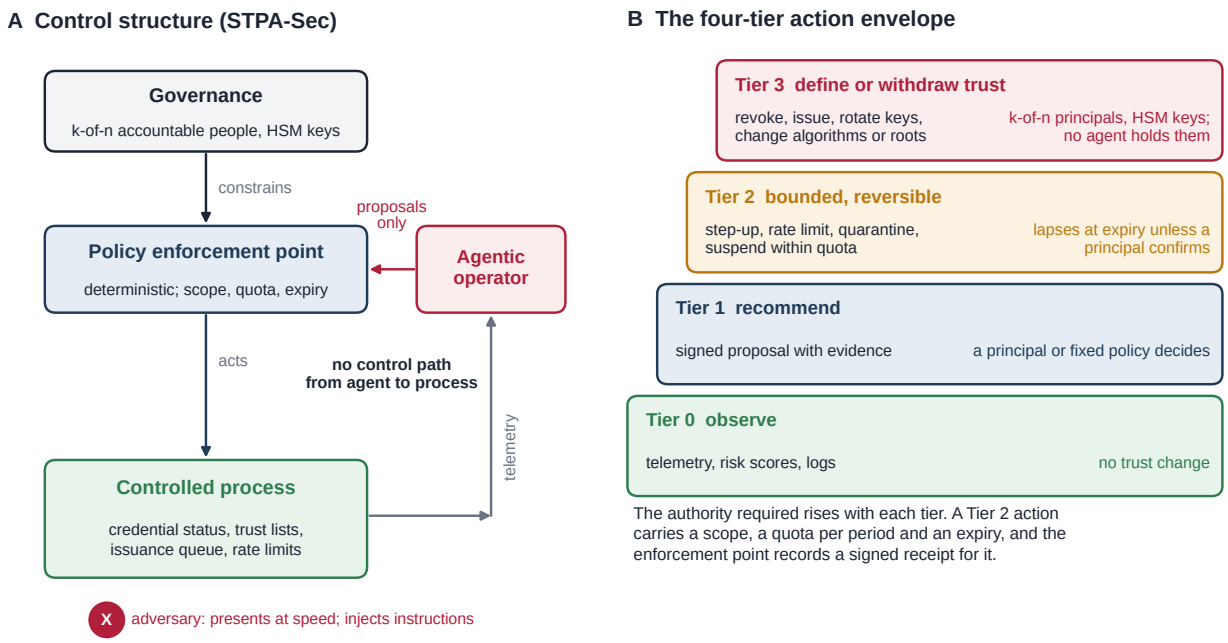



**Figure 20.** Bounded autonomy. (A) STPA-Sec control structure after Young and Leveson [193]: the agent receives telemetry and can only send proposals; there is no control path from the agent to the process. (B) The four tiers; no agent holds the keys that Tier 3 requires.

## 10.3 What could still go wrong: an STPA-Sec analysis

The envelope was derived with STPA-Sec, which treats security as a control problem and asks how each controller's actions could become unsafe, including through manipulated feedback [193, 194]. Four losses are to be prevented: acceptance of an unauthorised presentation (L1), disclosure beyond the presented attributes (L2), denial of service to legitimate holders (L3), and loss of accountability for a trust-relevant action (L4). Table 3 lists the unsafe control actions and the constraint each becomes.

**IN PLAIN TERMS**

Each of P1 to P5 holds under arbitrary behaviour of any agent in the operator role, provided the enforcement point admits actions only within the envelope and the cryptographic assumptions hold. A failing or manipulated agent can degrade two things: detection, so that an attack it should have flagged is missed, and availability, so that legitimate holders are suspended or rate-limited within the quota and expiry the envelope allows. Both are residual risks of a deployment, not failures of its security properties.

**WHAT THIS DOES NOT CLAIM**

This is a design claim about where security properties reside. It is not an empirical claim about the reliability of any agent, and it does not depend on one.

**Table 3.** Unsafe control actions (UCAs) and the constraints into which they invert.

| ID | Controller and action | Unsafe when | Constraint (enforced by) | Loss |
|---|---|---|---|---|
| UCA1 | Operator suspends a credential | holder is legitimate: false positive or manipulated telemetry | quota, automatic expiry, holder notification, principal confirmation (enforcement point) | L3 |
| UCA2 | Operator flags a Sybil enrolment | not provided during an automated campaign | issuance assurance that needs no detector; detection advisory (issuer policy) | L1 |
| UCA3 | Operator proposes revocation | on evidence injected through telemetry or content | k-of-n principals; evidence from a channel the agent does not control (governance) | L1, L3 |
| UCA4 | Operator changes an algorithm or parameter set | at any time | outside the action space; no agent holds an issuer or root key (enforcement point, HSM) | L1, L4 |
| UCA5 | Enforcement point admits an agent action | beyond scope, quota or expiry | deterministic, specified, tested policy; signed receipts (governance audit) | L3, L4 |
| UCA6 | Governance approves an agent proposal | without independent evidence | approval requires an independent feedback channel (governance procedure) | L1, L3 |
| UCA7 | Delegated agent presents a credential | beyond scope, or after the grant lapsed | verifier checks attenuation, expiry and status of every hop (verifier) | L1, L2 |
| UCA8 | Governance sets revocation freshness | window too long, or delegated to an agent | set and reviewed by principals; never adjusted by an agent (governance) | L1 |

## 11 What is in, what is out

An architecture that listed every available privacy-enhancing technology would not be implementable. Figure 21 sets out every mechanism considered and its status. The core is deliberately small: one key establishment mechanism, two signature parameter sets with a third confined to governance roots, one disclosure mechanism and one revocation mechanism. Homomorphic encryption, secure multi-party computation [180] and indistinguishability obfuscation are not on the synchronous verification path and are not needed for a conforming deployment. Excluded mechanisms are excluded for one of three reasons: they rest on assumptions the migration exists to retire, they have been broken [92], or they would make a security parameter depend on an AI component.




**Figure 21.** Mechanisms considered and their status in the architecture, following Table 4 of the published article.

## 12 What it means, and what remains open

The review shows a literature that is advanced within each domain and fragmented across them. Decentralised identity has matured in interoperability and governance but still rests on classical primitives [42, 44, 69]; post-quantum research has matured at the level of primitives and implementations [93, 101, 102] but not at the level of identity operations; privacy-enhancing computation is judged on efficiency rather than quantum resilience [166, 167]; and AI is studied as a defence more often than as an operating condition [2, 10]. The deployment profile is the integration this fragmentation leaves undone. Its costs are dominated by size, they surface in transport, ledger and delegation design rather than in computation, and several established mechanisms (pairing-based proofs, BBS-style presentation, accumulator revocation over algebraic groups) do not survive the change of assumption at all.

**Three implications for current deployments.** First, wallet ecosystems being rolled out now, including the European Digital Identity Wallet, which each Member State must provide by the end of 2026 [32, 52], will issue credentials whose validity runs into the deprecation window, and their presentation flows often rely on QR codes and proximity transports. Indirect presentation and status-list aggregation are cheap decisions now and expensive retrofits later. Second, the comparison with centralised systems is one of degree. Frameworks such as eIDAS 2.0 and Aadhaar can adopt post-quantum cryptography, but holding the algorithm choice in a single trust anchor and credential format means agility must be exercised by coordinated replacement rather than by negotiation between the parties; the paper does not claim that centralised systems cannot migrate. Third, agent identity is converging with personal identity. Agent payment and tool-access protocols already carry

delegated authority [8, 9], and the infrastructure that verifies people's credentials is the natural place to verify the authority under which agents act [7, 125]. At about 3.7 kB per hop, that authority argues for short chains, cached delegation credentials and indirect transport from the start.

**Five open problems.**

1. *Quantum-safe unlinkable presentation.* No standardised hash-based or lattice-based protocol provides unlinkability against colluding parties; the candidate routes [18, 19, 184] have not been evaluated at national scale.
2. *Hardware-bound holding.* ML-DSA-65 signing needs 84.8 KB peak memory, more than many secure elements have; either ML-DSA-44 becomes the issuer default for constrained ecosystems or secure elements must change.
3. *Compact signatures.* FN-DSA would relieve the transport constraint once standardised, at the cost of floating-point implementation hazards.
4. *Formal assurance.* The composed protocol of Section 7 and the enforcement point of Section 10 are small enough to be specified and verified formally, and should be, since the bounded-autonomy claim rests on the enforcement point being correct [183, 195].
5. *An empirical adversary model.* How fast an agentic adversary can drive a presentation interface, and how well it does against injection-resistant proofing, has not been measured; those quantities should set the freshness and timeout parameters the paper leaves to deployment.

### 12.1 Limitations

The paper is a review-based architectural study and presents no original experimental data. Its quantitative analysis consists of normative constants from the standards, quantities derived from them by stated arithmetic, and measurements from independently published benchmarks; none validates the composed architecture. The device measurements come from a single Cortex-M0+ implementation [164] and the transport measurements from one software stack [165], and neither extrapolates to the varied hardware of a national deployment without re-measurement. No prototype has been deployed, so end-to-end latency, throughput under load and interoperability across identifier methods and ledgers are unmeasured; claims about government, healthcare and financial deployment concern design properties, not demonstrated behaviour. FN-DSA and HQC were not final standards at the time of writing [89, 91]. The security argument rests on the hardness assumptions of the standardised schemes, and the composed protocol has not been formally verified. Unlinkability against colluding parties is identified rather than delivered, and the status of an anchored salted-hash commitment as personal data is left to deployment key management. The agentic threat model is qualitative and drawn from documented capabilities; no adversarial evaluation against an autonomous attacker was performed. The treatment of delegation concerns the authority under which agents present credentials; how an agent's identifier is bound to its developer, operator and model version, and who is liable for its actions, are outside scope. The programme of empirical work that deployment requires is defined by Table 2: a reference implementation binding the named parameter sets and identifiers to an existing credential stack, instrumented to report latency distributions rather than means, artefact and delegation-chain sizes on the wire, and interoperability against at least two independent identifier methods.

## 13 Conclusion

**THE RUNNING EXAMPLE**

*The car is rented. The rental desk learned that a valid licence exists and that the driver is over 21. It did not learn the birth date. It knows which agent acted and under whose authority, and that authority can be withdrawn without reissuing the licence. Every check would still hold against a quantum computer. The agent never held the person's key.*

The paper develops a standards-aligned deployment profile for decentralised identity in a setting where autonomous agents attack, act within and operate identity systems, and where the public-key cryptography beneath those systems has a fixed retirement date. It maps FIPS 203 to 205 onto identity operations at the level of parameter sets, wire formats and registered identifiers, with hash-based predicate proofs bound to ML-DSA-signed credentials so that the privacy layer adds no assumption a quantum adversary defeats (C1). It shows that the dominant cost of migration is size: identifier keys grow sixty-one-fold and credential signatures fifty-two-fold, which rules out single-QR transport and on-chain signature anchoring (C2). It specifies agility as a mechanism, through versioned verification methods, a six-state key lifecycle and a three-phase transition that migrates key establishment first (C3). It shows that accumulator revocation over RSA or pairing groups cannot be carried forward, that the signature rather than the bitstring dominates status-list cost, and that revocation is not erasure (C4).

The treatment of agentic AI is the most general finding (C5). As adversaries, agents set requirements. As delegates, they act under scoped, revocable delegation credentials rather than with the keys of the people they serve, at a quantified cost of 3.7 kB per hop. As operators, they observe, recommend and take bounded reversible actions, while every action that defines or withdraws trust is taken under keys held by accountable principals. No security property depends on an AI component, and agent failure is confined to stated losses of detection and availability.

**Agents propose; keys dispose.**

## Relationship to the published article

This document is an illustrated edition of a review article published in *Frontiers in Blockchain* under a CC BY 4.0 licence (doi:10.3389/fbloc.2026.1905568). It was prepared by the author. The changes are: the text has been rewritten and reordered for a wider technical readership, with a running example added; all figures are new; the published Tables 1, 3 and 5 are reproduced in condensed form, Table 2 is presented as Figure 11 and a caveat box, and Table 4 as Figure 21; reference errors corrected at proof stage appear in corrected form; and the publication of the composite ML-DSA specification as RFC 10073 is noted. The technical content, quantitative values and conclusions are those of the published article. Where the two differ in wording, the published article is the version of record.

## Data availability

No new data were generated or analysed in this study. All quantitative values are normative constants from the cited standards, values reported in the cited benchmark studies, or quantities derived from them by the arithmetic stated in the text. The code that renders the figures in this edition is available from the author on request.

## Funding

This work has been supported by the UK EPSRC [grant number EP/S035362/1], the Bill and Melinda Gates Foundation [reference code INV-057591] and SPRITE+ [funded under EPSRC (EP/W020408/1)].

## Acknowledgements

Eternal gratitude to the Fulbright Visiting Scholar Project.

## Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Generative AI statement

Generative AI was used in preparing both the published article and this edition. For the published article, Claude Opus 5.5 (Anthropic), accessed through Claude Code in November 2026, assisted with restructuring in response to editorial comments, language editing, checking the status of cited standards and references, and writing the code that renders the figures. For this edition, the same model, accessed in November 2026, assisted with adapting the text for a wider readership, designing the figures and writing the code that renders them, and checking every value against the published article. The author reviewed and verified all content, quantitative derivations and references, and takes full responsibility for this document.

## References

[1] Shavit Y, Agarwal S, Brundage M, et al. Practices for Governing Agentic AI Systems. OpenAI, 2023. https://cdn.openai.com/papers/practices-for-governing-agentic-ai-systems.pdf.

[2] Deng Z, Guo Y, Han C, et al. AI Agents Under Threat: A Survey of Key Security Challenges and Future Pathways. ACM Comput Surv 2025; 57: 182. DOI: 10.1145/3716628.

[3] ENISA. Remote ID Proofing — ENISA, https://www.enisa.europa.eu/publications/enisa-report-remote-id-proofing (2021, accessed 12 November 2026).

[4] Khanjani Z, Watson G, Janeja VP. Audio deepfakes: A survey. Front Big Data 2023; 5: 1001063.

[5] Fang R, Bindu R, Gupta A, et al. LLM Agents can Autonomously Exploit One-day Vulnerabilities. arXiv preprint arXiv:2404.08144, 2024. DOI: 10.48550/arXiv.2404.08144.

[6] Anthropic. Disrupting the first reported AI-orchestrated cyber espionage campaign. Anthropic, 2025. https://assets.anthropic.com/m/ec212e6566a0d47/original/Disrupting-the-first-reported-AI-orchestrated-cyber-espionage-campaign.pdf.

[7] South T, Marro S, Hardjono T, et al. Position: AI Agents Need Authenticated Delegation. In: Singh A, Fazel M, Hsu D, et al. (eds) Proceedings of the 42nd International Conference on Machine Learning. Proceedings of Machine Learning Research. PMLR, 2025, pp. 82211–82231. https://proceedings.mlr.press/v267/south25a.html.

[8] Model Context Protocol. Authorization. In: Model Context Protocol specification (revision 2025-06-18), https://modelcontextprotocol.io/specification/2025-06-18/basic/authorization (2025, accessed 12 November 2026).

[9] Parikh S, Surapaneni R. Powering AI commerce with the new Agent Payments Protocol (AP2). Google Cloud Blog, https://cloud.google.com/blog/products/ai-machine-learning/announcing-agents-to-payments-ap2-protocol (2025, accessed 12 November 2026).

[10] Aboukadri S, Ouaddah A, Mezrioui A. Machine learning in identity and access management systems: Survey and deep dive. Comput Secur 2024; 139: 103729.

[11] Shor PW. Algorithms for quantum computation: Discrete logarithms and factoring. Proceedings - Annual IEEE Symposium on Foundations of Computer Science, FOCS 1994; 124–134.

[12] Gidney C. How to factor 2048 bit RSA integers with less than a million noisy qubits. arXiv preprint arXiv:2505.15917, 2025. DOI: 10.48550/arXiv.2505.15917.

[13] Moody D, Perlner R, Regenscheid A, et al. Transition to post-quantum cryptography standards. NIST IR 8547. Gaithersburg, MD: National Institute of Standards and Technology, 2026. DOI: 10.6028/NIST.IR.8547.

[14] Mosca M. Cybersecurity in an Era with Quantum Computers: Will We Be Ready?. IEEE Security & Privacy 2018; 16: 38–41. DOI: 10.1109/MSP.2018.3761723.

[15] Taddeo M, McCutcheon T, Floridi L. Trusting artificial intelligence in cybersecurity is a double-edged sword. Nature Machine Intelligence 2019; 1(12): 557–560.

[16] Sommer R, Paxson V. Outside the closed world: On using machine learning for network intrusion detection. In: 2010 IEEE Symposium on Security and Privacy. Berkeley/Oakland, CA: IEEE, 2010, pp. 305–316. DOI: 10.1109/SP.2010.25.

[17] Wood T, Thomas K, Dean M, et al. Towards Post-Quantum Verifiable Credentials. In: Proceedings of the 19th International Conference on Availability, Reliability and Security. New York, NY, USA: ACM, 2024, pp. 1–10. DOI: 10.1145/3664476.3669932.

[18] Dutto S, Margaria D, Sanna C, et al. Toward a Post-Quantum Zero-Knowledge Verifiable Credential System for Self-Sovereign Identity. Cryptology ePrint Archive, 2022. https://eprint.iacr.org/2022/1297.

[19] Margaria D, Pino A, Vesco A, et al. Implementation of a Post-Quantum Anonymous Verifiable Credential Framework. In: 2025 IEEE Symposium on Computers and Communications (ISCC). IEEE, 2025, pp. 1–6. DOI: 10.1109/ISCC65549.2025.11325782.

[20] Solavagione A, Vesco A. Transition of Self-Sovereign Identity to Post-Quantum Cryptography. In: 2025 International Conference on Quantum Communications, Networking, and Computing (QCNC). IEEE, 2025, pp. 174–181. DOI: 10.1109/QCNC64685.2025.00035.

[21] National Institute of Standards and Technology. Module-lattice-based key-encapsulation mechanism standard. FIPS 203. Gaithersburg, MD: National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.FIPS.203.

[22] National Institute of Standards and Technology. Module-lattice-based digital signature standard. FIPS 204. Gaithersburg, MD: National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.FIPS.204.

[23] National Institute of Standards and Technology. Stateless hash-based digital signature standard. FIPS 205. Gaithersburg, MD: National Institute of Standards and Technology, 2024. DOI: 10.6028/NIST.FIPS.205.

[24] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ 2021; 372: n71. DOI: 10.1136/bmj.n71.

[25] Agrawal S, Banerjee S, Sharma S. Privacy and security of Aadhaar: a computer science perspective. Econ Polit Wkly 2017; 93–102.

[26] Tiwari PR, Agarwal D, Jain P, et al. India's "Aadhaar" Biometric ID: Structure, Security, and Vulnerabilities. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2022; 13411 LNCS: 672–693.

[27] Zyskind G, Nathan O, Pentland AS. Decentralizing privacy: Using blockchain to protect personal data. Proceedings - 2015 IEEE Security and Privacy Workshops, SPW 2015 2015; 180–184.

[28] Irvin-Erickson Y. Identity fraud victimization: a critical review of the literature of the past two decades. Crime Sci 2024; 13: 1–26.

[29] Kalvet T, Tiits M, Ubakivi-Hadachi P. Risks and Societal Implications of Identity Theft. Communications in Computer and Information Science 2019; 947: 67–81.

[30] Uludag U, Pankanti S, Prabhakar S, et al. Biometric cryptosystems: Issues and challenges. Proceedings of the IEEE 2004; 92: 948–959.

[31] Wang X, Wu YC, Zhou M, et al. Beyond surveillance: privacy, ethics, and regulations in face recognition technology. Front Big Data 2024; 7: 1337465.

[32] European Parliament and Council of the European Union. Regulation (EU) 2024/1183 of the European Parliament and of the Council of 11 April 2024 amending Regulation (EU) No 910/2014 as regards establishing the European Digital Identity Framework. Official Journal of the European Union, L series, 2024/1183, 2024.

[33] Aadhaar. Authentication Requesting Agency - Unique Identification Authority of India | Government of India, https://uidai.gov.in/en/ecosystem/authentication-ecosystem/authentication-requesting-agency.html (2022, accessed 12 November 2026).

[34] Mühle A, Grüner A, Gayvoronskaya T, et al. A survey on essential components of a self-sovereign identity. Comput Sci Rev 2018; 30: 80–86.

[35] Sedlmeir J, Smethurst R, Rieger A, et al. Digital Identities and Verifiable Credentials. Business and Information Systems Engineering 2021; 63: 603–613.

[36] Giannopoulou A. Digital Identity Infrastructures: a Critical Approach of Self-Sovereign Identity. Digit Soc 2023; 2: 1–19. DOI: 10.1007/s44206-023-00049-z.

[37] Wang F, De Filippi P. Self-Sovereign Identity in a Globalized World: Credentials-Based Identity Systems as a Driver for Economic Inclusion. Frontiers in Blockchain 2019; 2: 496586.

[38] Zeydan E, Blanco L, Mangues-Bafalluy J, et al. Blockchain-Based Self-Sovereign Identity: Taking Control of Identity in Federated Learning. IEEE Open Journal of the Communications Society. Epub ahead of print 2024. DOI: 10.1109/OJCOMS.2024.3449692.

[39] Ahmed MR, Islam AKMM, Shatabda S, et al. Blockchain-Based Identity Management System and Self-Sovereign Identity Ecosystem: A Comprehensive Survey. IEEE Access 2022; 10: 113436–113481.

[40] Stockburger L, Kokosioulis G, Mukkamala A, et al. Blockchain-enabled decentralized identity management: The case of self-sovereign identity in public transportation. Blockchain: Research and Applications 2021; 2: 100014.

[41] Glöckler J, Sedlmeir J, Frank M, et al. A Systematic Review of Identity and Access Management Requirements in Enterprises and Potential Contributions of Self-Sovereign Identity. Business and Information Systems Engineering 2023; 66: 421–440.

[42] Mazzocca C, Acar A, Uluagac S, et al. A Survey on Decentralized Identifiers and Verifiable Credentials. IEEE Commun Surv Tutor 2025; 27: 3641–3671. DOI: 10.1109/COMST.2025.3543197.

[43] Dunphy P, Petitcolas FAP. A first look at identity management schemes on the blockchain. IEEE Secur Priv 2018; 16: 20–29.

[44] Alizadeh M, Andersson K, Schelen O. Comparative Analysis of Decentralized Identity Approaches. IEEE Access 2022; 10: 92273–92283.

[45] Naik N, Jenkins P. Self-Sovereign Identity Specifications: Govern Your Identity through Your Digital Wallet using Blockchain Technology. Proceedings - 2020 8th IEEE International Conference on Mobile Cloud Computing, Services, and Engineering, MobileCloud 2020 2020; 90–95.

[46] Bochnia R, Richter D, Anke J. Self-Sovereign Identity for Organizations: Requirements for Enterprise Software. IEEE Access 2024; 12: 7637–7660.

[47] W3C. Decentralized Identifiers (DIDs) v1.0, https://www.w3.org/TR/did-core/ (2022, accessed 12 November 2026).

[48] World Wide Web Consortium. Verifiable credentials data model v2.0. W3C Recommendation. W3C, 2025. https://www.w3.org/TR/2025/REC-vc-data-model-2.0-20250515/.

[49] Fett D, Yasuda K, Campbell B. Selective disclosure for JSON web tokens. RFC 9901. Internet Engineering Task Force, 2025. DOI: 10.17487/RFC9901.

[50] OpenID. OpenID4VC: OpenID for Verifiable Credentials. EU: https://openid.net/sg/openid4vc/, 2024.

[51] Davie M, Gisolfi D, Hardman D, et al. The Trust over IP Stack. IEEE Communications Standards Magazine 2019; 3: 46–51.

[52] Lepore C, Laborde R, Eynard J. Aligning eIDAS and Trust Over IP: A Mapping Approach. ACM International Conference Proceeding Series. Epub ahead of print 30 July 2024. DOI: 10.1145/3664476.3670919.

[53] Badertscher C, Banfi F, Diaz J. What Did Come Out of It? Analysis and Improvements of DIDComm Messaging. In: Proceedings of the 2024 ACM SIGSAC Conference on Computer and Communications Security (CCS '24). New York: ACM, 2024, pp. 4732–4746. DOI: 10.1145/3658644.3690300.

[54] De Prisco R, Shevchenko S, Faruolo P. Enhancing OpenID Connect for Verifiable Credentials with DIDComm. Epub ahead of print 2024. DOI: 10.5220/0012850800003767.

[55] Rodriguez Garzon S, Natusch D, Philipp A, et al. DID Link: Authentication in TLS with Decentralized Identifiers and Verifiable Credentials, https://arxiv.org/abs/2405.07533v2 (2024, accessed 12 November 2026).

[56] Enge A, Satybaldy A, Nowostawski M. An offline mobile access control system based on self-sovereign identity standards. Computer Networks 2022; 219: 109434.

[57] Khovratovich D, Law J. Sovrin: digital identities in the blockchain era. Sovrin Foundation, 2017.

[58] Windley PJ. Sovrin: An Identity Metasystem for Self-Sovereign Identity. Frontiers in Blockchain 2021; 4: 626726.

[59] Reed D, Law J, Hardman D. The technical foundations of Sovrin: a white paper from the Sovrin Foundation. Sovrin Foundation, 2016.

[60] Camenisch J, Lysyanskaya A. An Efficient System for Non-transferable Anonymous Credentials with Optional Anonymity Revocation. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2001; 2045: 93–118.

[61] Shcherbakov A. Hyperledger Indy Besu as a permissioned ledger in Self-sovereign Identity. Lecture Notes in Informatics (LNI), Proceedings - Series of the Gesellschaft fur Informatik (GI) 2024; P-350: 127–137.

[62] Naik N, Jenkins P. Does Sovrin Network Offer Sovereign Identity? ISSE 2021 - 7th IEEE International Symposium on Systems Engineering, Proceedings. Epub ahead of print 13 September 2021. DOI: 10.1109/ISSE51541.2021.9582472.

[63] Simsek A, Tasci B, Yayla O. Implementation Analysis of Cryptography Toolbox in Hyperledger. 14th International Conference on Information Security and Cryptology, ISCTURKEY 2021 - Proceedings 2021; 179–184.

[64] Lundkvist C, Heck R, Torstensson J, et al. uPort: a platform for self-sovereign identity. uPort, 2017.

[65] Yang X, Li W. A zero-knowledge-proof-based digital identity management scheme in blockchain. Comput Secur 2020; 99: 102050.

[66] Muth R, Galal T, Heiss J, et al. Towards Smart Contract-Based Verification of Anonymous Credentials. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2023; 13412 LNCS: 481–498.

[67] Baniata H, Pflanzner T, Feher Z, et al. Latency Assessment of Blockchain-based SSI Applications Utilizing Hyperledger Indy. In: Proceedings of the 12th International Conference on Cloud Computing and Services Science (CLOSER 2022). SciTePress, 2022, pp. 264–271. DOI: 10.5220/0011082300003200.

[68] Siqueira A, Da Conceicao AF, Rocha V. Empirical Evaluation of Self-Sovereign Identity Technology. 2023 IEEE International Conference on Blockchain and Cryptocurrency, ICBC 2023. Epub ahead of print 2023. DOI: 10.1109/ICBC56567.2023.10174979.

[69] Grüner A, Mühle A, Lockenvitz N, et al. Analyzing and comparing the security of self-sovereign identity management systems through threat modeling. Int J Inf Secur 2023; 22: 1231–1248.

[70] Alangot B, Szalachowski P, Dinh TTA, et al. Decentralized Identity Authentication with Auditability and Privacy. Algorithms 2023; 16: 4. DOI: 10.3390/A16010004.

[71] Bhattacharya MP, Zavarsky P, Butakov S. Enhancing the security and privacy of self-sovereign identities on hyperledger indy blockchain. 2020 International Symposium on Networks, Computers and Communications, ISNCC 2020. Epub ahead of print 20 October 2020. DOI: 10.1109/ISNCC49221.2020.9297357.

[72] Shehu A-S. On the Compliance of Self-Sovereign Identity with GDPR Principles: A Critical Review. arXiv preprint arXiv:2409.03624, 2024.

[73] Shuhan MKB, Hasnayeen SM, Das TK, et al. Decentralised identity federations using blockchain. Int J Inf Secur 2024; 23: 2759–2782.

[74] Voigt P, von dem Bussche A. The EU General Data Protection Regulation (GDPR). Epub ahead of print 2017. DOI: 10.1007/978-3-319-57959-7.

[75] Hintze M, El Emam K. Comparing the benefits of pseudonymisation and anonymisation under the GDPR. Journal of Data Protection & Privacy 2018; 2: 145– 158.

[76] Rivest RL, Shamir A, Adleman L. A method for obtaining digital signatures and public-key cryptosystems. Commun ACM 1978; 21: 120–126.

[77] Chen L, Jordan S, Liu Y-K, et al. Report on post-quantum cryptography. NIST IR 8105. Gaithersburg, MD: National Institute of Standards and Technology, 2016. DOI: 10.6028/NIST.IR.8105.

[78] Alagic G, Alperin-Sheriff J, Apon D, et al. Status Report on the Second Round of the NIST Post-Quantum Cryptography Standardization Process. Gaithersburg. Epub ahead of print July 2020. DOI: 10.6028/NIST.IR.8309.

[79] Alagic G, Apon D, Cooper D, et al. Status report on the third round of the NIST post-quantum cryptography standardization process. NIST IR 8413. Gaithersburg, MD: National Institute of Standards and Technology, 2022. DOI: 10.6028/NIST.IR.8413-upd1.

[80] Bos J, Ducas L, Kiltz E, et al. CRYSTALS - Kyber: A CCA-Secure Module-Lattice-Based KEM. Proceedings - 3rd IEEE European Symposium on Security and Privacy, EURO S and P 2018 2018; 353–367.

[81] Ducas L, Kiltz E, Lepoint T, et al. CRYSTALS-Dilithium: A Lattice-Based Digital Signature Scheme. IACR Trans Cryptogr Hardw Embed Syst 2018; 2018: 238–268. DOI: 10.13154/tches.v2018.i1.238-268.

[82] Ducas L, Durmus A, Lepoint T, et al. Lattice signatures and bimodal Gaussians. In: Advances in Cryptology – CRYPTO 2013. Lecture Notes in Computer Science, vol. 8042. Berlin, Heidelberg: Springer, 2013, pp. 40–56. DOI: 10.1007/978-3-642-40041-4_3.

[83] Bernstein DJ, Hülsing A, Kölbl S, Niederhagen R, Rijneveld J, Schwabe P. The SPHINCS+ signature framework. In: Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security (CCS '19). New York: ACM, 2019, pp. 2129–2146. DOI: 10.1145/3319535.3363229.

[84] Regev O. On Lattices, Learning with Errors, Random Linear Codes, and Cryptography. Journal of the ACM 2009; 56(6): 34:1–34:40.

[85] Lyubashevsky V, Peikert C, Regev O. On Ideal Lattices and Learning with Errors Over Rings. In: Advances in Cryptology – EUROCRYPT 2010. Lecture Notes in Computer Science, vol. 6110. Berlin, Heidelberg: Springer, 2010, pp. 1–23. DOI: 10.1007/978-3-642-13190-5_1.

[86] Peikert C. A Decade of Lattice Cryptography. Foundations and Trends® in Theoretical Computer Science 2016; 10: 283–424. DOI: 10.1561/0400000074.

[87] Micciancio D, Regev O. Lattice-based Cryptography. In: Bernstein DJ, Buchmann J, Dahmen E (eds) Post-Quantum Cryptography. Berlin, Heidelberg: Springer, 2009, pp. 147–191. DOI: 10.1007/978-3-540-88702-7_5.

[88] Bernstein DJ. Post-Quantum Cryptography. In: van Tilborg HCA, Jajodia S (eds) Encyclopedia of Cryptography and Security. Boston, MA: Springer, 2011, pp. 949–950. DOI: 10.1007/978-1-4419-5906-5_386.

[89] Alagic G, Bros M, Ciadoux P, et al. Status Report on the Fourth Round of the NIST Post-Quantum Cryptography Standardization Process. NIST IR 8545. Gaithersburg, MD: National Institute of Standards and Technology, 2025. DOI: 10.6028/NIST.IR.8545.

[90] Hoffstein J, Howgrave-Graham N, Pipher J, et al. Practical Lattice-Based Cryptography: NTRUEncrypt and NTRUSign. Information Security and Cryptography 2009; 10: 349– 390.

[91] National Institute of Standards and Technology. FIPS 206 status update (NIST pqc-forum, 28 August 2025), https://groups.google.com/a/list.nist.gov/g/pqc-forum/c/1HXzjlMUU6Y (2025, accessed 12 November 2026).

[92] Castryck W, Decru T. An Efficient Key Recovery Attack on SIDH. In: Hazay C, Stam M (eds) Advances in Cryptology – EUROCRYPT 2023. Lecture Notes in Computer Science, vol. 14008. Cham: Springer Nature Switzerland, 2023, pp. 423–447. DOI: 10.1007/978-3-031-30589-4_15.

[93] Liu YK, Moody D. Post-quantum cryptography and the quantum future of cybersecurity. Phys Rev Appl 2024; 21: 040501.

[94] Castelvecchi D. The race to save the Internet from quantum hackers. Nature 2022; 602: 198–201.

[95] NIS Cooperation Group. A Coordinated Implementation Roadmap for the Transition to Post-Quantum Cryptography. European Commission, 2025. https://digital-strategy.ec.europa.eu/en/library/coordinated-implementation-roadmap-transition-post-quantum-cryptography.

[96] National Cyber Security Centre. Timelines for migration to post-quantum cryptography, https://www.ncsc.gov.uk/guidance/pqc-migration-timelines (2025, accessed 12 November 2026).

[97] Hamdi A, Fourati L, Ayed S. Vulnerabilities and attacks assessments in blockchain 1.0, 2.0 and 3.0: tools, analysis and countermeasures. Int J Inf Secur 2024; 23: 713–757.

[98] Fernandez V, Orue AB, Arroyo D. Securing Blockchain with Quantum Safe Cryptography: When and How? Advances in Intelligent Systems and Computing 2021; 1267 AISC: 371–379.

[99] Yunakovsky SE, Kot M, Pozhar N, et al. Towards security recommendations for public-key infrastructures for production environments in the post-quantum era. EPJ Quantum Technology 2021 8:1 2021; 8: 1–19.

[100] Alkim E, Ducas L, Pöppelmann T, et al. Post-quantum key exchange – a new hope. In: 25th USENIX Security Symposium (USENIX Security 16). Austin, TX: USENIX Association, 2016, pp. 327–343.

[101] Bellizia D, El Mrabet N, Fournaris AP, et al. Post-quantum cryptography: Challenges and opportunities for robust and secure HW design. In: 2021 IEEE International Symposium on Defect and Fault Tolerance in VLSI and Nanotechnology Systems (DFT). IEEE, 2021. DOI: 10.1109/DFT52944.2021.9568301.

[102] Kumar M. Post-quantum cryptography Algorithm's standardization and performance analysis. Array 2022; 15: 100242.

[103] Kuznetsov O, Zakharov D, Frontoni E. Deep learning-based biometric cryptographic key generation with post-quantum security. Multimed Tools Appl 2024; 83: 56909– 56938.

[104] He Y, Wang E, Rong Y, et al. Security of AI Agents, https://arxiv.org/abs/2406.08689v2 (2024, accessed 12 November 2026).

[105] OWASP Gen AI Security Project. Agentic AI – Threats and Mitigations. OWASP Foundation, 2025. https://genai.owasp.org/resource/agentic-ai-threats-and-mitigations/.

[106] OWASP Gen AI Security Project. OWASP Top 10 for Agentic Applications 2026. OWASP Foundation, 2025. https://genai.owasp.org/resource/owasp-top-10-for-agentic-applications-for-2026/.

[107] Greshake K, Abdelnabi S, Mishra S, et al. Not What You've Signed Up For: Compromising Real-World LLM-Integrated Applications with Indirect Prompt Injection. In: Proceedings of the 16th ACM Workshop on Artificial Intelligence and Security. New York, NY, USA: ACM, 2023, pp. 79–90. DOI: 10.1145/3605764.3623985.

[108] Goodfellow I, Pouget-Abadie J, Mirza M, et al. Generative Adversarial Networks. Commun ACM 2020; 63: 139–144. DOI: 10.1145/3422622.

[109] Karras T, Laine S, Aittala M, et al. Analyzing and Improving the Image Quality of StyleGAN. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2020, pp. 8110–8119.

[110] Wu Z, De Leon PL, Demiroglu C, et al. Anti-spoofing for text-independent speaker verification. IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP) 2016; 24: 768–783.

[111] Lee H, Lee C, Farhat K, et al. The Tug-of-War Between Deepfake Generation and Detection, https://arxiv.org/abs/2407.06174v4 (2024, accessed 12 November 2026).

[112] Cross C. Using artificial intelligence (AI) and deepfakes to deceive victims: the need to rethink current romance fraud prevention messaging. Crime Prevention and Community Safety 2022; 24: 30–41.

[113] Khan A, Malik KM. Securing Voice Biometrics: One-Shot Learning Approach for Audio Deepfake Detection. WIFS 2023 - IEEE Workshop on Information Forensics and Security. Epub ahead of print 5 October 2023. DOI: 10.1109/WIFS58808.2023.10374968.

[114] Caldwell M, Andrews JTA, Tanay T, et al. AI-enabled future crime. Crime Sci 2020; 9: 1– 13.

[115] Sarker IH, Furhad MH, Nowrozy R. AI-Driven Cybersecurity: An Overview, Security Intelligence Modeling and Research Directions. SN Comput Sci 2021; 2: 1–18.

[116] Douceur JR. The Sybil Attack. In: Peer-to-Peer Systems (IPTPS 2002). Lecture Notes in Computer Science, vol. 2429. Berlin, Heidelberg: Springer, 2002, pp. 251–260. DOI: 10.1007/3-540-45748-8_24.

[117] Yu H, Kaminsky M, Gibbons PB, et al. SybilGuard: Defending Against Sybil Attacks via Social Networks. IEEE/ACM Trans Netw 2008; 16: 576–589. DOI: 10.1109/TNET.2008.923723.

[118] Danezis G, Mittal P. SybilInfer: Detecting Sybil Nodes using Social Networks. In: 16th Annual Network and Distributed System Security Symposium (NDSS 2009). San Diego, CA: Internet Society, 2009.

[119] Alvisi L, Clement A, Epasto A, et al. SoK: The evolution of sybil defense via social networks. Proc IEEE Symp Secur Priv 2013; 382–396.

[120] Siddarth D, Ivliev S, Siri S, et al. Who Watches the Watchmen? A Review of Subjective Approaches for Sybil-Resistance in Proof of Personhood Protocols. Frontiers in Blockchain 2020; 3: 590171.

[121] Platt M, McBurney P. Sybil attacks on identity-augmented Proof-of-Stake. Computer Networks 2021; 199: 108424.

[122] Sánchez DC. Zero-Knowledge Proof-of-Identity: Sybil-Resistant, Anonymous Authentication on Permissionless Blockchains and Incentive Compatible, Strictly Dominant Cryptocurrencies. SSRN Electronic Journal. Epub ahead of print 22 May 2019. DOI: 10.2139/ssrn.3392331.

[123] Chan A, Kolt N, Wills P, et al. IDs for AI Systems. arXiv preprint arXiv:2406.12137, 2024. DOI: 10.48550/arXiv.2406.12137.

[124] Chan A, Wei K, Huang S, et al. Infrastructure for AI Agents. Transactions on Machine Learning Research. Epub ahead of print 2025.

[125] South T, Nagabhushanaradhya S, Dissanayaka A, et al. Identity Management for Agentic AI: The new frontier of authorization, authentication, and security for an AI agent world. OpenID Foundation, 2025. https://openid.net/wp-content/uploads/2025/10/Identity-Management-for-Agentic-AI.pdf.

[126] Hardt D. The OAuth 2.0 authorization framework. RFC 6749. Internet Engineering Task Force, 2012. DOI: 10.17487/RFC6749.

[127] Jones MB, Nadalin A, Campbell B, et al. OAuth 2.0 Token Exchange. RFC 8693. Internet Engineering Task Force, 2020. DOI: 10.17487/RFC8693.

[128] Rodriguez Garzon S, Vaziry A, Kuzu EM, et al. AI Agents with Decentralized Identifiers and Verifiable Credentials. In: Proceedings of the 18th International Conference on Agents and Artificial Intelligence. SCITEPRESS - Science and Technology Publications, 2026, pp. 252–259. DOI: 10.5220/0014234400004052.

[129] Huang K, Narajala VS, Yeoh J, et al. A Novel Zero-Trust Identity Framework for Agentic AI: Decentralized Authentication and Fine-Grained Access Control. In: 2026 International Conference on AI x Data and Knowledge Engineering (AIxDKE). IEEE, 2026, pp. 98–101. DOI: 10.1109/AIxDKE67294.2026.00024.

[130] Lauinger J, Ernstberger J, Regnath E, et al. A-PoA: Anonymous proof of authorization for decentralized identity management. IEEE International Conference on Blockchain and Cryptocurrency, ICBC 2021. Epub ahead of print 3 May 2021. DOI: 10.1109/ICBC51069.2021.9461082.

[131] Protogerou A, Papadopoulos S, Drosou A, et al. A graph neural network method for distributed anomaly detection in IoT. Evol Syst 2021; 12: 19–36. DOI: 10.1007/S12530-020-09347-0.

[132] Gronauer S, Diepold K. Multi-agent deep reinforcement learning: a survey. Artif Intell Rev 2022; 55: 895–943.

[133] Ahmadi C, Chen JL. Survey on Reinforcement Learning Techniques for Enhancing Security and Efficiency in Zero Trust Networks. Proceedings of the 2024 10th International Conference on Applied System Innovation, ICASI 2024 2024; 427–429.

[134] Shashkov A, Hemberg E, Tulla M, et al. Adversarial agent-learning for cybersecurity: a comparison of algorithms. Knowl Eng Rev 2023; 38: e3.

[135] Gheibi O, Weyns D, Quin F. Applying Machine Learning in Self-adaptive Systems. ACM Transactions on Autonomous and Adaptive Systems (TAAS); 15. Epub ahead of print 18 August 2021. DOI: 10.1145/3469440.

[136] Repetto M. Adaptive monitoring, detection, and response for agile digital service chains. Comput Secur 2023; 132: 103343.

[137] Ressi D, Romanello R, Piazza C, et al. AI-enhanced blockchain technology: A review of advancements and opportunities. Journal of Network and Computer Applications 2024; 225: 103858.

[138] National Institute of Standards and Technology. Artificial Intelligence Risk Management Framework (AI RMF 1.0). NIST AI 100-1. Gaithersburg, MD: National Institute of Standards and Technology, 2023. DOI: 10.6028/NIST.AI.100-1.

[139] Habbal A, Ali MK, Abuzaraida MA. Artificial Intelligence Trust, Risk and Security Management (AI TRiSM): Frameworks, applications, challenges and future research directions. Expert Syst Appl 2024; 240: 122442.

[140] Zhou S, Liu C, Ye D, et al. Adversarial Attacks and Defenses in Deep Learning: From a Perspective of Cybersecurity. ACM Comput Surv; 55. Epub ahead of print 23 December 2022. DOI: 10.1145/3547330.

[141] Shostack A. Threat modeling: designing for security. Indianapolis, IN: Wiley, 2014.

[142] Kingma DP, Welling M. Auto-Encoding Variational Bayes. In: 2nd International Conference on Learning Representations (ICLR 2014). 2014. https://arxiv.org/abs/1312.6114v11.

[143] Baltsou G, Sarridis I, Koutlis C, et al. SDFD: Building a Versatile Synthetic Face Image Dataset with Diverse Attributes, https://arxiv.org/abs/2404.17255v2 (2024, accessed 12 November 2026).

[144] Breidenbach L, Cachin C, Chan B, et al. Chainlink 2.0: Next Steps in the Evolution of Decentralized Oracle Networks. Chainlink Labs, 2021.

[145] Pasdar A, Lee YC, Dong Z. Connect API with Blockchain: A Survey on Blockchain Oracle Implementation. ACM Comput Surv; 55. Epub ahead of print 2 February 2023. DOI: 10.1145/3567582.

[146] Ezzat SK, Saleh YNM, Abdel-Hamid AA. Blockchain Oracles: State-of-the-Art and Research Directions. IEEE Access 2022; 10: 67551–67572.

[147] Belchior R, Vasconcelos A, Correia M, et al. Hermes: Fault-tolerant middleware for blockchain interoperability. Future Generation Computer Systems 2022; 129: 236–251.

[148] Fotiou N, Siris VA, Polyzos GC. Enabling self-verifiable mutable content items in IPFS using Decentralized Identifiers. 2021 IFIP Networking Conference, IFIP Networking 2021. Epub ahead of print 21 June 2021. DOI: 10.23919/IFIPNETWORKING52078.2021.9472820.

[149] Wood G. Ethereum: A secure decentralised generalised transaction ledger. Ethereum Project Yellow Paper, 2014. https://ethereum.github.io/yellowpaper/paper.pdf.

[150] Androulaki E, Barger A, Bortnikov V, et al. Hyperledger fabric: a distributed operating system for permissioned blockchains. In: Proceedings of the Thirteenth EuroSys Conference (EuroSys '18). New York: ACM, 2018, pp. 1–15. DOI: 10.1145/3190508.3190538.

[151] Kwiatkowski K, Kampanakis P, Westerbaan BE, et al. Post-Quantum Traditional (PQ/T) Hybrid Key Agreement Mechanisms for TLS 1.3. RFC 10024. Internet Engineering Task Force, 2026. DOI: 10.17487/RFC10024.

[152] Microsoft. Microsoft SEAL: Fast and Easy-to-Use Homomorphic Encryption Library, https://www.microsoft.com/en-us/research/project/microsoft-seal/ (2019, accessed 12 November 2026).

[153] Polyakov Y, Rohloff K, Ryan GW, et al. PALISADE lattice cryptography library user manual. Technical report. Cybersecurity Research Center, New Jersey Institute of Technology, 2017.

[154] Halevi S, Shoup V. Algorithms in HElib. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2014; 8616 LNCS: 554–571.

[155] Gentry C. Fully Homomorphic Encryption Using Ideal Lattices. In: Proceedings of the 41st Annual ACM Symposium on Theory of Computing (STOC '09). New York: ACM, 2009, pp. 169–178. DOI: 10.1145/1536414.1536440.

[156] Halevi S, Polyakov Y, Shoup V. An Improved RNS Variant of the BFV Homomorphic Encryption Scheme. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2019; 11405 LNCS: 83–105.

[157] Geelen R, Vercauteren F. Bootstrapping for BGV and BFV Revisited. Journal of Cryptology 2023; 36: 1–32.

[158] Munjal K, Bhatia R. A systematic review of homomorphic encryption and its contributions in healthcare industry. Complex and Intelligent Systems 2023; 9: 3759– 3786.

[159] Stebila D, Mosca M. Post-quantum Key Exchange for the Internet and the Open Quantum Safe Project. In: Avanzi R, Heys H (eds) Selected Areas in Cryptography – SAC 2016. Lecture Notes in Computer Science, vol. 10532. Cham: Springer International Publishing, 2017, pp. 14–37. DOI: 10.1007/978-3-319-69453-5_2.

[160] Albrecht MR, Player R, Scott S. On the concrete hardness of Learning with Errors. Journal of Mathematical Cryptology 2015; 9: 169–203. DOI: 10.1515/jmc-2015-0016.

[161] International Organization for Standardization. Information technology — Automatic identification and data capture techniques — QR Code bar code symbology specification. ISO/IEC 18004:2015. Geneva: ISO, 2015.

[162] Ethereum Improvement Proposals. EIP-2028: transaction data gas cost reduction, https://eips.ethereum.org/EIPS/eip-2028 (2019, accessed 12 November 2026).

[163] Ethereum Improvement Proposals. EIP-7623: increase calldata cost, https://eips.ethereum.org/EIPS/eip-7623 (2024, accessed 12 November 2026).

[164] Chhetri R. Benchmarking NIST-Standardised ML-KEM and ML-DSA on ARM Cortex-M0+: Performance, Memory, and Energy on the RP2040. arXiv preprint arXiv:2603.19340, 2026. DOI: 10.48550/arXiv.2603.19340.

[165] Raavi M, Khan Q, Wuthier S, et al. Security and Performance Analyses of Post-Quantum Digital Signature Algorithms and Their TLS and PKI Integrations. Cryptography 2025; 9: 38. DOI: 10.3390/cryptography9020038.

[166] El-Hajj M, Oude Roelink B. Evaluating the Efficiency of zk-SNARK, zk-STARK, and Bulletproof in Real-World Scenarios: A Benchmark Study. Information 2024; 15: 463. DOI: 10.3390/info15080463.

[167] Ernstberger J, Chaliasos S, Kadianakis G, et al. zk-Bench: A Toolset for Comparative Evaluation and Performance Benchmarking of SNARKs. LNCS 2024; 14973: 46–72.

[168] Ben-Sasson E, Chiesa A, Tromer E, et al. Succinct Non-Interactive Zero Knowledge for a von Neumann Architecture. In: 23rd USENIX Security Symposium (USENIX Security 14). San Diego, CA: USENIX Association, 2014, pp. 781–796.

[169] Ben-Sasson E, Chiesa A, Garman C, et al. Zerocash: Decentralized anonymous payments from bitcoin. Proc IEEE Symp Secur Priv 2014; 459–474.

[170] Bitansky N, Canetti R, Chiesa A, et al. From extractable collision resistance to succinct non-interactive arguments of knowledge, and back again. ITCS 2012 - Innovations in Theoretical Computer Science Conference 2012; 326–349.

[171] Bünz B, Bootle J, Boneh D, et al. Bulletproofs: Short Proofs for Confidential Transactions and More. In: 2018 IEEE Symposium on Security and Privacy (SP). San Francisco, CA: IEEE, 2018, pp. 315–334. DOI: 10.1109/SP.2018.00020.

[172] Ben-Sasson E, Bentov I, Horesh Y, et al. Scalable, transparent, and post-quantum secure computational integrity. 2018/046. Cryptology ePrint Archive, 2018. https://eprint.iacr.org/2018/046.

[173] Chiesa A, Manohar P, Spooner N. Succinct Arguments in the Quantum Random Oracle Model. In: Hofheinz D, Rosen A (eds) Theory of Cryptography. Lecture Notes in Computer Science, vol. 11892. Cham: Springer International Publishing, 2019, pp. 1– 29. DOI: 10.1007/978-3-030-36033-7_1.

[174] Chaum DL. Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms. Commun ACM 1981; 24: 84–90.

[175] Guo C, Lang F, Wang Q, et al. UP-SSO: Enhancing the User Privacy of SSO by Integrating PPID and SGX. 2021 International Conference on Advanced Computing and Endogenous Security, ICACES 2021. Epub ahead of print 2021. DOI: 10.1109/IEEECONF52377.2022.10013340.

[176] Hammann S, Sasse R, Basin D. Privacy-Preserving OpenID Connect. Proceedings of the 15th ACM Asia Conference on Computer and Communications Security, ASIA CCS 2020 2020; 277–289.

[177] Chaum D. Blind Signatures for Untraceable Payments. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Springer, Boston, MA, pp. 199–203.

[178] Prorock M, Steele O. ML-DSA for JSON Object Signing and Encryption (JOSE) and CBOR Object Signing and Encryption (COSE). RFC 9964. Internet Engineering Task Force, 2026. DOI: 10.17487/RFC9964.

[179] World Wide Web Consortium. Bitstring status list v1.0. W3C Recommendation. W3C, 2025. https://www.w3.org/TR/2025/REC-vc-bitstring-status-list-20250515/.

[180] Evans D, Kolesnikov V, Rosulek M. A Pragmatic Introduction to Secure Multi-Party Computation. Foundations and Trends® in Privacy and Security 2018; 2: 70–246.

[181] Goldreich O. Foundations of cryptography: volume 2, basic applications. Cambridge: Cambridge University Press, 2009.

[182] Grassi L, Khovratovich D, Rechberger C, et al. Poseidon: A New Hash Function for Zero-Knowledge Proof Systems. In: 30th USENIX Security Symposium (USENIX Security 21). USENIX Association, 2021, pp. 519–535. https://www.usenix.org/conference/usenixsecurity21/presentation/grassi.

[183] Liu J, Kretz I, Liu H, et al. Certifying Zero-Knowledge Circuits with Refinement Types. 2024 IEEE Symposium on Security and Privacy (SP) 2024; 1741–1759.

[184] Bootle J, Lyubashevsky V, Nguyen NK, et al. A Framework for Practical Anonymous Credentials from Lattices. In: Handschuh H, Lysyanskaya A (eds) Advances in Cryptology – CRYPTO 2023. Lecture Notes in Computer Science, vol. 14082. Cham: Springer Nature Switzerland, 2023, pp. 384–417. DOI: 10.1007/978-3-031-38545-2_13.

[185] Ounsworth M, Gray J, Pala M, Klaußner J, Fluhrer S. Composite Module-Lattice-Based Digital Signature Algorithm (ML-DSA) for use in X.509 Public Key Infrastructure. RFC 10073. Internet Engineering Task Force, 2026. DOI: 10.17487/RFC10073.

[186] National Institute of Standards and Technology. NIST selects HQC as fifth algorithm for post-quantum encryption, https://www.nist.gov/news-events/news/2025/03/nist-selects-hqc-fifth-algorithm-post-q

uantum-encryption (2025, accessed 12 November 2026).

[187] Helminger L, Kales D, Ramacher S, et al. Multi-party Revocation in Sovrin: Performance through Distributed Trust. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) 2021; 12704 LNCS: 527–551.

[188] Camenisch J, Lysyanskaya A. Dynamic Accumulators and Application to Efficient Revocation of Anonymous Credentials. In: Yung M (eds) Advances in Cryptology — CRYPTO 2002. Lecture Notes in Computer Science, vol. 2442. Berlin, Heidelberg: Springer Berlin Heidelberg, 2002, pp. 61–76. DOI: 10.1007/3-540-45708-9_5.

[189] Hoops F, Gebele J, Matthes F. CRSet: Private Non-Interactive Verifiable Credential Revocation. arXiv preprint arXiv:2501.17089, 2025. DOI: 10.48550/arXiv.2501.17089.

[190] Bansod S, Ragha L. Challenges in making blockchain privacy compliant for the digital world: some measures. Sadhana - Academy Proceedings in Engineering Sciences 2022; 47: 1–17.

[191] Wang J, Del Vecchio D, Humphrey M. Extending the security assertion markup language to support delegation for Web Services and Grid Services. Proceedings -2005 IEEE International Conference on Web Services, ICWS 2005 2005; 2005: 67–74.

[192] Radanliev P. The Rise of AI Agents: Integrating AI, Blockchain Technologies, and Quantum Computing. Pearson, 2025.

[193] Young W, Leveson NG. An integrated approach to safety and security based on systems theory. Communications of the ACM 2014; 57: 31–35. DOI: 10.1145/2556938.

[194] Leveson NG, Thomas JP. STPA Handbook. Cambridge, MA: MIT Partnership for Systems Approaches to Safety and Security, 2018. https://psas.scripts.mit.edu/home/get_file.php?name=STPA_handbook.pdf.

[195] Fett D, Kusters R, Schmitz G. The Web SSO Standard OpenID Connect: In-depth Formal Security Analysis and Security Guidelines. Proceedings - IEEE Computer Security Foundations Symposium 2017; 189–202.