

Supplementary Materials

Agentic Cryptographic Debt: Repository-Level Measurement of Post-Quantum Migration Regression Under Autonomous AI Software Development

Robert Campbell, Independent Researcher, Upper Marlboro, MD 20774, USA

This file accompanies the manuscript. All items listed here are also present in the archived replication package, whose location is given in the Data Availability Statement (Zenodo, DOI 10.5281/zenodo.21856155).

Table S1. Feasibility audit, per artifact. Each row records the artifact audited, its version, an install digest identifying the exact files inspected, what the library's own registries expose, and the verbatim result of each key-generation attempt. Install digests are the SHA-256 over the sorted per-file SHA-256 hashes of the package's Python sources, emitted by `audit_jose_support.py` and shown truncated to sixteen characters; the full values, the per-file hashes, and the complete generation transcripts are in the evidence bundle `jose-support-audit-2026-07-23.json`, audited 2026-07-23T03:57:03Z, which is 22 July 2026 in the author's local time zone and the date given in Table 2. Controls are run for the two libraries whose generation API accepts a key type at call time; the other three expose algorithms through a registry only and are read from it directly, so no control applies. The CIRCL entry gives the Go module content hash from `go.sum`, which is verifiable against the public Go checksum database and therefore identifies the audited artifact independently of this study; the Python entries give a digest computed locally over the installed sources, which identifies the files inspected but is not independently attested.

Artifact	Version	Install Digest (SHA-256)	Registries Exposed	RFC 9964 Generation Attempts	Control	Source Search
Authlib	1.7.2	f858b467a5d5e85e...	key types EC, OKP, RSA, oct; no AKP	<code>generate_key('ML-DSA', 'ML-DSA-44')</code> , <code>('ML-DSA-44', 44)</code> , <code>('AKP', 'ML-DSA-44')</code> , <code>('MLDSA', 44)</code> each raised <code>KeyError</code> naming the key type	<code>generate_key('OKP', 'Ed25519')</code> succeeded	0 hits for ML-DSA, MLDSA, AKP, Dilithium
joserfc	1.7.4	96e2a0f97ae96757...	17 JWS algorithms, all classical; key types EC, OKP, Oct, RSA; no AKP	<code>generate_key</code> for ML-DSA-44, ML-DSA-65, ML-DSA-87, AKP, ML-DSA each raised <code>InvalidKeyTypeError: Invalid key type</code>	<code>generate_key('OKP')</code> succeeded	0 hits
PyJWT	2.13.0	d89905e6fedb1608...	16 algorithms, all classical, with cryptography installed so that asymmetric algorithms register	no ML-DSA or AKP identifier registered	not applicable, registry read directly	0 hits
jwcrypto	1.5.8	d0acb9ab82dd5fc9...	43 registered algorithms, all classical	no ML-DSA or AKP identifier registered	not applicable, registry read directly	0 hits
python-jose	3.5.0	8076fd3e0c9eb7d9...	classical algorithm constants only	no ML-DSA or AKP constant	not applicable, registry read directly	0 hits
CIRCL sign/mldsa/mldsa44	v1.6.4	h1:p0XuDTCEYzydgUp0QCQz3 LsinKjiSk6nNP5Lt5K64U=	<code>GenerateKey</code> , <code>NewKeyFromSeed</code> , <code>SignTo</code> , <code>Verify</code> , <code>PublicKey</code> , <code>PrivateKey</code>	not applicable; support present	sign and verify round-trip succeeded; tampered message failed verification	n/a

A library is recorded as exposing support only on registry evidence or a successful generation. Control generations distinguish absent support from a malformed invocation. The audit is bounded to the artifacts named above.

Replication package contents. The package is indexed so that a claim in this paper can be traced to the artifact supporting it. It contains `MANIFEST.md`, listing every file with its size and SHA-256, and `RUN_INDEX.md`, one row per scored run giving repository, feasibility, model and checkpoint, condition, prompt variant, run identifier, the paths of the run record and transcript, the assembled prompt hash, the Layer 1 and Layer 2 findings, and the merged verdict where the record carries one. The index carries no Layer 3 column

because that layer was not invoked (Section 4.5). Both are generated by `build_manifest.py`, which also audits the package against the requirements below and reports what is present, partial, or absent.

Beyond the indexes the package contains: the harness and bounded agent loop; the scorer, all three layers as implemented, with its validation fixtures and the scoring specification that defines every finding code; the task specifications for all four primary conditions and the nested ablation; the seeded classical baselines; the run record and transcript for each of the 56 scored runs, including build output and, where captured, test output; the seven excluded records in a separate directory so the exclusions can be inspected; the thirty context-sensitivity ablation responses with their re-derivation script and its output; the model-provenance bundle from `capture_model_provenance.py` with manifest and weights digests, quantization, and context length; the feasibility-audit evidence bundle from `audit_jose_support.py` with per-library install digests, registry contents, and verbatim exceptions; environment lockfiles including pip freeze and the pinned toolchain versions; the correction chronology of Table 13; and the verification scripts `verify_allocation.py`, `verify_apparatus_currency.py`, `rescore_record.py`, `rescore_from_raw.py`, and `count_green_states.py`. A README, a licence, and a third-party licensing notice covering the two corpus repositories are included, with commands to reproduce a single run.

What the package does not contain. Three categories of material are absent and we state them rather than leave them to be discovered. The transcripts record the path and byte size of each file the agent wrote but not its contents, so an applied artifact can be identified but not reconstructed from the transcript alone; the applied working trees were temporary and were not retained, and where a run's archived raw response exists the returned files can be recovered from it. Pre- and post-run git-tree hashes were not recorded by the harness, so a reader cannot verify the repository state independently of the run record. And of the exploratory material of Section 6.3, the direct-question probes are not archived; the thirty context-sensitivity responses are archived as raw response files with the re-derivation script and its output, but not as scorer-format run records, and are not part of the 56-run dataset; the design of the systematic ablation specified to replace them is included.