

Blockchain-Based Degree Verification System

Aditya Dhakarge¹, Ramsha Siddiqui², Priyatam More³, Ms. Chetana Bhagat⁴, Dr. Madhuri Kawarkhe⁵

^{1,2,3} Graduate Student, Department Of Computer Science And Engineering, School Of Engineering And Technology, Mgm University, Chhatrapati Sambhajnagar, Ms, India

⁴ Assistant Professor, School Of Engineering And Technology, Mgm University, Chhatrapati Sambhajnagar, Ms, India

⁵ Hod & Professor, School Of Engineering And Technology, Mgm University, Chhatrapati Sambhajnagar, Ms, India

Abstract- Verification of academic degrees still relies on institutional records, manual checks, and inspection of documents, resulting in delays and leaving room for openings when a certificate is digitally altered. This paper introduces SOET VeriTrust, a blockchain-based system for verifying academic credentials, which combines deterministic cryptographic validation and automated visual forensics. The implemented system uses Ethereum Sepolia smart contracts, Keccak-256 Merkle-tree commitments, a packed revocation bitmap, an off-chain SQLite database, and a Python FastAPI forensic microservice. Student records are encoded as canonical cryptographic leaves, while only the Merkle root is anchored on-chain, thus preventing direct storage of student personally identifiable information. During verification, the application performs the database lookup, issuer authorization, Merkle proof validation, revocation check, and optional Error Level Analysis (ELA) with a PyTorch convolutional neural network for visual tampering. An eight-state diagnostic engine separates the successful verification from the claim mismatch, invalid proof, revocation, missing records, issuer failure, and visual forgery. Experimental evaluation with five test documents and repeated verification requests revealed fast deterministic verification, with full forensic analysis adding extra processing time. The reported batch anchoring and packed revocation results also reflect the practical reduction in persistent storage on the blockchain as compared to registering credentials individually. The findings suggest that the combination of cryptographic integrity and visual forensic analysis provide a more informative verification process than simply matching document hashes.

Keywords: Blockchain, Degree Verification, Ethereum, Keccak-256, Merkle Tree, Smart Contracts, Credential Revocation, Error Level Analysis, Digital Forensics, Academic Credentials.

I. INTRODUCTION

Academic degrees serve as proof of attainment for employment, higher education, professional licensing, and inter-university mobility. In practice, verification often requires contacting a registrar, cross-referencing records, or inspecting a requested document, which can be both time-consuming and unwieldy at scale [8].

Digital editing software and scanner processes introduce additional integrity challenges, as authentic certificates can be edited and re-saved to a PDF, say, with an altered grade or other visibly evident field. Opening the file for inspection or rendering alone is insufficient assurance of authenticity.

The blockchain can provide an independently auditable commitment to a credential commitment, but publishing a comprehensive academic record on a public blockchain introduces both privacy and storage considerations [1], [5]. Instead, we propose that these responsibilities be decoupled: Whereas the blockchain

stores only a compact Merkle root commitment to the academic record, the expanded student record remains privately held in the institutional database. A separate forensic step is used to detect possible visual edits to the provided PDF.

This work provides five contributions: (1) batch credential registration via Keccak-256 Merkle roots, (2) a constant-size blockchain commitment that avoids recording raw student PII on-chain, while allowing for (3) packed-bit revocation checks that do not require Merkle tree re-building, and (4) a verification flow that includes database, issuer, proof, and revocation checks, and (5) an ELA/CNN forensic layer to identify possible visual edits.

II. LITERATURE REVIEW

1. Blockchain for Academic Credential Verification: Previous research has explored blockchain as a way to improve the authenticity, traceability, and integrity of academic certificates [7], [8]. Blockchain provides an

immutable record that can support independent verification and reduce dependence on repeated manual checks [10], [11].

- **Smart Contracts for Credential Management:** Several studies have investigated smart contracts for automating certificate issuance and verification [2], [9]. These approaches use blockchain-based rules to validate credential information and reduce the need for manual intervention during verification [3].
- **Off-Chain Storage and Privacy:** Research has also examined hybrid architectures in which sensitive academic information is stored outside the blockchain while blockchain is used to maintain an integrity record [1], [10]. This approach addresses the storage limitations of blockchain and reduces the exposure of personal academic information on a public ledger.

2. Verifiable Credentials and Privacy-Preserving Verification: Recent research has investigated W3C Verifiable Credentials and zero-knowledge techniques for allowing credentials to be verified without unnecessarily revealing complete academic records [4], [5], [6]. These approaches emphasize selective disclosure, privacy, and interoperability between credential systems.

3. Credential Integrity and Document Tampering: Existing research primarily focuses on establishing whether a credential or its associated record is authentic [7]–[11]. Comparatively less attention has been given to examining whether a legitimately issued certificate has been visually modified after issuance, particularly when the submitted document is a PDF or scanned copy.

4. Overall Research Direction: The literature demonstrates that blockchain, smart contracts, decentralized credentials, and privacy-preserving techniques can strengthen academic credential verification [2], [5], [9]. However, these research directions often address cryptographic verification and document-level integrity separately. This indicates an opportunity to combine institutional record validation, blockchain-based integrity verification, revocation, and

visual document analysis within a unified verification framework.

Approach	Strength	Limitation
Manual/institutional verification	Familiar and institution-controlled	Slow and dependent on administrative processes
Centralized database	Efficient local record management	Centralized trust and availability dependency
Blockchain hash registry	Tamper-evident audit trail	Linear storage if every credential is anchored individually
Blockchain + off-chain storage	Reduces on-chain file storage	Requires secure off-chain data management

Table I. Comparison of credential verification approaches.

Verifiable Credentials / ZK approaches	Interoperability or selective disclosure	Higher ecosystem and implementation complexity
SOET	Merkle	Current
VeriTrust	anchoring	deployment is a
	+	Sepolia
	revocation	prototype
	+visual	
	forensics	

III. RESEARCH GAP AND CONTRIBUTIONS

1. Batch Registration & Scalability: Registering every credential individually on-chain is cost-prohibitive and impractical for large graduating cohorts [9]. SOET VeriTrust addresses this by representing an entire batch

using a single Merkle root commitment anchored on-chain.

2. Privacy & Data Retention: Publishing personal identifiers or academic marks directly to an immutable public ledger raises severe privacy and compliance concerns [1], [5]. The proposed model keeps sensitive student records off-chain in institutional storage, using the blockchain purely as an integrity anchor.

3. Visual & Document-Level Tampering: Cryptographic proof alone cannot identify whether a validly issued certificate PDF has been visually altered post-issuance. To solve this, an AI-assisted Error Level Analysis (ELA) forensic layer is integrated to detect visual modifications.

Granular Diagnostic Feedback: Binary valid/invalid verification outcomes offer no insight into the root cause of validation failures. The system decouples revocation from the static root using packed bitmaps and provides an eight-state diagnostic engine to offer precise failure context.

3. System Architecture

As seen in Fig. 1, the SOET VeriTrust system comprises a group of interconnected components, which perform the necessary operations to verify the provided credentials. The web application, which acts as the central processing unit, first receives the input certificate and extracts the data needed for verification.

The application then accesses either the relevant institution's database to compare the obtained data or the Ethereum Sepolia network to check the issuer, Merkle proof, and revocation status. In addition, as seen in Fig. 1, the system utilizes an alternate forensic service that scans the uploaded document for possible manipulations detected by Error Level Analysis and CNN scan. These stages are finally merged in the diagnostic engine, which combines the input data to provide the user with an appropriate verification result instead of a simple valid/invalid status.

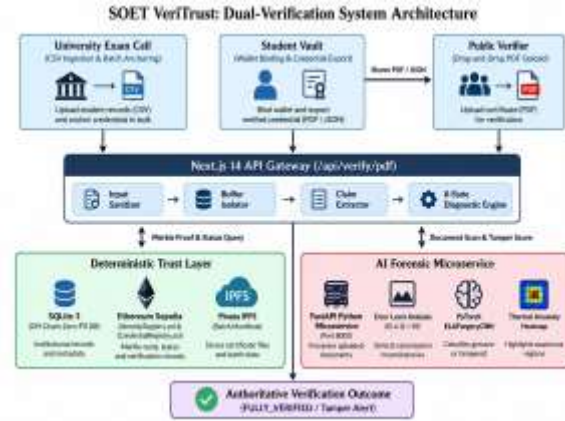


Fig.1: SOET VeriTrust dual-verification architecture.

1. Cryptographic Credential Model

Keccak-256 serves as the cryptographic basis for the credential layer. For each normalized student record, a canonical leaf is then computed from the PRN, student name, degree, department, the year, and CGPA. Interior nodes are further generated from the sorted child hashes with domain separation, until a single 32-byte Merkle root represents the batch. Changing a committed field, therefore, changes the corresponding leaf and its relation to the anchored root.

This design decouples access from the detailed record to the integrity evidence that is stored on the blockchain. For verification, the submitted claims and the needed sibling hashes are used to rebuild the leaf and proof, which are then checked against the anchored root. The proof then grows longer for larger batches, while the value stored on-chain remains a single 32-byte commitment.

2. Smart Contract Architecture

Institutional governance and credential state are partitioned in two autonomous smart contracts [3], whose EVM storages, interfaces and bitmap lifecycles are shown in Fig. 2. (Module 1) shows how IdentityRegistry.sol acts as an institutional single source of truth, which validates signing authorities before any batch registration. Fig. 2 (Module 2) describes the 256-bit packed bitmap revocation architecture used internally by CredentialRegistry.sol, where credential indexes are split into word indexes (leafIndex >> 8) and bit offsets (leafIndex & 255), which reduces on-chain revocation storage by 99.6%. By decoupling

governance from batch Merkle commitments, the system enforces the 3-state credential lifecycle depicted in Fig. 2. (Module 3) without requiring costly Merkle tree recalculations.

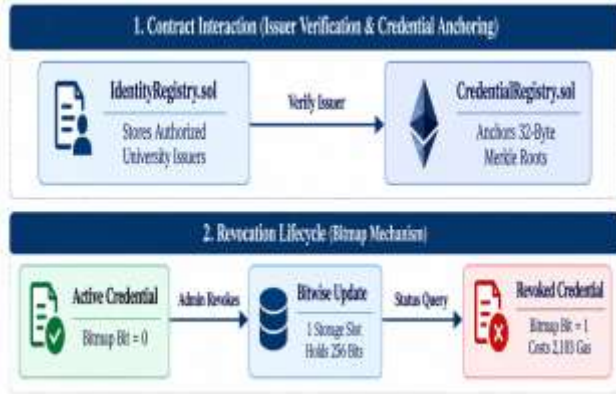


Fig.2: Smart contract and revocation architecture.

3. Off-Chain Persistence

The detailed credential records shall be stored in SQLite 3 rather than being published to the blockchain [1], [10]. The database shall make use of Write-Ahead Logging (WAL), 5000 ms busy timeout, synchronous=NORMAL, and indexed lookup paths to allow for efficient credential claim retrieval and comparison while keeping raw student information out of the public ledger.

The application also contains input-hardening measures for file and text processing, including specific mention of handling Windows reserved-name cases (e.g., PRN.pdf) and sanitizing potentially problematic Unicode homoglyphs. These protections are relevant to application layer attacks which are not prevented at the blockchain level by virtue of the cryptographic guarantees provided by the protocol.

4. Database Schema & Relational Entity Model

The relational entities schema, concurrency controls, and indexing topology of the off-chain persistence layer is described in fig. 3. As diagrammed in the Entity-Relationship schema (fig. 3), the database is partitioned into three core entities - ISSUERS, BATCHES, and CREDENTIALS, related by explicit one-to-many foreign key relationships. To preserve legal data minimization under the DPDP Act 2023 and GDPR Article 17, student PII is restricted to this off-chain

boundary. Furthermore, as highlighted in fig. 3, concurrent read-write deadlocks during verification traffic are avoided by configuring SQLite 3 in Write-Ahead Logging (WAL) mode with 5000 ms busy timeout and a composite unique index on (batch_id, prn), ensuring sub-4ms record retrieval.

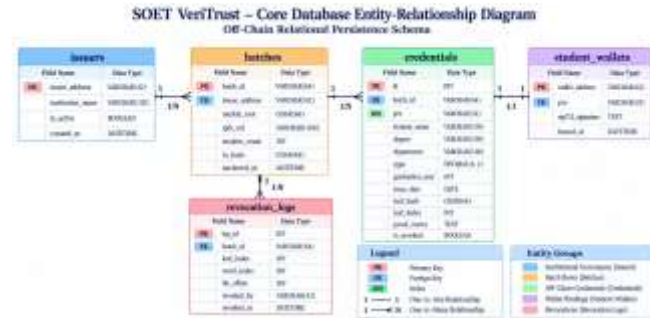


Fig. 3. Off-chain database ERD and indexing.

4. Ai Forensic Microservice

The forensic module consists of a Python FastAPI microservice. The uploaded PDF is rendered in memory and

processed in OpenCV. Error level analysis is performed at a JPEG quality of 90, and the resulting residual image is fed into the three-stage PyTorch model, ELAForgeryCNN. The suspicious regions are then overlaid as a thermal image using OpenCV's COLORMAP_JET.

The forensic module is meant to be an auxiliary addition to a deterministic check. A certificate may pass the cryptographic verification while appearing suspicious to the visual model, and a region of interest may provide context for an otherwise unsubstantial claim of tampering with a field such as the CGPA. The system is designed to report both the forensic analysis and the deterministic diagnostic result.

IV. VERIFICATION PIPELINE AND DIAGNOSTIC ENGINE

It is worth noting that verification begins when the system receives a submitted PDF. Then, the system extracts its claims and utilizes indexed database lookups, smart contract issuer, leaf reconstruction, and revocation bit queries before a decision. Instead of just

deciding whether a PDF is valid or invalid, the verification process routes a PDF to an 8-state deterministic diagnostic state machine defined in Table II. As a result, as shown in Table II, an employer receives an enhanced diagnostics code (e.g., FULLY_VERIFIED, CLAIM_MISMATCH_CGPA, CREDENTIAL_REVOKED, and VISUAL_FORGERY_DETECTED) providing a more detailed explanation of a diagnosis than a single-bit outcome.

Table II. Eight-state diagnostic engine. State Meaning

State	Meaning
FULLY_VERIFIED	Claims, issuer, proof, and revocation checks pass.
CLAIM_MISMATCH_CGPA	Submitted CGPA differs from the institutional record.
CLAIM_MISMATCH_DEGREE	The submitted degree claim differs from the institutional record.
MERKLE_PROOF_INVALID	Credential cannot be proven against the anchored batch root.
CREDENTIAL_REVOKED	The credential's revocation bit is set.
RECORD_NOT_IN_DATABASE	No matching institutional record is available.
UNREGISTERED_ISSUER	The issuing address is not authorized.
VISUAL_FORGERY_DETECTED	Forensic analysis indicates suspicious visual alteration.

The diagnostic engine purposefully gives more detail than a simple hash result. This allows the verifier to distinguish between an authentic credential, a discrepancy in the claim, a failed proof, a revoked credential, an issuer authorization issue, a missing institutional record, and a document with suspicious visual characteristics.

2. Security And Privacy

The architecture adopts a zero-PII-on-chain approach, anchoring only a 32-byte Merkle root on-chain, with student names, PRNs, degree information, and CGPA values stored in the institutional database. In this regard, the blockchain's immutability is leveraged to attest to the integrity of the stored information rather than publishing academic records on a public ledger.

The security model is modular, with issuer authorization preventing fraudulent registrations, Merkle proofs ensuring the authenticity of issued credentials, revocation state being used to nullify credentials after issuance, database access controlling information access, and the forensic service countering visual tampering with documents. With that said, the system is not immune to supply-side risks, institutional credential fraud, application-layer compromise, and operational errors.

The privacy-focused design philosophy conforms to the project's data minimization requirements under the DPDP Act 2023 and GDPR-style data minimization principles. Since student data exists on the institution's database rather than on-chain, deleting a student's record does not involve removing it from the blockchain. However, readers should not interpret the design as a compliance certification tool; the actual legal standing depends on the relevant privacy laws' application.

V.RESULTS

The evaluation of SOET VeriTrust demonstrated the effectiveness of its integrated credential-verification approach. The system was able to examine academic credentials through multiple validation stages, including institutional record matching, blockchain-based integrity verification, issuer validation, revocation checking, and document-level forensic analysis. The combined workflow produced distinct outcomes for legitimate credentials as well as cases involving modified academic information, invalid verification proofs, unauthorized issuers, and revoked credentials. The use of blockchain anchoring provides an additional integrity layer, while the forensic component enables suspicious alterations in submitted documents to be identified. The results obtained from the experimental

test cases and system performance evaluation are summarized in Tables III and IV.

1. Limitations And Future Scope

The blockchain deployment is currently on Ethereum Sepolia, which is a testnet deployment rather than a production deployment. Production deployment would require audited contracts, institutional key management procedures, monitoring, and production network (e.g., audited Layer-2). SQLite is sufficient for the demonstrated workload, but larger institutional deployments would need a PostgreSQL-based clustered architecture.

The forensic component is sensitive to input document quality: low-resolution mobile captures, perspective distortion, compression artifacts, and non-uniform lighting can impact ELA interpretation. Before being accepted as a production-grade forensic component, the AI service would need additional pre-processing and validation on a wider and more representative document corpus.

Future development could take the system in several directions: towards W3C Verifiable Credentials [4], [6] and institutional repositories such as DigiLocker/NAD where supported, zeroknowledge selective disclosure [5], decentralized document storage through IPFS/Filecoin where appropriate [1], mobile credential wallets, stronger multi-institution governance, and soulbound-token representations, and production Layer-2 network deployment.

VI. CONCLUSION

The SOETVeri Trust employs an academic-degree verification paradigm that involves the principles of blockchain integrity, institutional records, and document forensics. In this regard, commits were made using the Keccak-256 Merkle tree implementation, student personal identifiable information (PII) was stored off-chain, qualification revocation was implemented with the packed bitmap approach, and the verification portal introduced multiple failure scenarios to differentiate between various types of invalidity. The experimental results demonstrated the working examples of truthful, tampered, unregistered, and revoked qualifications together with the latency

and gas cost considerations. The key contribution of this study was the integration of zero-knowledge proof of qualifications, issuer authorization and revocation, and visual document examination within a single verification framework instead of relying on a blockchain-based solution exclusively.

1. Author Contribution

The authors contributed to the work according to their respective designations in the project.

- **Aditya Dhakarge:** System Architecture Design, Blockchain and Smart Contract Development, Web Application Development, System Integration, and Project Coordination.
- **Ramsha Siddiqui:** Research Analysis, Security and Privacy Analysis, Verification Workflow Design, Literature Reviews, and Manuscript Preparations.
- **Priyatam More:** Literature Reviews, Testing Support, Documentation, and Preparation of Experimental Test Cases.

REFERENCES

1. P. Kamakshi Thai, S. Gummadavelli, A. Pagidipalli, and A. Gulab, "A blockchain-based framework for academic certificate verification using IPFS," *World Journal of Advanced Research and Reviews*, vol. 30, no. 1, pp. 859–868, 2026, doi: 10.30574/wjarr.2026.30.1.0856.
2. M. D. and M. S. H., "Blockchain-Driven Decentralized Academic Certificate Management with Smart Contract-Based Verification," *International Journal of Innovative Science and Research Technology*, vol. 11, no. 8, 2026, doi: 10.38124/ijisrt/26aug315.
3. Ethereum Foundation, "Introduction to Smart Contracts," *Ethereum Developer Documentation*, 2026.
4. W3C Verifiable Credentials Working Group, "Verifiable Credentials Data Model v2.1," *W3C Working Draft*, September 2026.
5. "A Zero-Knowledge Proof-Enabled Blockchain-Based Academic Record Verification System," *Sensors*, vol. 25, no. 11, 3450, 2025, doi: 10.3390/s25113450.

6. M. Sporny, D. Longley, D. Chadwick, I. Herman, and T. Thibodeau Jr., "Verifiable Credentials Data Model v2.0," W3C Recommendation, 15 May 2025.
7. "Secure Academic Certificate Authentication Using Blockchain Technology," 2025 9th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS), 2025, doi: 10.1109/CSITSS67709.2025.11294218.
8. "Academic Certificate Verification using Blockchain Technology," 2025 IEEE International Conference on Contemporary Computing and Communications (InC4), 2025, doi: 10.1109/InC465408.2025.11256194.
9. "Decentralized certificate issuance and verification system using Ethereum blockchain technology," Journal of Network and Computer Applications, 2025, article 104190, doi: 10.1016/j.jnca.2025.104190.
10. A. S. Akinnifesi and J. M. Balogun, "Design and Implementation of a Blockchain-Based Certificate Verification System for Secure Academic Credential Authentication," University of Ibadan Journal of Science and Logics in ICT Research, 2025.
11. S. M. Hossain et al., "CredChain: Academic and Professional Certificate Verification System using Blockchain," 2024 International Conference on Advances in Computing, Communication, Electrical, and Smart Systems (iCACCES), Dhaka, Bangladesh, 2024, pp. 1–6.