Know Your Researcher Bio: Portable Authorization for AI-Bio Tools and Benchtop DNA Synthesizers[1]

Ashton Chew
Cornell University

With
Apart Research

Abstract

Sequence screening has matured faster than portable requester authorization. Whether a requester has been reviewed and remains authorized is still answered through ad hoc dossier review at every provider, AI-bio tool, and equipment vendor. KYR-Bio prototypes that missing layer: a reviewed, scoped, holder-bound credential that AI-bio tools, synthesis checkout flows, and benchtop DNA synthesizers verify locally without re-sharing the applicant dossier. Verifiers run signature, issuer governance, Bitstring Status List freshness, holder proof, and scope checks before a policy adapter applies local rules. Audit events carry a hash chain and rolling Merkle root, and exclude raw biological prompts, sequences, and reviewer notes. The synthetic evaluation passes 8 persona routing cases, 22 verifier cases, 18 AI-assistance behavioral checks, and 36 audit events with no forbidden content. Property-based and adversarial suites, an OPA policy sidecar, and W3C VC 2.0 and OpenID4VC export shapes widen that surface. KYR-Bio complements sequence screening as a valuable tool. The reusable artifact is the claim schema, scope vocabulary, verifier hard-check order, reason-code surface, and audit boundary.

[1] Research conducted at the AIxBio Hackathon, April 2026

1. Introduction

Recent biosecurity policy directs sequence screening at the content level: which sequences, which providers, which reporting (OSTP 2024 **Framework for Nucleic Acid Synthesis Screening**; HHS 2023 **Screening Framework Guidance**; NIH NOT-OD-25-012; the May 2025 executive order **Improving the Safety and Security of Biological Research**; and the proposed European Biotech Act, COM/2025/1022). Benchtop synthesis devices push that governance from a few centralized providers to many distributed instruments (NTI 2023; Langenkamp 2024). The U.S. policy environment is in active revision rather than settled.

A second question is largely unsolved by sequence screening alone: who is the requester, what scope of access have they been reviewed for, is that approval still active, and is the entity presenting the credential actually the reviewed holder? Today, that question is answered through ad hoc dossier review at each provider, AI-bio tool, or equipment vendor. The result is fragmented assurance, repeated dossier disclosure, and friction that disproportionately disadvantages independent, nontraditional, and resource-constrained researchers.

We provide that missing layer. **Know Your Researcher Bio** (KYR-Bio) packages a single human-reviewed access decision into a scoped, signed credential that a researcher's wallet can present at any participating relying party. The relying party verifies signature, issuer governance, status, holder proof, scope, and policy context locally, and writes a privacy-minimized hash-chained audit event without ingesting the underlying dossier or reviewer notes. Three relying-party demos exercise the loop end to end: a trusted AI-bio portal, a low-risk synthesis checkout, and a benchtop authorization page. KYR-Bio complements, not replaces, sequence screening.

Our main contributions are:

1. An end-to-end local trust loop spanning structured applicant onboarding, human reviewer decisions, scoped credential issuance, holder-bound wallet presentation, verifier policy decisions, and privacy-minimized audit events.
2. A privacy-minimized verifier interaction pattern that discloses only credential subject pseudonym, scope, expiry, holder-binding fingerprint, decision, and reason codes, never the full applicant dossier or reviewer notes.
3. A robustness layer over the workflow-invariant claim: a fast-check property suite over verifier hard-check invariants, adversarial onboarding and verifier suites, a Bitstring Status List freshness check, an OPA policy sidecar with TypeScript fallback, a hash-chained audit with a rolling Merkle root, and W3C VC 2.0 and OpenID4VC export shapes.

To our knowledge, public work has not yet demonstrated a portable, cross-verifier researcher-authorization flow that connects managed AI-bio access, synthesis checkout, and

benchtop authorized-user gating while preserving local verifier policy and minimizing dossier disclosure.

2. Related Work

Sequence screening focuses on hazard matching. SecureDNA describes a privacy-preserving federated screening system for global DNA synthesis (Baum et al., 2026; Sherman et al., 2025). IBBIS's Common Mechanism contributes open HMM-based biorisk tooling. NTI's 2023 *Benchtop DNA Synthesis Devices* report and Langenkamp (2024) frame benchtop devices as a distinct governance challenge requiring customer verification alongside sequence screening, and Wittmann et al. (*Science* 2025) and Edison, Toner, and Esvelt (*Nature Communications* 2026) argue for complementary controls against AI-designed variants and fragment-assembly threats.

For customer screening, IBBIS publishes operational resources and a 2025 white paper on emerging standards. RAND (2024, 2025) argues that customer screening is promising but providers lack mechanisms to share suspicious-customer information. NTI's 2026 *Framework for Managed Access to Biological AI Tools* calls for tiered access and legitimacy verification. Commercial provider-local compliance (e.g., Aclid; Bits in Bio, 2023) automates KYC inside one provider rather than across verifiers.

Concurrent independent work by Feldman, Feldman, and Anton (*Frontiers in Microbiology*, 2026) proposes a three-tier KYC framework for AI-bio access and notes that existing user-verification proposals "remain at the level of single-sentence recommendations rather than implementable architectures." KYR-Bio converges on the user-verification thesis but extends it as a signed, holder-bound, scoped credential that travels portably across independent verifiers, with cryptographic checks and privacy-minimized audit in place of institutional vouching. W3C VC 2.0, RFC 9901 SD-JWT, OpenID4VCI/VP, WebAuthn Level 3, and OPA collectively provide the standards substrate; KYR-Bio composes them. The gap is the missing portable, cross-verifier researcher-authorization artifact those threads imply.

Table 1. KYR-Bio compared with adjacent approaches.

Adjacent approach	Primary focus	How KYR-Bio differs
Sequence screening (SecureDNA, Common Mechanism)	Detects hazardous sequence content.	Authorization; checks the requester's scope and status.

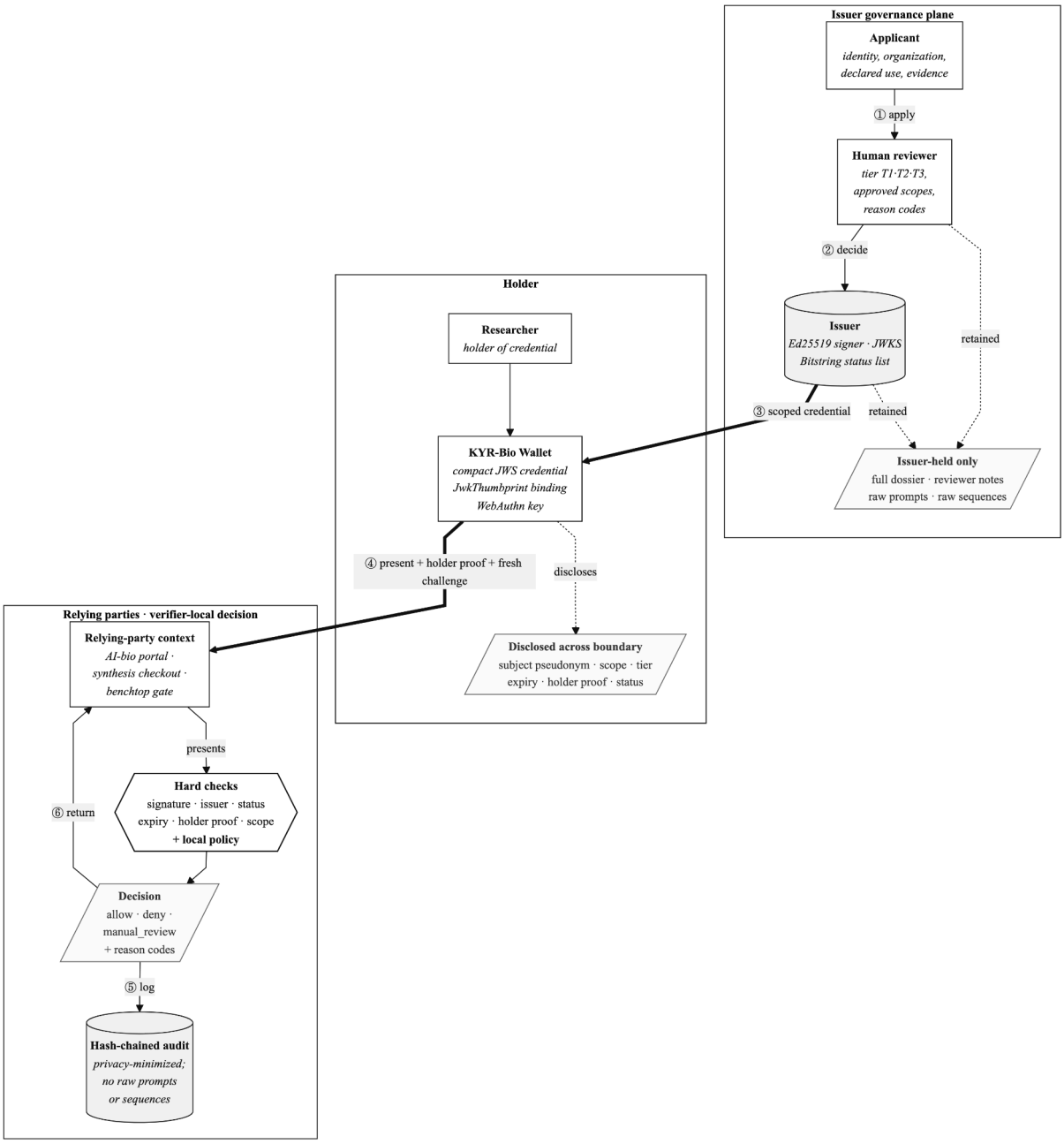
Adjacent approach	Primary focus	How KYR-Bio differs
SecureDNA Exemption Certification	Sequence- and SOC-specific exemption inside sequence-screening workflows.	Researcher-bound; portable across AI-bio, synthesis checkout, and benchtop.
Customer-screening guidance (IBBIS); managed access (NTI 2026)	Forms, standards, and tiered-access policy.	Turns a completed review into a machine-verifiable artifact with verifier-local policy.
Provider-local compliance (Aclid-style)	Sequence + KYC + legitimate-use inside one provider.	Cross-verifier; independent relying parties verify the same credential without each receiving the dossier.
KYR-Bio (this work)	**Portable, holder-bound researcher authorization with verifier-local policy and privacy-minimized audit.**	**Complements the above; replaces none.**

3. Methods

3.1 Architecture and trust loop

KYR-Bio prototypes a portable authorization plane: a reviewed, scoped, holder-bound researcher authorization that AI-bio tools, synthesis checkout flows, and benchtop DNA synthesizers test locally without re-sharing the applicant dossier (Figure 1).

Figure 1. Cross-stack KYR-Bio authorization plane.



Verifiers run hard checks (signature, issuer governance, status, holder proof, scope) and apply local policy. Each verifier sees only the subject pseudonym, scope, tier, expiry, holder proof, status, and reason codes; the full dossier, raw sequences, reviewer notes, and biological prompts stay at the issuer.

Portability has four properties: the same credential is presentable to multiple independent relying parties; the verifier never receives the full applicant dossier; the verifier applies its own local policy; and status and revocation are checked at presentation time. The issuer attests to reviewed facts and approved scopes; each verifier still applies local policy. The TypeScript monorepo signs

credentials with jose over an Ed25519 demo key, uses the Digital Bazaar Bitstring Status List library, SimpleWebAuthn for holder proof, and an optional Open Policy Agent sidecar; demo data is synthetic and deterministic.

3.2 Onboarding, review, issuance, wallet

Applicants flow through /apply/assistant for identity, organization, declared use, requested scopes, evidence, and a structured interview. A deterministic onboarding assistant extracts an editable draft from safe applicant notes, refuses unsafe content, fabrication, approval-odds probes, and prohibited screening prompts, and requires explicit confirmation. Reviewers work in /reviewer behind a demo cookie. A KYR-Bio Investigator module produces a non-binding brief with cited claims, confidence labels, an evidence-flow graph, and an alternative-evidence assessment; it never approves, predicts approval odds, fabricates citations, or uses ideology, religion, ethnicity, nationality, or other protected-class signals. Decisions emit stable reason codes, a trust tier (T1, T2, or T3), and approved scopes that subset the requested set.

The issuer signs a compact JWS/JWT over an Ed25519 demo key, exposes a JWKS endpoint, and emits a W3C VC 2.0 export with BitstringStatusListEntry status and JwkThumbprint2020 holder binding, plus OpenID4VCI/VP metadata at well-known endpoints. The compact JWS remains the verifier's authoritative input. The KYR-Bio Wallet ships as a Manifest V3 extension scoped to localhost:3000, with an in-page wallet fallback. Holder binding runs through one verifier interface with two paths: WebAuthn server verification (required user verification, RP ID, origin, signature counter) and a deterministic demo_simulated proof for synthetic personas.

3.3 Verifier, policy, audit

A relying-party page calls /api/verify/challenge for a fresh challenge bound to relying-party id, requested scope, credential JTI, and a context hash. /api/verify/presentation runs hard checks before policy: schema, JWS signature, issuer governance, credential lookup, Bitstring Status List freshness (ETag, 5-minute max-age, fail-closed), expiry/nbf, holder proof with challenge binding, and scope membership. Appendix E lists the 22 cases. Policy is policy-as-code with two interchangeable backends: a TypeScript engine that mirrors policies/kyr/policy.rego, and an OPA REST sidecar reachable via OPA_URL. The bundle is version-pinned as kyr-bio-policy-v2026-04, and tests/policy-opa.test.ts keeps the backends in lockstep. Decisions are allow, deny, manual_review, or non-blocking manual_review_signal.

Each verification writes one audit event with timestamp, relying-party id, requested scope, decision, reason codes, last-six display id of the credential JTI, and an optional sha256 content hash. Raw biological prompts, raw sequences, full dossiers, declared-use prose, and private reviewer notes are never written; a shared forbidden-terms constant and a privacy validator enforce this at the schema layer. Each event carries previousHash and eventHash over a

canonicalized payload, the chain head is a rolling Merkle root, and verifyAuditChain re-validates every link on load (Appendix B).

3.4 Evaluation and robustness

npm run eval exercises 8 synthetic personas, 22 verifier cases, and 18 AI-assistance behavioral cases, then runs a privacy audit. Beyond the eval, the verifier hard-check invariants run through a fast-check property suite over status (active, revoked, suspended, expired) and holder validity, plus adversarial onboarding and verifier suites covering jailbreak prompts, unsafe-content surfaces, and challenge-binding edge cases. Status-list freshness, OPA fallback parity, audit-chain integrity, and W3C VC 2.0 / OpenID4VC shape conformance each have a dedicated test. npm run submission:check runs lint, typecheck, tests, build, eval, extension scaffold validation, and the extension bridge smoke path.

4. Results

The evaluation reported in eval/results.md (generated 2026-04-27) is **Overall: PASS**.

Table 2. Synthetic evaluation summary.

Category	Cases	Pass	Notes
Persona routing	8	8	Spans legitimate startup, Global South lab with alternative evidence, established academic PI, independent researcher, an inconsistent applicant routed to escalate, plus three personas reused for status-deny paths.
Verification cases	22	22	Covers seven hard-deny invariants and four manual-review conditions; deny rows

Category	Cases	Pass	Notes
			surface the first failing invariant.
AI assistance cases	18	18	Refuses unsafe content, fabrication, approval-odds, prohibited screening; recognizes alternative evidence; preserves applicant_provided provenance; updates reviewer memory without echoing notes.
Privacy audit events	36	36	Zero forbidden raw-content matches across audit events, 8 interviews, 1 evidence graph, 1 privacy summary, and 1 monitoring signal.
Robustness layer	6 surfaces	PASS	Audit chain (Merkle root sha256:02d2fe7e5a668a8125e8ee84c6cc3d8ec4dccb194d510a616a5c4f960cfd97d0), property suite, 5 adversarial cases, status freshness, OPA TypeScript parity on bundle kyr-bio-policy-v2026-04, VC 2.0 / OpenID4VC interop.

Persona routing resolves the 8 personas to expected routes with codes such as alternative_evidence_accepted, manual_review_required, and inconsistent_organization_claims. The 22 verifier cases cover seven hard-deny invariants and four manual-review conditions (Appendix E); pass rows carry codes such as issuer_governance_trusted, signature_valid, and status_list_fresh, while deny rows surface the first failing invariant. The 18 AI-assistance cases verify refusal of unsafe content, fabrication, approval-odds, and prohibited screening; preservation of applicant_provided provenance; non-auto-rejection of nontraditional applicants; and reviewer-memory updates that never echo note text. The privacy auditor finds zero forbidden matches across 36 audit events, 8 interviews, 1 evidence graph, 1 privacy summary, and 1 monitoring signal.

The robustness layer extends scenario coverage to invariant coverage. The fast-check property suite holds the deny invariant whenever status is not active or holder proof is invalid, and the allow path otherwise. Adversarial onboarding and verifier suites exercise prompt-injection and challenge-binding edges without leaking forbidden content. The Bitstring Status List path is fail-closed under stale or unreachable lists, OPA and TypeScript backends agree on the bundle, and the audit-chain test reconstructs the rolling Merkle root and surfaces every mismatch as a typed error. VC 2.0 and OpenID4VC shape tests confirm the W3C v2 context, BitstringStatusListEntry status, JwkThumbprint2020 holder binding, the issuer credential offer with pre-authorized-code grant, the presentation definition, and EdDSA advertisement.

These are workflow-invariant checks, notably they are not real-world validation. They do not show that real applicants would be classified correctly, that the deterministic AI fallback substitutes for a live model, that synthetic adversarial inputs equal motivated red-teaming, or that the simulated holder-proof path substitutes for live passkey enrollment.

5. Discussion and Limitations

KYR-Bio sits beneath sequence screening and content moderation, not in place of them. Sequence screening asks whether requested content is allowed; KYR-Bio asks whether the reviewed holder retains scoped authorization to make the request at all. The prototype treats authorization as scoped, expiring, revocable, and auditable; a signed credential is not a verdict that a holder is globally trustworthy, and it does not eliminate insider threat, coercion, credential laundering, or post-review behavioral change. Synthetic-eval results are workflow-invariant outcomes; they are not real-world classification, legal compliance, or AI-safety guarantees.

Limitations

The WebAuthn path is supported in code and unit-tested but not validated end-to-end; the eval exercises the demo_simulated path. The Bitstring Status List is hosted locally rather than from a

distributed service, the W3C VC 2.0 and OpenID4VC artifacts are export shapes rather than wire-level interop, the hash chain is not anchored to an external timestamp, the store is in-memory, the reviewer role is a demo cookie, and revocation is local demo state. Procedural fairness is built into the design rather than claimed as a result: KYR-Bio surfaces alternative-evidence pathways, requires human review for ambiguous cases, narrows scope rather than auto-rejecting nontraditional applicants, and excludes ideology, religion, ethnicity, nationality, and other protected-class signals. Any deployment would still require reviewer training, appeals, and periodic disparate-impact audits.

Future Work

Future work prioritizes wiring live WebAuthn enrollment through the eval harness, moving signing keys to KMS, replacing the in-memory store with a durable database and an independently-hosted status service, anchoring the audit Merkle root to a notarization service, defining explicit issuer trust governance, adding OpenID4VCI/VP wire-level interop, and extending the relying-party set to CRO, cloud-lab, and reagent-seller verifiers.

6. Conclusion

If successful, KYR-Bio reduces duplicated legitimacy review, lowers friction for legitimate non-enterprise researchers, makes revocation and scope checks portable, and provides a privacy-minimized accountability surface across AI-bio and physical synthesis chokepoints. Sequence screening answers whether requested content is allowed; it does not answer whether the requester has been reviewed and remains authorized. KYR-Bio prototypes that missing layer as portable, scope-limited, holder-bound, revocable researcher credentials with privacy-minimized verifier interactions and tamper-evident audit. KYR-Bio complements, not replaces, sequence screening. The reusable artifact is the claim schema, scope vocabulary, verifier hard-check order, reason-code surface, and audit-minimization boundary.

Code and Data

- **Code repository**: https://github.com/ashtonchew/know-your-researcher-bio
- **Data/Datasets**: Synthetic evaluation outputs are in the following: eval/results.md, eval/results.json (committed in repository above).
- **Demo Video**: https://youtu.be/cvFaLHitLFY?si=iTdydyMKsrn6FzXS
- **Additional note:** All applicant personas and evidence are synthetic; no real biological sequences, raw research plans, or operational protocols are included.

New work during the hackathon. The KYR-Bio prototype, structured onboarding/reviewer flow, credential issuance and verifier services, wallet/demo presentation flow, AI-bio/synthesis/benchtop relying-party demos, synthetic evaluation harness, robustness layer,

audit-chain implementation, VC/OpenID export shapes, demo video, and this report were produced during the hackathon. Everything is created within the timeframe of the hackathon and can be audited by visiting the GitHub. Prior work used for context is cited in Related Work and References.

References

1. Executive Office of the President, Office of Science and Technology Policy. *Framework for Nucleic Acid Synthesis Screening.* April 29, 2024. https://bidenwhitehouse.archives.gov/ostp/news-updates/2024/04/29/framework-for-nucleic-acid-synthesis-screening/
2. U.S. Department of Health and Human Services. *Screening Framework Guidance for Providers and Users of Synthetic Nucleic Acids.* Federal Register, October 13, 2023. https://www.federalregister.gov/documents/2023/10/13/2023-22540/screening-framework-guidance-for-providers-and-users-of-synthetic-nucleic-acids
3. National Institutes of Health. *NOT-OD-25-012: Notification of NIH Requirements Regarding Procurement of Synthetic Nucleic Acids and Benchtop Nucleic Acid Synthesis Equipment.* October 25, 2024 (effective April 26, 2025). https://grants.nih.gov/grants/guide/notice-files/NOT-OD-25-012.html
4. The White House. *Improving the Safety and Security of Biological Research.* Executive Order, May 5, 2025. https://www.whitehouse.gov/presidential-actions/2025/05/improving-the-safety-and-security-of-biological-research/
5. European Commission. *Proposal for a Regulation strengthening the Union's biotechnology and biomanufacturing sectors (European Biotech Act).* COM/2025/1022 final. https://health.ec.europa.eu/biotechnology_en
6. Carter, S. R., Yassif, J. M., & Isaac, C. R. *Benchtop DNA Synthesis Devices: Capabilities, Biosecurity Implications, and Governance.* NTI, May 2023. https://www.nti.org/analysis/articles/benchtop-dna-synthesis-devices-capabilities-biosecurity-implications-and-governance/
7. NTI. *A Framework for Managed Access to Biological AI Tools.* January 28, 2026. https://www.nti.org/analysis/articles/a-framework-for-managed-access-to-biological-ai-tools/
8. International Biosecurity and Biosafety Initiative for Science. *Customer Screening Resources.* https://ibbis.bio/our-work/customer-screening/
9. International Biosecurity and Biosafety Initiative for Science. *Implementing Emerging Customer Screening Standards for Nucleic Acid Synthesis* (white paper). 2025. https://ibbis.bio/ibbis_whitepaper_2025_implementing-emerging-customer-screening-standards-for-nucleic-acid-synthesis/
10. International Biosecurity and Biosafety Initiative for Science. *Common Mechanism (commec).* https://ibbis.bio/our-work/common-mechanism/

11. Crawford, F. W., Webster, K., Epstein, G. L., Roberts, D., Fair, J., & Nevo, S. *Securing Commercial Nucleic Acid Synthesis. RAND Health Quarterly* 12(1), December 2024 (RAND research report RRA3329-1). https://www.rand.org/pubs/periodicals/health-quarterly/issues/v12/n1/04.html
12. Tarangelo, J. P., Attal-Juncqua, A., Somani, E., Roberts, D., & Webster, K. *Protecting Biological Materials and Services from Misuse: Opportunities for Access Monitoring and Control.* RAND research report RRA4067-1, 2025. https://www.rand.org/pubs/research_reports/RRA4067-1.html
13. Baum, C., Berlips, J., Chen, W., Cozzarini, H., Cui, H., Damgård, I., et al. *A system capable of verifiably and privately screening global DNA synthesis. National Science Review*, accepted manuscript, February 16, 2026. DOI: 10.1093/nsr/nwag103
14. SecureDNA. *Exemption Certification System: Safe Access to SOCs.* https://securedna.org/exemption_certification_system/
15. Wittmann, B. J., Alexanian, T., Bartling, C., Beal, J., Clore, A., Diggans, J., Flyangolts, K., Gemler, B. T., Mitchell, T., Murphy, S. T., Wheeler, N. E., & Horvitz, E. *Strengthening nucleic acid biosecurity screening against generative protein design tools. Science* 390(6768): 82–87, October 2, 2025. DOI: 10.1126/science.adu8578
16. National Academies of Sciences, Engineering, and Medicine. *The Age of AI in the Life Sciences: Benefits and Biosecurity Considerations.* National Academies Press, 2025. DOI: 10.17226/28868
17. World Wide Web Consortium. *Verifiable Credentials Data Model v2.0.* W3C Recommendation, 15 May 2025. https://www.w3.org/TR/vc-data-model-2.0/
18. Fett, D., Yasuda, K., & Campbell, B. *RFC 9901: Selective Disclosure for JSON Web Tokens.* IETF Proposed Standard, November 2025. https://www.rfc-editor.org/info/rfc9901
19. IETF OAuth Working Group. *SD-JWT-based Verifiable Digital Credentials.* Internet-Draft draft-ietf-oauth-sd-jwt-vc-16, April 2026. https://datatracker.ietf.org/doc/draft-ietf-oauth-sd-jwt-vc/
20. OpenID Foundation. *OpenID for Verifiable Credential Issuance 1.0* (Final, September 16, 2025) and *OpenID for Verifiable Presentations 1.0* (Final, July 9, 2025). https://openid.net/specs/openid-4-verifiable-credential-issuance-1_0.html ; https://openid.net/specs/openid-4-verifiable-presentations-1_0.html
21. World Wide Web Consortium. *Web Authentication: An API for accessing Public Key Credentials, Level 3.* Candidate Recommendation Snapshot, 13 January 2026. https://www.w3.org/TR/webauthn-3/
22. Langenkamp, M. *Securing Benchtop DNA Synthesizers.* Institute for Progress, December 10, 2024. https://ifp.org/securing-benchtop-dna-synthesizers/
23. Edison, K., Toner, R., & Esvelt, K. M. *Assembling unregulated DNA segments bypasses synthesis screening: regulate fragments as select agents. Nature Communications* 17:3189, 2026. DOI: 10.1038/s41467-025-67955-3

24. Pour Demain. *Guidelines for Nucleic Acid Synthesis Screening in Europe*. 2026. https://www.pourdemain.ngo/en/post/guidelines-for-nucleic-acid-synthesis-screening-in-europe
25. Feldman, J., Feldman, T., & Anton, A. I. *Know Your Scientist: KYC as Biosecurity Infrastructure. Frontiers in Microbiology*, 2026. DOI: 10.3389/fmicb.2026.1814993. Preprint: arXiv:2602.06172, February 5, 2026. https://www.frontiersin.org/articles/10.3389/fmicb.2026.1814993/full
26. Sandbrink, J. *Verified Researcher Credentials for Bio-Capable AI Models*. HackTalk, Apart Research AIxBio Hackathon, April 24-26, 2026. https://apartresearch.com/sprints/aixbio-hackathon-2026-04-24-to-2026-04-26
27. *Bits in Bio*. "Interview with Aclid" (Kevin Flyangolts, CEO). November 9, 2023. https://bitsinbio.substack.com/p/interview-with-aclid
28. Aclid. *Guide to the Screening Certification Process*. https://www.aclid.bio/resources/guide-to-the-screening-certification-process. *About* page: https://www.aclid.bio/about
29. World Wide Web Consortium. *Bitstring Status List v1.0*. W3C Recommendation, 15 May 2025. https://www.w3.org/TR/vc-bitstring-status-list/
30. Open Policy Agent project. https://www.openpolicyagent.org/
31. Sherman, A. T., Romanik Romano, J. J., Zieglar, E., Golaszewski, E., Fuchs, J. D., & Byrd, W. E. *Analysis of the Security Design, Engineering, and Implementation of the SecureDNA System*. arXiv:2512.09233, 2025; shorter version to appear in NDSS 2026. https://arxiv.org/abs/2512.09233

Appendix

Appendix A. Reason Codes (selected)

Positive: signature_valid, issuer_trusted, issuer_governance_trusted, credential_active, status_list_fresh, holder_bound, scope_valid, policy_allow, policy_backend_typescript, kyr-bio-policy-v2026-04.

Validation/security failures: invalid_verification_request, invalid_signature, issuer_untrusted, credential_not_active, credential_expired, credential_not_yet_valid, status_list_revoked, status_index_invalid, holder_proof_missing, holder_proof_invalid, relying_party_not_allowed, challenge_reused, challenge_relying_party_mismatch, challenge_scope_mismatch, challenge_credential_mismatch, challenge_context_mismatch, scope_not_approved.

Policy/manual-review: relying_party_scope_not_allowed, tier_too_low, synthesis_screening_context_required, soc_flagged_demo, enhanced_monitoring_required,

metadata_scope_escalation_pattern, review_only_scope_requires_manual_review, manual_review_required.

Appendix B. Audit Field Boundary

Field	**Logged?**	**Rationale**
Credential JTI display id (last 6)	Yes	Auditability without re-sharing the dossier
Issuer id, relying party id	Yes	Trust and accountability
Requested scope	Yes	Policy decision input
Decision (allow / deny / manual review)	Yes	Operational accountability
Reason codes	Yes	Stable, machine-readable
Optional sha256:<64 hex> content hash	Yes	Tamper-evidence without storing content
Audit previousHash and eventHash	Yes	Tamper-evident chain over minimized fields
Rolling Merkle root at chain head	Yes	Detects historical tampering on load
Raw sequence, design prompt, declared-use prose	**No**	Biosecurity and IP minimization
Full applicant dossier	**No**	Privacy and data minimization
Private reviewer notes	**No**	Review integrity and least disclosure

Appendix C. Open Trust-Governance Questions (production)

The current prototype assumes a single trusted issuer with a governance allowlist. Production deployment depends on questions this prototype does not answer: who is allowed to issue KYR-Bio credentials and on what authority; how revocation freshness is guaranteed across distributed verifiers; what appeals and correction processes look like in practice; how insider threat, credential laundering, and behavioral drift are detected, if at all, under privacy-preserving constraints; and

how cross-jurisdiction policy disagreements resolve when the same credential is presented in different regulatory regimes. Production also requires explicit verifier trust lists, status freshness requirements, issuer-compromise handling, and retention policy. None of these are cryptographic problems; all of them are governance problems. KYR-Bio surfaces them; it does not solve them.

Appendix D. Scope Vocabulary and Credential Subject Schema

Scope vocabulary (packages/shared/src/schemas.ts, scopeSchema):

- ai_bio_trusted_access: managed access to bio-capable AI tools.
- synthesis_checkout_low_risk: checkout for low-risk synthesis orders that have passed sequence screening.
- benchtop_authorized_user: authorization to operate or initiate sensitive benchtop synthesis workflows.
- soc_exemption_request_review_only: review-only request for sequence-of-concern exemption pathways; never auto-approves.

Credential subject (selected fields).

Field	**Type**	**Purpose**
subject_type	individual_researcher | organization_delegate	Who the holder represents
organization_id, organization_type	string	Reviewed organization handle
role	string	Reviewed role within the organization
trust_tier	T1 | T2 | T3	Reviewer-assigned tier
approved_scopes	array of Scope	Scopes the reviewer authorized (subset of requested)
assurance.identity / authenticator / federation	string	Identity/auth assurance metadata
review.reviewer_org, review.decision_id	string	Pointer back to the reviewing body and decision

Field	Type	Purpose
review.evidence_summary_hash	sha256:<64 hex>	Hash of evidence summary (no raw content)
review.alternative_evidence_used	boolean	Surfacing alternative-evidence pathways
review.monitoring_level	none | standard | enhanced | suspended	Ongoing-monitoring posture attached to credential
cnf.jwk_thumbprint	string	Holder binding (WebAuthn-shaped, exercised in eval through the simulated path)

These fields are the smallest set needed to support local verifier policy without requiring access to the underlying dossier.

Appendix E. Verification Case Matrix

The 22 verification cases below are all in eval/results.md (generated 2026-04-27T05:54:16.377Z, all pass). Rows that pass each hard-check stage carry the positive reason codes for that stage; security-failure rows expose the first failing invariant. Policy-routed rows additionally carry policy_backend_typescript and kyr-bio-policy-v2026-04.

Case	Outcome	Key reason codes (beyond the always-present set)
valid_startup_ai_access	allow	holder_bound, scope_valid, credential_active, policy_allow
valid_startup_synthesis_checkout	allow	holder_bound, scope_valid, credential_active, policy_allow
ai_portal_rejects_synthesis_scope	deny	relying_party_scope_not_allowed
synthesis_checkout_requires_screening_context	manual_review	synthesis_screening_context_required, manual_review_required

Case	Outcome	Key reason codes (beyond the always-present set)
revoked_denied	deny	status_list_revoked, credential_not_active
suspended_denied	deny	status_list_revoked, credential_not_active
expired_denied	deny	credential_expired
wrong_holder_denied	deny	holder_proof_invalid
unapproved_scope_denied	deny	scope_not_approved
soc_flagged_manual_review	manual_review	soc_flagged_demo, enhanced_monitoring_required, review_only_scope_requires_manual_review, manual_review_required
security_missing_holder	deny	invalid_verification_request, holder_proof_missing
security_invalid_signature	deny	invalid_signature
security_untrusted_issuer	deny	issuer_untrusted
security_cross_party_challenge	deny	challenge_relying_party_mismatch, holder_proof_invalid
security_challenge_scope_mismatch	deny	challenge_scope_mismatch, holder_proof_invalid
security_challenge_credential_mismatch	deny	challenge_credential_mismatch, holder_proof_invalid
security_challenge_context_mismatch	deny	challenge_context_mismatch, holder_proof_invalid
security_unknown_relying_party	deny	invalid_verification_request, relying_party_not_allowed, holder_proof_missing

Case	Outcome	Key reason codes (beyond the always-present set)
security_challenge_replay	deny	challenge_reused, holder_proof_invalid
security_credential_not_yet_valid	deny	credential_not_yet_valid
metadata_scope_escalation_policy_manual_review	manual_review	metadata_scope_escalation_pattern, manual_review_required
metadata_scope_escalation_signal	manual_review_signal	metadata_scope_escalation_pattern, scope_not_approved

Appendix F. Limitations and Dual-Use Considerations

The verifier denies on the first failing invariant and routes ambiguous cases to manual review rather than auto-approve, so the most common boundary error is a legitimate researcher routed to a human reviewer with reason codes attached. False positives, where unauthorized access is allowed, are bounded by scope expiry, status freshness, holder binding, and revocation. Real-world classification accuracy is an open question for any production deployment.

A signed credential is a scoped, expiring, revocable assertion of authorization, not a verdict that the holder is safe. Insider threat, coercion, credential laundering, and post-review behavioral change are addressed by adjacent controls (sequence screening, content moderation, monitoring, human review at the relying party) rather than by issuance, holder binding, or revocation alone.

The verifier sees only a subject pseudonym, scope, expiry, holder thumbprint, decision, and reason codes. The full applicant dossier, declared-use prose, raw sequences, and reviewer notes never leave the issuer surface. The audit log is hash-chained over a privacy-minimized payload, validated at the schema layer by a shared forbidden-terms constant, and re-checked end-to-end by tests/api-privacy.test.ts. Reviewer-facing AI assistance excludes ideology, religion, ethnicity, nationality, and other protected-class signals at the design level, surfaces alternative-evidence pathways, and never returns approval-odds estimates. The investigator brief is non-binding and human-review-labeled. Production deployment would still require reviewer training, an appeals process, and periodic disparate-impact audits.

All demo data is synthetic. The repository contains no real biological sequences, operational wet-lab instructions, sensitive research plans, or applicant dossiers. The issuer signing key is

checked in for reproducibility as a demo artifact; production deployment would migrate signing material to a KMS with rotation. Security-relevant findings against the prototype can be reported through repository issues; a production deployment would require a coordinated-disclosure path and an issuer-compromise playbook.

Appendix G. Extended Architecture and Claim/Evidence

G.1 Cross-stack scenario

A small therapeutics startup using AI-native design tools is approved once for ai_bio_trusted_access and synthesis_checkout_low_risk. The same credential is verified locally by the AI-bio portal, by the synthesis-checkout flow, and (when scope is later extended) by a benchtop authorized-user gate. A revoked or wrong-holder presentation fails at every relying party without any of those verifiers re-receiving the startup's dossier or reviewer notes. One review, three different gates, no dossier passing through the middle.

G.2 Theory of change

- **Problem.** Sequence screening is improving, but requester authorization is fragmented across providers and tools.
- **Intervention.** Issue portable, scoped researcher credentials after human legitimacy review.
- **Mechanism.** Verifiers check status, scope, and holder proof locally instead of redoing KYC or trusting unverified accounts.
- **Near-term users.** Trusted AI-bio access programs, synthesis checkout, benchtop authorized-user gates.
- **Risk reduction.** Fewer anonymous or weakly reviewed capability accesses; faster revocation; better attribution; lower friction for legitimate researchers; less dossier oversharing.
- **Long-term path.** Credential becomes a reusable trust primitive for CROs, cloud labs, reagents, equipment, and privacy-preserving monitoring.

G.3 Verifier hard-check order (full)

1. Schema validation of the presentation request.
2. Compact JWS verification.
3. Issuer trust list and governance check (issuer_governance_trusted, services/verifier/src/governance.ts).
4. Prototype-local credential record lookup or status-service response; production verifiers would rely on issuer metadata plus a distributed status mechanism rather than private database access.
5. Bitstring Status List freshness check (status_list_fresh), with deterministic hash-seeded index allocation, ETag-based caching, a 5-minute max-age, and verifier-side fail-closed

checks. Revoked or suspended indices both yield status_list_revoked in the demo's bitstring representation, where any non-active state sets the bit (services/status-list/src/index.ts).

6. Expiry and nbf checks.
7. Holder proof verification with challenge binding (relying party, scope, credential JTI, context hash).
8. Requested scope must be in approved_scopes.

When metadata suggests escalation but the requested scope is already separately denied, the verifier emits a non-blocking manual_review_signal outcome alongside the deny so reviewers still see the escalation pattern.

G.4 Claim/evidence matrix

Claim	**Evidence**	**Source**
End-to-end trust loop is implemented locally.	Routes /apply/assistant, /reviewer, /credential, /demo-ai-portal, /demo-synthesis-checkout, /demo-benchtop, /audit; services issuer, verifier, policy, status-list.	README.md, docs/specification.md
Verifier denies revoked, expired, wrong-holder, replayed, untrusted, and unapproved-scope presentations.	22/22 verification cases pass with the expected reason codes.	eval/results.md
Status checks use a Bitstring Status List with freshness and fail-closed behavior.	status_list_fresh and status_list_revoked in every applicable row; ETag and 5-minute freshness.	services/status-list/src/index.ts, tests/status-list.test.ts
Policy decisions come from policy-as-code with version pinning.	policy_backend_typescript and kyr-bio-policy-v2026-04 decorate every policy row; OPA sidecar agrees with the TypeScript backend.	services/policy/src/opa-client.ts, policies/kyr/policy.rego, tests/policy-opa.test.ts

Claim	Evidence	Source
Audit logs are tamper-evident and contain no raw biological content.	Hash-chain and Merkle root verify; 36 events scanned; 0 forbidden matches.	apps/web/lib/audit-chain.ts, tests/audit-chain.test.ts, eval/results.md
Holder binding has a WebAuthn verification path with a demo_simulated fallback for the eval.	verifyAuthenticationResponse with required user verification; deterministic demo proof for synthetic personas.	services/verifier/src/webauthn.ts, eval/results.md
Hard-check invariants hold beyond hand-curated cases.	Property suite over status x holder validity, plus adversarial onboarding and verifier suites.	tests/property-verifier.test.ts, tests/adversarial-*.test.ts
Reviewer notes do not appear in applicant, verifier, or audit surfaces.	API privacy tests; reviewer-memory case investigator_safe_reviewer_memory_used.	tests/api-privacy.test.ts, eval/results.md

G.5 Standardizable artifact

The reusable artifact of this work is not the demo app; it is the claim schema, the scope vocabulary, the verifier hard-check order, the reason-code surface, and the audit-minimization boundary documented in this paper and in packages/shared/src. Each piece is portable to other implementations under different stack choices (W3C VC 2.0, SD-JWT VC, OpenID4VC, OPA/Rego, durable status-list services), and each piece is independently testable.

LLM Usage Statement

We used LLMs for research ideation, coding, arguing direction, and compiling recent research. We checked every citation against the publisher's URL or DOI before submission.