IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, November 2026
Impact Factor: 8.2

AI-Based Fake Document Verification System

Tushar Manohar Patil[1], Aniket Ramchandra Patil[2], Pranav Bharat Sanabe[3], Bhagya Basappa Pattanshetti[4]
B.E. Students[1-4]
Dr. J. J. Magdum College of Engineering, Jaysingpur, Maharashtra, India

Abstract: *The rapid shift from paper-based administration to digital documentation has accelerated information sharing, but it has also made document forgery more likely. Altered certificates, identity proofs, academic records, invoices and other official documents, when accepted as authentic, can lead to financial loss, identity theft and administrative errors. Manual verification is slow, depends on the experience of the examiner and is difficult to scale. This paper presents the design of an Artificial Intelligence (AI)-based Fake Document Verification System that integrates machine learning, optical character recognition (OCR), image processing and rule-based validation to help identify suspicious documents. The main objectives are to extract information from documents, detect visual and structural irregularities, compare the extracted details with expected formats or trusted records, and produce a risk assessment that a human reviewer can understand. The scope covers uploaded PDF documents and scanned images that may be used in public service workflows, banking, insurance, hiring and educational institutions. Following a review of published research and technical standards, the paper proposes a seven-stage framework comprising document upload, preprocessing, text extraction, visual-forensic analysis, structural and semantic validation, trusted-source verification, and risk assessment with report generation. The advantages and limitations of automated verification are also discussed, including image quality, unseen forgery techniques, privacy, bias and incorrect decisions..*

Keywords: Fake Document Verification, Document Forgery Detection, Artificial Intelligence, Machine Learning, Optical Character Recognition, Image Forensics, Explainability, Privacy

I. INTRODUCTION

A. Background and Problem Statement

Documents are essential to contemporary government, business, social and educational activities. An identity card can be used to open a bank account, a degree certificate to apply for a job, a salary slip to obtain a loan, and an invoice to claim a payment. Such records convey the information on which an organization bases its decisions, and a decision founded on a falsified or altered document may itself be flawed.

The effort needed to modify document content has fallen considerably because of digital editing software, mobile scanning applications and generative tools. A fraudster may replace a name, change a date, alter a grade, insert a new photograph, imitate a signature, or combine genuine and fake parts. Some changes are clearly visible, while others preserve the overall layout and use matching fonts, colors and spacing. Deep-learning-based text-editing methods have even been shown to produce forged-and-recaptured identity document images that fooled some existing document authentication systems [1].

The research problem addressed in this paper is how an AI-assisted system can analyze a digital document and find signs of potential forgery in a consistent, efficient and explainable way. The system under consideration is not meant to declare every uploaded document legally authentic or fraudulent.

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/IJARSCT-40972

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, November 2026
Impact Factor: 8.2

B. General Theme

The overarching theme is the use of digital forensics and artificial intelligence to strengthen document authentication. AI can process large amounts of text and image data, prioritize the documents that require closer examination, and spot patterns that may be difficult to see by hand. Optical character recognition (OCR) transforms printed or handwritten visual text into machine-readable text.

C. Review of Related Work

Research in document forensics uses both active and passive approaches [2]. Active techniques make use of embedded data or security features such as digital signatures, watermarks, barcodes or issuer-controlled verification codes. Passive techniques examine a document after it has been created, without requiring a special mark to be inserted beforehand.

Qu et al. introduced the DocTamper dataset of 170,000 document images, together with a method for detecting tampered text in document images [3]. Their work illustrates how challenging it is to detect visually consistent changes in photographed documents and how important it is to train and assess models on a variety of examples. Bae et al. studied edge information as a forensic cue and proposed edge-attention and edge-concatenation layers that can be added to CNN and transformer-based classifiers, because edge-level information tends to weaken through repeated convolution and pooling [4].

Technical guidelines and standards provide another important foundation. NIST Special Publication 800-63A-4 explains identity proofing and evidence validation; it expects presented evidence to be confirmed as authentic, accurate and valid, and it recognizes automated document validation and cryptographic verification as acceptable techniques [5].

D. Motivation for the Study

The study is important because many organizations receive large volumes of paperwork and cannot examine every file thoroughly by hand. Manual reviews by staff members may differ from person to person and cause delays. Automated support can standardize preliminary checks, flag suspicious areas and provide a repeatable record of the reasons why a document was referred for review.

E. Objectives

The primary objective is to propose an AI-based framework that helps identify fraudulent or altered digital documents. The specific objectives are:

- to extract key information from uploaded documents using OCR;
- to detect visual and structural irregularities that may indicate alteration;
- to compare the extracted details with expected formats or trusted records; and
- to generate an explainable risk assessment that a human reviewer can understand.

F. Scope and Limitations

The proposed framework concentrates on digital copies of documents such as certificates, identity-related documents, academic records and administrative forms. Once a suitable dataset is prepared, it can be adapted to invoices, employment records and insurance documents. The framework does not claim to authenticate all forms of documents or to replace the legal authority of an issuing organization. Certain documents contain security features that cannot be assessed accurately from a compressed photograph.

G. Research Methodology

A literature review and conceptual design approach is adopted in this work. First, standards, technical specifications and pertinent research papers were examined. Second, typical document-forgery scenarios and verification requirements were identified. Third, a layered architecture covering upload, preprocessing, OCR, forensic analysis, rule-based checks, risk assessment and reporting was proposed.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, November 2026
Impact Factor: 8.2

II. PROPOSED FRAMEWORK

A. Description of the Research Problem

A digital document is typically represented as a collection of pixels, text, metadata and, occasionally, cryptographic data. A forged document may have one or more of these components altered, which gives rise to three challenges. The first is authenticity: the system must determine whether there are indications of alteration and whether the document appears to have come from the purported issuer. The second is accuracy: extracted data such as name, date, registration number and amount must be checked for internal consistency and, where feasible, cross-checked against a trusted source. The third is validity: even a genuine document may have expired, been revoked, or fallen outside the time frame accepted by the receiving organization.

B. Overview of the Proposed System

The proposed AI-based Fake Document Verification System has seven main stages: document upload and file validation, preprocessing, OCR text extraction, visual-forensic analysis, structural and semantic validation, cryptographic and trusted-source verification, and risk assessment with report generation. The overall workflow is shown in Fig. 1, and the function of each stage is summarized in Table I.

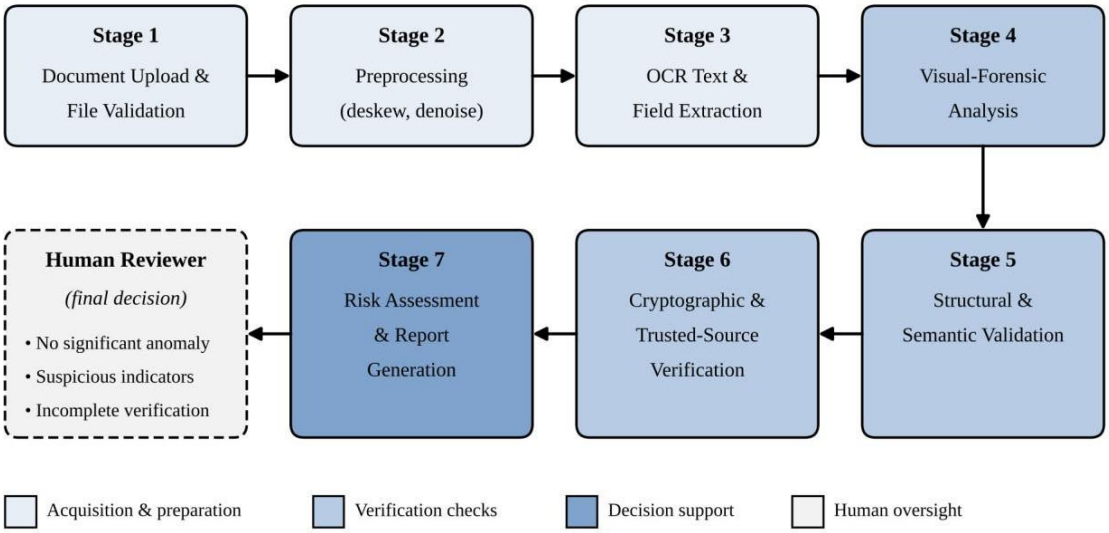


Fig. 1. Proposed seven-stage workflow of the AI-based fake document verification system.

TABLE I: STAGE-WISE SUMMARY OF THE PROPOSED FRAMEWORK

Stage	Main Function	Typical Techniques	Output
1	Upload and file validation	Checks of file size, extension, file signature, page count and readability	Accepted or rejected file
2	Preprocessing	Orientation correction, cropping, deskewing, denoising, contrast adjustment, resizing	Clean working copy; original preserved
3	OCR text extraction	OCR, layout-aware models, named-entity recognition, regular expressions, templates	Structured fields with OCR confidence
4	Visual-forensic analysis	Font, texture and compression analysis; copy-move detection; CNN and transformer classifiers; segmentation	Suspicious regions and anomaly score

Copyright to IJARSCT
www.ijarsct.co.in
DOI: 10.48175/IJARSCT-40972

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, November 2026
Impact Factor: 8.2

5	Structural and semantic validation	Date logic, ID patterns, arithmetic checks and cross-field comparison	List of inconsistencies
6	Cryptographic and trusted-source verification	Digital signatures, QR codes, issuer portals or APIs, verifiable credentials, provenance data	Issuer-validation result
7	Risk assessment and report generation	Weighted combination of signals with calibrated thresholds	Explainable report with a cautious outcome category

Stage 1: Document Upload and File Validation. A supported file, such as a JPG, PNG or PDF, is uploaded by the user. The system verifies the file size, extension, file signature, number of pages and basic readability. It should not rely on the file name alone, because an executable or corrupted file could be renamed with a document extension.

Stage 2: Preprocessing. Scanned documents may be skewed, rotated, blurred, shadowed or unevenly illuminated. Preprocessing therefore includes orientation correction, cropping, deskewing, denoising, contrast adjustment, resizing and conversion to a uniform color space. Each page of a multi-page PDF can be rendered at an appropriate resolution. The original file should be preserved, and a separate working copy should be produced.

Stage 3: OCR Text Extraction. OCR creates machine-readable characters from the text visible in a document image. Fields such as full name, date of birth, document number, institution name, course, grade, issue date and expiry date can then be identified from the extracted text. Field extraction may use layout-aware models, named-entity recognition, regular expressions and templates.

Stage 4: Visual-Forensic Analysis. This module looks for evidence of editing in the document image. Inconsistent font size, character spacing, alignment, color, background texture, edge sharpness, compression patterns and local noise are all possible warning signs. The texture or compression properties of a pasted area may differ from those of the surrounding areas, and regions that appear duplicated within the same image can be found using copy-move detection. A machine-learning classifier can examine image patches or the entire document to determine whether a region resembles known manipulated examples. Convolutional neural networks (CNNs) are capable of learning local visual features, while transformer-based models can capture more extensive relationships across a page; edge-focused layers can further strengthen sensitivity to boundary inconsistencies [4]. Segmentation models can highlight suspicious pixels or regions instead of merely returning a document-level label.

Stage 5: Structural and Semantic Validation. Even when a document has a consistent visual style, it may contain information that is impossible or contradictory. The system can verify that dates are arranged logically, that required fields are present, that an identification number follows the expected pattern, and that the sum of line items equals the stated amount. Semantic validation can also compare text fields with one another: a stated age should match the date of birth, a name appearing in two sections should not differ unexpectedly, and an issue date should not precede the date on which the institution was founded.

Stage 6: Cryptographic and Trusted-Source Verification. In accordance with the issuer's terms and the applicable law, the system should make use of any official verification portal, API, QR code or digitally signed credential that the issuer provides. A digital signature can be used to determine whether signed content has been altered and whether it was signed with the key of a trusted certificate. The W3C Verifiable Credentials Data Model 2.0 provides a model in which claims made by an issuer are presented by holders to verifiers, and cryptographic proofs can be used to detect manipulation of credential data [6]. The verifier nevertheless needs a trust policy to decide which issuers are accepted and for what purpose. Similarly, the specifications of the Coalition for Content Provenance and Authenticity (C2PA) support provenance data about the origin and history of digital media [7].

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, November 2026
Impact Factor: 8.2

Stage 7: Risk Assessment and Report Generation. The system integrates the results of rule checks, layout consistency, visual anomaly detection, OCR confidence and trusted-source validation. A conceptual decision model can place a document in one of three categories:

• incomplete verification;
• suspicious indicators found; or
• no significant anomaly detected.

The wording is purposefully cautious. In a basic conceptual risk model, independent signals such as visual-tampering probability, text-field inconsistency, metadata anomaly and issuer-validation result can be given weights. The total score should not be regarded as a universal mathematical truth; the weights must be learned or calibrated using a dataset that accurately represents the target documents and the operational environment. A weak image anomaly may be less significant than a strong issuer verification.

C. System Architecture and Implementation Factors

A workable web-based implementation may consist of a frontend for document submission and result display, a backend API for workflow management, protected object storage for original files, a processing service for OCR and image analysis, and a database for metadata and audit records. A queue allows large files to be handled asynchronously and keeps the web interface responsive. Only the data needed for verification should be kept in the database; possible records include a user or organization reference, document identifier, upload timestamp, file hash, document category, extracted fields, model version, check results, reviewer decision and retention date. Sensitive document images should be encrypted both in transit and at rest, retention should follow a specified policy, and access should be recorded.

D. Dataset Preparation and Model Training

A machine-learning detector needs examples of both authentic and altered documents. The dataset should cover several document categories, issuers, languages, scanners, mobile cameras, resolutions, lighting conditions and compression levels. Forgery examples may include text replacement, copy-move, region splicing, signature alteration, photo replacement, date modification and the creation of synthetic documents. Dataset leakage is a major risk: if nearly identical versions of the same document appear in both the training and test sets, a model may memorize the template instead of learning general forgery indicators. Splitting should therefore be done, where appropriate, by source document, issuer or template, and the test set should include realistic distortions and unseen templates.

E. Evaluation Methodology

The system should be assessed using precision, recall, accuracy, F1-score and a confusion matrix, as summarized in Table II. Precision is the proportion of documents reported as forged that are in fact forged, whereas recall quantifies the proportion of forged documents that are successfully identified. In a high-impact situation, false positives may unjustly delay legitimate applicants, while false negatives may allow fraudulent documents to pass. Processing time, memory usage, failure rate, OCR field accuracy and the proportion of cases that need manual review should also be measured.

TABLE II: MEASURES FOR EVALUATING THE VERIFICATION SYSTEM

Measure	What It Indicates	Practical Significance
Precision	Share of documents flagged as forged that are actually forged	A low value means genuine applicants are unjustly delayed
Recall	Share of forged documents that are correctly identified	A low value means forged documents pass undetected
Accuracy	Share of all documents classified correctly	Should be read together with precision and recall

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, November 2026
Impact Factor: 8.2

F1-score	Balance between precision and recall	Single summary of detection quality
Confusion matrix	Counts of true and false positives and negatives	Shows which kinds of error the system makes
Operational measures	Processing time, memory usage, failure rate, OCR field accuracy, share of cases sent to manual review	Shows whether the system is practical to deploy

III. ANALYSIS AND DISCUSSION

A. Analysis of the Proposed Framework

According to the literature and standards reviewed, no single verification method is adequate for every document. OCR can show what is written in a document, but it cannot determine whether the content is authentic. Image forensics can detect suspicious edits, although some legitimate processing operations leave similar traces. A template check can identify structural variations, although issuers may alter their formats over time. The first observation is therefore that verification should be organized around different questions: Is the file technically readable and safe? Does the document image appear to have been altered? Are the extracted details internally consistent? Does the issuer acknowledge the claims made in the document? Is the document up to date and appropriate for its intended use?

The second observation concerns the role of AI. Machine learning can be useful for identifying subtle patterns across large datasets in situations where manual review is repetitive. However, an AI model learns only from its training examples, and its reliability depends on how well those examples represent real documents.

The third observation is that explainability is a requirement in a high-impact verification workflow. A simple output such as “fake: 92%” does not tell a reviewer what was found, whether the score is calibrated, or what evidence supports the conclusion. A more responsible report identifies the suspicious area, the field that failed validation, the source consulted, and the degree of confidence or limitation of each check.

The fourth observation is the distinction between authenticity and truth. A document issued by a reputable organization may still contain inaccurate or out-of-date information [6]. Conversely, a document image may appear suspicious because of compression or angled photography even though the underlying document is authentic.

The fifth observation is that, where available, provenance data and cryptographic credentials can lessen the need for visual inspection. A digital signature can help identify modifications to signed data, and a verifiable credential can offer machine-checkable claims from an issuer. However, these mechanisms rely on clear governance, issuer trust lists, certificate validation, revocation processes and key management.

The sixth observation is that the operational design of the system can be just as significant as its model. Secure file handling, access control, encryption, audit trails, retention rules and incident response reduce the possibility that the verification service itself becomes a source of data exposure. Identity documents, in particular, can contain highly sensitive personal information.

B. Expected Benefits

The proposed system can help staff concentrate on high-risk cases, standardize the review process and reduce the time spent on routine preliminary checks. Automated extraction can decrease repeated manual data entry, and a visual heat map can direct an examiner to a suspicious area. By documenting which checks were completed, a consistent report can enhance auditability.

C. Risks and Challenges

False positives are a major concern. A legitimate document may be flagged if it is old, photocopied, compressed, stained or created from an unfamiliar template. False negatives are equally dangerous because sophisticated forgeries can mimic authentic layouts and security features; deep-learning-based text-editing attacks have already been shown to

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, November 2026
Impact Factor: 8.2

deceive some existing authentication systems [1]. The system may also perform differently across languages, scripts, document types and image-capture devices.

Obtaining trustworthy ground truth is another difficulty. A dataset labeled "genuine" should be supported by reliable provenance or issuer confirmation rather than by the appearance of a document alone, and a dataset labeled "fake" should have a documented history of known manipulation. Collecting such data can be challenging because real forged documents may be legally sensitive and institutions may not share their records.

D. Legal and Ethical Considerations

A verification system may affect access to public services, employment, education and finance, and an incorrect rejection can have serious consequences. Users should be able to submit additional evidence, correct extraction mistakes and request human review. Decisions should be proportionate to the risk and should not be based on irrelevant or protected personal characteristics. The use of third-party verification sources must also be transparent and lawful: an organization should not query issuer systems or scrape private databases without authorization. A QR code or URL printed on a document should be treated as untrusted input until its destination has been verified. Unless absolutely necessary, logs should not store complete identity numbers or document contents.

E. Proposed Pilot Study and Validation

Any future implementation should begin with a single document category, such as academic certificates from a small number of participating institutions. Before training, the team should prepare both genuine and manipulated examples, obtain consent and issuer-approved samples, and establish annotation rules. The first version may include OCR, template validation and a baseline image classifier. In addition to detection scores, the pilot should report the average processing time, the percentage of cases escalated to human review, the number of documents incorrectly flagged and the number of forged documents overlooked. Reviewers should evaluate how clear and helpful the explanations are, and performance should be tested on low-quality mobile photographs and unseen templates.

The choice of decision thresholds is another important problem. Although a model may produce a probability or confidence score, a threshold appropriate for one institution may not be appropriate for another. A financial institution, for instance, may require stronger issuer confirmation before approving a high-value transaction, whereas a university screening routine certificates may decide to send only borderline cases for manual review. Thresholds should be chosen by considering validation data and the effects of different errors. The system should also support an "insufficient evidence" conclusion: when the issuer service is unavailable or the image is unreadable, classifying every document as authentic or fraudulent would create unwarranted certainty.

Monitoring must continue during deployment. After training, document formats, institutional logos, security marks and typical forgery techniques may change. The system should therefore track changes in input quality, keep records of model versions and assess the frequency of human overrides. Subject to privacy and legal restrictions, an independent audit of a sample of accepted and rejected documents should be conducted regularly. The organization should be able to switch off automated decision-making, revert to manual review and retrain the model with fresh, validated examples when performance declines. Model updates should be tested and documented before release so that previous decisions can be reproduced.

The research contribution of this paper consists of a set of design principles for an engineering prototype and an integrated conceptual workflow. It combines active verification methods, semantic checks and passive forensic signals into a single decision-support procedure. The framework also highlights the difference between "not detected" and "proven authentic", which is crucial for responsible communication. Its usefulness will depend on the quality of the dataset, the availability of reliable issuer services and the organization's capacity to supervise human review. These requirements should be considered an integral part of the system rather than optional additions after model development.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, November 2026
Impact Factor: 8.2

IV. CONCLUSION

A. Conclusion Based on the Analysis

This study examined the problem of altered and fraudulent digital documents and proposed a multi-layered AI-based verification framework. The analysis shows that document authentication is a multifaceted process that involves issuer trust, image integrity, text extraction, structural consistency, accuracy of attributes and validity of the document. The proposed framework includes secure document upload, preprocessing, OCR, visual-forensic analysis, semantic and rule-based validation, trusted-source verification, risk assessment and report generation. Its primary contribution is the incorporation of these functions into an explainable workflow instead of dependence on a single AI prediction. It can assist organizations in prioritizing manual inspection, cutting down on repetitive tasks and keeping reliable verification records.

B. Perspectives and Future Scope

AI should be viewed as a tool that assists decision-making rather than as the final judge of the authenticity of a person or document. In high-impact cases the system should convey uncertainty, explain its results and let a qualified reviewer make the final decision. Future work could involve building a prototype for specific academic documents, creating a privacy-preserving dataset, supporting Indian languages, integrating authorized issuer APIs and evaluating contemporary deep-learning models for tamper localization. Research should also assess the fairness of error rates across document categories and capture conditions, as well as robustness against unseen forgery techniques.

ACKNOWLEDGMENT

The authors express their sincere gratitude to Dr. J. J. Magdum College of Engineering, Jaysingpur, for providing the academic support and facilities that made this work possible.

REFERENCES

[1] L. Zhao, C. Chen, and J. Huang, “Deep learning-based forgery attack on document images,” IEEE Trans. Image Process., vol. 30, pp. 7964–7979, 2021, doi: 10.1109/TIP.2021.3112048.
[2] R. Sukhija, M. Kumar, and M. K. Jindal, “Document forgery detection: A comprehensive review,” Int. J. Data Sci. Anal., vol. 20, no. 5, pp. 4385–4407, 2025, doi: 10.1007/s41060-025-00723-0.
[3] C. Qu, C. Liu, Y. Liu, X. Chen, D. Peng, F. Guo, and L. Jin, “Towards robust tampered text detection in document image: New dataset and new solution,” in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Vancouver, BC, Canada, 2023, pp. 5937–5946, doi: 10.1109/CVPR52729.2023.00575.
[4] Y.-Y. Bae, D.-J. Cho, and K.-H. Jung, “Enhancing document forgery detection with edge-focused deep learning,” Symmetry, vol. 17, no. 8, Art. no. 1208, 2025, doi: 10.3390/sym17081208.
[5] D. Temoshok et al., “Digital identity guidelines: Identity proofing and enrollment,” National Institute of Standards and Technology, Gaithersburg, MD, USA, NIST Special Publication 800-63A-4, Jul. 2025, doi: 10.6028/NIST.SP.800-63A-4.
[6] W3C, “Verifiable Credentials Data Model v2.0,” W3C Recommendation, May 15, 2025. [Online]. Available: https://www.w3.org/TR/vc-data-model-2.0/
[7] Coalition for Content Provenance and Authenticity (C2PA), “C2PA Technical Specification.” [Online]. Available: https://spec.c2pa.org/ (accessed Nov. 22, 2026).