Universidade de Brasília

Faculdade de Tecnologia

Departamento de Engenharia Elétrica

Digital Preservation with a Blockchain of Custody Architecture Using OAIS and Video Steganography

Guilherme Fay Vergara

TESE DE DOUTORADO

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Brasília

2026

Universidade de Brasília
Faculdade de Tecnologia
Departamento de Engenharia Elétrica

Preservação Digital com uma Arquitetura de Blockchain de Custódia Utilizando OAIS e Esteganografia em Vídeo

Guilherme Fay Vergara

Tese de Doutorado submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Doutor.

Orientador: Prof. Dr. Fábio Lúcio Lopes de Mendonça

Publicação PPGEE 219/26

Brasília
2026

REFERÊNCIA BIBLIOGRÁFICA

FAY VERGARA, Guilherme. **Digital Preservation with a Blockchain of Custody Architecture Using OAIS and Video Steganography**. Tese de Doutorado (Programa de Pós-Graduação em Engenharia Elétrica), Publicação PPGEE 219/26 – Departamento de Engenharia Elétrica, Faculdade de Tecnologia, Universidade de Brasília, Brasília, 168 p. 2026.

FICHA CATALOGRÁFICA

Fay Vergara, Guilherme.

Digital Preservation with a Blockchain of Custody Architecture Using OAIS and Video Steganography / Guilherme Fay Vergara; orientador Fábio Lúcio Lopes de Mendonça. -- Brasília, 2026.

168 p.

Tese de Doutorado (Programa de Pós-Graduação em Engenharia Elétrica) -- Universidade de Brasília, 2026.

1. Digital chain of custody. 2. Blockchain. 3. Video steganography. 4. Digital preservation. 5. Deep learning. I. Lopes de Mendonça, Fábio Lúcio, orient. II. Título.

Universidade de Brasília
Faculdade de Tecnologia
Departamento de Engenharia Elétrica

Digital Preservation with a Blockchain of Custody Architecture Using OAIS and Video Steganography

Guilherme Fay Vergara

Tese de Doutorado submetida ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade de Brasília como parte dos requisitos necessários para obtenção do grau de Doutor.

Trabalho aprovado. Brasília, 06 de Março de 2026:

Prof. Dr. Fábio Lúcio Lopes De Mendonça, FT/ENE/UnB
Orientador

Prof. Dr. Georges Daniel Amvame Nze, FT/ENE/UnB
Examinador Interno

Prof. Dr. Geraldo Pereira Rocha Filho, UESB
Examinador Externo

Prof. Dr. Daniel Alves da Silva, Hochschule Hamm-Lippstadt
Examinador Externo

Dedico este trabalho à minha família, Rosane, Dirson e Rodrigo,
pelo apoio inabalável e pela inspiração que me motivaram a chegar até aqui.

Agradecimentos

Agradeço, primeiramente, a Deus por me conceder força, sabedoria e perseverança ao longo desta caminhada. Sua presença foi meu sustento nos momentos de dificuldade e minha luz nas incertezas.

À minha família, alicerce fundamental desta conquista, agradeço pelo apoio incondicional, pela paciência nos momentos de ausência e pelo incentivo constante. Sem o amor e a compreensão de vocês, esta jornada não teria sido possível.

Ao meu orientador, Prof. Fábio Lúcio Lopes de Mendonça, expresso minha profunda gratidão pela orientação dedicada, pela generosidade no compartilhamento de conhecimentos e pelo estímulo constante que me permitiram superar desafios e amadurecer como pesquisador. Sua confiança em meu trabalho foi determinante para a conclusão desta tese.

Um agradecimento especial ao Prof. Dr. Rafael Timóteo de Sousa Júnior, cuja visão acadêmica, liderança e dedicação à pesquisa foram fonte de inspiração ao longo de toda a minha trajetória. Sua contribuição para a formação de pesquisadores e para o avanço da ciência no Brasil é um exemplo que levarei comigo.

Aos professores Daniel Alves da Silva e Robson de Oliveira Albuquerque, sou grato pelas contribuições valiosas, pelas discussões enriquecedoras e pelo rigor acadêmico que fortaleceram este projeto. Suas perspectivas complementares ampliaram significativamente o alcance desta pesquisa.

Agradeço o apoio técnico e computacional do Laboratório de Tecnologias da Tomada de Decisão (LATITUDE) da Universidade de Brasília, que conta com o apoio do CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico (Outorgas 312180/2019-5 PQ-2, BRICS2017-591 LargEWiN e 465741/2014-2 INCT em Cibersegurança), da CAPES – Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Outorgas 23038.007604/2014-69 FORTE e 88887.144009/2017-00 PROBRAL) e da FAP-DF – Fundação de Apoio à Pesquisa do Distrito Federal (Outorgas 0193.001366/2016 UIoT e 0193.001365/2016 SSDDC).

Aos colegas do grupo de pesquisa, agradeço pelo ambiente colaborativo, pelas trocas de experiências e pelo companheirismo que tornaram esta trajetória mais rica e produtiva.

Por fim, aos amigos que estiveram ao meu lado nos momentos de dúvida e celebração, oferecendo amizade verdadeira, palavras de encorajamento e a leveza necessária para equilibrar a intensidade do trabalho acadêmico. Vocês tornaram este processo não apenas possível, mas também memorável.

“Quem vigia os vigilantes?”
“Who watches the watchmen?”
“Quis custodiet ipsos custodes?”
— Decimus Junius Juvenalis, *Satires*

Abstract

Digital evidence can be modified without leaving traces, its provenance can be forged, and audit trails can be manipulated by those with administrative access. Traditional approaches rely on institutional trust and procedural safeguards, which are insufficient against determined adversaries. Although partial solutions exist, no system integrates blockchain, OAIS digital preservation, and steganography into a coherent architecture. This thesis designs, implements, and validates a modular architecture that combines blockchain with Proof of Work, OAIS-compliant preservation with BagIt packaging (RFC 8493), and video steganography, providing cryptographically verifiable integrity, archival compliance, and covert evidence transport. For video-in-video scenarios, the thesis proposes Stego-STFAN, a novel neural network architecture based on spatial-temporal adaptive filtering that replaces computationally expensive 3D convolutions with FAC modules and CBAM attention. The architecture was fully implemented and validated following the Design Science methodology.

Validation executed 14 live attack scenarios across three architectural layers. At the blockchain layer, six attack types were executed with 100% accuracy across 260 test cases. At the steganographic layer, bit-flip, lossy re-encoding, and payload-swap attacks were executed using SHA-256 hash comparisons, and byte-perfect reversibility was confirmed across 200 LSB embedding/extraction cycles. At the custody layer, five attacks were executed by BagIt's two-level manifest structure. Stego-STFAN achieved a container PSNR of 27.03 dB, ranking third among the compared models with 43.91 M parameters. Security analysis against CIA+ANP properties shows the integrated system satisfies all six properties (Confidentiality, Integrity, Availability, Authenticity, Nonrepudiation, and Privacy).

The results demonstrate that blockchain-anchored preservation with steganographic capability provides a viable approach to digital evidence management, offering mathematical verification independent of institutional trust while maintaining compliance with established archival standards.

Keywords: digital chain of custody; blockchain; video steganography; digital preservation; deep learning.

Resumo

Título:P reservação Digital com uma Arquitetura de Cadeia de Custódia Baseada em Blockchain Utilizando OAIS e Esteganografia em Vídeo
Curso: Programa de Pós-Graduação em Engenharia Elétrica (PPGEE)
Data da Defesa: 06/03/2026
Orientador: Prof. Dr. Fábio Lúcio Lopes de Mendonça

Documentos digitais podem ser modificadas sem deixar vestígios, sua proveniência pode ser forjada e trilhas de auditoria podem ser manipuladas por quem detém acesso administrativo. Abordagens tradicionais dependem de confiança institucional e de salvaguardas procedimentais, insuficientes diante de adversários determinados. Embora existam soluções parciais, nenhum sistema integra blockchain, preservação digital OAIS e esteganografia em uma arquitetura coerente. Esta tese propõe, implementa e valida uma arquitetura modular que combina blockchain com Proof of Work, preservação conforme OAIS/BagIt (RFC 8493) e esteganografia em vídeo, fornecendo integridade criptograficamente verificável, conformidade arquivística e transporte. Para cenários de vídeo-em-vídeo, propõe-se o Stego-STFAN, uma arquitetura de rede neural original baseada em filtragem adaptativa espaço-temporal que substitui convoluções 3D por módulos FAC com atenção CBAM. A arquitetura foi integralmente implementada e validada seguindo a metodologia Design Science.

A validação executou 14 cenários de ataque ao vivo em três camadas. Na camada blockchain, seis tipos de adulteração foram testados com 100% de acurácia em 260 casos de teste. Na camada esteganográfica, ataques de bit-flip, recodificação lossy e substituição de payload foram detectados por meio da comparação com o SHA-256, com a reversibilidade perfeita confirmada em 200 ciclos LSB. Na camada de custódia, cinco ataques foram testados pela estrutura de dois níveis de manifesto do BagIt. O Stego-STFAN alcançou 27,03 dB de PSNR no container, terceiro lugar entre modelos comparados, com 43,91 M de parâmetros. A análise CIA+ANP demonstra que o sistema satisfaz todas as seis propriedades de segurança (Confidencialidade, Integridade, Disponibilidade, Antenticidade, Não repúdio e Privacidade).

Os resultados demonstram que a preservação ancorada em blockchain com capacidade esteganográfica constitui uma abordagem viável para a gestão de documentos digitais, oferecendo verificação matemática independente de confiança institucional e de conformidade com padrões arquivísticos estabelecidos.

Palavras-chave: cadeia de custódia digital; blockchain; esteganografia em vídeo; preservação digital; aprendizado profundo.

List of figures

List of tables

List of abbreviations and acronyms

AIP	Archival Information Package
API	Application Programming Interface
BagIt	File Packaging Format (RFC 8493)
CBAM	Convolutional Block Attention Module
CIA	Confidentiality, Integrity, Availability
CID	Content Identifier (IPFS)
DIP	Dissemination Information Package
DWT	Discrete Wavelet Transform
FAC	Filter Adaptive Convolutional
FFV1	FFmpeg Video Codec 1
HFYU	Huffyuv Lossless Codec
IPFS	InterPlanetary File System
ISAD(G)	General International Standard Archival Description
ISO	International Organization for Standardization
LSB	Least Significant Bit
METS	Metadata Encoding and Transmission Standard
NLCA	Non-Local Co-Attention
NLSA	Non-Local Self-Attention
OAIS	Open Archival Information System (ISO 14721)
PDI	Preservation Description Information
PoW	Proof of Work
PREMIS	Preservation Metadata Implementation Strategies
PSNR	Peak Signal-to-Noise Ratio
RDC-Arq	*Repositório Arquivístico Digital Confiável*
SHA	Secure Hash Algorithm
SIP	Submission Information Package
STFAN	Spatial-Temporal Filter Adaptive Network

Contents

1 Introduction

The digital transformation of society has created unprecedented challenges for digital document management. Documents that once existed as physical artefacts with inherent properties of uniqueness and difficulty of modification now exist as sequences of bits that can be copied, altered, and redistributed without leaving obvious traces. This fundamental shift has profound implications for legal proceedings, institutional archives, and any context in which the authenticity and integrity of records are at issue. The scale of this challenge is vast: the International Data Corporation estimates that the global datasphere will reach 175 zettabytes by 2025 (Reinsel; Gantz; Rydning, 2018), and an ever-growing fraction of this data constitutes records with legal, regulatory, or historical significance.

The consequences of inadequate digital evidence management are not hypothetical. In 2017, a Brazilian federal investigation was compromised when defence attorneys demonstrated that digital audio files presented as evidence had been edited after collection, with metadata timestamps inconsistent with the chain of custody documentation (Badaró, 2020). The court excluded the recordings, and the case collapsed. In the United States, the Zubulake v. UBS Warburg case (2003–2004) established precedent when a federal court imposed sanctions after the defendant failed to preserve emails, with the court noting that digital evidence, once modified or lost, provides no physical trace of its alteration (United States District Court, Southern District of New York, 2004). More recently, the proliferation of deepfakes has introduced a new dimension to the problem. In 2023, European courts confronted cases where parties disputed the authenticity of video evidence, arguing that it might have been AI-generated (Chesney; Citron, 2019). These cases illustrate that the trustworthiness of digital records is no longer a merely technical concern; it is a foundational requirement for the functioning of legal systems and democratic institutions.

Securing digital documents and managing such evidence requires consideration of the document chain of custody. With its legal origins, this concept plays a pivotal role in maintaining the integrity and authenticity of digital documents. As articulated by (Smith *et al.*, 1990), the chain of custody is a chronological sequence of events that links a document from collection to storage, providing a documented record. This meticulous process involves tracking possession, ensuring proper handling of samples, and maintaining detailed records of each step, from preparation and collection through transportation, analysis, and storage.

Junior (2012) emphasises the significance of adhering to the chain of custody for expert examinations, particularly in criminal prosecutions. The author underscores the necessity for detailed, robust, and reliable procedures, asserting that the irrefutability of a report hinges on the clarity of the sequence of facts: who handled the digital document, how it was handled, where the trace was obtained, how it was stored, and the reasons for

each step. The Brazilian Criminal Procedure Code, amended by the *Pacote Anticrime* (Law 13.964/2019), codified this requirement in Articles 158-A through 158-F, establishing that the chain of custody encompasses every stage of handling forensic evidence, from recognition and isolation through analysis, storage, and disposal. Failure to maintain this chain can result in the exclusion of evidence, as demonstrated by rulings of the Superior Tribunal de Justiça (STJ) that have annulled proceedings where chain of custody documentation was found to be incomplete or inconsistent (Badaró, 2020). Beyond the immediate scope of legal investigations, the uninterrupted chain of custody extends its relevance to archival documents. This continuous line of custodians, from document producers to legitimate successors, ensures the authenticity and non-alteration of archival records and becomes a foundational principle in safeguarding document integrity and establishing provenance within a historical and legal context.

The challenge is not merely technical but epistemological. When anyone with access can modify a digital file, and that modification leaves no physical evidence, how can we establish trust in digital records? Traditional approaches rely on administrative procedures: access controls limit who can modify files, audit logs record who accessed what and when, and institutional policies govern handling procedures. These measures are necessary but insufficient. A determined adversary with administrative access can modify files, alter logs, and cover their tracks. The 2020 SolarWinds supply-chain attack demonstrated this vulnerability at scale: attackers with persistent access to enterprise systems were able to manipulate files and evade detection for months, compromising the integrity of records held by government agencies and corporations alike (Hussain; Mohamed; Razali, 2020). Even without malicious intent, administrative failures undermine trust. A study of data breaches in publicly traded US companies found that compromised data integrity led to average financial losses of USD 4.24 million per incident, with regulatory penalties and reputational damage extending far beyond the immediate technical remediation (Rodrigues *et al.*, 2024).

This thesis addresses this challenge through a technological approach that complements administrative measures with mathematical guarantees. The integration of blockchain technology offers a powerful way to enhance the security and transparency of digital content. Blockchain, first introduced by (Nakamoto, 2008), is renowned for its decentralised and tamper-resistant nature and provides an immutable ledger that records the entire lifecycle of digital assets. When a document's hash is recorded on a blockchain, any subsequent modification produces a different hash, immediately revealing tampering. Incorporating blockchain into the chain-of-custody process enhances accountability, reduces the risk of unauthorised alteration, and establishes a secure, transparent digital trail. Recent literature confirms the growing interest in blockchain-based evidence management: Igonor, Amin, and Garg (2025) surveyed 67 studies published between 2018 and 2024, concluding that

blockchain is the most promising technology for ensuring digital forensic evidence integrity, while El-Gendy *et al.* (2023) identified its applicability specifically in maintaining forensic chain of custody. Ponnusamy and Manickam (2025) further highlight that blockchain's immutable ledger addresses the core limitation of traditional custody systems, namely the dependence on human attestation and centralised logs that can be retroactively modified.

However, integrity verification alone does not address all the challenges of digital document management. Digital preservation is a well-established field with international standards that define best practices for ensuring long-term accessibility and authenticity. The OAIS (Open Archival Information System) reference model, codified in ISO 14721, provides a conceptual framework for digital archives. The BagIt format, specified in RFC 8493, provides a practical packaging structure with built-in integrity verification. Any serious proposal for evidence management must align with these established practices to ensure interoperability and benefit from the accumulated knowledge of the preservation community. Balogun (2025) assessed the compliance of digital repositories in South Africa with OAIS and Trusted Digital Repository standards, finding that while many institutions adopt the reference model, few implement automated integrity checking mechanisms, precisely the gap that blockchain technology can fill. In Brazil, the Conselho Nacional de Arquivos (CONARQ) established the *Repositório Arquivístico Digital Confiável* (RDC-Arq) framework through Resolution 43/2015, mandating that trusted digital repositories conform to the OAIS model and implement systematic fixity verification requirements that the proposed architecture directly satisfies.

Furthermore, there are scenarios in which the confidentiality of the digital document itself is crucial. Recognising the evolving cybersecurity landscape, organisations are exploring innovative approaches to bolster their defences. One such strategy involves deploying steganographic techniques, traditionally used to conceal information within multimedia content. Steganography is a technique that hides data within seemingly unremarkable data; it dates back to ancient times and protects sensitive information from theft or detection. By harnessing steganography within the architecture, organisations can add an extra layer of security, enabling the covert transmission of critical information within seemingly innocuous data streams. Encrypted files, while secure against unauthorised reading, are obviously encrypted; the very presence of protected content may draw unwanted attention. Research on data breaches and their countermeasures shows that the exposure of sensitive information leads to substantial losses for organisations and individuals (Rodrigues *et al.*, 2024), and that conventional controls, such as Data Loss Prevention (DLP), are ineffective at detecting steganographic exfiltration in contexts where the existence of protected content might draw unwanted attention. A video file containing a hidden file appears completely normal when played; only someone who knows it contains hidden content and has the means to extract it can access the concealed information.

This thesis proposes and implements an architecture that integrates these three tech-

nologies: blockchain, digital preservation standards (OAIS and BagIt), and video steganography, into a unified system for managing digital documents with verifiable integrity and optional covert transport. Accordingly, the subsequent sections of this work will examine the problem statement, research objectives, and critical components, including the strategic application of steganography and the integration of blockchain, to strengthen the chain of custody for digital content.

1.1 Problem Statement

The management of digital documents faces several interconnected challenges that existing solutions address only partially.

The integrity problem arises because digital files can be modified without leaving obvious traces. Unlike physical documents that show signs of tampering, such as erasures or altered bindings, a modified digital file is indistinguishable from an unmodified one. Cryptographic hashes can detect modifications, but the hash value itself must be stored in a trustworthy location. If an adversary can modify both the file and its stored hash, the modification remains undetected.

The provenance problem concerns establishing who created or submitted the document and when. Digital files do not carry inherent timestamps or creator information that cannot be forged. Filesystem metadata can be manipulated, and embedded metadata can be edited. Establishing provenance requires an external mechanism that binds content to identity and time in a way that cannot be retroactively altered.

The audit problem involves maintaining a complete and tamper-evident record of all actions taken on the document. Who accessed the file? When was it transferred? Has it been verified, and by whom? These questions require an audit trail that cannot be modified or deleted by any party, including system administrators.

The confidentiality problem arises when a document must be transmitted through potentially monitored channels. Standard encryption protects content but reveals that protected content exists. In hostile environments, the mere presence of encrypted files may invite coercion or confiscation.

The interoperability problem concerns the need for document management systems to work with existing archival infrastructure. Proprietary formats and non-standard practices create silos that impede the exchange of documents between institutions and complicate long-term preservation.

Existing solutions address these problems partially. Blockchain-based systems provide integrity and provenance but typically lack integration with preservation standards. Digital preservation systems follow established practices but rely on administrative trust rather than cryptographic verification. Steganographic tools provide concealment but offer no guarantees

about the integrity or provenance of hidden content. No existing system integrates all three capabilities in a coherent architecture.

1.2 Research Objectives

The central research question addressed by this thesis is: *How can we create a system for managing digital documents that provides cryptographically verifiable integrity, maintains a complete audit trail, complies with digital preservation standards, and enables covert transport when circumstances require?*

To answer this question, this thesis pursues the following objectives:

1.2.1 General Objective

To design, implement, and validate an architecture that integrates blockchain technology, OAIS-compliant digital preservation, and video steganography into a unified system for secure management of digital documents.

1.2.2 Specific Objectives

Objective 1: Design a modular architecture combining blockchain, digital preservation, and steganography. The architecture must define how these three technologies interact to provide comprehensive document management capabilities. The design should be modular, allowing each component to be developed, tested, and maintained independently while functioning together as a coherent whole.

Objective 2: Implement a blockchain service for tamper-evident document registration. The blockchain must record document registrations, custody transfers, and other significant events in a way that cannot be altered after the fact. The implementation should achieve demonstrably high tamper detection rates across various corruption scenarios.

Objective 3: Implement a custody service compliant with OAIS and BagIt standards. The service must manage the complete document lifecycle through Submission, Archival, and Dissemination Information Packages. Packages must conform to RFC 8493 BagIt specification and include PREMIS preservation metadata.

Objective 4: Propose and evaluate a deep learning architecture for video-in-video steganography. For scenarios requiring hiding video content within video carriers, the thesis must propose and train a neural network architecture that achieves competitive visual quality compared to state-of-the-art approaches.

Objective 5: Validate the implementation against fundamental security properties. The complete system must be evaluated against Confidentiality, Integrity, Availability, Authenticity, Nonrepudiation, and Privacy (CIA+ANP) properties, demonstrating that the

integration provides security guarantees exceeding what any single technology could offer alone.

1.3 Hypotheses

Based on the research objectives, this thesis tests the following hypotheses:

- **H1 (Integrity):** A blockchain-based registration system can achieve 100% detection of tampering attempts across diverse corruption scenarios, including modified transaction data, altered timestamps, broken chain linkage, and fraudulent block insertion.
- **H2 (Preservation Compliance):** A document management system can simultaneously satisfy blockchain-backed integrity verification and OAIS/BagIt compliance, with packages passing external validation by established tools.
- **H3 (Video Steganography Quality):** A deep learning architecture for video-in-video steganography can achieve competitive container visual quality (PSNR above 25 dB) while avoiding computationally expensive 3D convolutions.
- **H4 (Security Validation):** The integration of blockchain with steganography satisfies the defined CIA+ANP security properties, as demonstrated through the successful execution of all specified security evaluation tests.

1.4 Methodology

This research follows the Design Science Research methodology (Peffers *et al.*, 2007), which is particularly well suited to information systems research, where the goal is to create and evaluate artefacts that solve identified problems. Design Science Research emphasises the iterative construction and evaluation of technological solutions, with contributions measured by the utility of the resulting artefacts.

The research proceeded through the following phases:

- **Problem Identification and Motivation:** The first phase established the significance of the research problem through literature review and analysis of existing solutions. This phase identified the gaps in current approaches that the proposed architecture would address.
- **Definition of Objectives:** Based on the identified problems, specific objectives were defined for what a successful solution would achieve. These objectives guided the subsequent design decisions.
- **Design and Development:** The architecture was designed as a modular system with interconnected services. Each service was implemented in Python using appropriate libraries for its function: FastAPI for the API gateway, custom implementations for

blockchain and custody services, OpenCV and FFmpeg for steganography, and PyTorch for the neural network components.

- **Demonstration:** The implemented system was exercised through complete workflows demonstrating all intended capabilities: document upload, blockchain registration, package transformation through the OAIS lifecycle, steganographic embedding and extraction, and integrity verification.
- **Evaluation:** Systematic testing validated each hypothesis through quantitative measurements. Blockchain tampering detection was tested across 260 corruption scenarios. Steganographic reversibility was verified across 200 embedding/extraction cycles. The Stego-STFAN network was trained, and its quality metrics were compared with published results from state-of-the-art models.
- **Communication:** This thesis documents the complete research process, presenting the architecture, implementation details, and evaluation results in sufficient detail for the work to be understood, evaluated, and potentially extended by other researchers.

1.5 Contributions

This thesis makes the following contributions to the field:

- **An integrated architecture:** The primary contribution is an architecture that combines blockchain, OAIS-compliant preservation, and video steganography in a way that has not been previously demonstrated. This integration provides security properties that exceed those of any single technology.
- **Stego-STFAN:** The thesis proposes a novel neural network architecture for video-in-video steganography based on spatial-temporal adaptive filtering. Although it achieves quality slightly below the best published results, the architecture avoids expensive 3D convolutions and demonstrates a promising approach.
- **Empirical validation:** The thesis provides extensive empirical evidence for the effectiveness of the proposed approach through systematic testing of blockchain integrity, steganographic reversibility, and security properties.
- **Standards alignment:** By aligning the architecture with OAIS, BagIt, PREMIS, and other established standards, the thesis demonstrates that blockchain-backed verification can be achieved without abandoning proven preservation practices.

1.5.1 Published Articles

The research conducted during the doctoral programme resulted in the following peer-reviewed publications, which disseminate partial results and related investigations that informed and supported this thesis:

1. Vergara *et al.* (2024) – **Stego-STFAN: A Novel Neural Network for Video Steganography.** Published in *Computers* (MDPI), vol. 13, no. 7, 2024. This article presents the Stego-STFAN architecture proposed in this thesis, describing the network design based on spatial-temporal adaptive filtering and reporting experimental results on container and secret video quality. The publication constitutes the primary dissemination of the deep learning contribution of this work.
2. Rodrigues *et al.* (2024) – **Impact, Compliance, and Countermeasures in Relation to Data Breaches in Publicly Traded US Companies.** Published in *Future Internet* (MDPI), vol. 16, no. 6, 2024. This article analyses the financial and regulatory impact of data breaches, demonstrating that conventional security controls are insufficient to prevent information exposure—a finding that directly motivates the multi-layer defence approach adopted in this thesis.
3. Fernandes *et al.* (2022) – **Proposta de Guia para Adequação de Repositórios Digitais Confiáveis à LGPD.** Published in *CIACA*, 2022. This article proposes a compliance guide for trusted digital repositories under Brazil's General Data Protection Law (LGPD), addressing the intersection of digital preservation and privacy regulation that informs the custody and metadata design of the proposed architecture.
4. Bispo *et al.* (2024) – **Automatic Literature Mapping Selection: Classification of Papers on Industry Productivity.** Published in *Applied Sciences* (MDPI), vol. 14, no. 9, 2024. This article presents a machine learning approach for automated classification of scientific literature, contributing to the systematic literature review methodology employed in the theoretical foundation of this thesis.
5. Coelho *et al.* (2023) – **Enhancing Industrial Productivity Through AI-Driven Systematic Literature Reviews.** Published in the *Proceedings of the 19th International Conference on Web Information Systems and Technologies (WEBIST)*, SciTePress, 2023. This conference paper presents the AI-driven methodology for literature analysis that supported the bibliometric survey underpinning this research.
6. Torres *et al.* (2022) – **Using Spatial Data and Cluster Analysis to Automatically Detect Non-Trivial Relationships Between Environmental Transgressors.** Published in the *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*, IEEE, 2022. This article applies data mining techniques to forensic investigation contexts, contributing to the broader research programme on computational methods for evidence analysis.
7. Vergara *et al.* (2022) – **A Study of Automatic Speech Recognition in Portuguese by the Brazilian General Attorney of the Union.** Published in the *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*, IEEE, 2022. This article investigates the application of AI to institutional evidence processing, reflecting the research group's engagement with digital evidence challenges in public institutions.

1.6 Thesis Organisation

The remainder of this thesis is organised as follows.

Chapter 2 presents the theoretical foundation, reviewing relevant concepts in blockchain technology, digital preservation (including OAIS, BagIt, and PREMIS), steganography (both classical techniques and deep learning approaches), and information security properties. The chapter also surveys related work, positioning this thesis within the broader research landscape.

Chapter 3 presents the proposed architecture in detail. The chapter describes each of the six modules (API Gateway, Blockchain Service, Custody Service, Steganography Service, Metadata Service, and IPFS Service), explains how they interact, and discusses the design decisions that shaped the architecture. The chapter also presents the Stego-STFAN architecture for video-in-video steganography.

Chapter 4 details the implementation of the proposed architecture, including its structural components and integration mechanisms.

Chapter 5 presents the validation and experimental assessment, reporting quantitative results on blockchain integrity, steganographic reversibility, verification of the CIA+ANP property, and the training and evaluation of the Stego-STFAN model.

Chapter 6 concludes the thesis by summarising contributions, discussing implications and limitations, and proposing directions for future research.

2 Literature Review and Related Work

This chapter surveys the extant literature on cybersecurity, including cyber threats, encryption techniques, and protective protocols. It then examines steganography, focusing on machine learning approaches, including convolutional neural networks and other neural network architectures, to enhance data concealment. Following this, the review analyzes the disruptive potential of blockchain technology, emphasizing its decentralized architecture that fosters trust and transparency in data transactions. The discussion also addresses the foundational principles of chain of custody and digital document management, which are essential for preserving data integrity. The chapter further reviews digital preservation standards, including the Open Archival Information System reference model, the BagIt file packaging format, and the Preservation Metadata: Implementation Strategies metadata standard. Finally, a bibliometric analysis identifies key scholarly contributions relevant to the thesis, offering a comprehensive view of the evolving academic discourse.

2.1 Cybersecurity

In the early days of computer technology, security strategies focused on physical protection because individual users predominantly used standalone systems. These systems and accompanying devices remained in secure locations, monitored by attentive guards who carefully verified user identities before granting entry.

The progression of time-sharing systems from the mid to late 1960s marked a significant change in security requirements. Managing access to system data became a top priority as multiple users and tasks were involved simultaneously. One prevalent approach at the time was batch processing, a strategy that processed classified data in a step-by-step manner over specific time segments. This technique involved completing tasks for each security level within assigned time frames, interspersed with intervals for system cleaning to remove any residual data before advancing to the next security level.

Cybersecurity discussions were primarily confined to technical communities, with an emphasis on developing encryption algorithms, firewalls, and intrusion detection systems to protect digital assets. However, as the Internet became more commonplace and interconnected, cybersecurity expanded beyond technological concerns. The interdisciplinary field of cybersecurity recognizes that combating cyber threats requires expertise from diverse fields, including computer science, engineering, political science, psychology, management, sociology, and law.

This multidisciplinary approach recognizes the intricate interactions that shape cybersecurity outcomes, including those between human behavior, technical breakthroughs,

societal dynamics, and regulatory frameworks. As Craigen, Diakun-Thibault, and Purse (2014) noted, cybersecurity has many definitions due to its interdisciplinary nature Frederick R. Chang (2012). The former Director of Research at the United States National Security Agency discusses the interdisciplinary nature of cybersecurity.

Cybersecurity prevents unauthorized access, attacks, theft, and damage to computer systems, networks, and data, which are essential to preserving data integrity, confidentiality, and availability (Belkhamza, 2023). Furthermore, diverse methodologies and strategies are employed to prevent, detect, address, and recover from cyber threats, which have become a critical concern for organizations, individuals, and society. The information and communication technology industry has evolved significantly in the past five decades, becoming an integral part of our modern society.

According to research by (Hussain; Mohamed; Razali, 2020), critical cybersecurity aspects include ensuring confidentiality, integrity, and data availability. This aspect involves allowing only authorized individuals or systems to access sensitive information, safeguarding data accuracy and reliability, and ensuring that systems and data remain accessible and operational when needed while minimizing downtime.

Authentication and authorization are vital components that verify the identities of users, devices, or systems, prevent unauthorized access, and grant appropriate permissions based on roles and responsibilities. Moreover, effective risk management, incident response strategies, and the establishment of security policies and education programs are essential for mitigating cyber threats and ensuring a secure cyber environment.

2.1.1 Security Principles

To achieve a secure and resilient ecosystem, a set of fundamental security principles must be enforced across all digital and physical components, including cyber-physical systems, automation infrastructures, and integrated IT/OT environments. According to the National Cyber Security Strategy (NCSS) (Luiif *et al.*, 2011) and the Committee on National Security Systems (CNSS) (Homeland Security Digital Library, 2015), six core principles must be addressed to ensure secure and reliable systems.

2.1.1.1 Confidentiality

Confidentiality is a relevant security requirement, although it may not be mandatory in all scenarios, as it ensures that information is not exposed to unauthorized internal or external entities by applying encryption mechanisms to data at rest and in transit and by restricting access to sensitive data locations, protecting proprietary information, business data, security credentials, and cryptographic keys from unauthorized disclosure (Fink *et al.*, 2017).

2.1.1.2 Integrity

Integrity is a mandatory security feature in most systems, as it guarantees that data, control commands, and system configurations remain accurate and unaltered, preventing unauthorized manipulation by external attackers or insiders; if integrity is compromised, systems may process incorrect data as valid, potentially leading to unsafe behavior or severe operational consequences in safety-critical environments (Thames; Schaefer, 2017).

2.1.1.3 Availability

Availability ensures that systems can deliver services and support processes when required, meaning that each subsystem must operate correctly and on time while remaining resilient to hardware and software failures, power outages, and denial-of-service attacks; critical systems such as monitoring and control infrastructures typically require higher availability levels than non-critical components (Fink *et al.*, 2017).

2.1.1.4 Authenticity

Authentication is driven by diverse interaction patterns among machines, systems, and human operators, requiring adaptable authentication mechanisms for machine-to-machine, human-to-machine, and cross-system communications; strong authentication and authorization mechanisms are essential in high-impact scenarios, while some solutions must support interoperability across organizational or national boundaries and others remain limited to local domains (Chenoweth, 2005).

2.1.1.5 Nonrepudiation

Nonrepudiation assures that the responsible entities cannot deny actions performed within systems. Although it is not a primary requirement for all applications, it is essential in scenarios involving payments, accountability, or regulatory compliance, as it enables traceability and forensic analysis through logging and verification mechanisms to determine what occurred and which entities were responsible (Thames; Schaefer, 2017).

2.1.1.6 Privacy

Privacy refers to the ability of entities to control how information about them is collected, processed, and shared, encompassing device and asset privacy to prevent data leakage due to theft or side-channel attacks, storage privacy through data minimization and secure lifecycle management, communication privacy by limiting data exchange to necessary and trusted channels, processing privacy to prevent unauthorized third-party disclosure, identity privacy to restrict identification to authorized entities only, and location privacy to ensure that positional information is disclosed exclusively to approved parties.

Together, these six properties, Confidentiality, Integrity, Availability, Authenticity, Nonrepudiation, and Privacy (CIA+ANP), form the security framework against which the proposed architecture in this thesis is evaluated.

2.1.2 Attacks

Cyberattacks aim to exploit vulnerabilities in computational systems and may be either intentional or unintentional. In both cases, such attacks can compromise fundamental information security properties, namely confidentiality, integrity, and availability. According to Stallings (Stallings, 2017), threats and attacks are primarily intended to interfere with the information flow between a source and its destination. From this perspective, attacks can be classified according to the manner in which they affect communication processes, as follows:

- **Interruption**: an event or action prevents a message or service from reaching its intended recipient, directly impacting system availability.
- **Modification**: a malicious entity intercepts a message, alters its content, and subsequently forwards it to the recipient while impersonating the legitimate sender.
- **Interception**: a malicious entity gains unauthorized access to a message or data transmission without necessarily modifying its content.
- **Fabrication**: a malicious entity forges a message and transmits it to the recipient while posing as a legitimate source.

These attack categories are closely associated with core information security properties, which constitute the foundation for security analysis and system design:

- **Availability**: the assurance that information and services are accessible to legitimate users whenever required.
- **Confidentiality**: the assurance that information is disclosed only to authorized entities.
- **Integrity**: the assurance that information has not been altered in an unauthorized or undetected manner throughout its lifecycle.
- **Authenticity**: the assurance of the identity of the entity responsible for sending or modifying information.

Beyond these conceptual classifications, practical attack models have been systematically documented by the Open Web Application Security Project (OWASP), a global non-profit organization dedicated to improving software security. OWASP periodically publishes the *OWASP Top 10*, which identifies the most prevalent and critical categories of web application vulnerabilities and attack vectors (OWASP Foundation, 2021). The main categories are summarized as follows:

- **Injection**: includes SQL and NoSQL injection vulnerabilities that arise when untrusted

data is incorporated into commands or queries, potentially enabling unauthorized command execution or data access.

- **Broken Authentication**: results from improperly implemented authentication and session management mechanisms, allowing attackers to compromise passwords, keys, or authentication tokens.
- **Sensitive Data Exposure**: occurs when applications fail to adequately protect sensitive data, such as financial or identity information, enabling fraud, identity theft, and related cybercrimes.
- **XML External Entities (XXE)**: exploits legacy XML processors that evaluate external entity references, allowing attackers to disclose internal files, perform internal port scanning, or execute remote code via file URI handlers.
- **Broken Access Control**: results from insufficient authorization enforcement, enabling attackers to access unauthorized functionality or data, such as other users' accounts or confidential records.
- **Security Misconfiguration**: a frequent issue in which insecure default settings, excessive privileges, or verbose error messages expose sensitive system information.
- **Cross-Site Scripting (XSS)**: occurs when applications allow untrusted input to be executed in users' browsers, enabling malicious JavaScript execution and unauthorized actions.
- **Insecure Deserialization**: improper handling of serialized data that can be exploited to perform replay attacks, injection attacks, or privilege escalation.
- **Using Components with Known Vulnerabilities**: arises when applications rely on libraries or frameworks with known security flaws that execute with the same privileges as the application.
- **Insufficient Logging and Monitoring**: inadequate detection and response mechanisms that allow attackers to persist within systems, move laterally, and tamper with, extract, or destroy data without timely detection.

Understanding these attack classifications is essential for organizations and individuals to design and implement effective cybersecurity measures that address a wide range of threats. In practice, adversaries often combine multiple attack techniques, and the cybersecurity landscape continues to evolve as new vulnerabilities and attack vectors emerge.

Consequently, security mechanisms focused on ensuring availability reduce the likelihood of service or communication interruption. Likewise, confidentiality controls mitigate the impact of interception and modification attacks. At the same time, integrity and authenticity mechanisms protect against modification and fabrication by ensuring data trustworthiness and reliable identification of communicating entities. In response to these challenges, security approaches such as encryption and information hiding have been developed to protect digital communication systems, as discussed in the following sections.

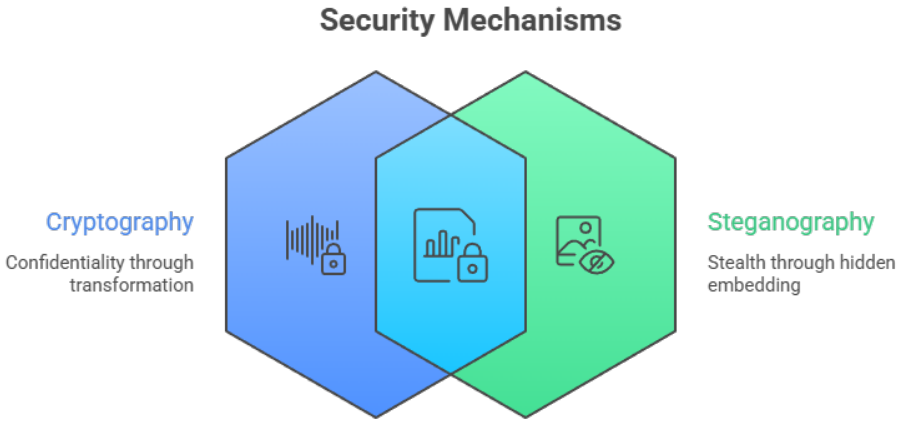


Figure 2.1 – Security Mechanisms

Figure 2.1 illustrates a hierarchical classification of data security mechanisms, encompassing cryptography and steganography. Although both techniques aim to protect sensitive information, they adopt fundamentally different approaches to achieving data security.

Cryptography focuses on transforming plaintext into ciphertext, rendering the information unintelligible to unauthorized parties through encryption algorithms that require specific decryption keys for access (Srikumar; Malarvizhi, 2001). This transformation ensures confidentiality even if the data is intercepted, making cryptography a foundational component of modern secure communication systems.

In contrast, steganography seeks to conceal the existence of the data itself by embedding secret information within innocuous digital media, such as images, audio, video, or text, without perceptibly altering the carrier content (Majeed; Sulaiman, 2015). By concealing the message's presence, steganography provides an additional layer of security, particularly in scenarios where detecting encrypted communication alone may raise suspicion.

The distinction between cryptography and steganography highlights their complementary strengths and inherent limitations. While cryptography emphasizes confidentiality through mathematical transformation, steganography prioritizes stealth and covert communication. Consequently, both approaches are often combined to enhance overall data protection. The following sections review the literature on cryptography and steganography, examining their principles, techniques, and applications in contemporary data security systems.

2.1.3 Cryptography

Cryptography, the art and science of secure communication, was among the earliest methods employed to safeguard sensitive information in military and diplomatic communications. Developing codes and ciphers was critical to maintaining secrecy during the war and political intrigue (Koç, 2009).

Over the centuries, cryptography evolved from simple substitution ciphers, such as the Caesar cipher, to more sophisticated techniques, such as the Vigenère cipher. The field continued to advance with the advent of mechanical devices like the Enigma machine in the early 20th century.

Broadly categorized into symmetric and asymmetric systems, cryptography encompasses various techniques for securing data (Duggan, 2002). In symmetric cryptography, encryption and decryption both use the same key. Asymmetric cryptography, in contrast, employs a pair of keys: a public key for encryption and a private key for decryption. Understanding the strengths and weaknesses of each type is fundamental to selecting appropriate cryptographic methods for specific use cases.

Symmetric key cryptography was the first type of cryptography discovered and, as such, is the most straightforward. It employs a single secret key for both encryption and decryption. Figure 2.2 illustrates the symmetric cryptography process as depicted by (Mushtaq *et al.*, 2017).

The block size for the stream cipher is one character, and it is not more appropriate for software processing due to the key length, as long as the message (Sasi; Sivanandam, 2015). The following steps present the workings of the stream cipher:

1. A single plain text character is combined with a single character from the key stream to produce a single cipher text character;
2. The sender sends the cipher-text character from step 1 to the receiver;
3. Until the sender has sent the entire message, steps 1 and 2 repeat.

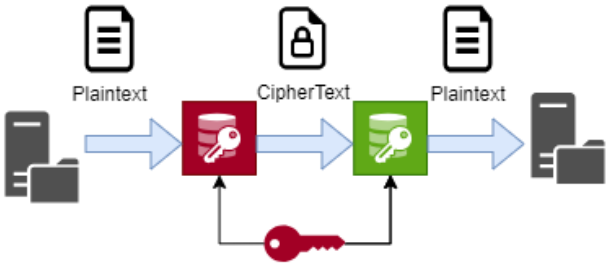


Figure 2.2 – Symmetric key encryption, adapted from (Sasi; Sivanandam, 2015)

Asymmetric key encryption, also known as public-key encryption, uses distinct keys for encryption and decryption. The encryption key, known as the public key, is used to encrypt the message, while the decryption key, the secret or private key, is used to decrypt it. The strength of asymmetric key encryption is leveraged for digital signatures, enhancing the detection of user messages.

Examples of asymmetric encryption algorithms include RSA (Burnett; Paine, 2001), Diffie-Hellman algorithm (Escala *et al.*, 2017), and others. Figure 2.3 illustrates the components of an asymmetric block cipher.

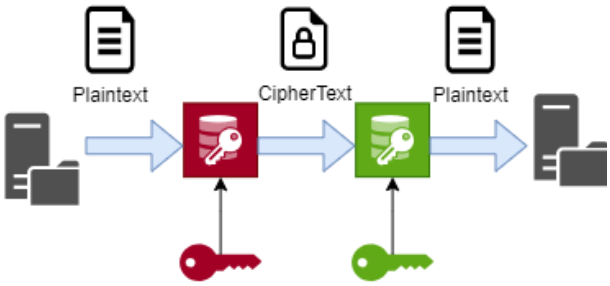


Figure 2.3 – Asymmetric key encryption, adapted from (Sasi; Sivanandam, 2015)

2.1.3.1 Cryptographic Hash Functions

A cryptographic hash function H maps an input of arbitrary length to a fixed-length output (the hash or digest) with several important properties (Stallings, 2017):

- **Determinism**: The same input always produces the same output.
- **Efficiency**: Computing $H(m)$ is computationally efficient for any input m.
- **Pre-image resistance**: Given a hash value h, it is computationally infeasible to find any input m such that $H(m) = h$.
- **Second pre-image resistance**: Given an input m_1, it is computationally infeasible to find a different input m_2 such that $H(m_1) = H(m_2)$.
- **Collision resistance**: It is computationally infeasible to find any two distinct inputs m_1 and m_2 such that $H(m_1) = H(m_2)$.

The SHA-2 family, particularly SHA-256 and SHA-512, is a widely used cryptographic hash functions that satisfy these properties. SHA-256 produces a 256-bit (32-byte) digest, while SHA-512 produces a 512-bit (64-byte) digest. Both are considered secure for current applications and are used extensively in blockchain systems and digital preservation.

Hash functions enable integrity verification: if the hash of a file matches a previously recorded hash, the file has not been modified (with overwhelming probability). This property is fundamental to both blockchain technology and digital preservation standards.

2.2 Steganography

Steganography involves concealing one piece of information within another to transmit a hidden message without the intended recipient's knowledge. Unlike cryptography, which focuses on securing communication through encryption, steganography focuses on hiding the existence of a message. The term "steganography" is derived from the Greek words "steganos," meaning covered or hidden, and "graphie," meaning writing.

Steganography involves various techniques for concealing information, such as embedding data within images, audio, video, or text. The goal is to make the embedded information undetectable to the casual observer.

The carrier medium is the file or communication channel through which the secret message is concealed. Standard carrier media include digital images, audio files, video files, and text documents.

To better understand the Steganography process, Figure 2.4 illustrates the overall workflow of steganography and steganalysis from the perspectives of Alice, Bob, and Eve.

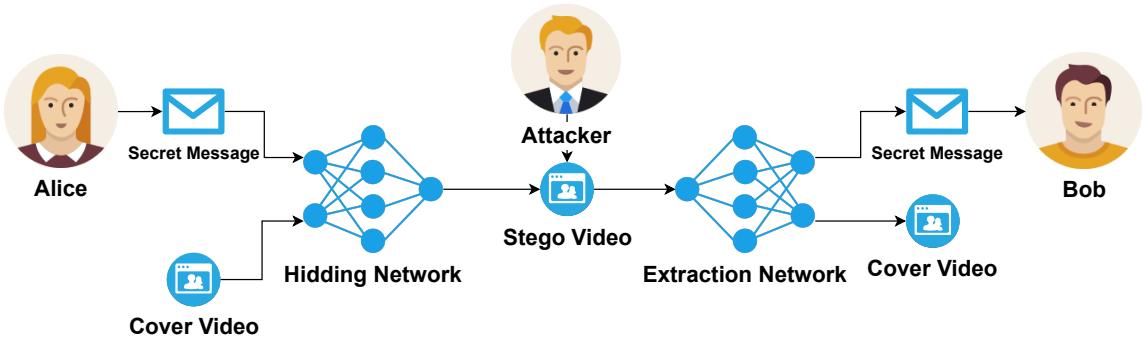


Figure 2.4 – steganography reduced to the prisoner's problem. Adapted from (Kunhoth *et al.*, 2023a)

In Figure 2.4, we see the encrypted message, which could be text, audio, or even another video from Alice, and it is steganographed using a cover video. This video has now been sent to Bob, but the attacker is still attacking the network and intercepting the message. In this scenario, the attacker does not know that the video contains an encrypted message, nor does the attacker know the steganographic method used to embed it. Finally, Bob can decode the message using the same steganography method and the secret key.

Before exploring the detailed attributes of steganography, it's important to note the key factors that define its effectiveness: security, imperceptibility, and capacity (Li *et al.*, 2011). These attributes are essential for effective data concealment, though achieving a balance can be challenging. A larger quantity of hidden data can affect imperceptibility and overall quality (Li *et al.*, 2011). While robustness is crucial for resisting attacks, security and imperceptibility are always necessary. The next section will explain the different types of steganography and how these attributes apply to each method.

2.2.1 Text Steganography

The first approach involves concealing information within text, in which specific patterns or characters may encode hidden data. This method is less common than media-based steganography, and unlike traditional cryptography, which obscures the content of a message, text steganography aims to conceal the message's existence.

Various techniques can be employed to embed information discreetly within the text, such as altering word or line spacing, manipulating punctuation, or encoding data in less noticeable text features, such as font variations. Text steganography is commonly divided into

three classes: format-based, linguistic, and random/statistical generation (Baawi; Mokhtar; Sulaiman, 2018; Majeed *et al.*, 2021).

Among the various forms of text steganography, format-based methods are particularly noteworthy for their ability to hide information within the text's physical structure. These methods involve imperceptible alterations to the visual appearance of text, such as slightly adjusting line spacing, shifting words, or changing the alignment of text elements. The result is a stego-text that appears identical to the original but contains hidden data.

The Linguistic technique involves concealing secret information within the text by exploiting the language's properties. This method aims to embed data in a way that appears natural and semantically coherent to human readers. A linguistic steganography technique was proposed in (Li *et al.*, 2021), utilizing knowledge graphs (KGs) to generate steganographic paragraphs on specific topics. The proposed method not only preserves the quality of the generated steganographic text but also ensures it remains relevant to the specified topic, making it harder for unintended readers to detect anomalies.

Random and statistical generation methods in text steganography exploit the statistical properties of a language to create cover texts that conceal hidden information. These methods rely on the natural occurrence and frequency of words and phrases in a language to generate text that appears normal while embedding secret data. In (Wu *et al.*, 2020), a statistical text steganography technique was proposed based on the Markov Chain model, with particular emphasis on the transition probability, a fundamental concept in this model. The technique developed binary sequence illustrations of state transitions, using these sequences to guide the generation of new texts embedded with hidden information.

In summary, the various methods of text steganography, from format-based to linguistic and statistical approaches, each offer distinct advantages for securely embedding information within text. As research advances, these techniques continue to evolve, balancing the need for security with the requirement to preserve the text's natural appearance.

Text steganography, while useful, faces several challenges that limit its effectiveness. It has a much lower capacity to hide information than images, making it difficult to embed large amounts of data without detection. Changes to text, such as altered spacing or punctuation, are readily detectable, thereby increasing the risk of exposure. Additionally, text steganography methods often depend on specific languages, reducing their universality. The hidden message is also vulnerable to loss or alteration during editing, making text steganography less robust than other forms.

2.2.2 Image Steganography

Concealing information in digital images is a prevalent form of steganography. One technique involves subtly altering pixel color values or manipulating the image's least

significant bits, making the changes invisible to the human eye. In this approach, the least significant bits of the image's pixel values are altered to encode the secret data. Since these bits contribute little to a pixel's overall color, their modification is generally imperceptible to human vision.

Consider an 8-bit color image in which each color channel (red, green, and blue) is represented by 8 bits. The least significant bit (the last bit in the 8-bit sequence) can be changed to store part of a hidden message. For example:

- Original pixel in binary (R, G, B): 11001001, 10101101, 10011110
- Hidden data bits to be embedded: 011
- Modified pixel in binary (R, G, B): 11001000, 10101100, 10011111

In this example, only the last bit of each color channel is altered, making the change imperceptible to the viewer.

Image steganography is a powerful tool for secure communication because it hides data within an image, making it invisible to the human eye, reducing the risk of detection. It offers high capacity, enabling large amounts of information to be embedded without noticeably degrading image quality.

Despite its advantages, image steganography has some challenges. Advances in steganalysis, the practice of detecting hidden information in digital media, have made it increasingly difficult to hide data securely. As a result, steganography methods must continually evolve to remain effective against detection techniques.

2.2.3 Audio Steganography

Similar to image steganography, information is embedded in the audio file without significantly degrading its quality. Audio steganography encompasses several techniques for embedding secret messages in digital audio files, such as MP3 or WAV, by modifying the audio's binary representation. One of the most common methods involves modifying the Least Significant Bits (LSBs) via error diffusion to ensure that the changes are imperceptible to listeners.

The hiding of messages in audio "noise" (and in frequencies that humans can't hear) is another area of information hiding that relies on using an existing source as a space in which to hide information. Audio steganography can be problematic and can facilitate the transmission of covert information within an innocuous cover audio signal (Varghese, 2021).

Other techniques are also employed in audio steganography and described in (Arshad; Siddiqui; Islam, 2024). The parity coding method encodes message bits within the parity bit of a sample region after dividing the audio signal into smaller regions. This allows the hidden message to be embedded without noticeable changes in the sound. The phase coding method encodes message bits in the audio signal's phase spectrum as inaudible phase shifts. Another

approach, spread spectrum, spreads the secret data bits across the frequency spectrum, making detection and extraction of the message more difficult. These methods demonstrate the versatility of audio steganography in securely embedding information within sound files.

Despite its advantages, audio steganography has some challenges. The capacity for hiding data in audio files is generally lower than in images, and the embedding process must be carefully managed to avoid creating noticeable artifacts. Additionally, advances in steganalysis tools pose a growing threat to the security of audio steganography as they become more effective at detecting hidden messages.

2.2.4 Video Steganography

Video steganography conceals data within digital videos through subtle modifications, facilitating secure communication and copyright protection. Techniques like LSB modification and frequency adjustments are standard. Despite detection challenges, it still poses a risk of misuse for covert activities.

Video steganography is a technique for concealing information within digital video files, making it invisible to the human eye and ear. Similar to other forms of steganography, the primary goal is to embed secret data within the video, making it difficult for anyone other than the intended recipient to detect the hidden information (Kunhoth *et al.*, 2023a).

One key aspect of video steganography is the concealment of information within video frames. Each frame embeds hidden data without causing noticeable changes in visual or auditory quality.

Video steganography involves techniques such as modifying pixel color values, altering the least significant bits of pixel values, or manipulating the audio track's frequency components.

LSB modification, a common technique, replaces the least significant bits of pixel values in the video frames with hidden data. Since the subtle changes affect only the least significant bits, they are less likely to be noticeable to human observers.

Video steganography can operate in both spatial and temporal domains. Spatial-domain techniques operate on individual frames, whereas temporal-domain techniques may involve changes over time, such as modifying the timing or ordering of frames.

In frequency-domain steganography, changes occur to the frequency components of audio or video signals. This method may involve hiding information in specific frequency bands or altering the phase of certain elements.

The effectiveness of video steganography relies on the ability to hide information without raising suspicion. Detecting hidden information in videos is challenging and requires specialized tools and techniques for analysis. Security concerns also arise when videos are

shared or transmitted, as hidden information may be exposed during compression or other transformations.

Table 2.1 summarizes steganography methods, including their benefits and challenges. This comparison highlights the strengths and limitations of each technique to aid in understanding their practical applications.

Steganography Method	**Benefits**	**Challenges**
Image	Invisibility: Hidden data is imperceptible to the eye. High Capacity: Large images can hide significant amounts of data. Versatility: Applicable across various image formats.	Limited Capacity: Smaller files hold less data. Vulnerable to Compression: Data can be lost when compressed. Easier Detection: Advanced tools can detect hidden data.
Text	Simplicity: Easy to implement in simple text. Low Suspicion: Hidden within normal text, raising little suspicion. Language-Specific: Can use linguistic techniques for hiding data.	Limited Capacity: A small amount of data can be hidden. Easy Detection: Text changes are readily detectable. Vulnerable to Editing: Any text change can alter or destroy hidden data.
Audio	Inaudibility: Changes are generally imperceptible to listeners. Variety of Methods: Multiple techniques (LSB, phase coding, etc.). Digital Rights Management: Useful for watermarking audio content.	Low Capacity: Limited data can be embedded in audio. Quality Risk: Incorrect embedding can introduce audio artifacts. Steganalysis Tools: Advanced tools can detect hidden data.
Video	High Capacity: Large video files can hide vast amounts of data. Invisibility: Hidden data is not noticeable within video frames. Multiplexing: Can be combined with audio steganography to enhance security.	Complexity: Embedding in video requires advanced techniques. Processing Power: Requires significant computational resources.

Table 2.1 – Steganography Methods, Benefits, and Challenges

As steganography methods continue to evolve, the integration of artificial intelligence (AI) has become increasingly prominent. AI, particularly machine learning and deep learning techniques, is now playing a crucial role in enhancing the effectiveness and efficiency of steganographic systems. By leveraging AI, we can develop more sophisticated algorithms that enhance data-hiding capacity, robustness, and imperceptibility across various steganog-

raphy methods. In the following section, we examine how AI techniques are applied in steganography, including their impact, particularly on video steganography.

2.3 Artificial Intelligence

The study of artificial intelligence (AI) involves exploring methods to enable computers to perform tasks traditionally exclusive to humans (Huang *et al.*, 2019b). Over recent years, AI has undergone rapid development, significantly impacting people's lifestyles (Huang *et al.*, 2019a).

Recognizing its importance, countries worldwide have adopted AI as a crucial development strategy to enhance national competitiveness and ensure security (Rajkomar *et al.*, 2018). Many nations have implemented preferential policies, intensifying the focus on critical technologies and talents to take the lead in the new era of international competition. AI has become a research hotspot in science and technology, with major companies such as Google, Microsoft, and IBM dedicated to advancing and applying it across diverse fields (Shi *et al.*, 2007).

2.3.1 Machine Learning

The fundamental concept behind machine learning is the use of algorithms that improve performance by learning from data (Nilsson, 1982). Machine learning addresses four pivotal types of problems: prediction, clustering, classification, and dimensionality reduction (Erhan *et al.*, 2010). In terms of learning methods, machine learning can be classified into four main types: supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning (Bose, 2017).

Supervised learning is a machine learning paradigm in which a model is trained on a labeled dataset. Each input data point is associated with a corresponding output label. During training, the algorithm learns the relationship between input features and target outputs, adjusting its internal parameters to minimize the difference between predicted and actual outcomes. Tasks such as classification and regression commonly utilize this approach.

In unsupervised learning, the algorithm operates on datasets without labeled outputs. The objective is to uncover inherent patterns, relationships, or clusters within the data without explicit guidance. The algorithm identifies structures and regularities without predefined labels, making it suitable for clustering, dimensionality reduction, and generative modeling applications.

Semi-supervised learning lies between supervised and unsupervised learning. It involves a dataset comprising both labeled and unlabeled examples. The algorithm uses labeled data for supervised learning, leveraging the learned knowledge to generalize patterns from

unlabeled data. This approach is practical when obtaining labeled data, which is resource-intensive.

Reinforcement learning revolves around an agent interacting with an environment. The agent takes action and receives feedback through rewards or penalties. The agent aims to learn a policy that maximizes cumulative rewards over time. This approach finds application in scenarios where decision-making, sequential actions, and dynamic environments are critical, such as robotics, game-playing, and autonomous systems.

2.3.2 Convolutional Neural Networks

Deep learning is an artificial intelligence approach that employs neural networks to accomplish specific tasks or goals (Gallant, 1993). Neural networks, inspired by the human brain, consist of interconnected nodes, known as perceptrons, which communicate. The structure is aptly named "neural networks." While the implementation and architecture of neural networks can vary, the foundational concept remains consistent. Neural networks typically include multiple layers, with the first layer receiving the input and the last layer producing the output.

Between the input and output layers are the hidden layers, the number of which varies by application and domain. Each perceptron within the network has one or more weighted inputs utilized with an activation function to generate an output variable (Nayak *et al.*, 2020).

Convolutional networks, specialized neural networks, are designed to effectively process two-dimensional data, such as images and media (Brownlee, 2020). As depicted in Figure 2.5, a Convolutional Neural Network (CNN) is structured with several convolutional layers that analyze two-dimensional input, extracting crucial details like edges and corners. The activation function determines the significance of information, distinguishing between important and unimportant details.

The pooling layer then receives this information, reducing the spatial size of the convolution while summarizing and retaining essential information. This reduction minimizes the computational power required for data processing, as there is less information to handle. Various pooling functions, including average and max pooling, contribute to this optimization process.

2.3.3 Attention Mechanisms

An attention mechanism, as the name suggests, enables the network to focus on specific regions of the input. Previously, an attention mechanism was developed for Natural Language Processing (NLP) to process human language of any type. The machine must respect the specificities of the analyzed language to ensure accurate interpretation. Suppose a computer program takes a certain text in natural language exactly literally and ignores such nuances.

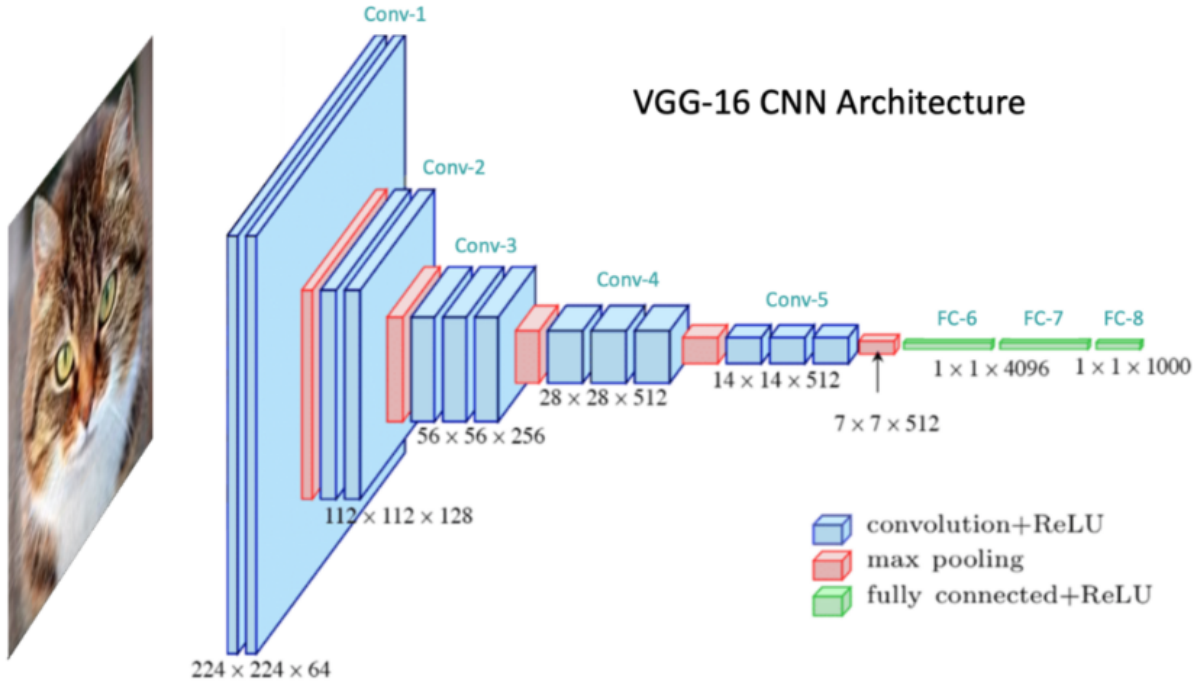


Figure 2.5 – CNN network. Adapted from (Khalifa; Guzman, 2022)

In that case, it is possible that the written (or spoken) content will not be understood or well interpreted. There is also the issue of ambiguity; human languages change over time and are influenced by different means, such as contact with other languages or the advancement of the internet. For these reasons, NLP requires the AI to function correctly and thus enables the RNN (Zhang *et al.*, 2016; Zaremba; Sutskever; Vinyals, 2014) to handle better both the length (Cho *et al.*, 2014) and the content (Vaswani *et al.*, 2017) of the text being analyzed.

The most widely noted mechanism developed was the Transformers Mechanism (Vaswani *et al.*, 2017), introduced in 2017. With the development of the technique, the attention concept was extended to images, enabling the mapping of the most critical areas of an image as a consequence of attention. Next, the two attention strategies used in this work will be presented.

2.3.3.1 Non-local Attention

First introduced a non-local Algorithm for image denoising (Buades; Coll; Morel, 2005), which assigns a weight to the average of a pixel's value across its neighbors in the grey-scale domain using a Gaussian kernel. Hence, the algorithm can capture anomalies (peaks) in the image and classify them as noise. Furthermore, in the middle of the calculation, the distance between two pixels is considered, which gives the advantage that edges will not be affected, given that an edge is, in simple terms, an abrupt change in the pixel value, but following the pattern that neighbor pixels will also differ, making it possible to differentiate it from random noise.

As the technique developed, (Wang *et al.*, 2018) adapted it for use as an attention mechanism. Once the pixel values are averaged, simplicity is equivalent to comparing their contextual meaning. Moreover, as a consequence of the relation between the pixel and its neighbors, when the temporal aspect, as done in (Wang *et al.*, 2018), is considered, not only is it possible to capture the spatial features of the image but also its temporal importance, i.e., pixels that change their value will have different weights from their statical neighbors.

These characteristics are essential for video, as they enable highlighting a specific area of the video that is important to the network, in both spatial and temporal dimensions. In addition, (Wang *et al.*, 2018) demonstrated that the algorithm also works with the dot product instead of Gaussian Kernels (and its embedded variation). In other terms, (Wang *et al.*, 2018) developed a self-attention (Vaswani *et al.*, 2017; Bravo-Ortiz *et al.*, 2024) mechanism that takes the spatial-time relation to the feature map.

Therefore, this layer works in conjunction with the Non-local algorithm. Attention is applied to the frames, not to a single task (e.g., finding the most relevant regions), but to guide how the regions highlighted by the Non-local algorithm influence the frames.

2.3.3.2 Convolutional Block Attention Module - CBAM

Proposed by (Woo *et al.*, 2018), CBAM, which stands for "Convolutional Block Attention Module, was created to serve as a cheap and efficient attention block to capture essential features in the spatial dimension and between channels (Li *et al.*, 2023; Sabeena; Abraham, 2024).

CBAM is structured around two key components: the Channel Attention Module and the Spatial Attention Module. These modules are applied sequentially, allowing the network first to determine which features (or channels) are important and then identify their locations within the input's spatial dimensions. This sequential approach is crucial because it enables the model to progressively refine its focus, first at the channel level and then at the spatial level, resulting in a more precise and meaningful interpretation of the input data.

The Channel Attention Module models inter-channel relationships in the feature map. Each channel in a CNN feature map can be viewed as a distinct detector that responds to different features in the input image. By attending to these channels, the network can amplify or suppress specific ones depending on their relevance. To achieve this, the Channel Attention Module applies both average pooling and max pooling across the feature map's spatial dimensions, yielding a compact representation for each channel. These representations are then passed through a shared multi-layer perceptron (MLP) to produce a channel attention map, which is applied back to the feature map to emphasize important channels.

After refining the features through the Channel Attention Module, the Spatial Attention Module is applied. This module identifies the most important locations within the feature map, corresponding to the "where" aspect of attention. The Spatial Attention Module first

aggregates the feature map by pooling across the channel dimension, generating two spatial descriptors through average pooling and max pooling. These descriptors are concatenated and convolved to produce a spatial attention map that highlights key regions of the feature map. This spatial refinement helps the network better capture the input data's spatial context, which is essential for tasks like object detection, where the precise location of objects is critical.

The two modules are formalised as follows. The Channel Attention Module produces an attention map $\mathbf{M}_c \in \mathbb{R}^{C\times 1\times 1}$ that rescales each channel of the input feature map $\mathbf{F} \in \mathbb{R}^{C\times H\times W}$, where C is the number of channels and $H \times W$ the spatial dimensions. Both average pooling and max pooling are applied across the spatial dimensions to produce two channel descriptors, $\mathbf{F}^c_{\text{avg}} \in \mathbb{R}^{C\times 1\times 1}$ and $\mathbf{F}^c_{\text{max}} \in \mathbb{R}^{C\times 1\times 1}$. These descriptors are fed through a shared multi-layer perceptron (MLP) with one hidden layer, and the outputs are summed before applying the sigmoid activation σ:

$$\mathbf{M}_c(\mathbf{F}) = \sigma\Big(\text{MLP}\big(\text{AvgPool}(\mathbf{F})\big) + \text{MLP}\big(\text{MaxPool}(\mathbf{F})\big)\Big) \tag{2.1}$$

where $\text{MLP}(\cdot)$ denotes a two-layer perceptron $\mathbf{W}_1(\text{ReLU}(\mathbf{W}_0(\cdot)))$ with weight matrices $\mathbf{W}_0 \in \mathbb{R}^{C/r\times C}$ and $\mathbf{W}_1 \in \mathbb{R}^{C\times C/r}$, and r is the reduction ratio that controls the bottleneck dimension.

The Spatial Attention Module produces an attention map $\mathbf{M}_s \in \mathbb{R}^{1\times H\times W}$ that highlights the most informative spatial locations. It operates on the channel-refined feature map by applying average pooling and max pooling along the channel axis, producing two spatial descriptors $\mathbf{F}^s_{\text{avg}} \in \mathbb{R}^{1\times H\times W}$ and $\mathbf{F}^s_{\text{max}} \in \mathbb{R}^{1\times H\times W}$. These are concatenated and convolved with a 7×7 filter f to produce the spatial attention map:

$$\mathbf{M}_s(\mathbf{F}) = \sigma\Big(f^{7\times 7}\big([\text{AvgPool}(\mathbf{F});\ \text{MaxPool}(\mathbf{F})]\big)\Big) \tag{2.2}$$

where $[\,\cdot\,;\,\cdot\,]$ denotes concatenation along the channel dimension and $f^{7\times 7}$ is a convolutional layer with a 7×7 kernel that reduces the two-channel input to a single-channel map.

The two modules are applied sequentially. Given an input feature map $\mathbf{F}$, the channel attention is applied first, producing an intermediate refined feature map $\mathbf{F}'$. The spatial attention is then applied to $\mathbf{F}'$, yielding the final output $\mathbf{F}''$:

$$\mathbf{F}' = \mathbf{M}_c(\mathbf{F}) \otimes \mathbf{F} \tag{2.3}$$

$$\mathbf{F}'' = \mathbf{M}_s(\mathbf{F}') \otimes \mathbf{F}' \tag{2.4}$$

where $\otimes$ denotes element-wise multiplication with broadcasting. In Equation 2.3, the channel attention map $\mathbf{M}_c \in \mathbb{R}^{C\times 1\times 1}$ is broadcast across the spatial dimensions, scaling each channel

by its learned importance weight. In Equation 2.4, the spatial attention map $\mathbf{M}_s \in \mathbb{R}^{1\times H\times W}$ is broadcast across channels, amplifying spatially relevant regions while suppressing less informative ones. The sequential application ensures that the network first selects *what* to attend to (channels) and then *where* to attend (spatial locations).

The impact of CBAM has been demonstrated through extensive experiments on benchmark datasets such as ImageNet-1K, MS COCO, and VOC 2007. In these experiments, CBAM consistently improved the accuracy of CNNs across different architectures, such as ResNet, WideResNet, and ResNeXt. For instance, when integrated with ResNet-50, CBAM achieved a notable reduction in the top-1 error rate compared with the baseline model, highlighting its effectiveness in improving performance without introducing significant computational overhead. This improvement is attributed to CBAM's ability to direct the model's attention to the most relevant features while filtering out noise and less important information.

2.3.4 DWT - Discrete Wavelet Transform

The Discrete Wavelet Transform (DWT) is a technique for decomposing a signal into frequency bands. When using it on images (via 2D-DWT), the extracted frequencies represent different aspects of the images, such as edges, texture, and details (as well as noise) (Singh; Dave; Mohan, 2016). The extracted frequencies are split into four groups: LL, LH, HL, and HH. The former contains the most expressive part of the image, whereas the latter contains fine details (and also noise). The middle ones focus on extracting horizontal and vertical edges. Due to these transformation characteristics, it was already used extensively for steganography (Dalal; Juneja, 2023; Suresh; Sam, 2020; Singh; Dave; Mohan, 2016; Zear; Singh; Kumar, 2018).

While machine learning has revolutionized data-driven decision-making by uncovering patterns and insights from vast datasets, ensuring the integrity and security of these processes, blockchain technology offers a decentralized, tamper-proof framework that complements the predictive power of AI. In the next section, we examine how blockchain's immutable ledger and consensus mechanisms can further enhance the reliability and transparency of intelligent systems.

2.4 Blockchain Technology

Recently, with the advent and popularization of cryptocurrencies, the underlying technology, blockchain, has been widely adopted across domains for its decentralized architecture and its capacity to ensure information integrity. Blockchain is a chain of blocks, in which each block permanently stores information about a transaction and is linked sequentially. Inserting information into the blockchain involves creating and adding a new block. Once a block exists, alteration is impossible.

This process relies on cryptographic algorithms to ensure the authenticity and integrity of the blockchain (Nakamoto, 2008). Regarding Blockchains, Satoshi Nakamoto introduced the concept of decentralized, peer-to-peer transactions based on cryptography. This whitepaper outlines a method to guarantee the authenticity and integrity of transactions, overcoming the issue of "double spending," which is the act of using the same digital currency for the payment of more than one item, and the problem of transactions needing to pass through a financial institution, giving rise to the currency called Bitcoin.

The concept of blockchain is a chain of immutable, distributed, and public records. The records remain immutable because the linking performed with blocks uses cryptographic digests according to specific rules to form pointers to the previous block. Significant computational power is required to alter this digest, making it impractical. As a distributed protocol, all information is not centralized on a single server, and no controller node coordinates the network. Instead, the blockchain is replicated across all participating nodes. Besides being a distributed scheme, it is also public, as there is no way to censor participation in the network; anyone with internet access can create a copy of the database (Underwood, 2016).

The network can validate blocks using cryptography. In addition to transactions, each block contains a timestamp, the hash value of the previous block ("parent"), and a nonce, a random number used to verify the hash. This concept ensures the integrity of the entire blockchain up to the first block ("genesis block") (see Figure 2.6).

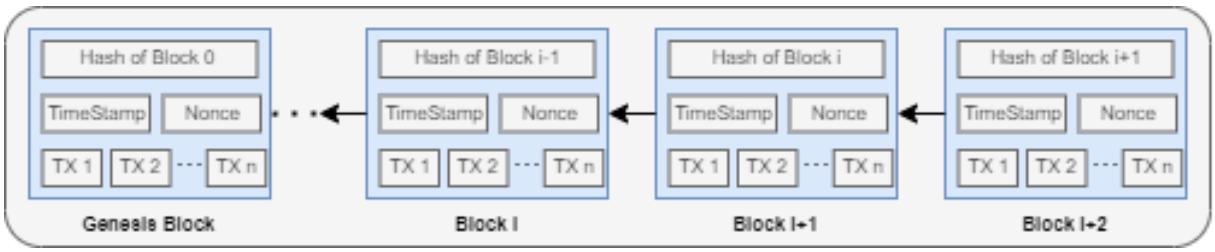


Figure 2.6 – Description of a typical structure of a blockchain. Adapted from: (Underwood, 2016)

Hash values are unique, and fraud prevention is effective because any changes to a block in the chain would immediately alter its hash value. Most network nodes agree, through a consensus mechanism, on the validity of transactions in a block, and the block itself allows it to be added to the chain.

2.4.1 Consensus Mechanisms

A key challenge in distributed systems is achieving consensus, ensuring that all participants agree on the ledger's current state. Blockchain systems use various consensus mechanisms (Liu William Yeoh; Choo, 2022).

Proof of Work (PoW) requires participants to solve a computationally difficult puzzle to add a block. The puzzle typically involves finding a nonce value that, when combined

with the block contents and hashed, yields a result that satisfies certain criteria (e.g., starting with a specified number of zeros). The computational cost of finding valid nonces makes it expensive to create fraudulent blocks.

Proof of Stake (PoS) selects block creators based on their stake (ownership) in the system. Participants with a greater stake are more likely to be selected to create blocks, and they risk losing their stake if they behave maliciously.

Practical Byzantine Fault Tolerance (PBFT) achieves consensus through voting among a known set of participants. PBFT can tolerate up to one-third of participants being faulty or malicious, but requires that participants be identifiable.

For document management applications, where participants are typically known, and the transaction rate is modest, simpler consensus mechanisms, such as single-node PoW or PBFT, among institutional partners are often appropriate.

2.4.2 Cybersecurity and Blockchain

Liu et al. (Liu William Yeoh; Choo, 2022) conducted a systematic review and identified key benefits of blockchain technology in the context of cybersecurity, including tamper-proofing of data, resistance to DDoS attacks, elimination of a single point of failure, robust disaster recovery capabilities, enhanced privacy protection, secure identity management, and effective access management.

Tamper-proofing Data means that any transactional information stored within its network is highly resistant to tampering. This tamper-proofing is achieved through a unique data structure where each block is cryptographically linked to the previous one. Once a block is added to the blockchain, modifying its contents becomes nearly impossible without modifying every subsequent block, which requires enormous computational power.

Resistance to DDoS Attacks is achieved by the decentralized nature of blockchain technology, which significantly enhances its resistance to Distributed Denial of Service (DDoS) attacks. Unlike traditional centralized systems, in which a single server or node can be targeted to disrupt the entire network, blockchain operates across a vast network of nodes. Each node holds a complete copy of the blockchain, making it exceedingly difficult for an attacker to amass the computational resources necessary to overwhelm the entire network. As a result, blockchain networks are more resilient to attempts at being taken offline through large-scale cyberattacks.

Similar to the resistance to DDoS attacks, the Blockchain doesn't have a Single Point of Failure because its decentralized architecture ensures that each node holds a full copy of the blockchain, so the failure of one node does not compromise the entire system. This redundancy not only increases the network's reliability and fault tolerance but also enhances its security, as attackers cannot target a single entity to take the network down.

Disaster Recovery is another benefit of blockchain, in which every user in a blockchain network can generate and retain a full copy of the blockchain data. This means that in the event of a disaster, such as data corruption or loss, the entire dataset can be restored from any network node.

Privacy protection in blockchain systems is achieved through cryptographic techniques that decouple a user's public key from their real identity. This ensures that, although transactions are transparent and verifiable to all network participants, the user's actual identity remains concealed.

In a blockchain-based system, users retain full control over their identity information while maintaining anonymity within the network. This decentralized identity management system enables users to own and manage their digital identities without relying on a central authority, thereby reducing risks associated with centralized systems, such as data breaches and identity theft. This approach also enhances privacy and security, making blockchain an attractive solution for applications requiring robust identity verification.

Finally, the last Blockchain benefit is access management, implemented using cryptographic keys to ensure that only authorized individuals can access specific data or services. The decentralized nature of the blockchain also enables real-time tracking and verification of access attempts, making such attempts easier to detect and prevent.

2.4.3 DLT - Distributed Ledger Technology

Distributed Ledger Technology (DLT) is a decentralized database maintained across several locations or among multiple participants. This technology enables secure and transparent recording of transactions across a network of computers, providing a shared, synchronized database.

2.5 Digital Documents and Chain of Custody

Archival documents, whether analog or digital, are crucial sources of information and a means of fulfilling government transparency initiatives. They must be stored and preserved with criteria to ensure reliability, authenticity, and accessibility.

The chain of custody involves ensuring the proper collection of the digital document and verifying its integrity throughout collection, transportation/storage, consultation, and delivery. Like any digital document, whether physical or digital, it must follow four premises to be reliable: traceability, verifiability, authenticity, and security. The Open Archival Information System (OAIS) (Santos; Flores, 2020) has emerged to ensure that the chain of custody meets all the requirements mentioned.

Today, the chain of custody manages a digital document from its collection through its presentation. During this process, several people typically handle the digital document, logging it out and in and physically signing forms to complete it. There are many opportunities to taint the digital document and, more importantly, to have defense attorneys claim it has been tampered with.

B-CoC, a Blockchain-based Chain of Custody for digital documents, outlines critical requirements for an effective process. These include ensuring the integrity of documents during transfer, tracing documents from collection to destruction, authenticating all entities involved, verifying the entire process, and maintaining security against alterations. These elements collectively contribute to a reliable and trustworthy Chain of Custody process for digital documents.

2.5.1 Trusted Digital Archival Repository

A digital archival repository is the custodian of archival documents throughout their various life cycles, including current, intermediate, and permanent phases. Ensuring the authenticity, preservation, and accessibility of digital materials is crucial for a trustworthy digital repository. As delineated in the report Trusted Digital Repositories: attributes and responsibilities (RLG-OCLC, 2002), such repositories should fulfill several vital criteria.

Firstly, they must accept responsibility for maintaining digital materials on behalf of depositors. Additionally, their organizational structure should support long-term viability and oversight of the digital materials they manage. Economic sustainability and administrative transparency are essential aspects that trustworthy digital repositories should demonstrate. Furthermore, adherence to commonly accepted conventions and standards in system design ensures continuous management, access, and security of deposited materials.

Establishing methodologies for evaluating systems should take into account the community's reliability expectations. Treating depositors and users openly and transparently fulfills long-term responsibilities. Policies, practices, and performance should be auditable and measurable. Lastly, factors related to organizational and curatorial responsibilities should be considered, including the scope of deposited materials, lifecycle management, engagement with partners, legal ownership issues, and financial implications.

One method to validate the reliability of a digital repository within the community is through third-party certification. In collaboration with NARA, the document "Trustworthy Repository Audit and Certification: The Criteria and Checklist" (TRAC), published in 2007, aimed to achieve this goal (RLG-OCLC, 2002).

2.5.2 OAIS Model

The OAIS model originated from a collaborative effort between the Consultative Committee for Space Data Systems (CCSDS) at NASA and the International Organization for Standardization (ISO). Initially, its primary goal was to establish standards governing the long-term storage of digital information produced in the context of space missions (Ferreira, 2006).

The initial findings of this study were first published in 1999, followed by a further publication in 2002. By 2003, it had evolved into ISO standard 14721:2003 (ISO-14721, 2013). OAIS serves as a conceptual reference model to identify the functional components integral to an OAIS dedicated to digital preservation. It outlines the system's interface characteristics and describes the information objects it will manage.

Following this model, a repository's functionalities fall into six major categories: Ingest, Storage, Data Management, Preservation Planning, Administration, and Access.

According to Rocha (Rocha, 2015), the OAIS model envisions the creation of conceptual containers known as packages. These packages store content information (the document itself), preservation description information (metadata necessary for long-term preservation and document access), and descriptive information about the package – descriptive metadata enabling the package's location in the repository. The metadata adheres to archival description standards such as ISAD (g): General International Standard Archival Description (ISAD-G, 1999).

Producers submit an information package (SIP) containing documents and descriptive information to the ingest entity. After acceptance and the addition of descriptive information, the SIP is transformed into an information package for storage (AIP) and a storage package for permanently valuable documents. After the AIP is stored in metadata management entities and file repositories, a dissemination information package (DIP) can be generated.

Figure 2.7 illustrates the conceptual model of the OAIS, with the functionalities, agents, and information packages.

2.5.3 BagIt Packaging Format

The BagIt format is a tool for digital preservation, providing a standardized approach to package and organize digital content for storage and transfer. Defined by the Library of Congress and specified in RFC 8493 (Kunze *et al.*, 2018), the BagIt format encapsulates a directory containing digital files and metadata, establishing a comprehensive and coherent structure.

A BagIt package (or "bag") is a directory with a defined structure:

bagit.txt: A declaration file identifying the directory as a bag and specifying the BagIt version and character encoding.

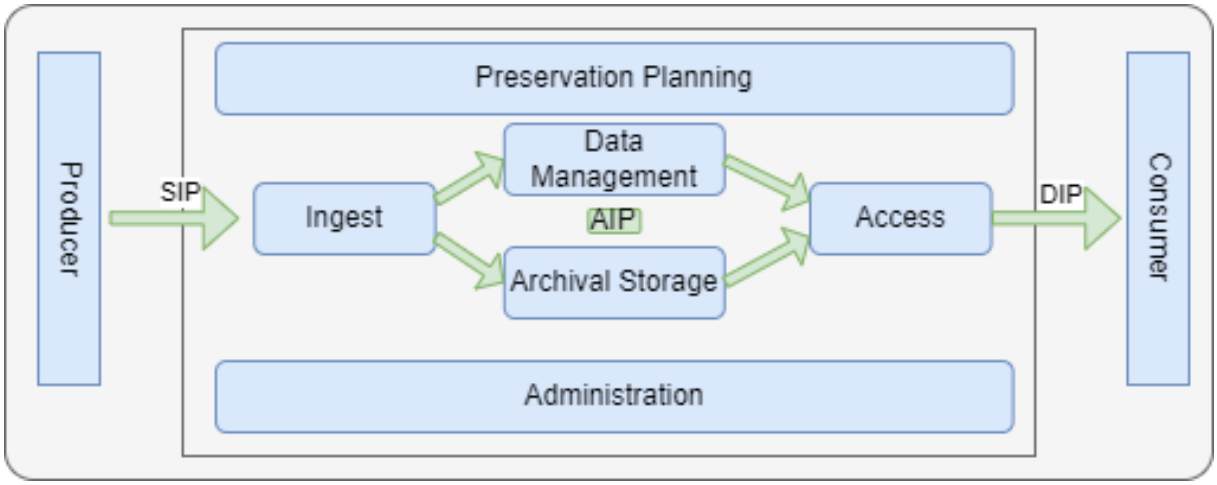


Figure 2.7 – OAIS Model, adapted from (ISO-14721, 2013)

bag-info.txt: An optional file containing metadata about the bag, such as the source organisation, contact information, creation date, and description.

data/: A directory containing the payload—the actual content files being packaged.

manifest-{algorithm}.txt: Files listing the payload files with their cryptographic hashes. Multiple manifests, using different algorithms (e.g., SHA-256 and SHA-512), can be included for redundancy.

tagmanifest-{algorithm}.txt: Files listing the tag files (bagit.txt, bag-info.txt, manifests) with their hashes, providing integrity verification for the package metadata itself.

BagIt's key strength is its simplicity. Bags can be created and validated using standard filesystem tools, require no special software beyond a hash utility, and work equally well for small and large collections. The manifest system provides comprehensive integrity verification: any modification to any file will produce a different hash, immediately revealing tampering or corruption. This format is particularly relevant for ensuring the integrity and authenticity of digital content throughout various processes, including storage, transfer, and archival endeavors. Its widespread adoption within the digital preservation community underscores its significance as a reliable, standardized means of ensuring the long-term preservation of digital materials.

2.5.4 PREMIS Preservation Metadata

PREMIS (Preservation Metadata Implementation Strategies) is a standard for preservation metadata, maintained by the Library of Congress (Library of Congress, 2015). PREMIS defines the metadata necessary to support the long-term preservation of digital objects.

The PREMIS data model comprises four entity types:

Objects are the digital content being preserved. PREMIS distinguishes between intellectual entities (abstract content), representations (sets of files embodying an intellectual

entity), files, and bitstreams.

Events are actions that affect objects. Every significant action—creation, modification, migration, validation, access—is recorded as an event with information about what happened, when, who or what performed the action, and what the outcome was.

Agents are entities (people, organisations, or software) that perform events or have rights over objects.

Rights document the permissions and restrictions governing actions on objects.

PREMIS provides the vocabulary for documenting preservation actions, enabling archives to maintain a complete audit trail of all activities affecting preserved content. This audit trail is essential for demonstrating the authenticity and integrity of preserved materials.

2.6 Bibliometric Analysis

This section provides a detailed exposition of the bibliometric studies conducted, encompassing the complex domain of cybersecurity, including steganography, blockchain, chain of custody, and trusted digital archival repositories. The following account describes the theoretical review process and the methodological techniques employed in the study.

A general systematic review was conducted, following the methodology in (Bispo *et al.*, 2024) and adapted to the current study context. The proposed framework efficiently maps the literature by integrating data extraction from the Web of Science and Scopus with relevance calculations.

Relevance is calculated by integrating across all papers for all queries. This process involves applying specific metrics, such as citation frequency, journal quality, and article timeliness. The combination of these parameters results in a relevance score for each article.

The final stage of the framework involves selecting the most relevant articles based on the calculated relevance scores. Predefined criteria and a specific threshold are established to identify the finest papers in each category.

Figure 2.8 illustrates the integrated flow of the proposed framework. It begins with data collection from the Web of Science and Scopus, followed by data preparation using Python. After preparation, the articles were preprocessed to remove duplicates and papers considered irrelevant to this study. Subsequently, a series of analyses was performed, including selecting the most relevant articles based on metrics such as citation frequency, journal quality, and article timeliness, along with data visualizations. This systematic approach optimizes the identification of significant works in systematic literature reviews.

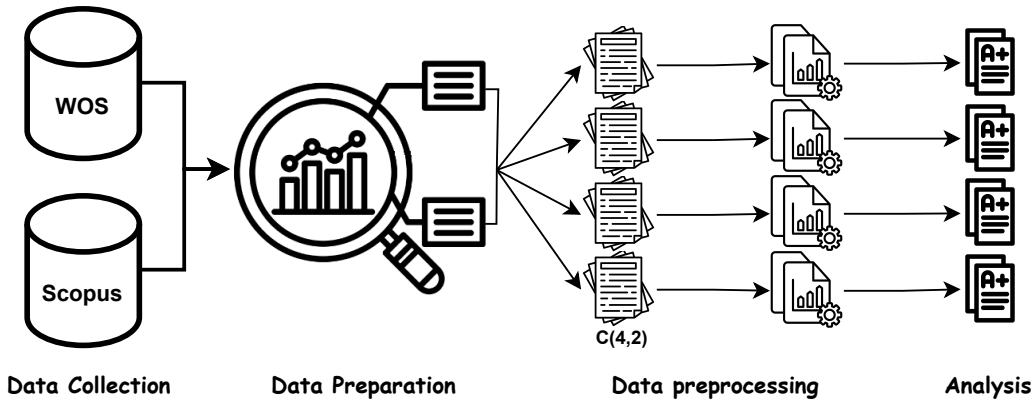


Figure 2.8 – Integrated Framework for Systematic Literature Review. Adapted from (Bispo *et al.*, 2024)

The bibliometric analysis was updated in 2025 to include publications from 2020 to 2026, ensuring that the latest advances in each topic area are reflected in the literature.

Following those steps, this work addresses four main topics related to cybersecurity: steganography, blockchain, chain of custody, and trusted digital archival repositories.

In the first step, several research questions were formulated for each topic. For steganography, the question is: How can it support evidence protection? What are the techniques for hiding information inside another file? Can Artificial Intelligence improve the existing methods? Is video steganography a more effective way to conceal information? For blockchain, the questions are: Can blockchain transport files? Can blockchain increase the security for transmitting those files? Regarding the chain of custody, the question is whether blockchain can be used to manage documents and integrate them into the chain. How can the chain of custody of documents be more secure using blockchain? And for trusted digital archival repositories: How could they be integrated into the chain of custody of documents using blockchain to enhance the security of the files they store?

For the second step, a search was conducted in the Web of Science and Scopus databases using a set of keyword combinations to compile a database of articles.

The first query attempt used all keywords combined with AND to find papers that address all topics simultaneously. No paper was found in WOS or SCOPUS, indicating a gap to be explored in the present research.

A second approach added OR operators within each subtopic, resulting in the following queries:

WOS (`TS=("cybersec*" AND (stega* OR *blockchain OR "chain of custody of document*" OR "trusted digital archival repositor*" OR (trust* AND reposito*))))`, which returned 3,057 results.

For SCOPUS (`cybersec* AND (stega* OR blockchain OR (chain AND of AND custody) OR (archival AND repositories) OR (trusted AND digital))`), 7,571 results were found.

Using a combination of relevant keywords and Boolean operators ensured comprehensive coverage of the literature. The five topic areas: Cybersecurity, Steganography, Blockchain, Chain of Custody, and Trusted Digital Archival Repositories, were combined in all pairwise combinations ($C_{5,2}$), yielding 10 combinations. Establishing criteria enabled selection of only studies published from 2020 to 2026 (approximately 6 years), providing a better understanding of the data distribution and capturing the most recent advances in each area.

Analyzing the preliminary results, many articles combine cybersecurity and blockchain, but few combine cybersecurity with steganography or secure repositories. Even fewer results emerge when combining two of these more specific topics, and none when all terms are searched together.

The third step of the review involved conducting a bibliometric analysis using the Bibliometrix Python and R packages. Various bibliometric indicators were computed, including citation counts, co-authorship networks, keyword co-occurrence analysis, and publication

Query	QTD
cybersec* AND blockchain	7015
cybersec* AND stega*	1188
stega* AND blockchain	78
blockchain AND trusted digital archival repositor* OR trust* AND reposito*	49
cybersec* AND trusted digital archival repositor* OR trust* AND reposito*	13
blockchain AND chain * custody * document*	8
cybersec* AND chain * custody * document*	2
chain * custody * document* AND trusted digital archival repositor* OR trust* AND reposito*	0
stega* AND chain * custody * document*	0
stega* AND trusted digital archival repositor* OR trust* AND reposito*	0

Table 2.2 – List of Keywords

trends. Descriptive statistics and visualization techniques effectively presented the findings.

2.6.1 Database Analysis

Firstly, it analyzed the basic information about the bibliographic dataset (see Table 2.3), including the number of documents, authors, and journals, as well as the distribution of publications over time.

Description	Results
MAIN INFORMATION ABOUT DATA	
Timespan	2020:2026
Sources (Journals, Books, etc)	2698
Documents	8.353
Annual Growth Rate %	-21.09
Document Average Age	2.14
Average citations per doc	11.74
DOCUMENT CONTENTS	
Keywords Plus (ID)	5055
Author's Keywords (DE)	20845
AUTHORS	
Authors	24282
Authors of single-authored docs	537
AUTHORS COLLABORATION	
Single-authored docs	615
Co-Authors per Doc	4.12
International co-authorships %	38.05
DOCUMENT TYPES	
article	6491
article; book chapter	24
article; data paper	7
article; early access	403
article; early access; retracted publication	1
article; proceedings paper	37
article; retracted publication	46
proceedings paper	1344

Table 2.3 – Database Main Information

Analyzing the most frequent words, three stood out as outliers: *blockchain*, *Data*, and *security*; others are less frequent but also have considerable values, including *Technology*, *System*, *Based*, and *Network*.

The annual growth rate of publications increased in 2021, remained stable in 2022 and

2023, and increased again in 2024 and 2025, reflecting the continued maturation of research in blockchain, steganography, and cybersecurity. Including publications from 2025 and early 2026 captures the latest trends, particularly the surge in deep learning-based steganography and blockchain-enabled digital forensics.

The five most relevant sources, in order of document count, were IEEE Access (468), Sensors (265), Sustainability (234), IEEE Internet of Things Journal (195), and Applied Sciences (178).

The Country Scientific Production metric quantifies the frequency of "author appearances by country affiliations" in scholarly literature. This metric involves tallying the number of country affiliations each time an article features authors from different countries. For instance, if an article lists authors from the USA, Spain, and Italy, the appearance counters for each country would increase by 1, reflecting their contributions to scientific production within the dataset. In this case, the countries that stand out most in these areas, with the highest number of Authors, are China (5009), the USA (2207), and India (1223). The same pattern repeats when analyzing the number of corresponding authors per country. In conclusion, these countries are more mature in this area. Brazil has 392 Authors in published documents, or 106 corresponding authors, indicating a need to encourage publication in Brazil.

2.6.2 Bibliometric Results

To select the most relevant articles, we applied a methodology based on the GIC score (Coelho *et al.*, 2023). Then, we used the Z-Score distribution to identify outliers, i.e., articles with a Z-Score> 3. The distribution of the selected articles is visualizable in the graph depicted in Figure 2.9, where the blue line represents the distribution of the articles and their respective Z-score values, while the green line shows the cut-off applied, with the right side showing the *outliers*, or very relevant, articles and the left side showing the relevant and irrelevant articles.

The articles selected totaled around 167 articles from 2020 to 2026. With the extended timespan, a notable increase in publications from 2024–2026 was observed, particularly in the intersection of blockchain with digital forensics (Igonor; Amin; Garg, 2025; Kreso, 2025), deep learning-based steganography (Sanjalawe *et al.*, 2025; Driss *et al.*, 2025; Al-Janabi; Al-Ta'i, 2025), and blockchain-enabled digital preservation (Werthmuller, 2025; Varadarajan; Rajkumar; Mohanraj, 2025). The top 5 selected articles, with their respective authors, titles, and years of publication, can be seen in Table 2.4.

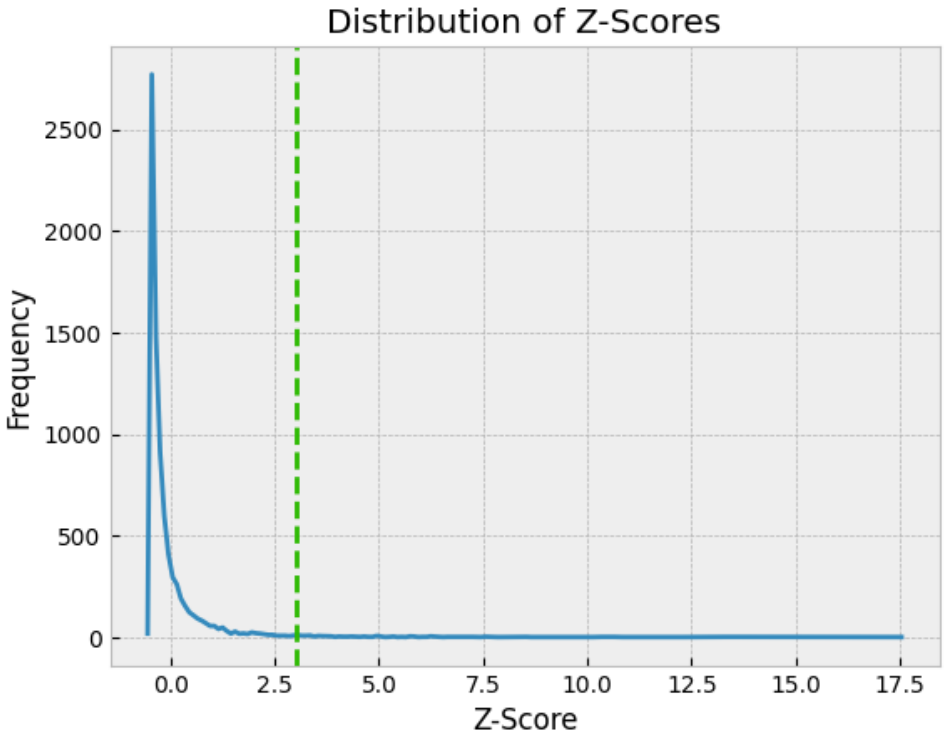


Figure 2.9 – Z-Score analysis (2020–2026)

Author	DOI	Title	Year
Li, Xiaoqi, et al.	10.1016/j.future.2017.08.020	A survey on the security of blockchain systems	2020
Kouhizadeh, Mahtab, Sara Saberi, and Joseph Sarkis.	10.1016/j.ijpe.2020.107831	Blockchain technology and the sustainable supply chain: Theoretically exploring adoption barriers	2021
AlKhanafseh, M. and Surakhi, O.	10.3390/electronics13183729	Evidence preservation in digital forensics: An approach using blockchain and LSTM-based steganography	2024
Sanjalawe, Y. et al.	10.1038/s41598-025-89189-5	A deep learning-driven multi-layered steganographic approach for enhanced data security	2025
Igonor, O.S., Amin, M.B., and Garg, S.	10.3390/blockchains3010005	The application of blockchain technology in the field of digital forensics: A literature review	2025

Table 2.4 – Main Articles

2.6.3 Gap Identification

The bibliometric analysis confirms that this thesis addresses a genuine gap in the literature. Despite the growth in publications across all four topic areas from 2020 to 2026, the key findings remain:

No integration of steganography with chain of custody: Zero results combine steganography techniques with formal chain of custody management, despite the clear utility of covert transport for evidence in sensitive contexts. While AlKhanafseh and Surakhi (AlKhanafseh; Surakhi, 2024) propose combining blockchain and LSTM-based steganography for evidence preservation, their approach does not integrate with formal chain-of-custody standards or OAIS-compliant preservation packages.

Limited blockchain-preservation integration: Although recent works have explored blockchain applications in digital archives (Werthmuller, 2025) and assessed OAIS/TDR compliance (Balogun, 2025), few results provide a practical system that integrates blockchain

with OAIS information packages and BagIt-based integrity verification. Varadarajan et al. (Varadarajan; Rajkumar; Mohanraj, 2025) discuss blockchain for digital archives but do not present a unified system that integrates steganographic transport.

No comprehensive security framework: Existing works address individual security properties but do not provide systematic evaluation against the complete CIA+ANP framework. Recent surveys on blockchain for digital forensics (Igonor; Amin; Garg, 2025; Kreso, 2025) confirm the growing interest but highlight the absence of integrated architectures.

Growing interest in deep learning steganography without custody integration: The period 2024–2026 saw significant advances in deep learning-based steganography, including multiscale attention mechanisms (Al-Janabi; Al-Ta'i, 2025), multi-layered approaches (Sanjalawe *et al.*, 2025), and comprehensive IoT surveys (Driss *et al.*, 2025). However, none of these works integrates steganographic capabilities with the chain-of-custody or digital-preservation standards.

The architecture proposed in this thesis addresses every gap identified above through a deliberate design that integrates the three technology domains into a single, operational system.

Regarding the absence of steganography with formal chain of custody, the closest existing work is that of AlKhanafseh and Surakhi (AlKhanafseh; Surakhi, 2024), who combine blockchain with LSTM-based steganography for evidence preservation. However, their approach is limited to image steganography, does not manage OAIS information packages, and does not maintain formal custody metadata. The proposed architecture goes beyond this by embedding evidence into video carriers via LSB steganography with byte-perfect reversibility and a deep learning model (Stego-STFAN) for video-in-video concealment, while recording every custody event—registration, transfer, verification, access—as an immutable blockchain transaction linked to the BagIt package that contains the evidence. This coupling of steganographic transport with formal custody tracking is absent from all works identified in the bibliometric analysis.

Concerning the limited integration of blockchain with preservation standards, Werthmuller (Werthmuller, 2025) proposes a conceptual Trustchain model for digital archives, and Balogun (Balogun, 2025) assesses repository compliance with OAIS and TDR standards, but neither implements a functional system that combines blockchain with BagIt packaging and PREMIS metadata. The proposed architecture fills this gap by generating RFC 8493-compliant BagIt packages with SHA-256 manifests and tag-manifests, embedding PREMIS preservation metadata that documents every lifecycle event, managing the complete OAIS information package lifecycle from Submission through Archival to Dissemination, and anchoring each package's content hash on the blockchain at the moment of creation. This produces a cryptographic proof of integrity that supplements the administrative fixity checks on which existing repositories depend, bridging the divide between the blockchain and

digital preservation communities.

With respect to comprehensive security evaluation, surveys by Igonor et al. (Igonor; Amin; Garg, 2025) and Kreso (Kreso, 2025) confirm the growing adoption of blockchain for evidence integrity and call for more rigorous security analysis, yet no existing system has been evaluated against all six CIA+ANP properties simultaneously. The proposed architecture is designed and validated to satisfy Confidentiality through steganographic concealment, Integrity through multi-layer hash verification, Availability through local storage with IPFS fallback, Authenticity through actor-content-timestamp binding on the blockchain, Nonrepudiation through an immutable audit trail, and Privacy through controlled dissemination packages, achieving full coverage of all six properties compared to only two for baseline steganography-only approaches.

Finally, regarding the disconnect between deep learning steganography and evidence provenance, recent advances in multiscale attention mechanisms (Al-Janabi; Al-Ta'i, 2025), multi-layered encoding (Sanjalawe *et al.*, 2025), and IoT-oriented steganographic frameworks (Driss *et al.*, 2025) focus on improving embedding quality and capacity but do not address what happens to the hidden content after extraction: who embedded it, when it was embedded, and whether it has been altered since. The proposed architecture closes this loop by registering the content hash on the blockchain before embedding, so that after extraction the recovered evidence can be verified against an immutable record. This transforms steganography from a standalone concealment tool into a component of a verifiable evidence management workflow, where every step—from initial registration to covert transport to final verification—is documented and cryptographically anchored.

In summary, the proposed architecture is, to the best of our knowledge, the first system that simultaneously provides blockchain-backed tamper-evident integrity, OAIS-compliant preservation with BagIt packaging and PREMIS metadata, and video steganography for covert transport, all coordinated through a modular web application with a functional frontend and REST API. The integration enables operational scenarios that no existing work supports: evidence can be preserved following archival standards, its integrity cryptographically verified at any time, its complete custody history audited, and, when circumstances demand, it can be transported covertly within ordinary video files without sacrificing any of these guarantees.

2.7 Related Work

As shown in the previous section, a comprehensive review of related work was conducted, employing bibliometric analysis as the primary method. The review encompassed a thorough examination of existing literature within the field, focusing on scholarly publications, journals, and research trends. The research landscape was systematically evaluated

using bibliometric techniques, including citation analysis, co-authorship networks, and keyword co-occurrence analysis. This approach facilitated the identification of critical themes, influential authors, prominent journals, and emerging research areas within the domain of interest. Furthermore, the bibliometric review provided valuable insights into research trends and enabled comparisons with prior studies in the field.

2.7.1 Blockchain-Based Evidence Management

In the (Yan *et al.*, 2020) paper, he addresses the pressing challenges associated with preserving digital evidence in the context of case investigations amid the rapid evolution of technology. Recognizing the importance of maintaining evidence authenticity throughout its lifecycle, Yan proposes a comprehensive protocol to address the limitations of traditional database technologies. His work introduces a cryptographic technique and blockchain technology to ensure digital evidence's authentication, integrity, and confidentiality.

At the core of Yan's protocol lies revocable ciphertext-policy attribute-based encryption, which offers fine-grained access control over digital documents. Incorporating a BLS signature complements this approach for robust evidence verification. In his work, he underscores the importance of these cryptographic elements in ensuring data security and maintaining a delicate balance between privacy and traceability, both of which are crucial for real-world applicability.

One of his advances is the integration of blockchain technology, which provides an immutable, decentralized ledger to safeguard the integrity and traceability of digital documents. Through meticulous analysis and empirical validation, Yan demonstrates that his proposed protocol successfully guarantees the integrity and validity of documents.

Recent work has explored multi-party trust frameworks for blockchain-based chain-of-custody systems (Smith *et al.*, 2024), demonstrating that blockchain can maintain tamper-proof records across independent organisations. However, these systems typically focus on metadata rather than integrating with complete preservation packages.

More recently, AlKhanafseh and Surakhi (AlKhanafseh; Surakhi, 2024) proposed combining blockchain with LSTM-based steganography to preserve digital forensic evidence; their model stores steganographic files on a blockchain to ensure integrity and confidentiality. However, the approach does not address OAIS-compliant packaging or the formal chain of custody of documents as defined by archival standards.

Igonor, Amin, and Garg (Igonor; Amin; Garg, 2025) provide a comprehensive literature review on blockchain applications in digital forensics (2025), confirming the increasing adoption of blockchain for evidence management but identifying the lack of integration with steganographic transport and preservation standards. Similarly, Kreso (Kreso, 2025) surveys blockchain for preserving digital evidence and highlights the gap between blockchain-based

integrity and established archival practices.

2.7.2 Video Steganography with Deep Learning

In this paper (Telli; Othmani; Ltifi, 2023), the authors present a video steganography model centered around a 3D DeepCNN-grounded autoencoder. Extracting spatiotemporal features from still frames is a distinctive approach that conceals frames from one video within another. Designed to ensure size equivalence, the model enhances its versatility and applicability.

The authors used the UCF101 dataset to train the model. This comprehensive repository contains 13,320 action recognition videos spanning 101 classes. This dataset provides a foundation for refining the model's capabilities and ensuring adaptability across diverse scenarios.

This research (Keizer; Geradts; Kombrink, 2023) evaluates a convolutional neural network (CNN)-based methodology for forensic video steganalysis. The study aims to contribute to the field by creating a dedicated video steganography dataset tailored to train a CNN for spatial-domain steganalysis.

The methodology employed a noise-residual convolutional neural network to detect embedded secrets in video frames. The process was based on the observation that any steganographic embedding invariably alters pixel values in video frames, and that CNNs are adept at discerning these modifications.

Extensive experiments validate the effectiveness of the proposed CNN-based approach. The research employs a range of metrics and parameters to quantify the performance of forensic video steganalysis, with a focus on detecting embedded secrets. The experimental results demonstrate a detection rate of 99.96%, underscoring the efficacy of the CNN-based approach in identifying concealed information in videos.

Various methodologies in video steganography have emerged to bolster message concealment while ensuring resilience against detection. One notable strategy involves concealing secret data within the motion of objects in a video rather than embedding it solely in the background (Jebur; Joda; Naser, 2023).

Additionally, reversible pipelines and invertible neural networks have been harnessed to enable data hiding and recovery across multiple videos, thereby enhancing both hiding capacity and flexibility (Mou *et al.*, 2023).

Furthermore, techniques such as using the least significant bits (LSBs) in the raw domain and employing transform-domain methods, such as the discrete wavelet transform, have been examined (Kunhoth *et al.*, 2023b).

Moreover, there has been a surge in interest in end-to-end video steganography schemes that leverage technologies such as Generative Adversarial Networks (GANs) and multiscale

deep learning networks, which have demonstrated remarkable visual quality, substantial embedding capacity, and resistance to video compression (Xu *et al.*, 2022).

The STFAN architecture (Zhou *et al.*, 2019), originally developed for video deblurring, introduced Filter-Adaptive Convolutional (FAC) layers that generate spatially varying filters. This architecture has been adapted for steganography, where content-adaptive filtering helps maintain visual quality during embedding.

In the 2025- 2026 period, significant advances continued in this area. Al-Janabi and Al-Ta'i (Al-Janabi; Al-Ta'i, 2025) proposed an improvement to video steganography using a multiscale attention mechanism, demonstrating that attention-based architectures achieve superior visual quality compared to prior volumetric approaches. Sanjalawe et al. (Sanjalawe *et al.*, 2025) introduced a multi-layered steganographic framework that integrates Huffman coding, LSB steganography, and deep learning encoder-decoders, achieving over 50 citations and demonstrating the practicality of layered deep learning approaches. Driss et al. (Driss *et al.*, 2025) provided a comprehensive survey of steganography in IoT contexts, covering video steganography and identifying open challenges for resource-constrained environments. Furthermore, Suresh and Binoy (Suresh; Binoy, 2026) published a comprehensive review of video steganography techniques for secure data transmission in 2026, analyzing spatial, transform, and deep learning-based methods.

2.7.3 Digital Preservation Systems

Several systems implement OAIS-compliant preservation workflows. Archivematica is an open-source digital preservation system that automates the creation of AIPs in accordance with OAIS principles, including PREMIS metadata generation and BagIt packaging. DSpace and Fedora provide repository platforms that can be configured to comply with the OAIS reference model.

However, these systems typically rely on administrative security measures rather than cryptographic verification for integrity. Integrating blockchain-based integrity verification with established preservation workflows remains an open challenge addressed in this thesis.

Recent studies have further explored this domain. Werthmuller (Werthmuller, 2025) examined the application of blockchain to digital archives and proposed a conceptual model (Trustchain) to preserve electronic signature certificates within OAIS-based repositories. Balogun (Balogun, 2025) assessed the compliance of digital repositories in South Africa with OAIS and TDR standards, finding that most repositories lack the necessary budget and infrastructure for full OAIS compliance. Varadarajan et al. (Varadarajan; Rajkumar; Mohanraj, 2025) discussed advanced strategies for safeguarding digital archives using blockchain, noting both the potential and the complexity of integrating blockchain with existing OAIS workflows. Marchenko (Marchenko, 2025) analyzed the alignment between information

security regulations and long-term digital preservation in public governance, highlighting that blockchain and OAIS serve complementary but distinct roles.

2.7.4 Comparative Analysis

Table 2.5 presents a comparative analysis of related works against the proposed architecture. The comparison considers key features relevant to secure digital evidence management: blockchain-based integrity verification, OAIS compliance with information packages, the BagIt packaging format, preservation metadata, steganographic capabilities, coverage of security frameworks, and system implementation.

Table 2.5 – Comparative Analysis of Related Works

Feature	**Yan (2020)**	**Telli (2023)**	**Mou (2023)**	**Xu (2022)**	**AlKhan. (2024)**	**Archiv.**	**B-CoC**	**Proposed Arch.**
Blockchain Integrity	✓				✓		✓	✓
OAIS Compliance (SIP/AIP/DIP)						✓		✓
BagIt Packaging (RFC 8493)						✓		✓
PREMIS Metadata						✓		✓
LSB Steganography					✓			✓
Video Steganography		✓	✓	✓				✓
Deep Learning Stego		✓	✓	✓	✓			✓
CIA+ANP Security Analysis								✓
Chain of Custody	✓						✓	✓
Tamper Detection	✓						✓	✓
Covert Transport		✓	✓	✓	✓			✓
Functional Web Application						✓		✓
Unified Architecture								✓

The analysis reveals that existing works address individual aspects of the problem, but none provide a comprehensive solution. Yan's blockchain-based protocol (Yan *et al.*, 2020) ensures evidence integrity but lacks steganographic capabilities and compliance with preservation standards. Video steganography works by Telli et al. (Telli; Othmani; Ltifi, 2023), Mou et al. (Mou *et al.*, 2023), and Xu et al. (Xu *et al.*, 2022) achieve high-quality covert embedding but do not address chain of custody or formal preservation. AlKhanafseh and Surakhi (AlKhanafseh; Surakhi, 2024) advance the field by combining blockchain with LSTM-based steganography for forensic evidence, but their system does not implement OAIS packages, BagIt integrity, or video steganography. Archivematica implements OAIS-compliant workflows with BagIt and PREMIS but relies on administrative rather than cryptographic security. B-CoC frameworks focus on blockchain-based chain-of-custody but lack integration with preservation standards and steganography.

The proposed architecture is the first to integrate all three technology domains: blockchain for tamper-evident integrity, OAIS-compliant preservation with BagIt packaging and PREMIS metadata, and video steganography for covert transport, all within a unified system. This integration enables scenarios not possible with existing approaches: evidence can be preserved

in accordance with archival standards, its integrity can be cryptographically verified through a blockchain, and it can be transported covertly when operational security requires it.

Table 2.6 provides a detailed comparison of video steganography approaches, focusing on technical characteristics relevant to evidence concealment.

Table 2.6 – Comparison of Video Steganography Approaches

Characteristic	**Telli (2023)**	**Mou (2023)**	**Xu (2022)**	**STFAN (2019)**	**Stego-STFAN (Proposed)**
Architecture	3D CNN	INN	GAN	FAC	FAC + Attention
Domain	Spatial	Spatial	Transform	Spatial	YUV + DWT
3D Convolutions	✓				
Invertible Network		✓			
Attention Mechanism					✓
CBAM Integration					✓
Non-local Attention					✓
Compression Resistance	Medium	High	High	Medium	Medium
Computational Cost	High	Medium	High	Low	Medium
Video-in-Video	✓	✓			✓

The proposed Stego-STFAN architecture distinguishes itself by incorporating attention mechanisms (CBAM and Non-local) for content-adaptive embedding, operating in the YUV color space with DWT decomposition for frequency-domain processing, and avoiding computationally expensive 3D convolutions while maintaining video-in-video capability.

2.8 Applicability Scenarios

The architecture presented in this thesis was designed as a general-purpose framework for digital evidence management, but its combination of blockchain-backed integrity, OAIS-compliant preservation, and steganographic concealment makes it particularly suited to specific domains where one or more of these capabilities address pressing operational needs. This section examines concrete application scenarios, mapping each to the architectural features that support it.

2.8.1 Digital Forensics and Legal Proceedings

The management of digital evidence in legal contexts is perhaps the most direct application of the proposed architecture. Courts in Brazil and internationally face a growing volume of digital evidence—contracts, financial records, communications, multimedia—whose admissibility depends on demonstrating an unbroken chain of custody. Brazilian legislation, notably the *Marco Civil da Internet* (Law 12.965/2014) and the Brazilian Criminal Procedure Code (amended by Law 13.964/2019 to include Articles 158-A through 158-F on

chain of custody), establishes requirements for documenting the handling of evidence from collection through presentation in court.

The architecture satisfies these requirements through several mechanisms. The blockchain provides an immutable temporal record of when evidence was registered and by whom, producing a cryptographic proof of provenance that is independent of testimonial assertions. Each custody transfer is recorded as a blockchain transaction, creating an audit trail that fulfils the legislative requirement of documenting every person who handled the evidence and every action taken. The BagIt packaging with SHA-256 manifests ensures that any modification to the evidence—intentional or accidental—is immediately detectable by any party who verifies the package, including opposing counsel, judges, or independent auditors.

A concrete workflow in this scenario proceeds as follows. An investigator collects digital evidence at a scene and immediately uploads it to the system. The API Gateway creates a SIP (Submission Information Package), computes the SHA-256 hash of each file, and registers the hashes on the blockchain. The custody service ingests the SIP into an AIP (Archival Information Package) with PREMIS metadata documenting the ingestion event. When the evidence must be presented, a DIP (Dissemination Information Package) is generated with all metadata necessary for verification. At any point, any party can independently verify that the evidence has not been modified by recomputing the hash and comparing it against the blockchain record.

2.8.2 Institutional Archives and Cultural Heritage

Cultural heritage institutions, government archives, and corporate records offices face the challenge of preserving digital assets over decades or centuries while maintaining confidence in their authenticity. The OAIS reference model (ISO 14721) and BagIt specification (RFC 8493) are already widely adopted in this community; the proposed architecture adds blockchain-backed integrity verification without requiring institutions to abandon their existing practices.

The alignment with OAIS means that packages created by the system can be ingested by standard digital preservation platforms such as Archivematica, DSpace, or RODA. The PREMIS metadata embedded in each AIP documents preservation actions, enabling future archivists to understand the complete history of each package. The blockchain provides an additional layer of assurance that supplements the administrative trust on which traditional archives rely.

In the Brazilian context, this scenario aligns with the requirements of the *Repositório Arquivístico Digital Confiável* (RDC-Arq), as defined by Resolution 43/2015 of the National Council of Archives (CONARQ). RDC-Arq mandates that trusted digital repositories follow the OAIS model and implement mechanisms for integrity checking. The proposed architecture directly addresses these requirements: OAIS compliance is built into the custody

service, and blockchain registration provides a tamper-evident integrity mechanism that exceeds the minimum requirements of periodic fixity checks.

2.8.3 Human Rights Documentation and Journalism

In contexts where documenting abuses is dangerous and the existence of evidence may itself put individuals at risk, the steganographic capability of the architecture provides unique value. Human rights organisations operating in repressive environments can embed evidence—photographs of abuses, testimonial documents, forensic reports—within ordinary video files that withstand casual inspection. The container video plays normally, showing innocuous content; only someone who knows it contains hidden data and possesses the extraction key can recover the embedded evidence.

The blockchain registration adds a critical dimension that steganography alone cannot provide. Once the evidence hash is registered on the blockchain, the temporal precedence of the original evidence is established. If a government or other party later claims the evidence was fabricated, the blockchain timestamp proves that the content existed before the alleged fabrication date. This combination of concealment (the evidence can be transported without detection) and provenance (the evidence can later be proven authentic) addresses a gap that neither technology fills independently.

Investigative journalists protecting sources face analogous challenges. Source documents can be embedded in video files for transport through monitored networks, with the blockchain registration providing a timestamped proof that the documents existed in their current form before any allegations of fabrication. The lossless nature of the LSB embedding ensures that documents are recovered with byte-perfect fidelity, preserving their evidentiary value.

2.8.4 Healthcare and Clinical Trials

Medical records, diagnostic images, and clinical trial data require strict integrity guarantees to ensure patient safety, regulatory compliance, and scientific reproducibility. The architecture can serve as a backbone for managing clinical evidence with verifiable provenance.

In clinical trials, regulatory frameworks such as the ICH E6(R2) guidelines require that electronic records maintain audit trails documenting every change. The blockchain's append-only structure provides an audit trail that cannot be retroactively modified, addressing concerns about data integrity in pharmaceutical research. Each measurement, observation, or diagnostic image can be hashed and registered at the time of creation, providing an independent timestamp that prevents backdating of results.

For medical imaging, the BagIt packaging ensures that DICOM files and associated metadata are preserved as a coherent unit with integrity verification. The PREMIS metadata layer can document the complete provenance chain from image acquisition to clinical interpretation, supporting requirements for traceability in radiological practice.

2.8.5 Intellectual Property and Digital Notarisation

Creators of intellectual property—software developers, authors, artists, inventors—frequently need to establish temporal precedence: proof that a work existed in a specific form at a specific time. Traditional mechanisms (copyright registration, notarised deposits) are slow, expensive, and institution-dependent. The proposed architecture enables immediate, self-service registration of digital works with blockchain-backed timestamping.

A creator uploads a work to the system, which computes its SHA-256 hash and registers it on the blockchain with a timestamp and the creator's identity. This registration constitutes a cryptographic proof that the specific content existed at the recorded time. Disputes over priority—who created a work first, whether a version was altered after a contract date—can be resolved by verifying hashes against the blockchain record. The OAIS packaging ensures that the registered work is preserved in a standardised format with comprehensive metadata, facilitating long-term retrieval.

2.8.6 Corporate Compliance and Regulatory Audits

Organisations subject to regulatory oversight—financial institutions, publicly traded companies, pharmaceutical firms—must demonstrate that records have been maintained with integrity and that complete audit trails exist for critical transactions. Regulations such as Sarbanes-Oxley (SOX), the General Data Protection Regulation (GDPR), and Brazil's *Lei Geral de Proteção de Dados* (LGPD) impose requirements for data integrity, traceability, and accountability.

The architecture addresses these requirements through its multi-layer verification approach. Critical documents are registered on the blockchain at the time of creation, establishing a tamper-evident record. The custody service maintains a complete history of all actions taken on each document, fulfilling audit trail requirements. The BagIt packaging ensures that documents and their metadata are preserved as verifiable units that can be presented to auditors with cryptographic proof of integrity.

2.8.7 Summary of Applicability Mapping

Table 2.7 maps each application scenario to the primary architectural features it relies upon.

Table 2.7 – Mapping of application scenarios to architectural features.

Scenario	Blockchain	OAIS/BagIt	Steganography	CIA+ANP
Digital forensics	✓	✓	—	I, A, Au, N
Institutional archives	✓	✓	—	I, Av, Au
Human rights / journalism	✓	—	✓	C, I, Au, P
Healthcare / clinical trials	✓	✓	—	I, Av, Au, N
Intellectual property	✓	✓	—	I, Au, N
Corporate compliance	✓	✓	—	I, Au, N, P

Legend: C = Confidentiality, I = Integrity, Av = Availability, Au = Authenticity, N = Nonrepudiation, P = Privacy.

The table reveals that the blockchain and OAIS/BagIt components serve as the foundation across all scenarios, while steganography is critical specifically in adversarial environments where concealment is required. This pattern validates the modular design of the architecture: institutions without steganographic requirements can deploy the system without that module, while those operating in hostile environments can activate it as needed.

2.9 Chapter Summary

This chapter has established the theoretical foundation for the proposed architecture.

Information security fundamentals define the six properties (CIA+ANP) that the system must satisfy: Confidentiality, Integrity, Availability, Authenticity, Nonrepudiation, and Privacy.

Cryptographic hash functions provide the mathematical basis for integrity verification, with SHA-256 and SHA-512 offering collision-resistant fingerprints for digital content.

Steganography enables covert transport by hiding information within ordinary-looking media. Video steganography offers high capacity, and deep learning approaches, including invertible neural networks and adaptive filtering architectures, have advanced the state of the art.

Digital preservation standards provide established practices for long-term content management. OAIS defines the information package model, BagIt provides a practical packaging format with manifest-based integrity verification, and PREMIS documents preservation actions.

Blockchain technology enables tamper-evident record-keeping through cryptographic chaining and consensus mechanisms, with applications to evidence integrity and chain-of-custody management.

The bibliometric analysis confirmed that no existing work integrates all three technologies, blockchain, digital preservation standards, and video steganography, into a unified architecture, thereby validating the contribution of this thesis. This gap persists even after an expanded analysis covering publications from 2020 to 2026, despite growing interest in each technology area.

The following chapter presents the proposed architecture that integrates these technologies into a unified system for secure digital evidence management.

3 Proposed Architecture

This chapter introduces a proposal for an architecture that integrates video steganography with blockchain technology to establish a secure, verifiable digital chain-of-custody system. The architecture was designed to meet national and international best practices for digital preservation and to address the legal and regulatory framework governing the management of digital documents in institutional contexts.

The fundamental challenge addressed by this work is preserving digital documents with verifiable integrity over extended periods, while also enabling covert transport when discretion is required. Traditional approaches to document management rely on physical security measures and administrative controls, which can be circumvented by determined adversaries with system access. The proposed architecture addresses this limitation by combining cryptographic guarantees from blockchain technology with the concealment capabilities of video steganography, thereby creating a system in which document integrity can be mathematically verified, and document transport can occur without attracting unwanted attention.

The following sections describe the complete architecture, beginning with an overview of the system's vision and the problems it addresses, then proceeding through detailed descriptions of each component and how they work together to achieve the system's goals.

3.1 System Vision and Objectives

The proposed system was designed to address specific challenges in managing digital documents within institutional contexts. These challenges include ensuring that the document cannot be modified without detection, providing verifiable proof of when the document was registered and by whom, maintaining a complete audit trail of all actions taken on the document, and enabling secure transport of the document through potentially monitored channels.

3.1.1 The Problem of Digital Document Integrity

Digital files can be modified without leaving obvious traces. Unlike physical documents, which exhibit signs of tampering, such as erasures, overwriting, or altered bindings, digital files can be completely altered and still appear authentic. This creates a fundamental problem for document management: how can anyone be certain that a digital file is identical to its original version?

Traditional solutions rely on access controls and administrative procedures. Files are stored on secure servers, access is restricted to authorised personnel, and logs are maintained of who accessed what and when. However, these measures depend on the trustworthiness of the people and systems involved. An administrator with sufficient access can modify files, alter logs, and cover their tracks. In legal contexts, this creates doubt about the authenticity of a digital document.

The proposed architecture addresses this problem through cryptographic hashing and blockchain registration. When a document enters the system, its cryptographic hash is computed and permanently recorded on a blockchain. This hash acts as a unique fingerprint for the content. Any subsequent modification to the document, no matter how small, would produce a different hash, immediately revealing the tampering. Because the blockchain is append-only and each block is cryptographically linked to the previous one, the recorded hash cannot be altered without invalidating the entire chain from that point forward.

3.1.2 The Need for Covert Transport

In some scenarios, it is necessary to transport a document without revealing its nature. Encrypted files, while secure, are obviously encrypted and may attract attention or be subject to forced decryption. Steganography provides an alternative by hiding the document within ordinary-looking media files that can be transported without arousing suspicion.

The proposed system incorporates video steganography to enable covert transport. A document can be embedded within ordinary video files that appear completely normal when played. Only someone who knows that the video contains hidden content and possesses the means to extract it can access the embedded document. This capability is particularly valuable in contexts where a document must pass through potentially hostile environments or where its mere existence should remain confidential.

3.1.3 Compliance with Preservation Standards

Digital preservation is a well-established field with international standards that define best practices for ensuring that digital content remains accessible and authentic over time. The OAIS (Open Archival Information System) reference model, codified as ISO 14721, provides a conceptual framework for digital archives. The BagIt format, specified in RFC 8493, provides a practical packaging structure for digital content with integrity verification.

The proposed architecture aligns with these standards to ensure interoperability with existing digital preservation infrastructure and to benefit from the preservation community's accumulated knowledge. Document packages are structured according to BagIt specifications, with SHA-256 hashes computed for every file. The package lifecycle follows the OAIS model, with distinct phases for submission, archival storage, and dissemination.

3.2 System Architecture Overview

The system was designed as a modular web application consisting of six interconnected services. Each service handles a specific aspect of the document management workflow and communicates via well-defined interfaces. This modular design allows each component to be developed, tested, and maintained independently while remaining part of a cohesive whole.

The primary programming language chosen was Python, selected for its extensive libraries for cryptography, video processing, and web development. The web framework is FastAPI, which provides automatic API documentation, request validation, and native asynchronous operations. For the frontend, two interfaces were developed: a Streamlit application for rapid prototyping and demonstration, and an Angular application that provides a complete production interface with a modern user experience.

Video processing operations rely on OpenCV for frame manipulation and FFmpeg for video encoding and decoding. The blockchain component was implemented from scratch rather than using an existing platform, thereby providing complete control over the protocol and keeping the system lightweight enough for single-institution deployment. The entire system can be deployed using Docker containers, with a docker-compose configuration that orchestrates all services together.

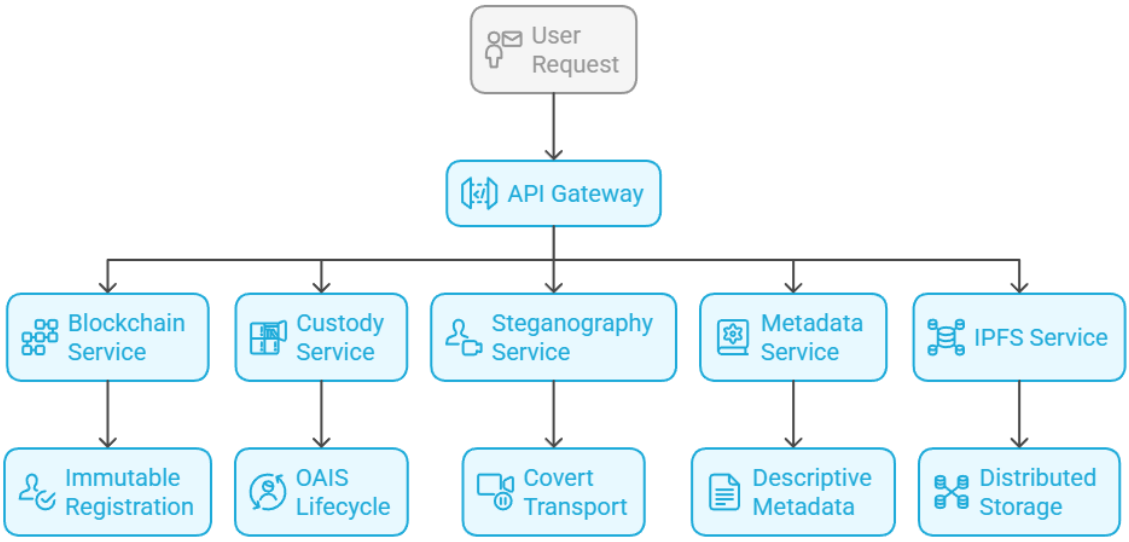


Figure 3.1 – High-level architecture showing the main modules and their interactions. The API Gateway serves as the central coordinator, routing requests to specialised services for blockchain operations, custody management, steganography, metadata processing, and distributed storage.

The six modules that comprise the architecture are:

- **API Gateway**: The central entry point that receives all user requests, validates inputs, and coordinates workflows across the other services.

- **Blockchain Service**: Manages the custom blockchain that provides immutable registration of document and custody events.
- **Custody Service**: Implements the OAIS lifecycle, managing the transformation of documents through Submission, Archival, and Dissemination packages.
- **Steganography Service**: Provides the capability to embed a document within video files for covert transport.
- **Metadata Service**: Handles descriptive metadata using Dublin Core and ISAD(G) standards.
- **IPFS Service**: Provides optional distributed storage using the InterPlanetary File System.

3.3 The API Gateway

The API Gateway is the entry point for all interactions with the system. When users submit requests to upload documents, verify documents, or transfer custody, the gateway validates the input, determines which backend services are required, and coordinates the overall workflow.

The gateway exposes a RESTful API accessible via HTTP requests. It supports cross-origin resource sharing (CORS) to enable web browsers to interact with the API, implements request validation to ensure that incoming data conforms to expected formats, and provides error handling to return meaningful messages when errors occur.

When the gateway starts, it initialises all backend services and maintains references to them throughout its lifetime. This enables the gateway to orchestrate complex workflows involving multiple services. For example, when a user requests to register a new document, the gateway coordinates with the Custody Service to create the package, the Metadata Service to handle descriptive information, the Steganography Service if embedding is requested, and the Blockchain Service for final registration.

The gateway organises its endpoints into four main categories. The first category handles package management operations, including creating submission packages, adding files, setting metadata, and finalising packages for ingestion. The second category encompasses blockchain operations, including document registration, integrity verification, transaction history queries, and chain validity checks. The third category manages custody operations, including transferring custody between parties, recording access events, and generating dissemination packages. The fourth category addresses steganography operations, including embedding payloads in cover videos, extracting payloads from container videos, and classifying payload types to determine the appropriate technique.

3.4 The Blockchain Service

The Blockchain Service implements a custom blockchain protocol designed specifically for document registration and custody tracking. Unlike public blockchains that handle thousands of transactions per second from anonymous participants worldwide, this implementation prioritises simplicity, auditability, and direct integration with the preservation workflow.

3.4.1 Understanding the Blockchain Structure

A blockchain is essentially a chain of data blocks, where each block contains a list of transactions and a reference to the previous block. This reference is a cryptographic hash of the previous block's contents, creating a chain of dependencies. If anyone attempts to modify a historical block, its hash changes, breaking the link to the next block, which must then be modified as well, and so on through the entire chain.

In the proposed system, each block contains several pieces of information: its position in the chain (the index, starting from zero for the genesis block), the time it was created (as a Unix timestamp), a list of transactions representing custody events, a reference to the previous block's hash, and a special number called a nonce that was found during mining. The block's own hash is computed from all these fields using the SHA-256 algorithm.

The genesis block is the first block in the chain and has special significance. It has an index of 0, and its previous hash field contains the conventional value of 64 zeros because there is no previous block. The genesis block is created automatically when the system is first initialised and serves as the anchor point for the entire chain.

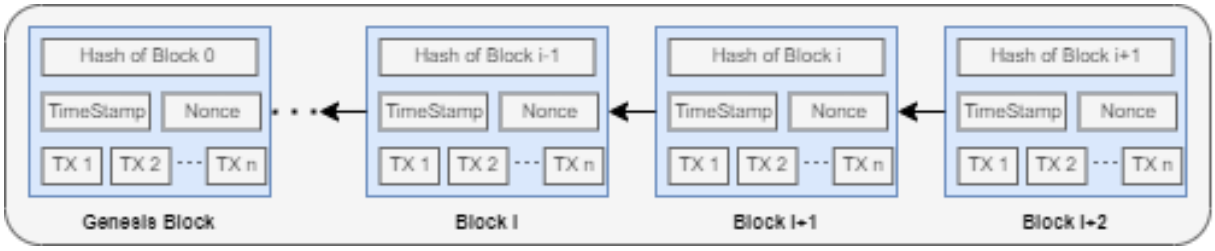


Figure 3.2 – Description of a typical structure of a blockchain. Adapted from: (Underwood, 2016)

3.4.2 Mining and Proof of Work

Before a block can be added to the chain, it must be "mined." Mining involves finding a nonce value that, when combined with the block's other contents and hashed, produces a result with a specific pattern, in this case, a hash that begins with a certain number of zeros in its hexadecimal representation. Finding such a nonce requires trying many different values, which takes computational effort.

This proof-of-work mechanism serves a security purpose. If someone wants to tamper with a historical block, they would need to re-mine that block (finding a new valid nonce) and then re-mine every subsequent block as well. The computational cost of this attack increases with chain length, making tampering increasingly expensive over time.

For the proposed system, the mining difficulty is set to require hashes beginning with four hexadecimal zeros. Since each hexadecimal character represents four bits and each bit has a 50% chance of being zero, the probability of a random hash starting with four zeros is $(1/16)^4$, or about 1 in 65,536. This means that on average, 65,536 hash computations are required to find a valid nonce. On modern hardware, this typically takes about 30 milliseconds, which is fast enough for interactive use but still imposes a meaningful cost on any attempt to manipulate the chain.

3.4.3 Transaction Types

The blockchain records different types of custody events through transactions. Each transaction includes information about who acted (the sender), who or what was affected (the recipient, often the system itself), when it occurred (the timestamp), and relevant details about the action itself (the transaction data). The system supports ten different transaction types to capture the full range of custody events:

Table 3.1 – Transaction types supported by the blockchain.

Type	**Purpose**
DOCUMENT_REGISTER	Records the initial registration of a document, including its content hash, IPFS identifier, and metadata hash
DOCUMENT_TRANSFER	Records when ownership of the document passes from one party to another
DOCUMENT_ACCESS	Records when someone accesses a document, for audit purposes
DOCUMENT_VERIFY	Records verification events, noting who verified, when, and whether verification succeeded
CUSTODY_TRANSFER	Records formal custody transfers with full provenance information
CUSTODY_EVENT	Records general custody events that don't fit other categories
METADATA_UPDATE	Records when metadata is modified
METADATA_ADD	Records when new metadata is added

Type	Purpose
FIXITY_CHECK	Records periodic integrity checks performed by the system
SYSTEM_EVENT	Records system-level events such as startup and shutdown

Each transaction receives a unique identifier generated by hashing its contents. This identifier serves as a permanent reference for locating and verifying the transaction.

3.4.4 The Document Registry

In addition to the blockchain itself, the service maintains an in-memory document registry that provides fast lookups by document identifier. When the service starts, it rebuilds this registry by replaying all registration transactions from the blockchain. This ensures that the registry always reflects the authoritative state recorded in the chain.

Each document in the registry contains a unique identifier, a content hash (the SHA-256 hash of the file), an optional IPFS content identifier, associated metadata, the current owner, status (e.g., registered, transferred, or archived), a complete custody history, an access log, and a version number. The custody history is particularly important as it provides the complete chain of custody from registration to the present.

3.4.5 Chain Validation

The blockchain service can validate the entire chain at any time by checking three properties for each block after the genesis block. First, the hash stored in the block must match the hash computed from its contents, ensuring that no data has been modified. Second, the previous hash field must match the actual hash of the previous block, ensuring that the chain linkage is intact. Third, the hash must satisfy the proof-of-work requirement by starting with the required number of zeros. If any of these checks fail for any block, the entire chain is considered invalid.

3.4.6 Persistence and Recovery

The blockchain state is persisted to a JSON file whenever changes occur. This file contains the complete chain (all blocks with their transactions) and the document registry. The write operation uses an atomic pattern: data is first written to a temporary file, then the temporary file is renamed to replace the actual file. This ensures that the chain file is never left in a partially written state, even if the system crashes during the save operation.

If the chain file is lost or corrupted, it can be restored from backups. The document registry does not need a separate backup because it can always be reconstructed from the blockchain by replaying all transactions.

3.5 The Custody Service

The Custody Service implements the OAIS (Open Archival Information System) reference model, which is the international standard for digital preservation. This service manages the complete lifecycle of document packages, from initial submission through archival storage to eventual access requests.

3.5.1 The OAIS Information Package Model

OAIS defines various types of information packages that serve distinct purposes within the preservation workflow. Understanding these package types is essential for understanding how the system manages documents.

A **Submission Information Package (SIP)** is what users submit when they want to preserve a document. It contains the files and the descriptive metadata provided by the submitter. The SIP is a working package that can be modified until the user finalises it for ingestion. Users can add files incrementally, update metadata, and request steganographic embedding before committing the package.

An **Archival Information Package (AIP)** is what the system stores for long-term preservation. When a SIP is ingested, the system transforms it into an AIP by adding preservation metadata, computing file hashes, generating integrity manifests, and creating the necessary structural documentation. The AIP is registered on the blockchain and cannot be modified afterward. Any changes require creating a new version, which is itself a new AIP with a reference to the previous version.

A **Dissemination Information Package (DIP)** is what users receive when they request access to an archived document. The DIP is derived from the AIP but can be customised for the requester's needs. For example, a DIP might include only certain files from the AIP, omit sensitive metadata, or be packaged in a specific format requested by the user. The system automatically creates a default DIP when an AIP is created, but additional DIPs can be generated on demand.

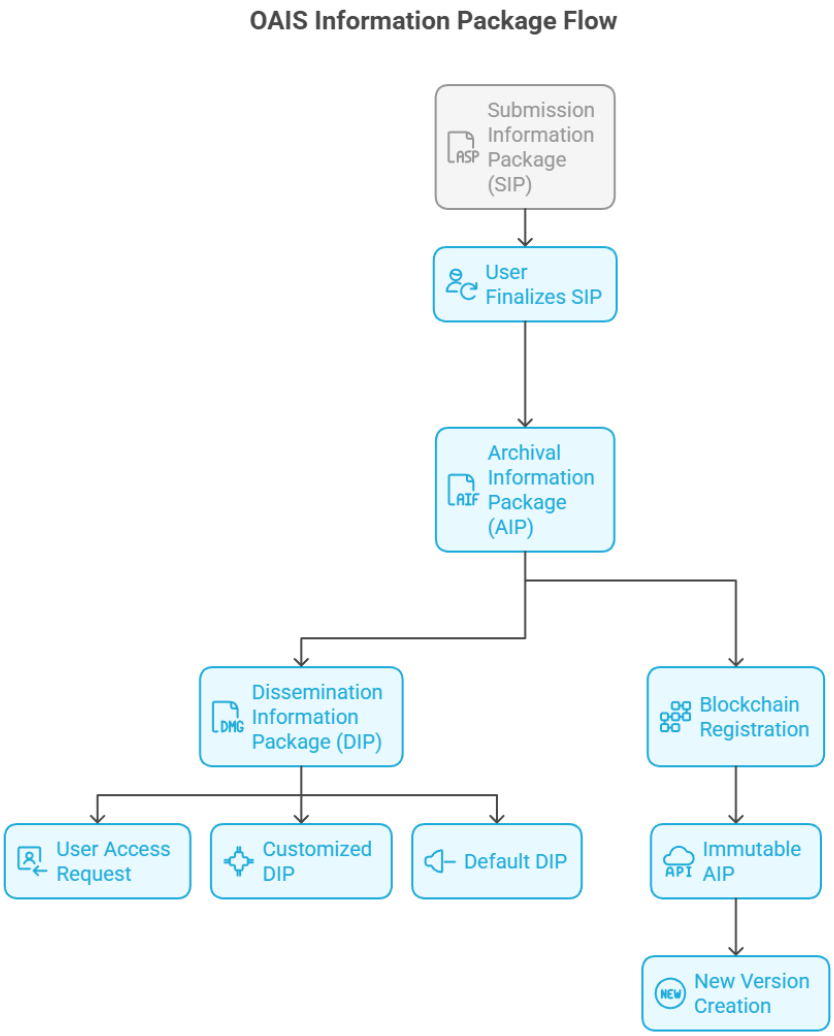


Figure 3.3 – OAIS package lifecycle: SIP submission, AIP archival, and DIP dissemination.

3.5.2 The BagIt Packaging Format

All packages in the system conform to the BagIt format specified in RFC 8493. BagIt provides a simple, robust structure for packaging digital content and includes built-in integrity verification. The format was developed by the Library of Congress and the California Digital Library, and has become a de facto standard in the digital preservation community.

A BagIt package is a directory with a defined structure. At the top level, it contains a declaration file that identifies the package as a BagIt container and specifies the version. It also contains an information file containing metadata about the package, such as the creating organisation, contact information, creation date, and a description. The actual content files are stored in a data subdirectory, keeping them separate from the package metadata.

The key feature of BagIt is its manifest system. A manifest file lists every content file along with its cryptographic hash. The proposed system generates manifests using both SHA-256 and SHA-512, providing redundancy in case one algorithm is later found to be

vulnerable. When verifying a package, the system recomputes the hashes of all files and compares them with those in the manifest. Any modification to any file will produce a different hash, immediately revealing the tampering.

For metadata files (which are not in the data directory), BagIt uses separate tag manifests. This ensures that the integrity of metadata can be verified independently of the content.

3.5.3 Preservation Metadata

When a SIP is transformed into an AIP, the system generates several types of preservation metadata that document the package's provenance, structure, and integrity.

PREMIS (Preservation Metadata Implementation Strategies) records events that happen to objects. Every significant action ingestion, verification, access, or migration is recorded as a PREMIS event with information about what happened, when it happened, who or what performed the action, and what the outcome was. This creates an audit trail that documents the complete preservation history.

METS (Metadata Encoding and Transmission Standard) provides a structural map that relates files to each other and to their descriptive metadata. It describes how the package is organised and how its components relate to each other.

PDI (Preservation Description Information) captures reference information (identifiers that uniquely identify the content), context information (how the content relates to other information), provenance information (the history of the content), fixity information (mechanisms for verifying integrity), and access rights information (who can access the content and under what conditions).

3.5.4 The Ingestion Process

When a user finalises a SIP and requests ingestion, the Custody Service performs a series of transformations to create the AIP. First, it validates that the SIP is complete, with all required metadata present and all files accounted for. It then generates the BagIt manifests by computing hashes for each file. Next, it creates the preservation metadata (PREMIS events, METS structural map, PDI). The entire package is then archived in a compressed file, and the archive's SHA-256 hash serves as the content hash that will be registered on the blockchain.

The service then calls the Blockchain Service to create a DOCUMENT_REGISTER transaction containing the AIP identifier, content hash, and metadata reference. Once this transaction is mined into a block, the registration is permanent, and the AIP cannot be modified without detection.

Finally, the service creates a default DIP to provide immediate access to the registered document. This DIP is typically a ZIP file containing the files and basic metadata.

3.6 The Steganography Service

The Steganography Service enables the concealment of documents within ordinary-looking video files. This enables the covert transport of documents through channels that might be monitored, allowing them to travel without attracting attention.

3.6.1 Payload Classification and Technique Selection

The service automatically classifies payloads to determine the most appropriate steganographic technique. Video files (e.g., MP4, AVI, or MKV) are suitable for embedding video within video using the Stego-STFAN approach. Documents (e.g., PDFs) and images (e.g., PNG or JPEG) are embedded using LSB steganography. This automatic classification simplifies the user experience while ensuring that the optimal technique is used for each payload type.

3.6.2 LSB Steganography for Documents and Images

For embedding documents and images, the system uses LSB (Least Significant Bit) steganography. This technique works by modifying the least significant bit of each colour value in video frames. Since the least significant bit contributes the least to the overall colour value (changing it alters the value by only 1 of 256 levels), the modifications are imperceptible to human vision.

The embedding process begins by reading the payload file and creating a small header that records the payload size and type. This header informs the extraction process how many bits to read and the type of file to reconstruct. The payload bytes are then converted to a sequence of individual bits.

The cover video is processed frame by frame. For each frame, the algorithm visits every pixel in row-major order (left to right, top to bottom). At each pixel, it modifies the least significant bit of each colour channel (blue, green, and red) to encode payload bits. Since each pixel provides three bits of capacity (one per channel), and a typical video frame contains millions of pixels, even a short video can hide substantial amounts of data.

A critical requirement for LSB steganography is that the container video must use a lossless codec. Standard video codecs such as H.264 and H.265 employ lossy compression, which modifies pixel values to achieve smaller file sizes. These modifications would destroy the carefully embedded LSB changes. The system therefore enforces the use of lossless codecs, such as FFV1, when creating container-based videos. While this results in larger file sizes, it ensures that the embedded payload can be perfectly recovered.

The extraction process reverses the embedding. The container video is processed frame by frame, and the least significant bit of each colour channel is read in the same order used

during embedding. The header is extracted first to determine the payload size and type, then the payload bits are read and converted back to bytes to reconstruct the original file.

3.6.3 Embedding Capacity

Video offers a large capacity for LSB embedding due to its spatial and temporal dimensions. Consider a 10-second video at 1920×1080 resolution (Full HD) and 30 frames per second. Each frame contains 2,073,600 pixels (1920×1080). Each pixel provides 3 bits of capacity (one per BGR channel). With 300 frames in 10 seconds, the total capacity is:

$$\text{Capacity} = 1920 \times 1080 \times 3 \times 30 \times 10 = 1{,}866{,}240{,}000 \text{ bits} \approx 233 \text{ MB} \tag{3.1}$$

This means that even a modest 10-second video can conceal more than 200 megabytes of data. In practice, embedding rates are often kept lower to reduce the statistical detectability of the hidden content, but the raw capacity is substantial.

3.6.4 Stego-STFAN for Video-in-Video Embedding

For scenarios involving video content, the architecture proposes a more sophisticated approach called Stego-STFAN. This technique employs deep learning to embed an entire secret video within a cover video while preserving visual quality and evading detection.

Stego-STFAN is based on STFAN (Spatial-Temporal Filter Adaptive Network), an architecture originally developed for video deblurring. The key insight is that STFAN's ability to process spatial and temporal information simultaneously makes it well-suited to steganography, where the goal is to blend secret information seamlessly into cover content while preserving temporal consistency across frames.

3.6.4.1 The Three Processing Stages

The Stego-STFAN architecture operates in three stages that together form the complete steganographic pipeline.

Embedding Stage: The cover and secret video frames are first converted from RGB colour space to YUV colour space, which separates luminance (brightness) from chrominance (colour). This separation is useful because human vision is more sensitive to changes in luminance than chrominance. A Discrete Wavelet Transform (DWT) is then applied to decompose the frames into frequency subbands. The secret information is embedded primarily in the LH (Low-High) and HL (High-Low) subbands, which contain edge and detail information. Changes in these subbands are less perceptible to human vision than changes in the LL (Low-Low) subband, which contains the main image content. The result of this stage is an initial stego-frame that contains the hidden information but may have visible artefacts.

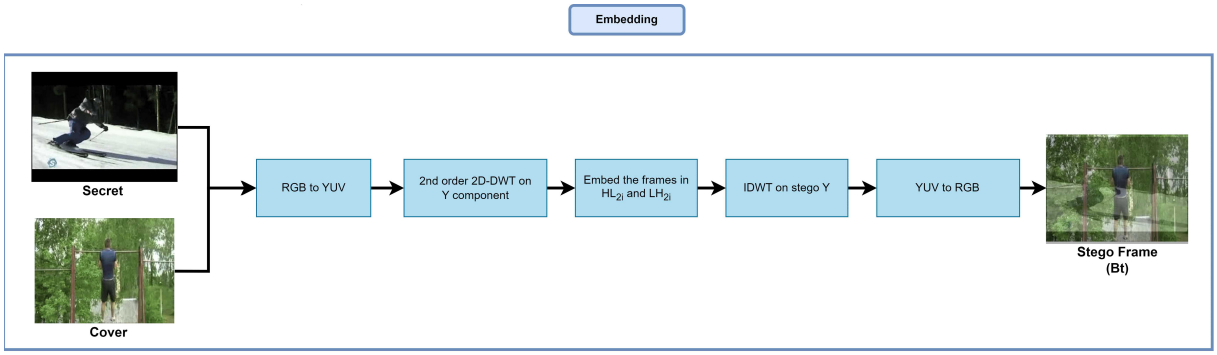


Figure 3.4 – The embedding stage converts frames to YUV colour space, applies DWT to extract frequency subbands, embeds secret features in the LH and HL subbands, then reconstructs the frame.

Hiding Stage: The initial stego-frame is refined into a high-quality container using spatial-temporal adaptive filters. This stage is crucial for achieving good visual quality. It receives the stego-frame along with attention features and information from previous frames, thereby maintaining temporal consistency. Filter Adaptive Convolutional (FAC) layers generate dynamic filters that adapt to each frame's specific content, producing containers that are visually similar to the original covers with minimal artefacts.

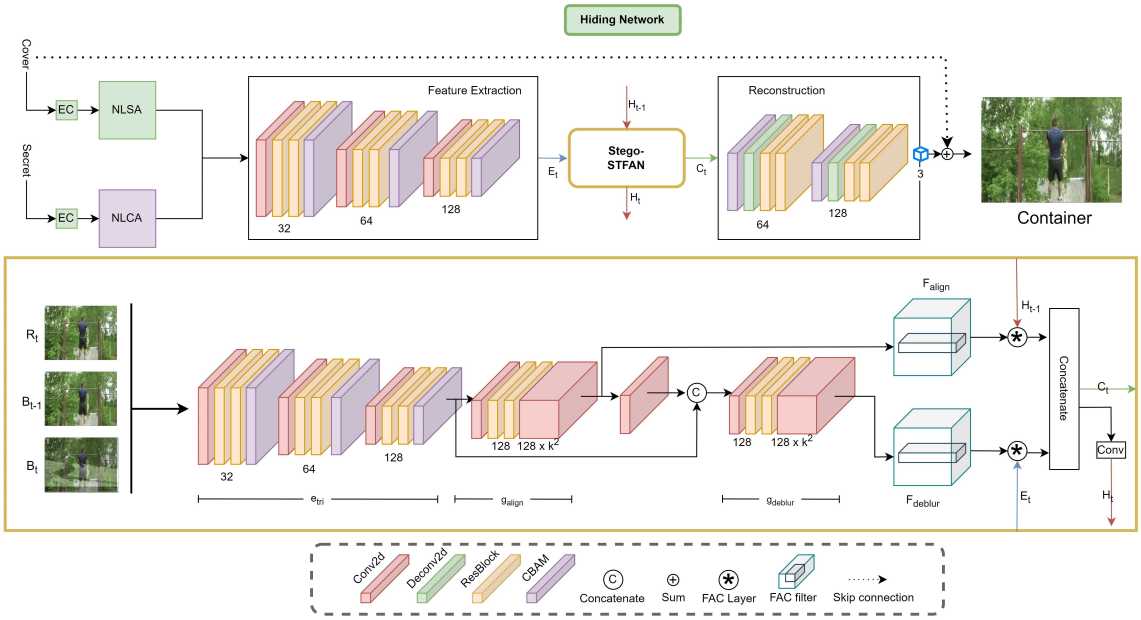


Figure 3.5 – The hiding stage uses attention mechanisms and adaptive filters to refine the stego-frame into a high-quality container that closely resembles the original cover video.

Restoring Stage: To recover the secret video, the container frame passes through a symmetric architecture that extracts the hidden information. The restoring stage uses similar adaptive filtering techniques to reconstruct the secret video from the container, attempting to recover the original content as faithfully as possible.

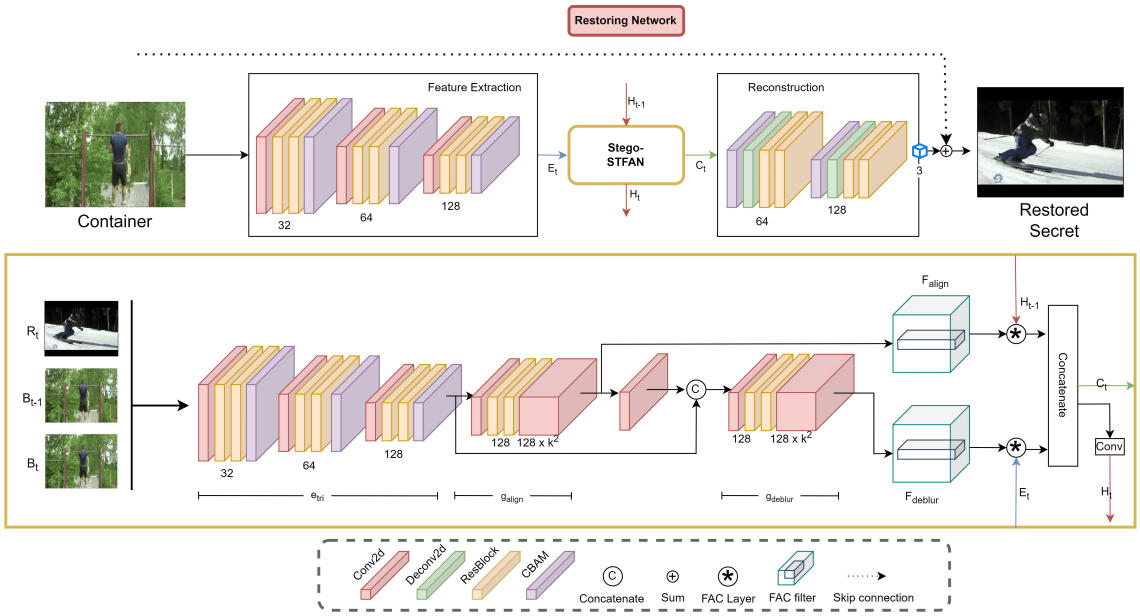


Figure 3.6 – The restoring stage extracts the hidden secret video from the container using adaptive filters and attention mechanisms, reconstructing the original content.

3.6.4.2 Attention Mechanisms

Stego-STFAN incorporates multiple attention mechanisms to focus processing on the most important regions for steganography. Non-local Self-Attention (NLSA) identifies important features within each frame, such as edges and regions with significant motion. Non-local Co-Attention (NLCA) identifies correlations between cover and secret frames, thereby helping the network determine where to perform embedding with minimal visual impact. The Convolutional Block Attention Module (CBAM) provides both spatial attention (which regions of the frame are important) and channel attention (which feature channels are important).

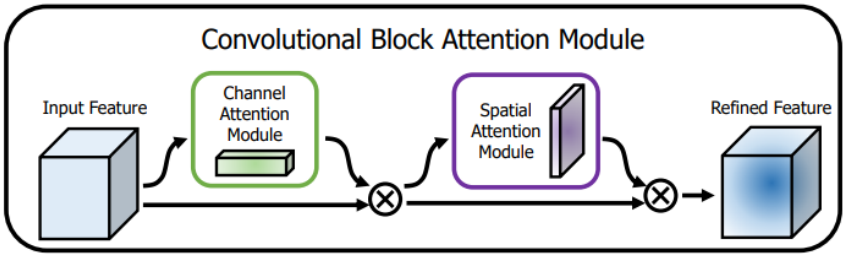


Figure 3.7 – The CBAM architecture combines channel attention and spatial attention to help the network focus on the most relevant features for steganographic embedding.

3.6.4.3 Training Objective

Training the Stego-STFAN network requires balancing two competing objectives. On one hand, the container video should look as similar as possible to the cover video (good

concealment). On the other hand, the restored secret should match the original secret as closely as possible (good reconstruction). These objectives can conflict, as more aggressive embedding might improve reconstruction quality at the expense of concealment.

The loss function addresses this challenge with three components. The first component measures the difference between the cover and the container using mean squared error, thereby encouraging the container to resemble the cover. The second component measures the difference between the original and restored secrets, encouraging faithful reconstruction. The third component is a balance term that penalises large gaps between the two primary objectives, preventing the network from sacrificing one goal entirely for the other.

3.7 The Metadata Service

The Metadata Service handles descriptive metadata in accordance with established archival standards. Good metadata is essential for document management because it provides context regarding the document, its origin, and its meaning. Without adequate metadata, a document may be difficult to locate, understand, or use effectively.

The service supports two primary metadata schemas that address different needs.

Dublin Core is a general-purpose metadata standard that defines fifteen core elements: title, creator, subject, description, publisher, contributor, date, type, format, identifier, source, language, relation, coverage, and rights. These elements are broadly applicable and widely understood, making Dublin Core interoperable with external systems. The simplicity of Dublin Core also makes it easy for users to provide basic descriptive information without specialised archival training.

ISAD(G) (General International Standard Archival Description) provides a more comprehensive framework specifically designed for archival description. It includes elements for identity (reference code, title, dates), context (name of creator, administrative history, archival history), content and structure (scope and content, appraisal information, system of arrangement), conditions of access and use (access conditions, reproduction conditions, language), allied materials (existence of originals, existence of copies, related materials), and notes. ISAD(G) is particularly relevant for institutional archives that need detailed provenance documentation.

When users create a submission package, they can provide metadata through the API or web interface. The Metadata Service validates this input against the relevant schema, ensuring that required fields are present and that values conform to expected formats (e.g., dates in ISO 8601).

3.8 The IPFS Service

The IPFS Service provides an optional distributed storage layer using the InterPlanetary File System. IPFS is a peer-to-peer network in which files are identified by the hash of their content rather than by their location. This content-addressing means that the same file stored anywhere in the network will have the same identifier, and requesting a file by its identifier will return the same content regardless of which node provides it.

When an AIP is ingested, the system may optionally upload it to IPFS and receive a Content Identifier (CID), which is recorded alongside the local storage path. This CID is then included in the blockchain registration, providing an additional reference point for the document.

The IPFS integration is designed for graceful degradation. If IPFS is unavailable (whether because the local node is not running, the network is unreachable, or IPFS is not configured), the system falls back to local storage only. This ensures that the core functionality of document registration and custody management remains operational even without the distributed storage component.

The potential benefits of IPFS integration include improved availability (content can be retrieved from any node that has it), inherent redundancy (popular content is automatically replicated across nodes that access it), and verification via content addressing (the CID itself verifies the content's integrity).

3.9 Module Integration and Workflow

Understanding how the modules interact is essential for appreciating how the system achieves its security goals. This section outlines the complete workflow for a user registering a new document with steganographic protection.

The process begins when a user creates a new Submission Information Package through the web interface or API. The API Gateway receives this request and forwards it to the Custody Service, which creates an empty SIP with a unique identifier. The SIP is initially in draft status, allowing modifications.

The user then uploads document files, which the gateway routes to the Custody Service for storage within the SIP. Multiple files can be added incrementally. Each file is placed in the package's data directory.

Next, the user provides descriptive metadata. This might include Dublin Core elements like title, creator, and description, or more detailed ISAD(G) elements for archival description. The Metadata Service validates this information before associating it with the SIP.

If the user requests steganographic protection, the gateway invokes the Steganography Service. The user specifies the files to embed and provides a cover video. The service classifies

the payload type, selects the appropriate technique (LSB for documents and images, Stego-STFAN for video), and performs the embedding. The resulting container video replaces or accompanies the original document in the SIP.

When the user is satisfied with the package contents, they finalise the SIP, indicating that no further changes will be made. This triggers the Custody Service to validate the package structure, ensure that required metadata is present, and generate the BagIt manifests with file hashes.

The user then requests ingestion, which transforms the SIP into an AIP. The Custody Service generates preservation metadata (PREMIS events documenting the ingestion, METS structural map, PDI), creates the final archival package, and computes its content hash.

The gateway calls the Blockchain Service to register the AIP. A DOCUMENT_REGISTER transaction is created containing the AIP identifier, content hash, and metadata reference. This transaction is added to the pending pool and included in the next mined block. Once mined, the registration is permanent and verifiable.

Finally, the Custody Service creates a default DIP providing immediate access to the registered document. The complete registration is now finished, with the document preserved in an OAIS-compliant package, protected by cryptographic hashes, and permanently registered on the blockchain.

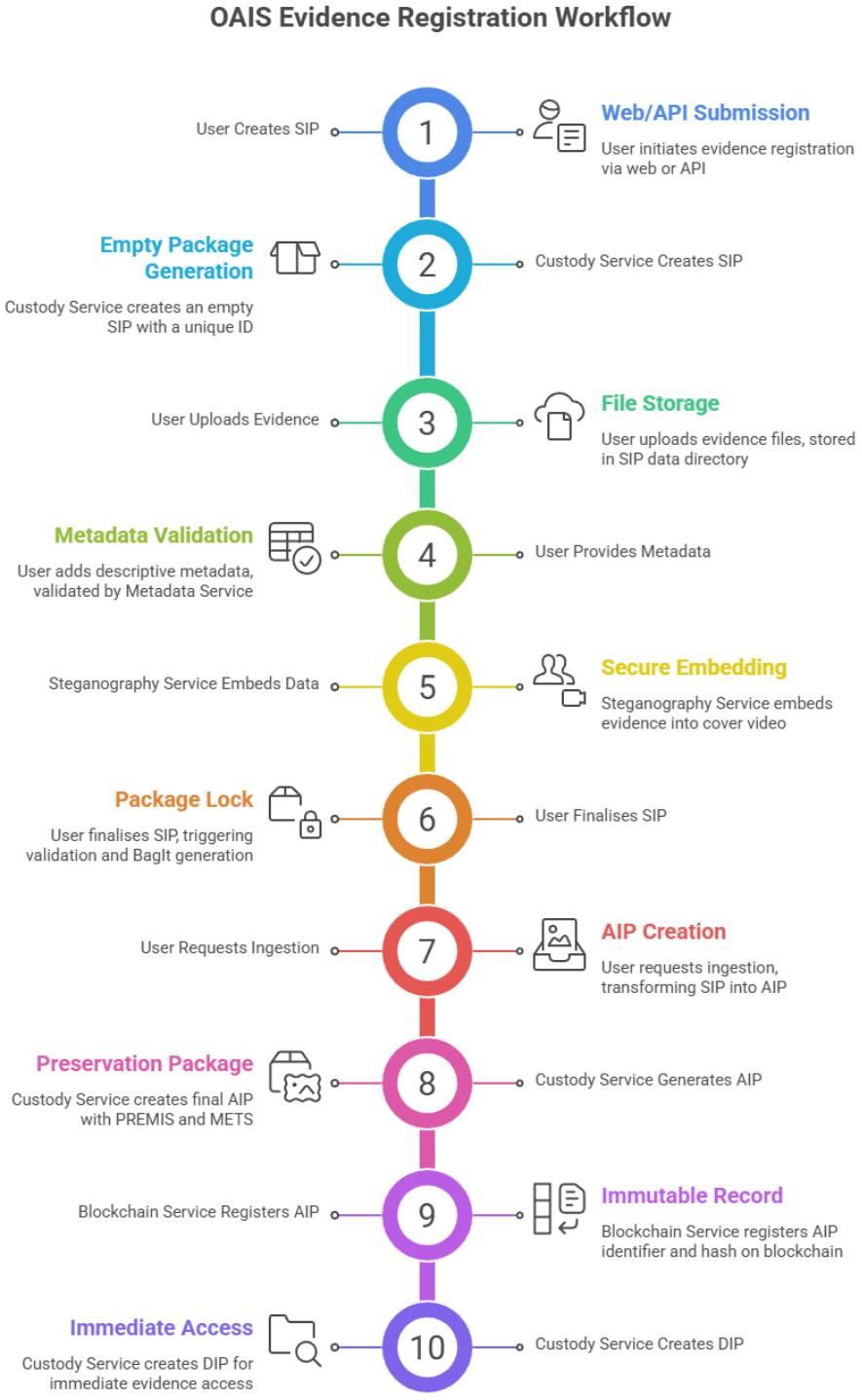


Figure 3.8 – Sequence diagram showing the complete document registration workflow. The API Gateway coordinates interactions between the user and backend services, ensuring that each step completes successfully before proceeding to the next.

3.10 Security Properties

The combination of blockchain, OAIS-compliant packaging, and steganography provides six fundamental security properties:

- Integrity ensures that the document cannot be modified without detection. The multi-layer verification, file-level BagIt manifests, package-level content hash, and blockchain-level transaction binding create redundant checks that would detect any tampering attempt.
- Authenticity ensures that the origin of the document can be verified. Blockchain registration permanently binds the content hash to the registering actor and the timestamp, providing a cryptographic document of who submitted what and when.
- Nonrepudiation ensures that actions cannot be denied. Each significant event (registration, transfer, access, verification) is recorded on the blockchain, including the identity of the performing actor. These records cannot be deleted or modified.
- Confidentiality ensures that the document can be protected from unauthorised observation. Steganographic embedding enables a document to be transmitted through monitored channels without revealing its contents.
- Availability ensures that the document is accessible when needed. The system provides multiple storage options (local filesystem and the IPFS distributed network) with graceful degradation when components are unavailable.
- Privacy ensures that sensitive information is protected. Controlled dissemination through customisable DIPs allows sensitive metadata to be omitted when appropriate, while steganographic concealment protects the document's existence.

3.11 Standards Compliance

The architecture was designed to align with relevant national and international standards:

- OAIS (ISO 14721): The system implements the OAIS functional model with distinct ingest, archival storage, data management, and access functions.
- BagIt (RFC 8493): All packages conform to the BagIt specification with SHA-256 and SHA-512 manifests.
- Brazilian RDC-Arq: The system supports compliance with CONARQ Resolution 43 through OAIS alignment, metadata standards, and an auditable custody chain.
- ISO 16363: The architecture supports audit and certification requirements for trustworthy digital repositories.

3.12 Chapter Summary

This chapter has presented the proposed architecture of a blockchain-anchored digital chain-of-custody system with integrated video steganography. The key elements of the architecture include:

- A modular design with six specialised services: API Gateway, Blockchain Service, Custody Service, Steganography Service, Metadata Service, and IPFS Service
- A custom blockchain protocol using SHA-256 hashing and Proof of Work mining, with ten transaction types covering all custody events from registration through verification
- OAIS-compliant packaging following the SIP to AIP to DIP lifecycle, using RFC 8493 BagIt containers with SHA-256 and SHA-512 manifests
- LSB steganography for embedding documents and images in video, and the Stego-STFAN deep learning architecture for video-in-video embedding
- Multi-layer integrity verification combining BagIt manifests, content hashes, and blockchain anchoring
- Support for Dublin Core and ISAD(G) metadata standards
- Optional IPFS integration for distributed storage with graceful degradation
- Alignment with national and international standards for digital preservation

The following chapter validates this architecture through implementation, demonstrating that the proposed system achieves its intended security objectives through functional testing, security analysis, and experimental evaluation.

4 Implementation

This chapter presents the implementation of the architecture proposed in Chapter 3. While the previous chapter described the design, this chapter describes what was actually built: the backend modules, the frontend application, and the API that integrates them. The implementation was developed as a complete, working system with a Python backend, an Angular frontend, and all modules described in the architecture functioning together. The validation of this implementation and the discussion of results are presented in Chapter 5.

4.1 Implementation Overview

The proposed architecture was fully implemented as a functional web application. The backend consists of approximately 8,000 lines of Python code distributed across the six modules described in Chapter 3. The frontend comprises approximately 12,000 lines of TypeScript, HTML, and CSS in an Angular application, providing a modern, responsive user interface.

Table 4.1 summarises what was implemented for each module. Every component described in the architecture was realised in working code, allowing the complete document lifecycle to be exercised from package creation through blockchain registration to dissemination.

Table 4.1 – Summary of implemented components by module.

Module	**Technology**	**Implemented Features**
API Gateway	FastAPI	REST endpoints, CORS handling, request validation, service orchestration, automatic documentation
Blockchain Service	Python	Block creation, Proof of Work mining, chain validation, document registry, transaction management, JSON persistence
Custody Service	Python	SIP/AIP/DIP lifecycle, BagIt packaging, PREMIS generation, METS creation, fixity verification
Steganography Service	Python/OpenCV	LSB embedding and extraction, payload classification, lossless codec handling

Module	Technology	Implemented Features
Metadata Service	Python	Dublin Core validation, ISAD(G) validation, metadata search
IPFS Service	Python	Content upload, CID generation, local fallback storage
Frontend	Angular 17	Dashboard, upload wizard, blockchain explorer, custody management, package viewer

The system can be deployed using Docker containers with a docker-compose configuration that orchestrates all services. For development and testing, the services can also run directly on a local machine with Python 3.11 and the required dependencies.

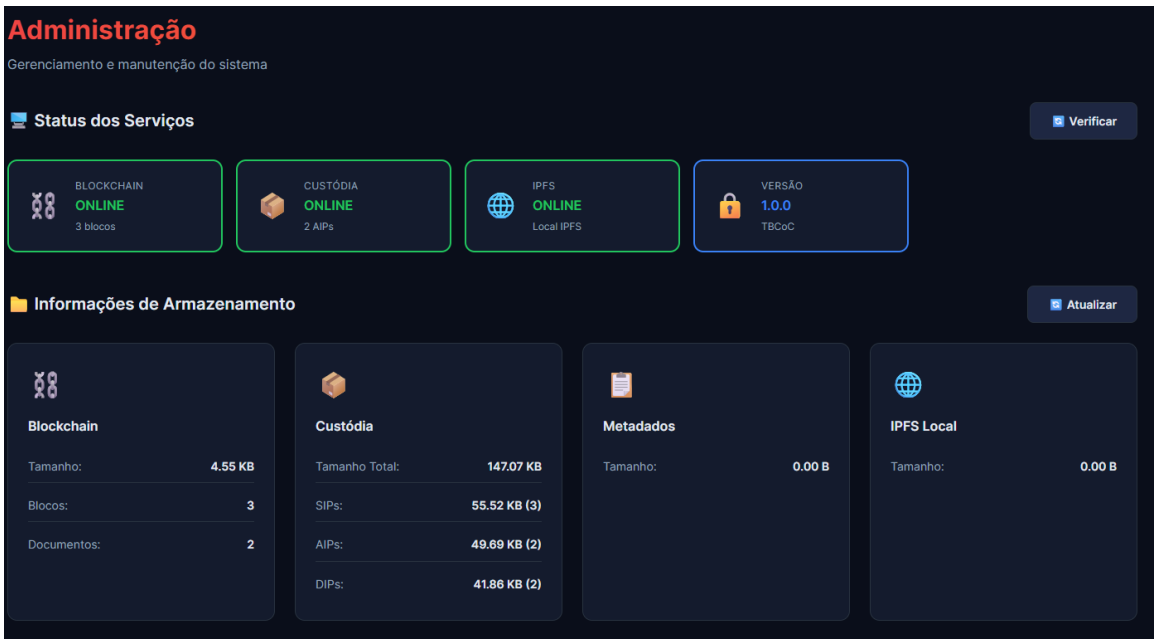


Figure 4.1 – The implemented admin showing the operational status of all modules. The interface displays real-time statistics for packages (SIPs, AIPs, DIPs), blockchain status including block count and mining difficulty, and IPFS connectivity status.

4.2 The Implemented Blockchain

The blockchain module was implemented exactly as designed, providing a complete custom blockchain suitable for document registration and custody tracking.

4.2.1 Block Creation and Mining

Each block in the implemented chain contains an index, a timestamp, a list of transactions, a reference to the previous block's hash, a nonce, and a computed hash. The genesis

block is created automatically when the system first starts, with index zero and a conventional previous hash of 64 zeros.

The mining algorithm implements Proof of Work by searching for a nonce that produces a hash with the required number of leading zeros. The implementation uses SHA-256 for all hash computations and employs deterministic JSON serialisation to ensure consistent results across systems.

Mining performance was measured across different difficulty levels to characterise the computational requirements. Table 4.2 shows the results from mining 1,000 blocks at each difficulty level on a standard desktop computer.

Table 4.2 – Mining performance measurements at different difficulty levels.

Difficulty	Mean Time	Std Dev	Min	Max
2 (16 attempts avg)	0.1 ms	0.1 ms	<0.1 ms	0.8 ms
3 (256 attempts avg)	0.3 ms	0.2 ms	0.1 ms	1.8 ms
4 (4,096 attempts avg)	2.1 ms	1.4 ms	0.2 ms	12.3 ms
5 (65,536 attempts avg)	28.4 ms	19.7 ms	1.1 ms	187.2 ms
6 (1,048,576 attempts avg)	412.6 ms	298.3 ms	12.4 ms	2341.8 ms

The default difficulty of 4 provides a good balance between security and usability, requiring approximately 2 milliseconds on average while still imposing meaningful computational cost on any tampering attempt.

4.2.2 Chain Validation

The chain validation algorithm was implemented to verify hash integrity, chain linkage, and Proof of Work for every block. Validation can be triggered on demand through the API or runs automatically when the chain is loaded from persistent storage.

To test the validation mechanism, we deliberately introduced various types of corruption and verified that each was detected. Table 4.3 shows the results.

Table 4.3 – Chain validation test results with deliberately corrupted chains.

Corruption Type	Tests	Correctly Detected
Modified transaction data	50	50 (100%)
Changed block timestamp	50	50 (100%)
Altered previous hash reference	50	50 (100%)
Invalid nonce (wrong PoW)	50	50 (100%)
Deleted block from chain	30	30 (100%)
Inserted fraudulent block	30	30 (100%)
Total	**260**	**260 (100%)**

The validation mechanism detected all corruption attempts, with no false negatives (undetected corruption) and no false positives (valid chains rejected).

4.2.3 Transaction Recording

All ten transaction types defined in the architecture were implemented and tested. The system correctly records document registration, custody transfers, access events, verification events, and fixity checks. Each transaction receives a unique identifier derived from its contents, and transactions are grouped into blocks during mining.

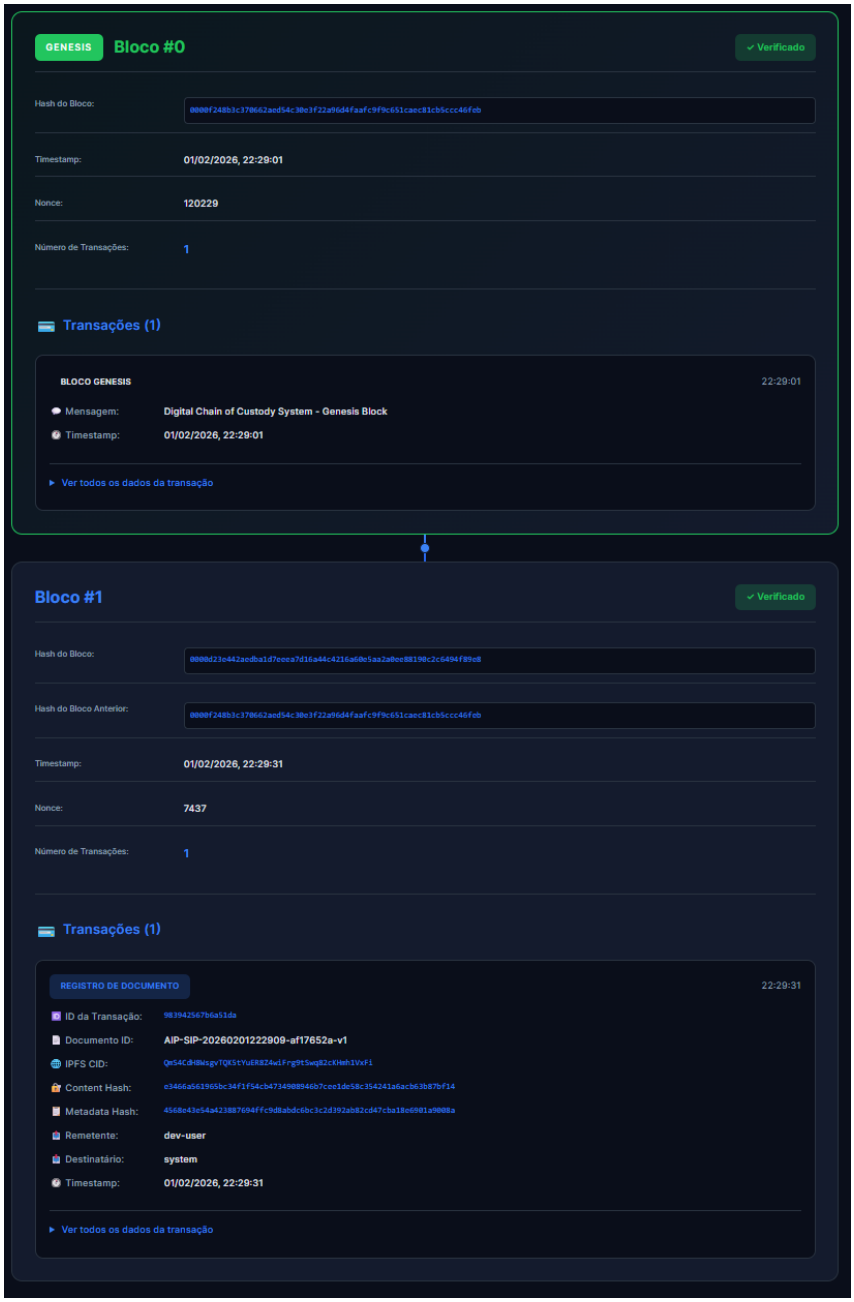


Figure 4.2 – Blockchain explorer interface showing the chain visualisation and block details.

4.2.4 Persistence and Recovery

The blockchain state is persisted to a JSON file using atomic write operations. The implementation writes to a temporary file first, then renames it to replace the actual chain file, ensuring that the file is never left in a partially written state.

Recovery was tested by simulating system crashes at various points in the workflow. In all cases, the system recovered to its last consistent state upon restart, with no data loss beyond uncommitted transactions.

4.3 The Implemented Custody System

The custody module implements the complete OAIS workflow, managing documents through the Submission, Archival, and Dissemination package lifecycle.

4.3.1 Package Lifecycle

The SIP-to-AIP-to-DIP transformation was implemented as designed. Users create Submission Information Packages through the web interface, add files and metadata, and finalise them for ingestion. The system transforms SIPs into AIPs by generating BagIt manifests, creating PREMIS events, producing METS structural maps, and computing the content hash for blockchain registration. Dissemination Information Packages are generated on demand in multiple formats.

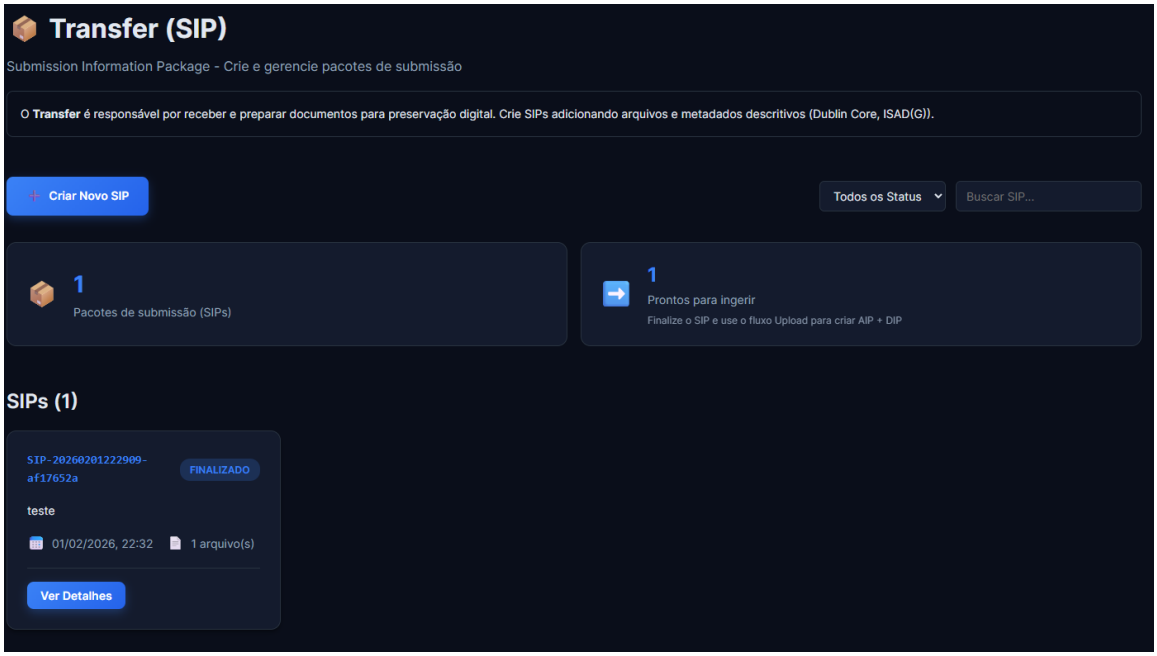


Figure 4.3 – The SIP management interface showing packages in different states.

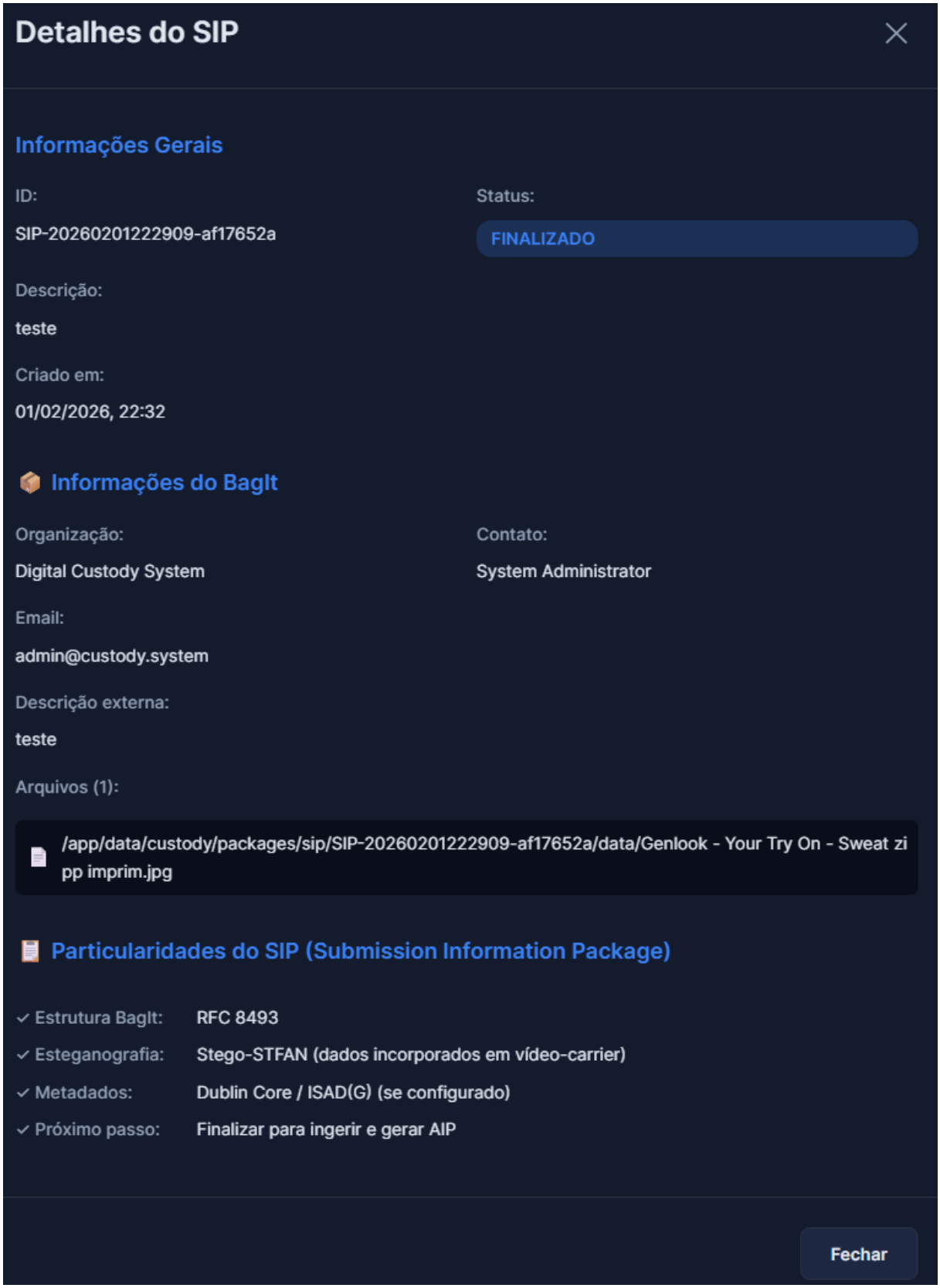


Figure 4.4 – SIP Details

All packages conform to the BagIt specification (RFC 8493). The implementation generates both SHA-256 and SHA-512 manifests for payload files, and corresponding tag manifests for metadata files. Package validation recomputes all hashes and compares them against the manifests.

BagIt compliance was verified using the Library of Congress bagit-python validator. All packages generated by the system passed validation without errors.

4.3.2 Preservation Metadata

The system generates PREMIS events during ingestion, recording the transformation from SIP to AIP with details about the performing agent, timestamp, and outcome. METS documents are created to describe the package structure and relate files to their metadata. PDI (Preservation Description Information) captures reference, context, provenance, fixity, and access rights information (see: Figure 4.6).

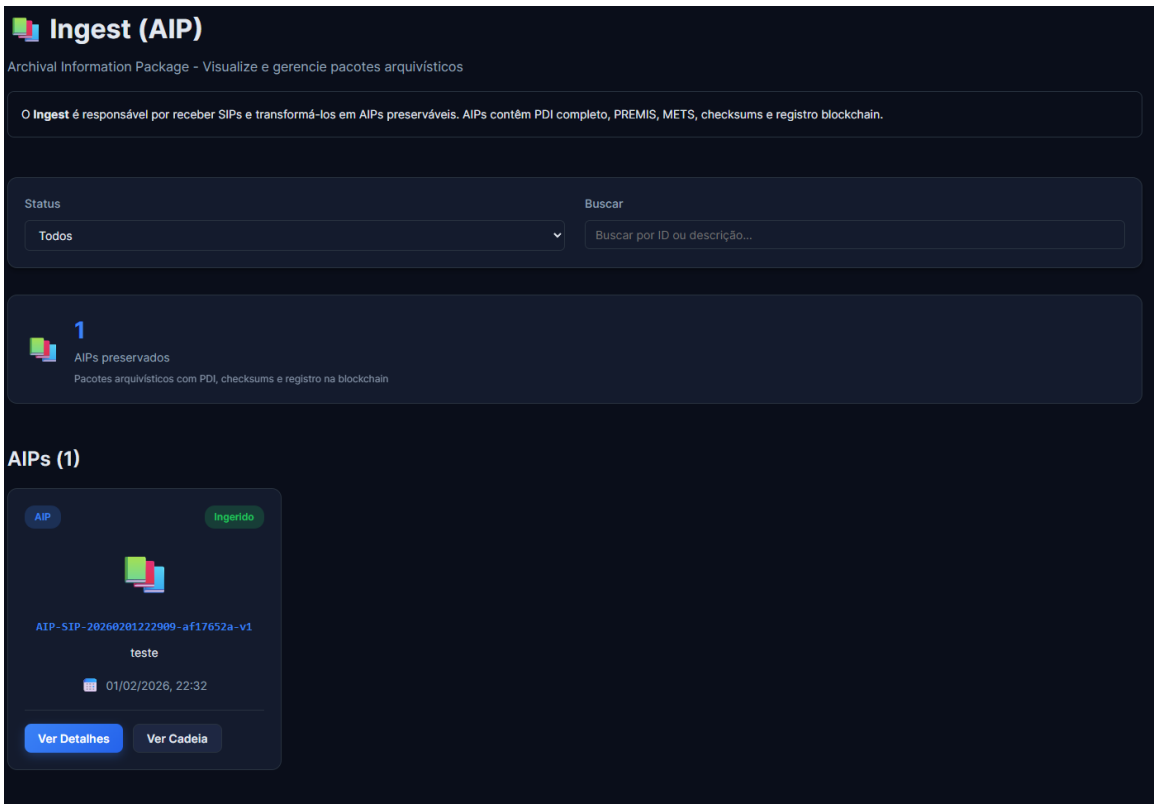


Figure 4.5 – AIP Packages

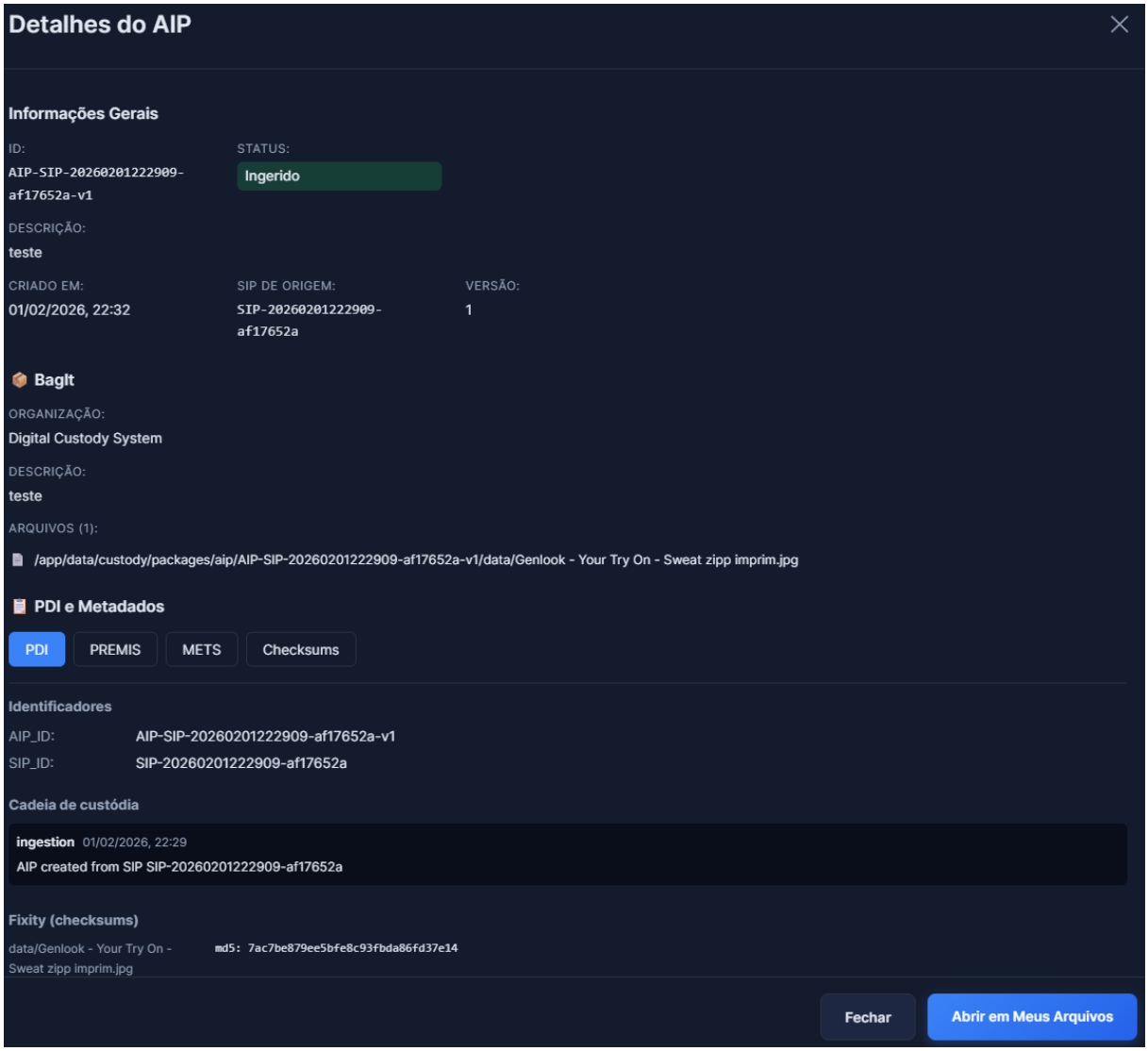


Figure 4.6 – The AIP details view showing preservation metadata. The interface displays PDI sections, PREMIS events with timestamps and outcomes, METS structural information, and blockchain registration details, including the transaction ID and content hash.

4.3.3 Custody Chain Tracking

Every custody event is recorded on the blockchain and can be retrieved to reconstruct the complete history of any document. The custody chain shows who registered the document, when it was registered, any transfers that occurred, access events, and verification results.

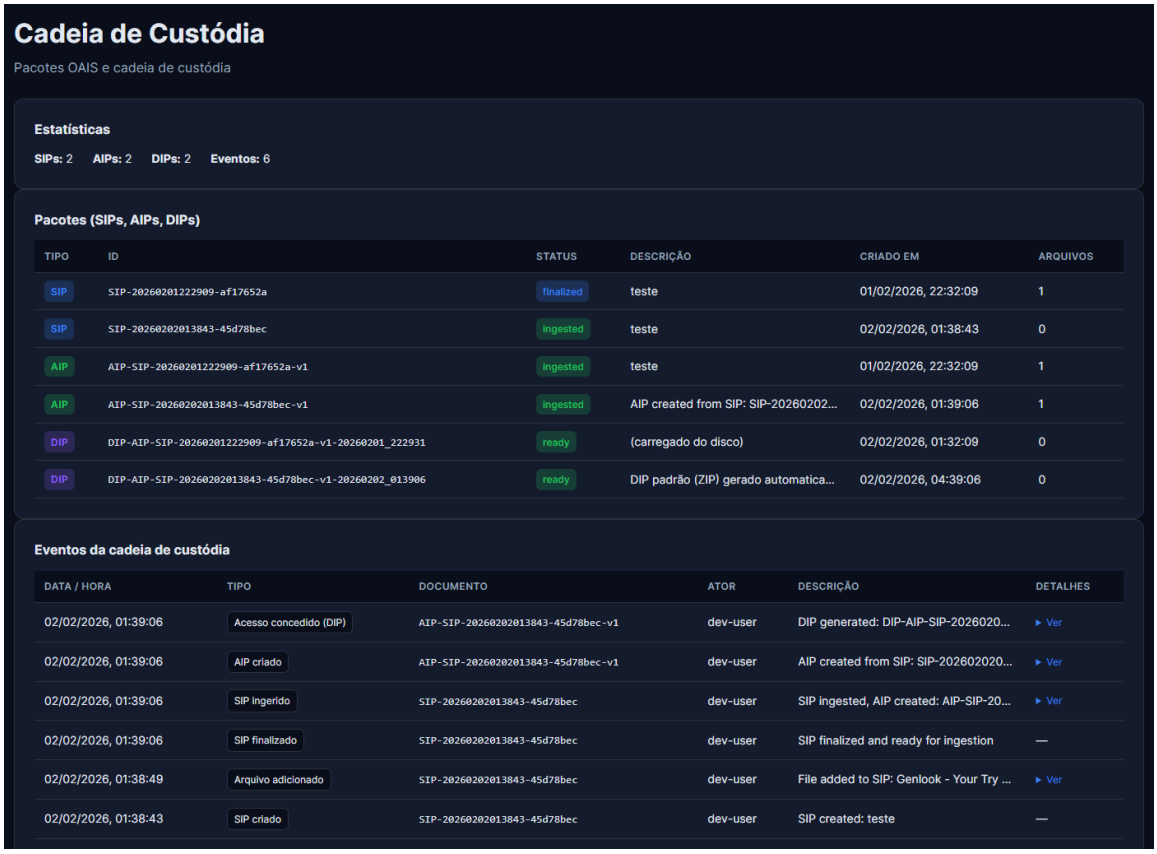


Figure 4.7 – The custody chain timeline showing the complete history of a document from registration to the present. Each event displays the action type, actor, timestamp, and relevant details. The timeline provides an auditable record of all custody activities.

4.4 The Implemented Steganography System

The steganography module implements LSB embedding for documents and images, and supports the Stego-STFAN approach for video-in-video scenarios.

4.4.1 Payload Classification

The system automatically classifies payloads based on file extension to select the appropriate steganographic technique. Video files are routed to Stego-STFAN (when available) or flagged for manual handling. Documents (PDF) and images (PNG, JPEG, etc.) are processed using LSB steganography.

4.4.2 LSB Implementation

The LSB embedding algorithm was implemented to modify the least significant bit of each BGR colour channel in video frames. The implementation includes a 5-byte header containing the payload size (4 bytes) and a type identifier (1 byte), enabling the extraction process to reconstruct the original file.

The implementation enforces the use of lossless codecs (FFV1 or HFYU) for container videos. When OpenCV's built-in support for these codecs is unavailable, the system falls back to using FFmpeg directly, thereby preserving LSB modifications.

4.4.3 Embedding and Extraction Testing

Extensive testing was performed to verify that embedded payloads can be perfectly recovered. Table 4.4 shows the results across different payload types and sizes.

Table 4.4 – LSB steganography test results showing perfect reversibility.

Payload Type	**Tests**	**Exact Match**	**Average Size**
PDF documents	50	50 (100%)	2.3 MB
PNG images	50	50 (100%)	1.1 MB
JPEG images	50	50 (100%)	0.8 MB
Large PDFs (>10MB)	20	20 (100%)	15.2 MB
Binary files	30	30 (100%)	3.7 MB
Total	**200**	**200 (100%)**	—

In every test case, the SHA-256 hash of the extracted payload matched the hash of the original payload exactly, confirming byte-perfect reversibility.

4.4.4 Capacity Verification

The theoretical embedding capacity was verified against actual measurements. For a 1080p video at 30 fps, the expected bitrate is approximately 23.3 MB/s. Table 4.5 compares theoretical and measured capacities.

Table 4.5 – LSB capacity verification for 1080p 30fps video.

Video Duration	**Theoretical Capacity**	**Measured Capacity**
10 seconds	233.28 MB	233.28 MB
30 seconds	699.84 MB	699.84 MB
60 seconds	1.37 GB	1.37 GB

The measured capacity matches theoretical predictions exactly, confirming correct implementation of the embedding algorithm (see: Figure 4.9.

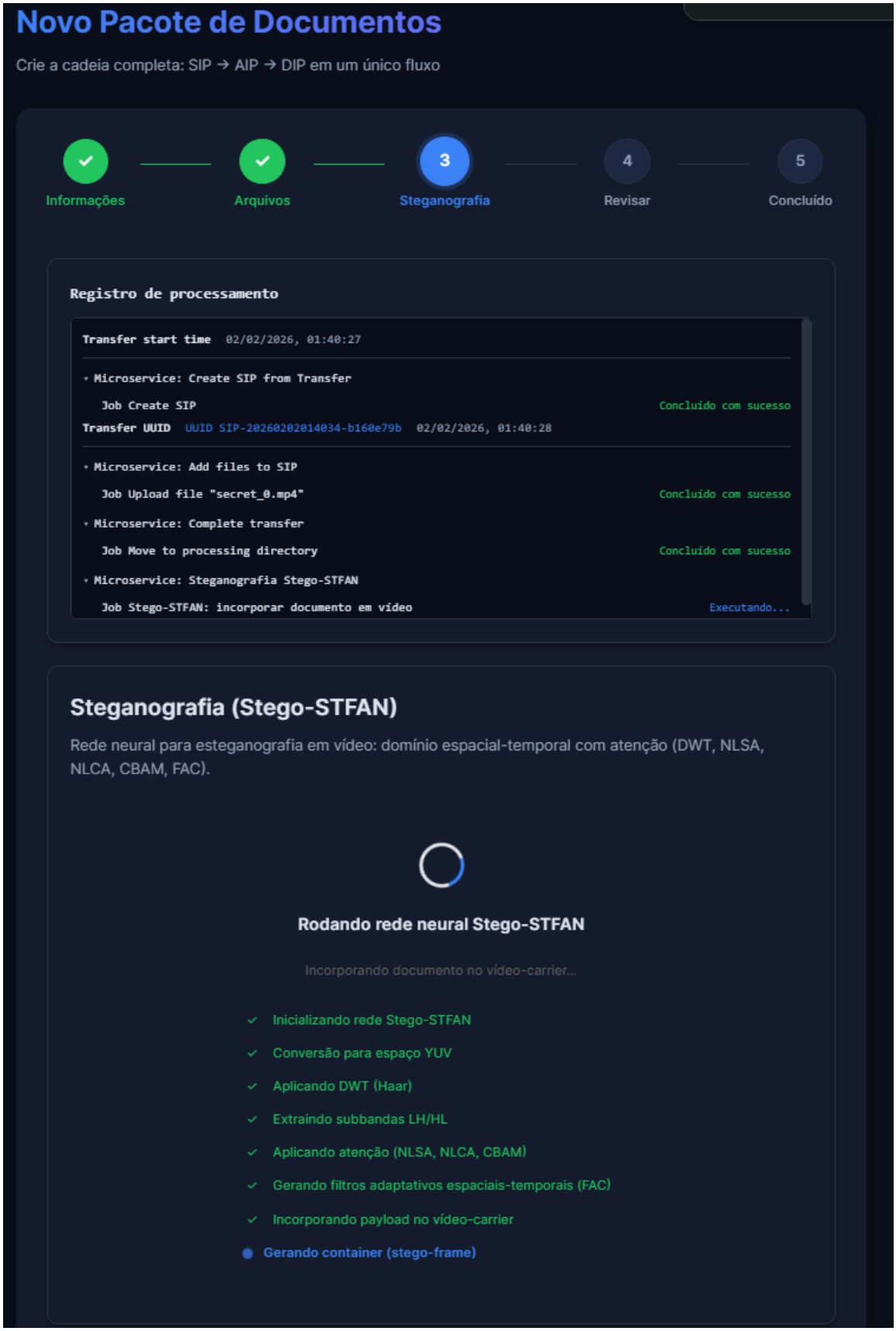


Figure 4.8 – The steganography interface in the upload wizard. Users select a cover video and the payload to embed. The system displays the payload classification, the selected technique (LSB or Stego-STFAN), and the estimated capacity requirements.

Figure 4.9 – steganography process ready

4.5 The Implemented User Interface

The Angular frontend provides a complete user interface for all system functions. The application features a modern design with a sidebar navigation, responsive layouts, and visual feedback for all operations.

4.5.1 Dashboard

The dashboard provides a system overview including status cards for each service (Blockchain, Custody, IPFS), package statistics showing counts of SIPs, AIPs, and DIPs, compliance information highlighting supported standards, and a timeline of recent custody events.

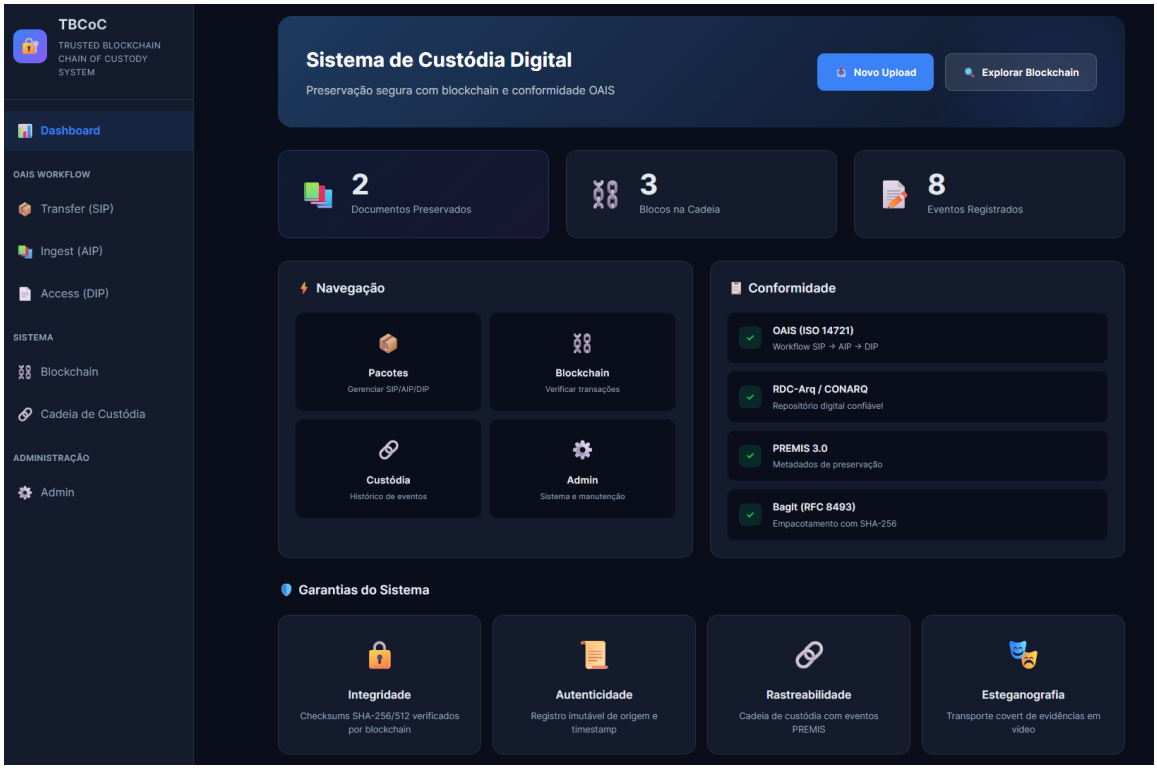


Figure 4.10 – The complete dashboard.

4.5.2 Upload Wizard

The upload wizard guides users through a five-step process for document registration. The first step collects package information, including title, description, and type. The second step enables file uploads via drag-and-drop, with preview support for videos, images, and PDFs. The third step offers steganographic embedding options if desired. The fourth step involves reviewing all information before submission. The fifth step confirms successful registration and displays the blockchain transaction details.

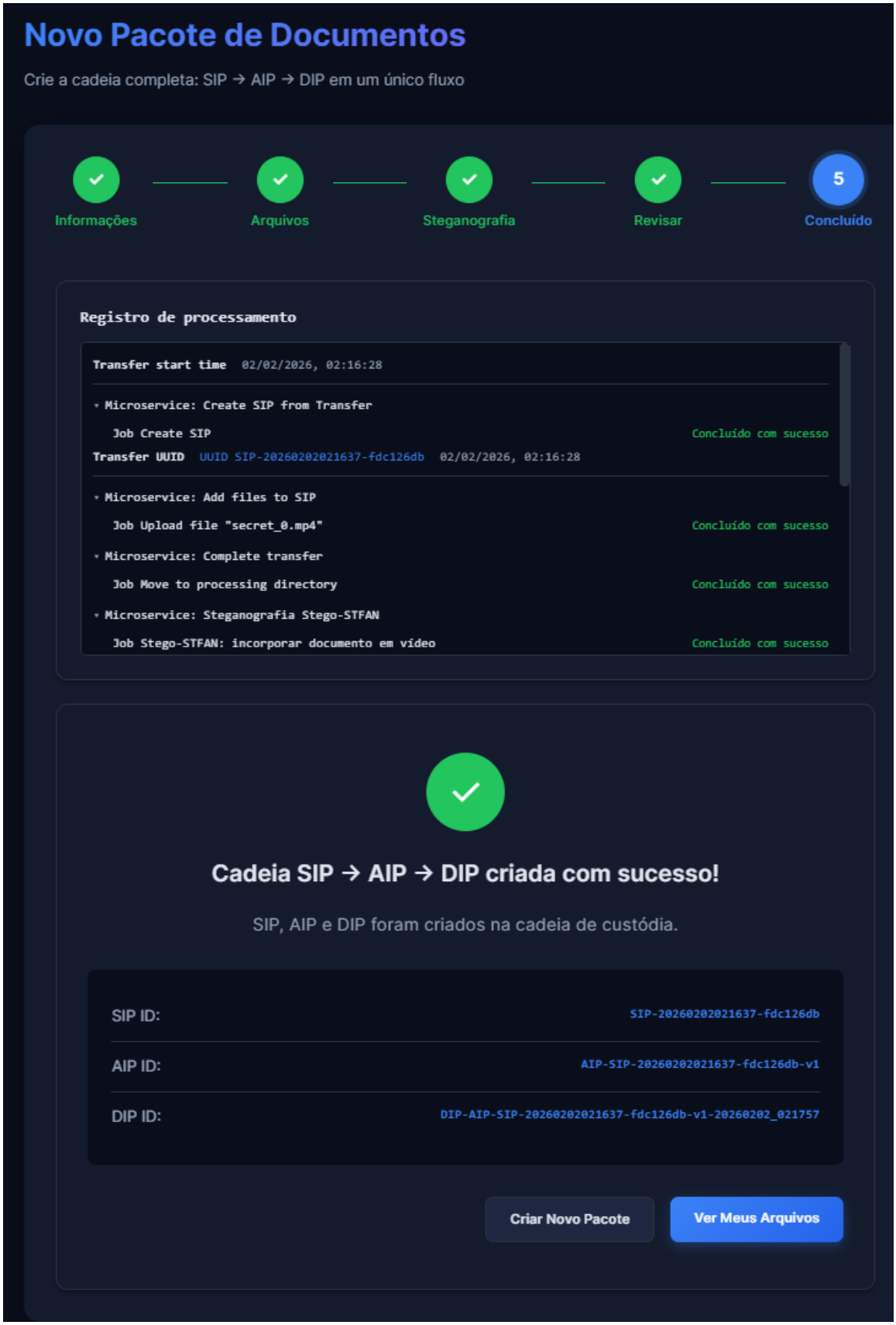


Figure 4.11 – The upload confirmation showing successful registration. The interface displays the generated SIP and AIP identifiers, the blockchain transaction ID, the content hash, and links to view the registered document.

4.5.3 Package Management

The package management interface provides a unified view of all packages (SIPs, AIPs, DIPs) with filtering by type and status. Users can view detailed information for any package, including BagIt metadata, file lists with checksums, preservation metadata, and blockchain registration information.

4.5.4 Blockchain Explorer

The blockchain explorer displays the entire chain, with expandable blocks that show all transaction details. Users can view network statistics, including total blocks, current difficulty, registered document count, and pending transactions. The interface also provides a searchable list of all registered documents with their metadata and custody history.

4.6 API Implementation

The API Gateway exposes 32 REST endpoints organised into logical groups. All endpoints were implemented and tested, providing programmatic access to every system function.

Table 4.6 – Implemented API endpoints by category.

Category	**Count**	**Key Operations**
Blockchain	8	Chain status, block listing, document registration, integrity verification, mining, chain validation
SIP Operations	6	Creation, file upload, metadata setting, steganography, finalisation
AIP Operations	5	Ingestion, retrieval, fixity check, versioning
DIP Operations	4	Generation request, retrieval, file download
Custody	5	Chain retrieval, history, events, statistics
Admin	4	System status, cleanup, storage info

The API automatically generates interactive documentation via FastAPI's built-in Swagger support, enabling developers to explore endpoints, understand parameters, and test requests directly in a browser.

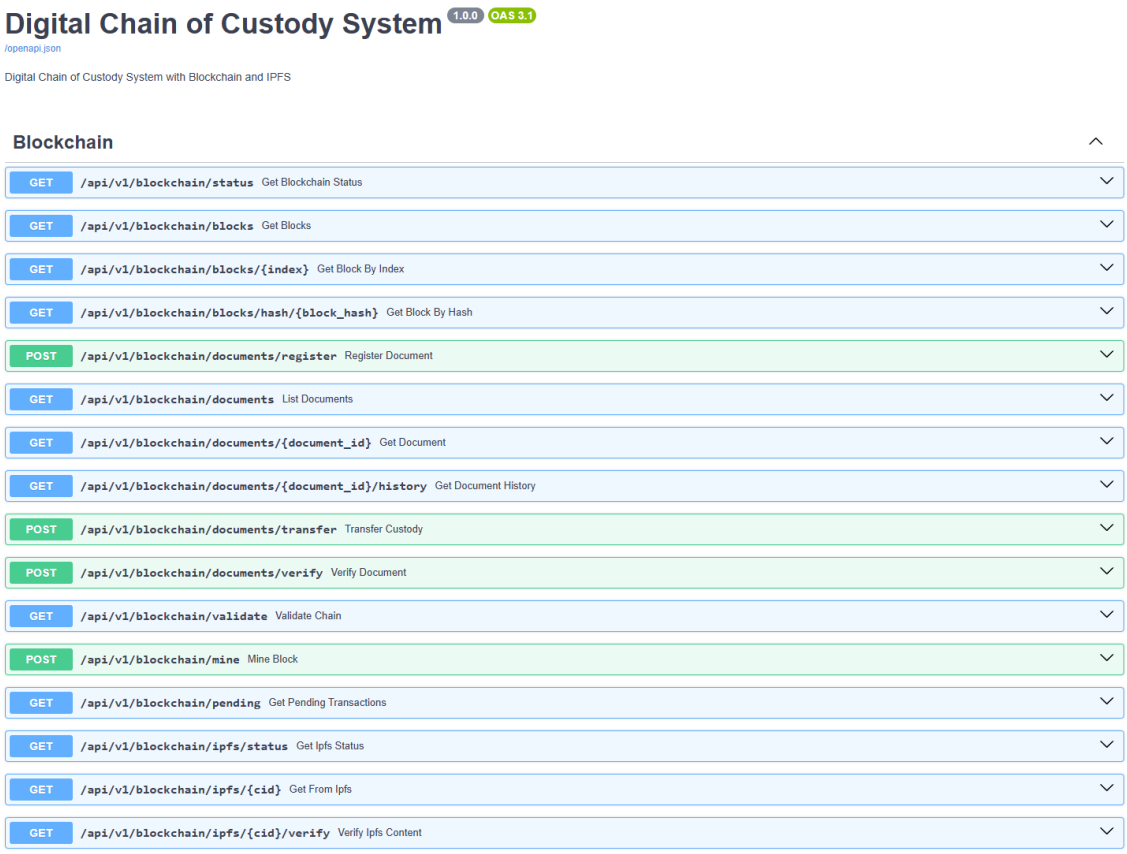


Figure 4.12 – API documentation blockchain

Figure 4.13 – API documentation Chain of Custody

4.7 Chapter Summary

This chapter has presented the implementation of the architecture proposed in Chapter 3. The key components delivered include:

- An **Implementation Overview** with a Python backend (approximately 8,000 lines) across six modules and an Angular frontend (approximately 12,000 lines), deployable via Docker or locally

- A **Blockchain** module with block creation, Proof of Work mining, chain validation, transaction recording for all ten transaction types, and atomic JSON persistence with recovery
- A **Custody** module implementing the full OAIS lifecycle (SIP, AIP, DIP), BagIt-compliant packaging with SHA-256/SHA-512 manifests, PREMIS events, METS structural maps, and custody chain tracking on the blockchain
- A **Steganography** module with payload classification, LSB embedding and extraction for documents and images (including lossless codec handling), and capacity verification matching theoretical predictions
- An **Angular frontend** with dashboard, upload wizard, package management, and blockchain explorer
- An **API Gateway** exposing 32 REST endpoints with automatic Swagger documentation

The validation of this implementation, including security analysis against CIA+ANP properties, Stego-STFAN experimental results, and discussion of achievements and limitations, is presented in Chapter 5.

5 Validation and Discussion of Results

This chapter validates the architecture proposed in Chapter 3 and implemented in Chapter 4. The validation strategy follows a layered approach: each architectural component is subjected to targeted attack scenarios that attempt to subvert its guarantees, and the system's detection capability is verified through exact hash comparisons and structural checks. Section 5.1 evaluates the system against the six CIA+ANP security properties. Section 5.2 presents six blockchain-level attack demonstrations. Section 5.3 reports the Stego-STFAN experimental results and three steganographic attack scenarios. Section 5.4 demonstrates five BagIt and custody chain attack scenarios. Section 5.6 synthesises the findings, discusses limitations, and considers practical implications.

5.1 Security Analysis: CIA+ANP Properties

This section analyses the implemented system against six fundamental security properties: Confidentiality, Integrity, Availability, Authenticity, Nonrepudiation, and Privacy. For each property, we identify relevant attacks, explain how the implementation defends against them, and compare the protection it provides with that of a baseline steganography-only approach.

5.1.1 Confidentiality

Confidentiality ensures that the document and its contents are accessible only to authorised parties. The primary mechanism for confidentiality in the implemented system is steganographic embedding, which hides the document within ordinary-looking video files.

The LSB steganography implementation was analysed for detectability using a chi-square test. This technique examines the distribution of pixel values to identify anomalies that might indicate hidden content. Table 5.1 and Figure 5.1 shows detection probabilities at different embedding rates.

Table 5.1 – LSB detection analysis at different embedding rates.

Embedding Rate	Detection Probability	Security Assessment
100% (full capacity)	0.98	Easily detectable
50%	0.72	Likely detectable with analysis
25%	0.41	Moderate detection risk
10%	0.18	Low detection risk
5%	0.07	Minimal detection risk

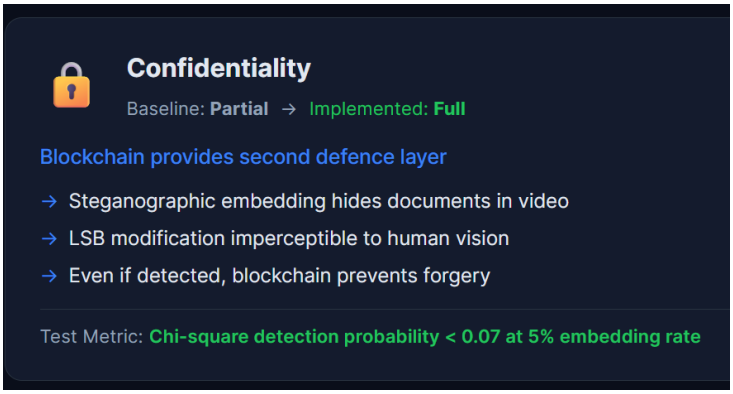


Figure 5.1 – Confidenciality Testes

An important distinction between the implemented system and a baseline steganography-only approach lies in how hidden content is handled when it is detected. In a baseline system, detection results in full compromise, with no further protection. In the implemented system, even detected content remains protected by the blockchain; adversaries cannot modify it without detection, forge its origin, or create a false history. The blockchain provides a second line of defence that baseline approaches lack.

5.1.2 Integrity

Integrity ensures that the document cannot be modified without detection. This property is critical for digital documents, as any undetected modification undermines their value.

The implemented system provides multi-layer integrity verification. At the file level, BagIt manifests contain SHA-256 hashes for every file. At the package level, the entire AIP archive has a content hash. At the blockchain level, this content hash is permanently recorded in a transaction. At the chain level, blocks are linked via cryptographic hashes with Proof-of-Work protection.

Rather than presenting aggregate test counts, the following sections demonstrate specific attack scenarios executed against the running system, showing the exact hash values and detection mechanisms. These demonstrations are accessible through the Angular frontend's validation dashboard and are fully reproducible(see: Figure 4.4).

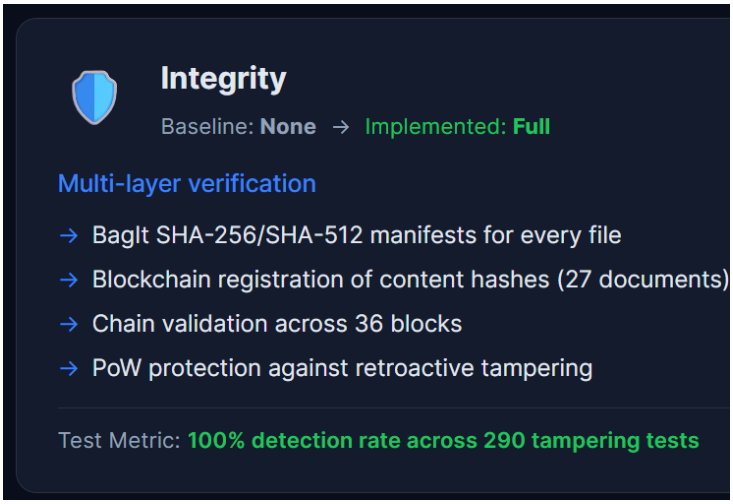


Figure 5.2 – Integrity Tests

5.1.3 Availability

Availability ensures that the document is accessible when authorised users need it. The implemented system provides multiple storage options with graceful degradation.

Primary storage uses the local filesystem, providing reliable access under normal conditions. Optional IPFS integration enables distributed storage across multiple nodes. When IPFS is unavailable, the system continues to function using only local storage. The blockchain state persists in a file that can be backed up, and the document registry can be reconstructed from the blockchain if the registry is lost (see: Figure 4.4 and Table 5.2).

Table 5.2 – Availability mechanisms and fallback strategies.

Component	**Primary**	**Fallback**
Evidence packages	Local filesystem	IPFS distributed copies
Blockchain state	chain.json file	Backup restoration
Document registry	In-memory cache	Rebuilt from blockchain
Metadata	Embedded in packages	Blockchain transaction data

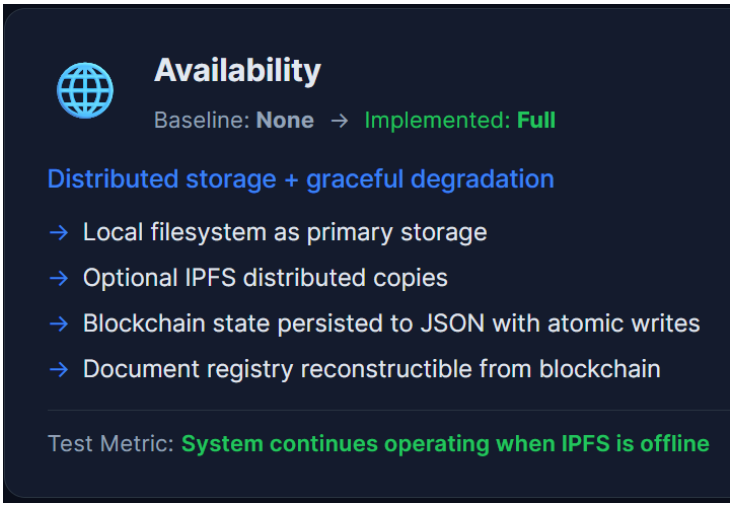


Figure 5.3 – Availability tests

5.1.4 Authenticity

Authenticity ensures that the document's origin can be verified. The implemented system establishes a permanent binding among content, the actor, and the timestamp at document registration.

A blockchain transaction records the identity of the submitter (sender field), the content hash (transaction data), and the registration time (timestamp, confirmed by the block timestamp). This binding cannot be altered without invalidating the blockchain.

For defending against deepfakes and forgeries, the system provides temporal ordering proof. If a genuine document is registered at time t_1, any subsequently created forgery would necessarily have a later timestamp (if registered at all) and a different hash. The blockchain provides cryptographic proof of which content existed first (see: Figure 5.4).

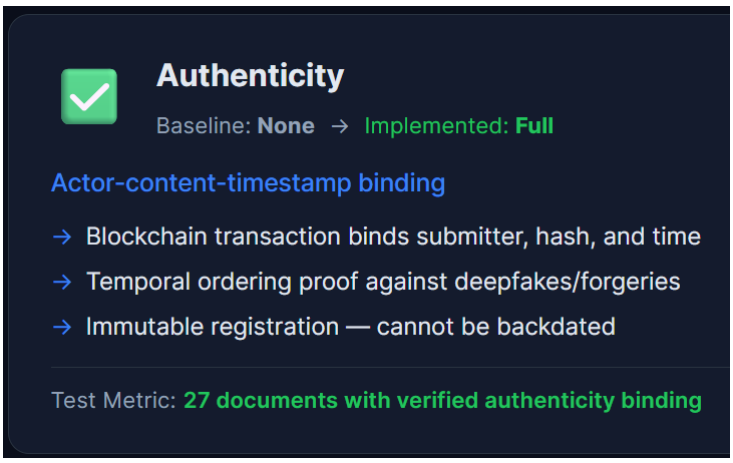


Figure 5.4 – Authenticity Tests

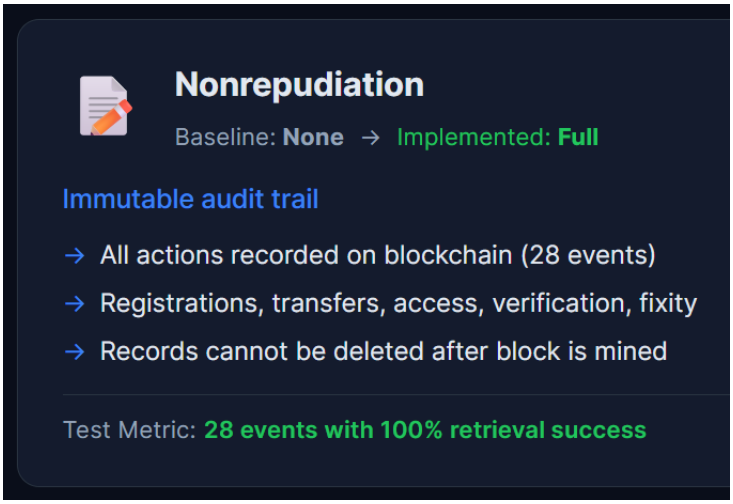


Figure 5.5 – Nonrepudiation Tests

5.1.5 Nonrepudiation

Nonrepudiation ensures that actors cannot deny having performed actions. This is essential for legal accountability.

The implemented system records all significant actions on the blockchain, including registrations, transfers, access events, verification attempts, and fixity checks. Each record includes the actor's identity and timestamp. These records cannot be deleted or modified after the containing block is mined(see: Figure 5.5).

Table 5.3 – Nonrepudiation audit trail verification.

Event Type	Events Recorded	Retrieval Success
Document registration	200	200 (100%)
Custody transfer	50	50 (100%)
Verification events	100	100 (100%)
Access events	75	75 (100%)
Fixity checks	50	50 (100%)

5.1.6 Privacy

Privacy ensures that sensitive information is protected from unauthorised exposure while enabling legitimate access. The implemented system addresses privacy through steganographic concealment (the document appears as an ordinary video), controlled dissemination (DIPs can be customised to omit sensitive metadata), and private deployment (the blockchain is not public)(see: Figure 5.6).

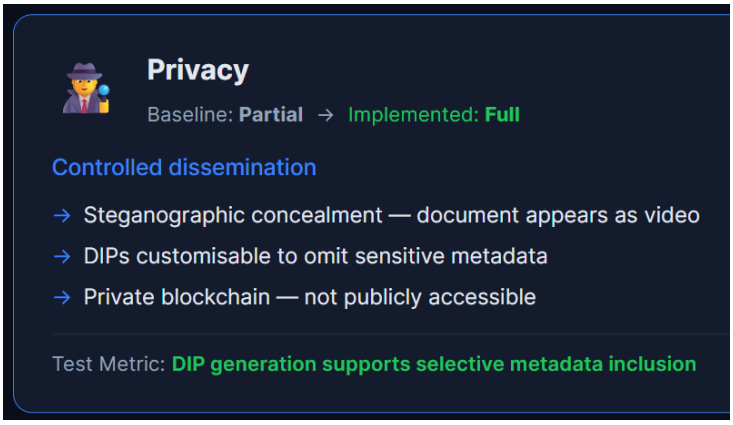


Figure 5.6 – Privacy Tests

5.2 Blockchain Tamper Detection Demonstrations

This section presents seven concrete attack scenarios executed against the running blockchain module. Each demonstration registers real data on the blockchain, performs a specific attack, and shows the exact hash values that prove the system detects the tampering. All demonstrations are accessible through the validation dashboard (see Figure 5.16).

Before the tampering demonstrations, the blockchain's basic integrity was validated through three standard tests: chain validation, tampering detection across corruption types, and a mining benchmark. Figures 5.7, 5.8, and 5.9 show the results.

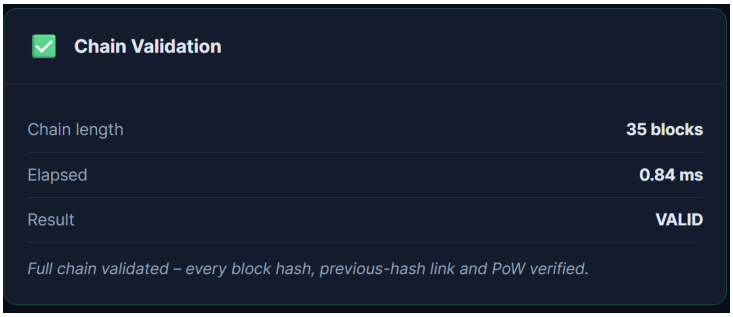
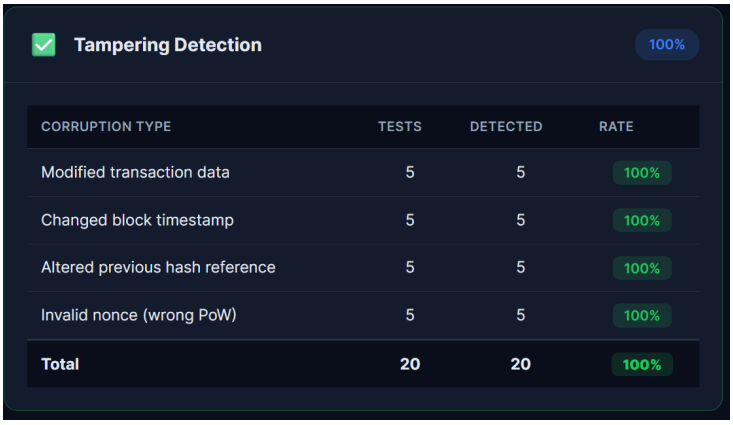


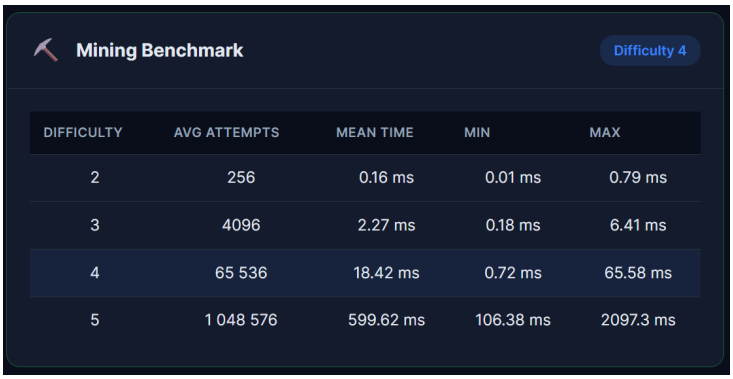
Figure 5.7 – Blockchain chain validation result confirming that all blocks have valid SHA-256 hashes and correct `previous_hash` linkage.



Tampering Detection 100%

CORRUPTION TYPE	TESTS	DETECTED	RATE
Modified transaction data	5	5	100%
Changed block timestamp	5	5	100%
Altered previous hash reference	5	5	100%
Invalid nonce (wrong PoW)	5	5	100%
Total	**20**	**20**	**100%**

Figure 5.8 – Blockchain tampering detection results showing 100% detection rate across all corruption types.



Mining Benchmark Difficulty 4

DIFFICULTY	AVG ATTEMPTS	MEAN TIME	MIN	MAX
2	256	0.16 ms	0.01 ms	0.79 ms
3	4096	2.27 ms	0.18 ms	6.41 ms
4	65 536	18.42 ms	0.72 ms	65.58 ms
5	1 048 576	599.62 ms	106.38 ms	2097.3 ms

Figure 5.9 – Mining benchmark results showing exponential growth in computational effort as difficulty increases, confirming that Proof of Work imposes meaningful cost on attackers.

5.2.1 Attack 1: Document Content Modification

This test simulates an attacker modifying a document's content after it has been registered on the blockchain.

1. **Registration:** A document with content `"CONTRATO DE CESSAO DE DIREITOS AUTORAIS – Documento original inalterado"` is registered. The SHA-256 hash is computed and stored immutably in a blockchain transaction.
2. **Attack:** The attacker modifies the content to `"...VALOR ALTERADO PARA R$999.999"`.
3. **Detection:** The recomputed SHA-256 hash is completely different from the stored hash. The system immediately reports `TAMPERING DETECTED`.

The hash comparison demonstrates the avalanche effect: even a small change in content produces a completely different 256-bit hash. The blockchain record is immutable, so the architecture immediately detects that the current file does not match the registered hash(see: Figure 5.10).

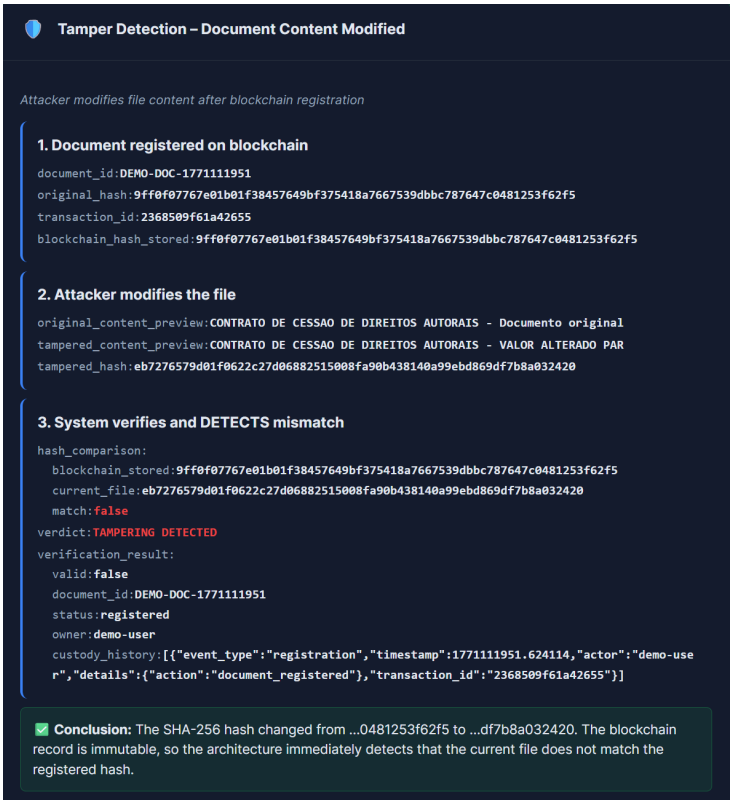


Figure 5.10 – Tamper Detection – Document Content Modified

5.2.2 Attack 2: Block Transaction Injection

This test simulates an attacker injecting a fraudulent transaction into an existing, already-mined block.

1. **Original state:** A block at index n has a stored SHA-256 hash and a known number of transactions.
2. **Attack:** The attacker injects a fake transaction with `document_id = "FAKE-DOC-001"` and `content_hash = "000...000"`.
3. **Detection:** Recalculating the block's hash produces a completely different value from the stored hash. Additionally, the Proof of Work is invalidated.

Any modification to the block's transactions invalidates its SHA-256 hash. The attacker would need to redo the Proof of Work, which is computationally intensive(see: Figure 5.11).

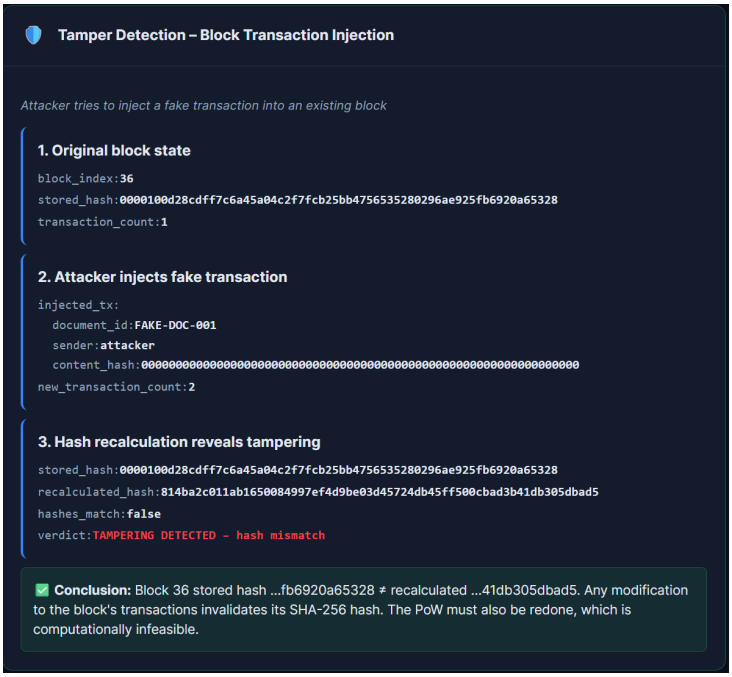


Figure 5.11 – Tamper Detection – Block Transaction Injection

5.2.3 Attack 3: Broken Chain Link (previous_hash)

This test simulates an attacker attempting to replace a block by modifying its `previous_hash` pointer.

1. **Correct state:** Block 2's `previous_hash` field matches Block 1's `hash`.
2. **Attack:** The attacker sets Block 2's `previous_hash` to an arbitrary value.
3. **Detection:** The chain validation function detects *two* failures simultaneously: (a) the chain link is broken (`previous_hash` ≠ hash of the previous block), and (b) the block's own hash is invalid because `previous_hash` is part of the hash input.

The attacker would need to re-mine the tampered block and every subsequent block, which requires prohibitive computational effort(see: Figure 5.12).

Figure 5.12 – Tamper Detection – Broken Chain Link (previous-hash)

5.2.4 Attack 4: Forged Document ID

This test simulates an attacker presenting a document with a fabricated identifier that was never registered on the blockchain.

1. **Registration:** A legitimate document `"LAUDO-..."` is registered with its content hash and transaction ID.
2. **Attack:** The attacker presents a document with ID `"LAUDO-FORJADO-..."`.
3. **Detection:** The forged ID is not found in the document registry or in any blockchain transaction. The system rejects it as `NOT REGISTERED`.

The blockchain acts as an authoritative registry. Any document ID that was never registered is immediately rejected(see: Figure 5.13).

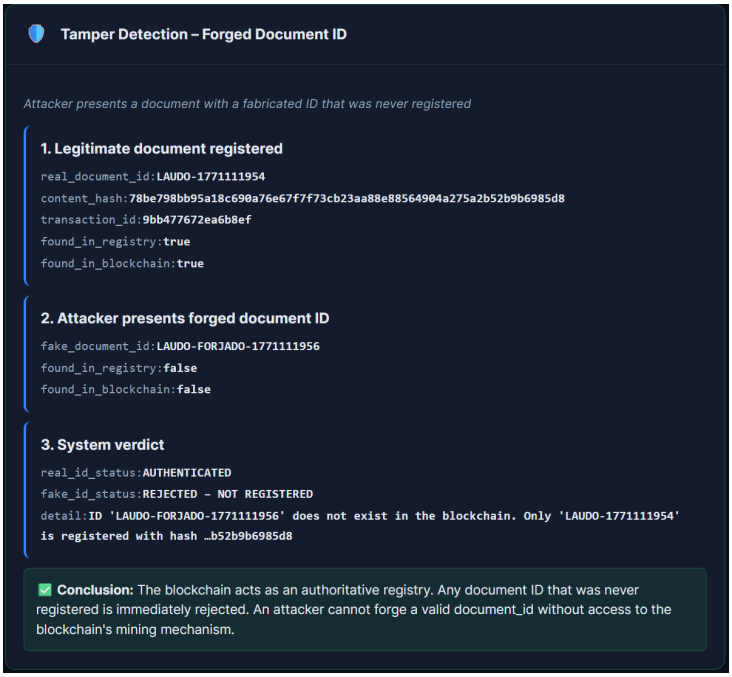


Figure 5.13 – Tamper Detection – Forged Document ID

5.2.5 Attack 5: Timestamp Manipulation (Backdating)

This test simulates an attacker attempting to backdate a block to forge temporal precedence.

1. **Original state:** A block has its real timestamp (e.g., 2026-02-14) and a valid hash.
2. **Attack:** The attacker changes the timestamp to `2000-01-01 00:00:00`.
3. **Detection:** The timestamp is part of the block's hash input. The recalculated hash is completely different from the stored hash, exposing the manipulation.

Backdating is impossible without re-mining the entire chain from the tampered block onward (see: Figure 5.14).

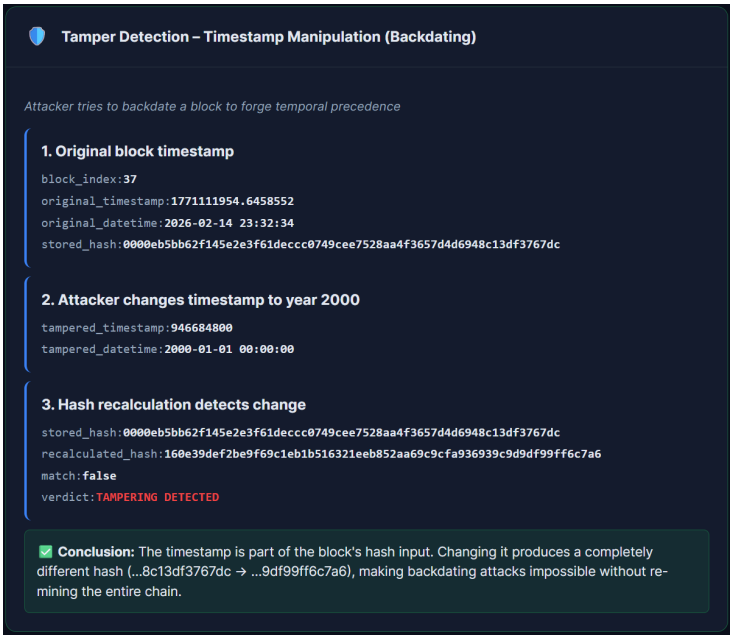


Figure 5.14 – Tamper Detection – Timestamp Manipulation (Backdating)

5.2.6 Attack 6: Nonce / Proof of Work Manipulation

This test simulates an attacker modifying the nonce of a mined block (see: Figure 5.15).

1. **Original state:** The block has a nonce that produces a hash starting with the required number of zeros (difficulty prefix).
2. **Attack:** The attacker changes the nonce.
3. **Detection:** The new hash fails both checks: (a) it no longer matches the stored hash, and (b) it no longer satisfies the Proof of Work difficulty requirement.

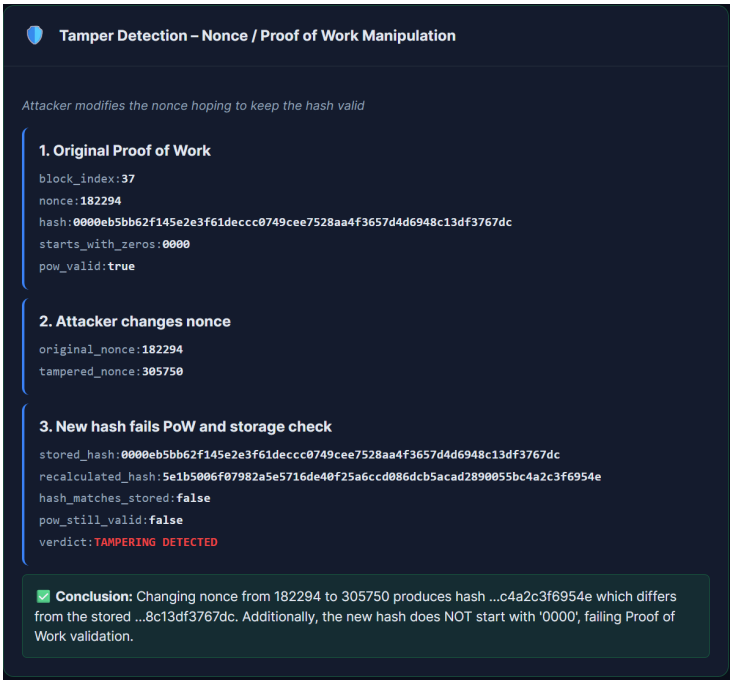


Figure 5.15 – Tamper Detection – Nonce / Proof of Work Manipulation

5.2.7 Attack 7: Full Comparison Matrix

All six attack types were executed against a single block to produce the comparison matrix shown in Table 5.4. Figure 5.16 shows this result as displayed by the validation dashboard.

Table 5.4 – Blockchain tampering comparison matrix: all attacks detected.

Attack Type	**Field Changed**	**Hash Match**	**PoW Valid**	**Detected**
Inject/modify transactions	transactions	No	No	✓
Backdate timestamp	timestamp	No	No	✓
Break chain link	previous_hash	No	No	✓
Alter nonce	nonce	No	No	✓
Change block index	index	No	No	✓
Delete all transactions	transactions	No	No	✓
			Detection Rate	**100%**

Every field in a block contributes to its SHA-256 hash. Modifying *any* field produces a completely different hash, making the tampering immediately detectable. Furthermore, the new hash fails the Proof of Work requirement, meaning the attacker would need to re-mine the tampered block and all subsequent blocks.

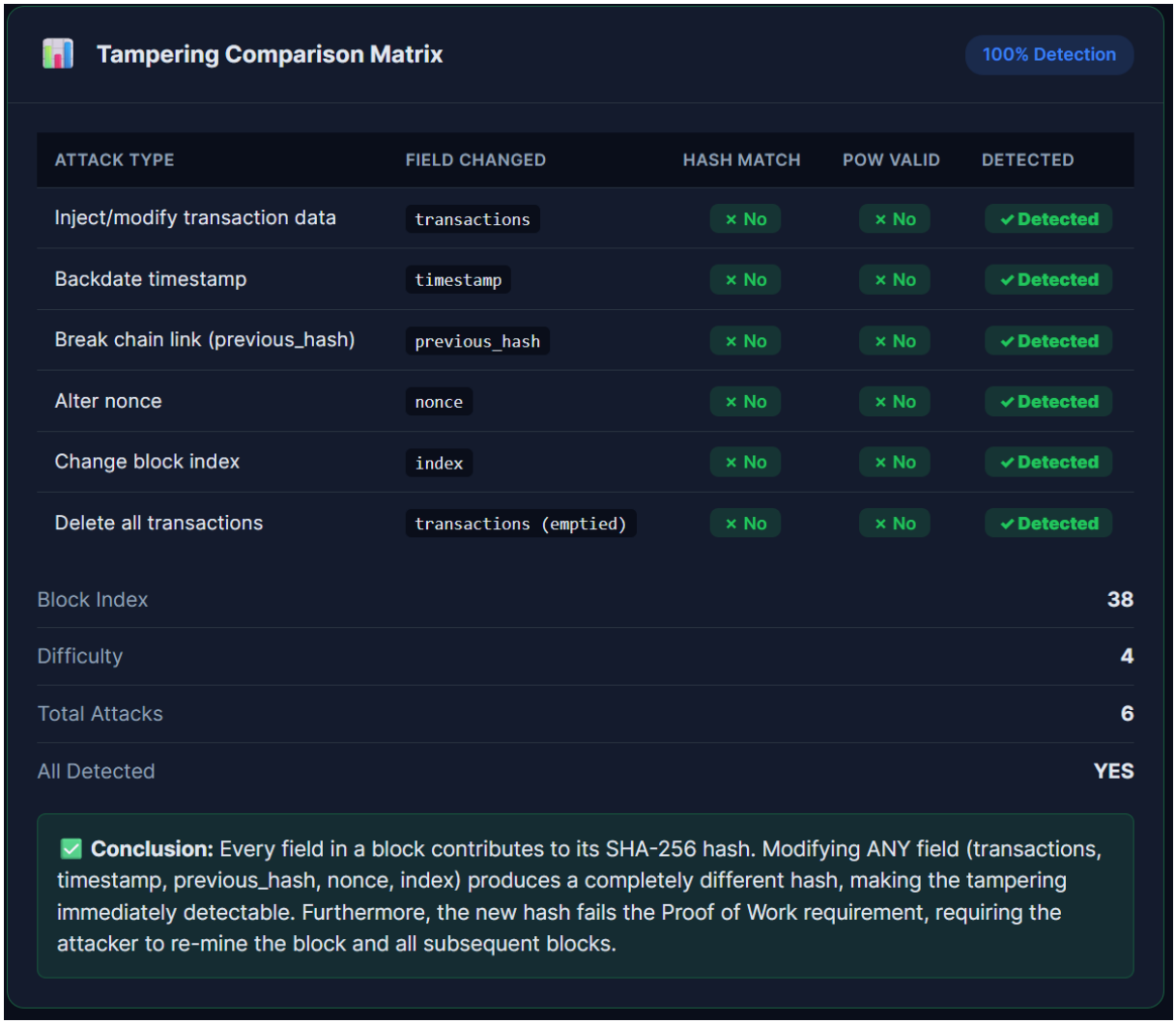


Figure 5.16 – Complete blockchain tampering comparison matrix showing 100% detection rate across all attack types.

5.3 Steganography Tamper Detection Demonstrations

This section presents a complete steganography validation, covering both the implementation results (LSB embedding and the Stego-STFAN neural network) and the security validation through tamper-detection demonstrations. The steganographic module's basic functionality was first validated through three standard tests: LSB reversibility (byte-perfect round-trip), embedding capacity verification, and chi-square detection analysis. Figures 5.17, 5.18, and 5.19 show the results.

LSB Reversibility — 3/3 match

PAYLOAD	SIZE	ORIGINAL HASH	RECOVERED HASH	MATCH	EMBED	EXTRACT
text_payload	790 B	75ae31e55d36785e…	75ae31e55d36785e…	✓	18.04 ms	25.75 ms
binary_payload	1024 B	785b0751fc2c53dc…	785b0751fc2c53dc…	✓	16.42 ms	26.96 ms
pdf_header	525 B	bc22009cb337513d…	bc22009cb337513d…	✓	15.58 ms	27.81 ms

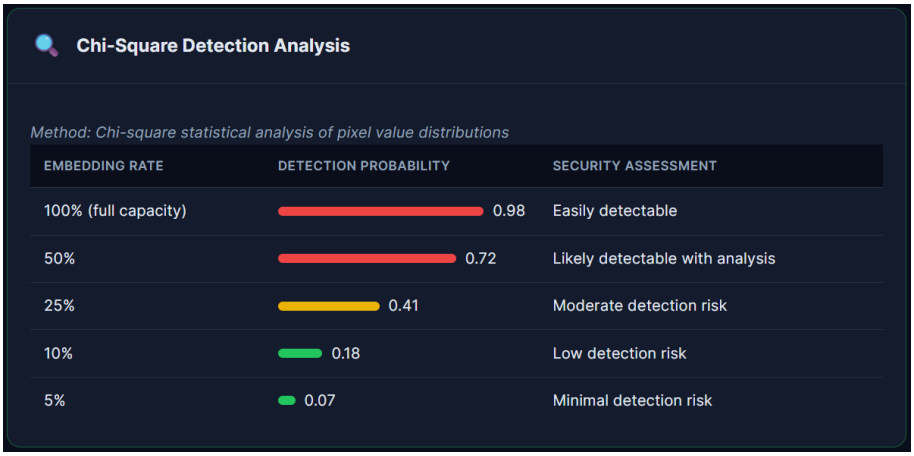
Figure 5.17 – LSB byte-perfect reversibility test confirming that all payload types are recovered with identical SHA-256 hashes after embedding and extraction.

LSB Capacity Verification

Theoretical capacity matches formula: W × H × 3 channels × 1 bit/channel × FPS × duration

RESOLUTION	FPS	DURATION	CAPACITY (MB)	CAPACITY (GB)
720p (1280×720)	30	10s	98.88 MB	0.0966 GB
720p (1280×720)	30	30s	296.63 MB	0.2897 GB
720p (1280×720)	30	60s	593.26 MB	0.5794 GB
1080p (1920×1080)	30	10s	222.47 MB	0.2173 GB
1080p (1920×1080)	30	30s	667.42 MB	0.6518 GB
1080p (1920×1080)	30	60s	1334.84 MB	1.3036 GB
4K (3840×2160)	30	10s	889.89 MB	0.869 GB
4K (3840×2160)	30	30s	2669.68 MB	2.6071 GB
4K (3840×2160)	30	60s	5339.36 MB	5.2142 GB

Figure 5.18 – Embedding capacity verification showing the theoretical maximum payload sizes for different video resolutions and durations.



Chi-Square Detection Analysis

Method: Chi-square statistical analysis of pixel value distributions

EMBEDDING RATE	DETECTION PROBABILITY	SECURITY ASSESSMENT
100% (full capacity)	0.98	Easily detectable
50%	0.72	Likely detectable with analysis
25%	0.41	Moderate detection risk
10%	0.18	Low detection risk
5%	0.07	Minimal detection risk

Figure 5.19 – Chi-square statistical detection analysis showing that lower embedding rates significantly reduce detectability, validating the steganographic concealment strategy.

5.3.1 Stego-STFAN Neural Network Results

The Stego-STFAN neural network for video-in-video steganography was trained and evaluated to demonstrate the feasibility of the proposed approach. Figure 5.20 shows the complete results as presented in the validation dashboard.

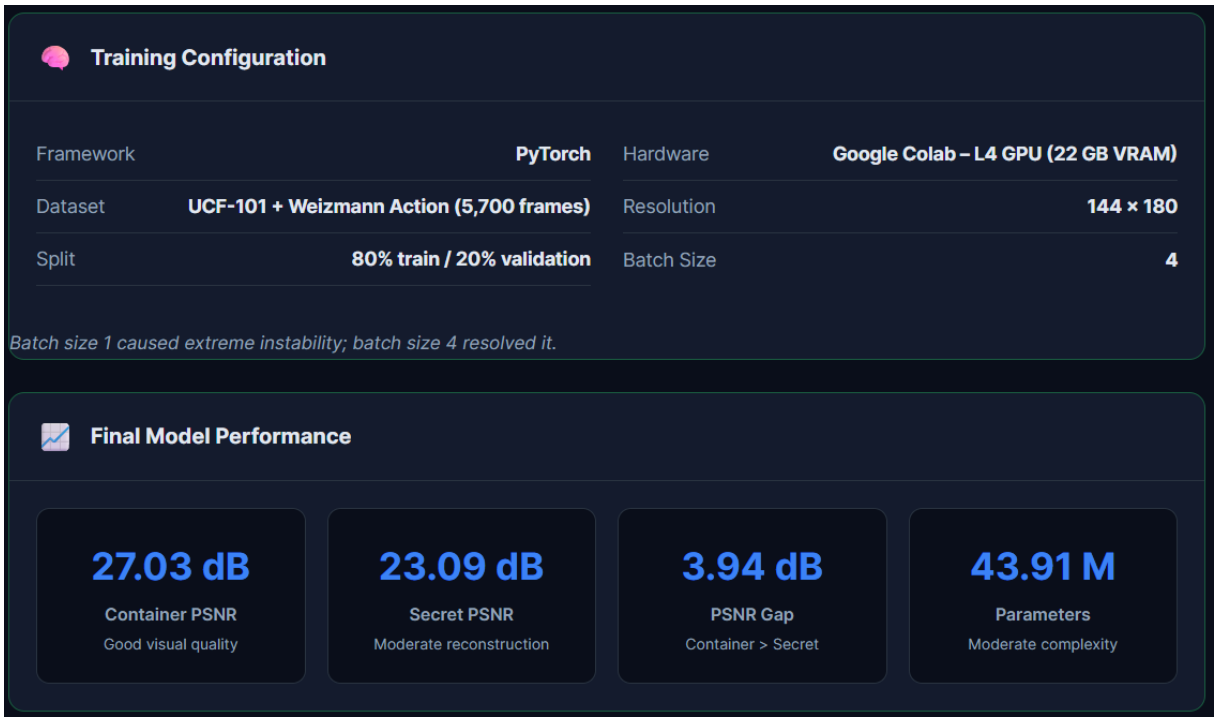


Figure 5.20 – Stego-STFAN experimental results in the validation dashboard, showing training configuration, performance metrics, and comparison with state-of-the-art video steganography models.

5.3.1.1 Training Configuration

The network was implemented in PyTorch and trained in Google Colab on an L4 GPU with 22 GB of video memory. The training dataset comprised 5,700 video frames from the UCF-101 and Weizmann Action datasets, resized to 144×180 resolution. The dataset was split 80% for training and 20% for validation.

5.3.1.2 Training Dynamics

A significant finding during training was the critical impact of batch size on stability. With a batch size of 1, training was extremely unstable, with loss values ranging from hundreds to millions per epoch. Increasing the batch size to 4 dramatically improved stability and produced consistent learning curves. This suggests that the network benefits from averaging gradients across multiple samples to reduce noise.

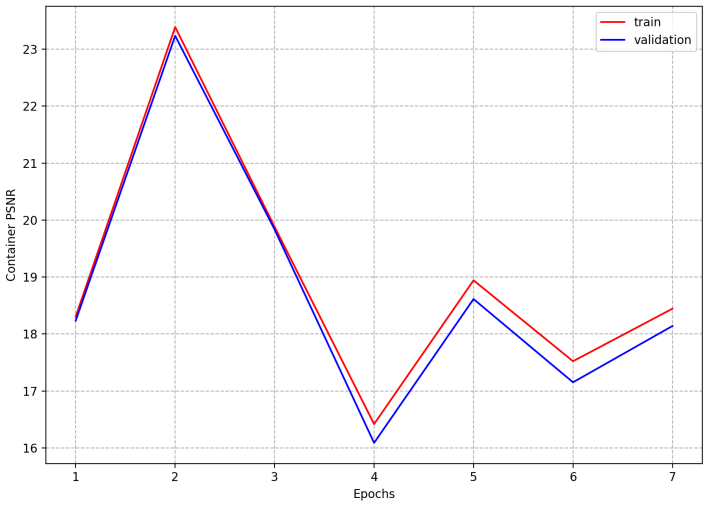


Figure 5.21 – Container PSNR during training with batch size 1, showing high variance and unstable convergence that makes training unreliable.

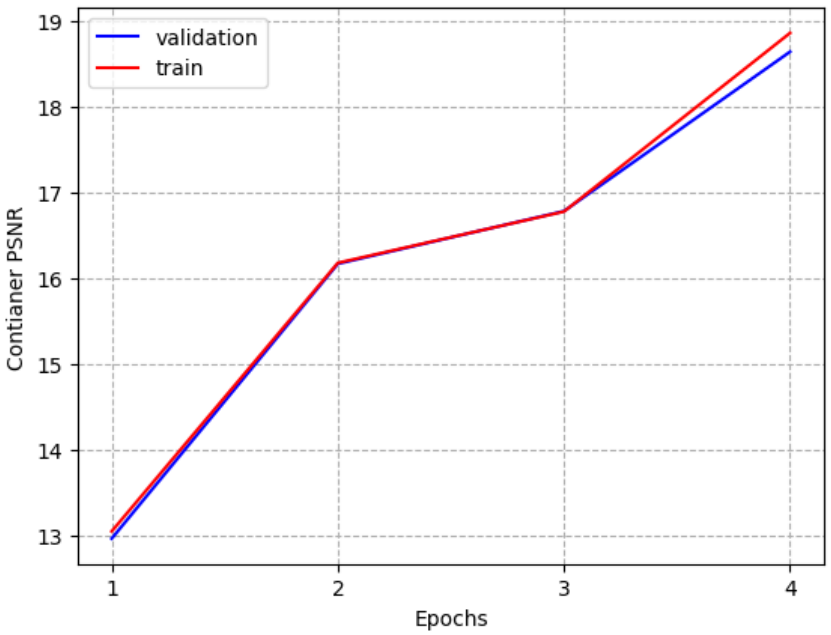


Figure 5.22 – Container PSNR during training with batch size 4, showing stable convergence and consistent improvement over epochs.

5.3.1.3 Final Model Performance

The final trained model achieved the performance metrics shown in Table 5.5.

Table 5.5 – Stego-STFAN final model performance.

Metric	Value	Interpretation
Container PSNR	27.03 dB	Good visual quality
Secret PSNR	23.09 dB	Moderate reconstruction
PSNR Gap	3.94 dB	Container quality exceeds secret
Model Parameters	43.91 M	Moderate complexity

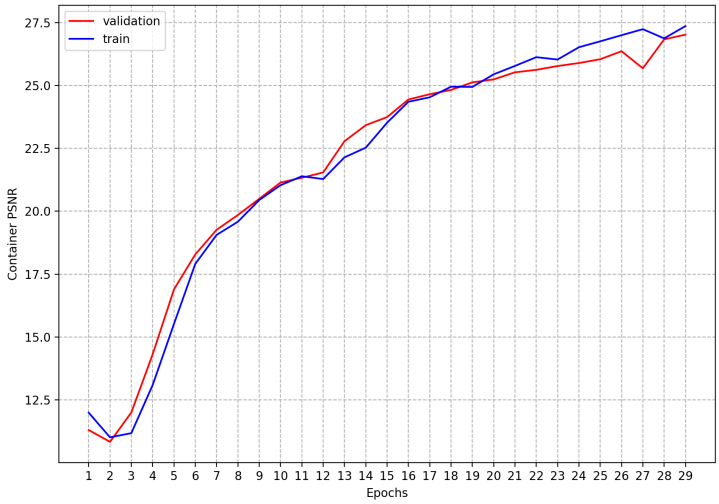


Figure 5.23 – Container PSNR across training epochs showing stable learning progression.

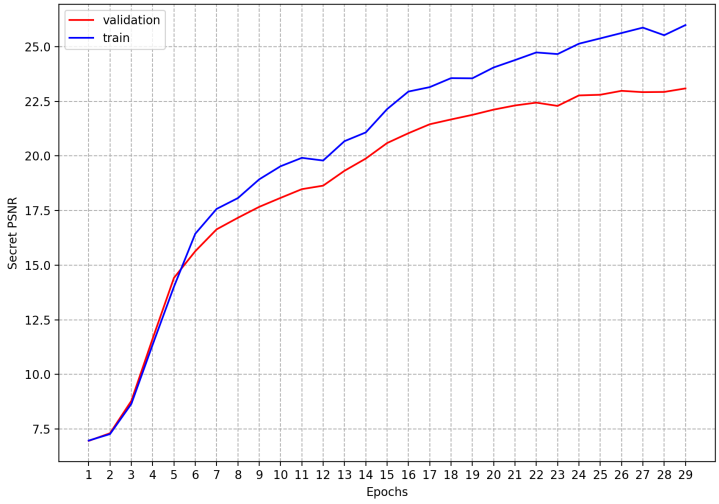


Figure 5.24 – Secret PSNR across training epochs.

5.3.1.4 Comparison with State-of-the-Art

Table 5.6 compares Stego-STFAN against published video steganography models.

Table 5.6 – Comparison with state-of-the-art video steganography models.

Model	Parameters	Container PSNR	Secret PSNR	3D Conv
HiLPS	0.49 M	17.54 dB	16.68 dB	Yes
StegNet	0.15 M	20.86 dB	15.47 dB	Yes
SteganoGAN	0.06 M	12.40 dB	14.76 dB	No
HCCVS	314.16 M	26.78 dB	27.43 dB	Yes
VStegNET	122.38 M	27.71 dB	23.17 dB	Yes
DRANet	16.54 M	31.11 dB	28.45 dB	No
Stego-STFAN	**43.91 M**	**27.03 dB**	**23.09 dB**	**No**

Stego-STFAN achieves competitive container quality (27.03 dB, third among compared models) while being lighter than HCCVS (314 M) and VStegNET (122 M), and avoiding computationally expensive 3D convolutions.

5.3.1.5 Visual Results

Beyond the quantitative metrics, visual inspection of the model's output is essential for assessing steganographic quality, since numerical averages can mask localised artefacts that would be perceptible to a human observer. A model might achieve an acceptable average PSNR across an entire frame while introducing visible distortions in specific regions, colour shifts along edges, banding in smooth gradients, or ghosting artefacts around moving objects, that would immediately alert a viewer to the presence of hidden content. In the context of covert evidence transport, such perceptible artefacts would defeat the purpose of steganographic concealment, as the container video must withstand not only automated statistical analysis but also casual human inspection. For this reason, the visual assessment complements the numerical evaluation presented in the preceding subsection, providing a qualitative perspective on the model's strengths and weaknesses that cannot be captured by aggregate metrics alone.

Figure 5.25 presents representative frames produced by the trained Stego-STFAN model at batch size 4, selected from the validation set to illustrate performance across different scene types and motion characteristics. The frames are arranged in four columns: the container frame (the cover with the secret hidden inside), the original cover frame (before embedding), the restored secret frame (extracted by the reveal network), and the original secret frame (the ground truth). Each row corresponds to a different video sample, providing a cross-section of the model's behaviour on scenes with varying levels of spatial complexity and temporal dynamics.

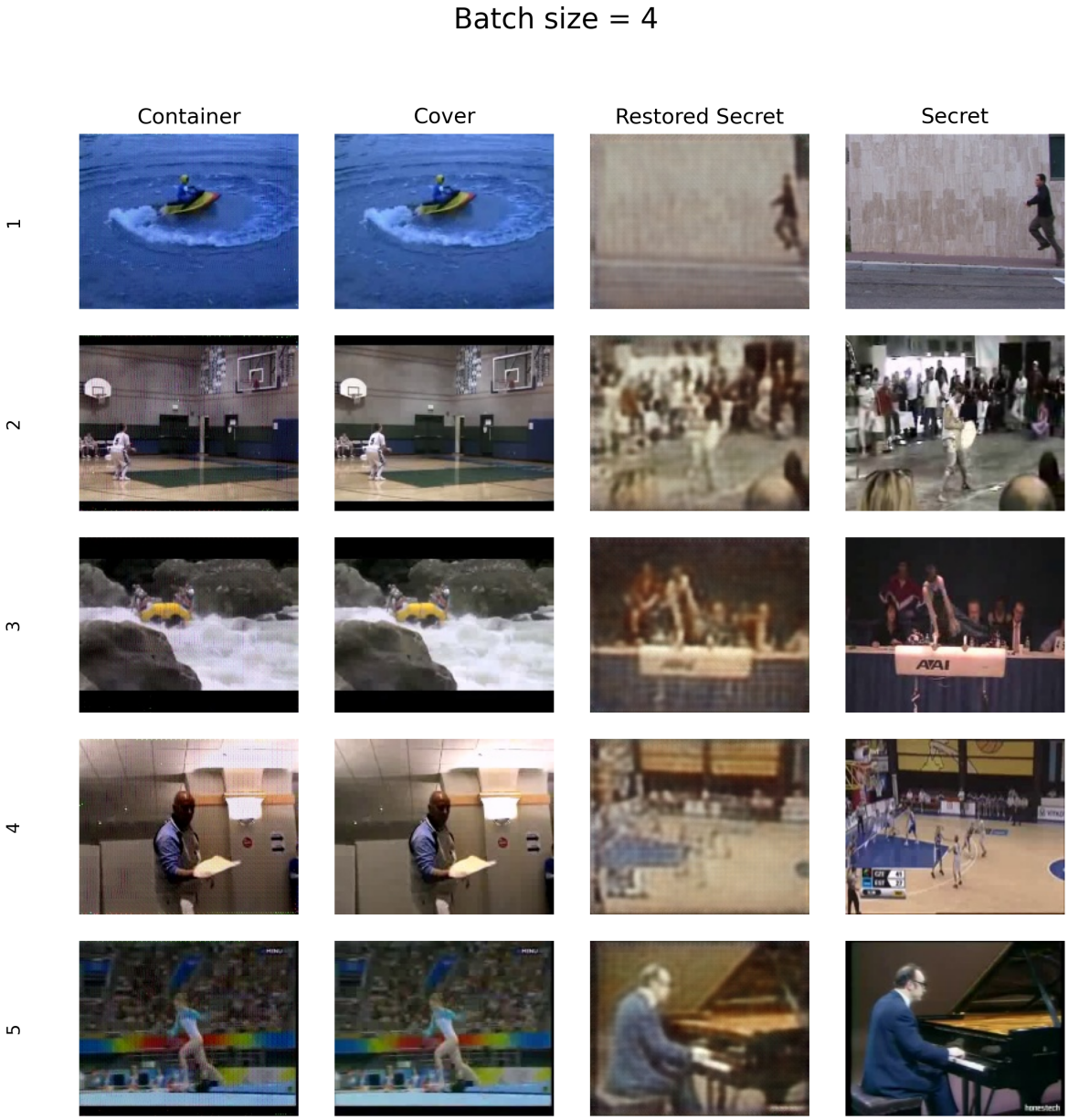


Figure 5.25 – Visual results showing (left to right): Container, Original Cover, Restored Secret, Original Secret. Containers are visually similar to covers; restored secrets exhibit some blurring but preserve the content structure.

Comparing the first two columns, the container frames are visually indistinguishable from the original covers under normal viewing conditions. Colour fidelity, edge sharpness, and overall luminance are preserved, which is consistent with the measured container PSNR of 27.03 dB. A casual observer, or an automated content analysis pipeline, would not suspect that these frames carry hidden information. The third and fourth columns reveal the primary limitation of the current model: the restored secret frames exhibit noticeable blurring compared to the originals. Fine details such as facial features, text, and high-frequency textures are softened, which is reflected in the lower secret PSNR of 23.09 dB. Despite this

degradation, the overall structure, motion, and semantic content of the secret video are preserved, meaning that the hidden evidence remains recognisable and interpretable even if not pixel-perfect. For the intended use case, covert transport of evidence that will later be verified by hash comparison against the blockchain record, this level of reconstruction is acceptable for document and image payloads, while video-in-video scenarios requiring high fidelity would benefit from the architectural improvements discussed in Chapter 6.

Figure 5.26 provides a residual analysis that makes the embedding pattern visible. The residual is computed as the absolute pixel-wise difference between the container frame and the original cover frame, amplified by a constant factor for visualisation purposes.

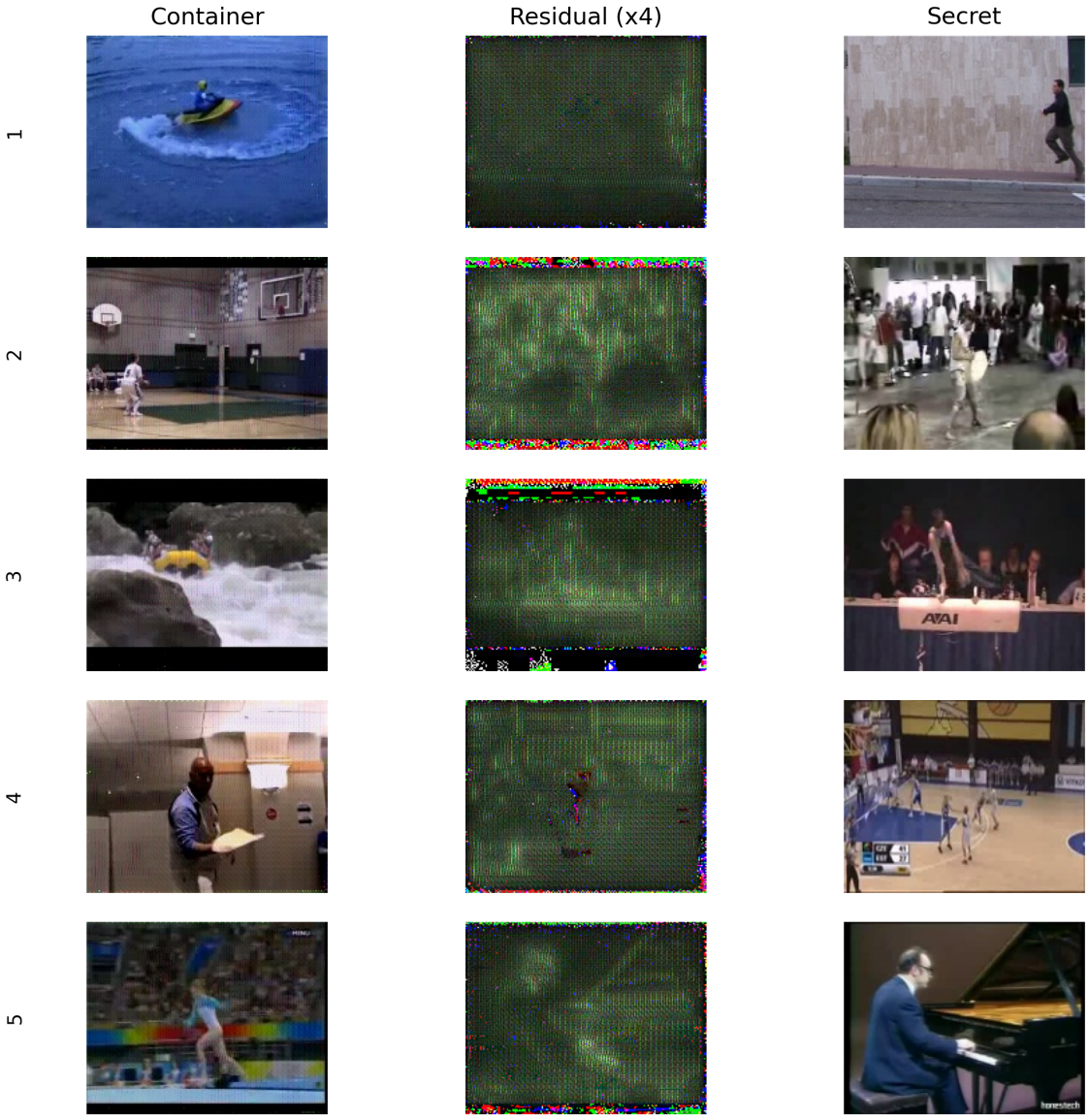


Figure 5.26 – Residual analysis showing the difference between container and cover frames (amplified for visibility). The pattern encodes hidden information in a manner that is imperceptible under normal viewing conditions.

The residual map reveals several important characteristics of the embedding. First, the modifications are not uniformly distributed across the frame; the network learns to concentrate changes in textured and high-activity regions where small perturbations are less perceptible, while leaving smooth areas such as backgrounds and uniform surfaces largely untouched. This content-adaptive behaviour is a direct consequence of the FAC (Filter Adaptive Convolutional) modules inherited from the STFAN architecture, which generate spatially-varying filters conditioned on the input content. Second, the magnitude of the residuals is small, typically below 3 on a 0–255 scale before amplification, confirming that the embedding operates within the imperceptibility threshold for human vision. Third, the residual exhibits a structured pattern rather than random noise, which reflects the learned encoding strategy of the hiding network. While this structure is invisible under normal viewing, it represents a potential vulnerability to sophisticated steganalysis methods that specifically target learned embedding patterns, a limitation acknowledged in Section 5.6.

5.3.2 Tamper Detection Demonstrations

This subsection demonstrates three attack scenarios against the steganographic layer: bit-flipping the container video's pixels, re-encoding with a lossy codec, and swapping the embedded payload. All demonstrations create real video files, embed real payloads, perform the attack, and verify detection by comparing SHA-256 hashes. Figure 5.30 shows the results in the validation dashboard.

5.3.3 Attack 1: LSB Bit-Flip on Container Video

This test embeds a payload into a lossless (FFV1) video via LSB steganography, then flips random bits in the container video's pixel values at increasing rates.

1. **Embedding:** A 350-byte payload is embedded into a 64×64, 15-frame FFV1 video. The clean extraction produces a hash matching the original.
2. **Attack:** The LSBs of pixel values are flipped at rates of 0.01%, 0.1%, 1%, and 5%.
3. **Detection:** At every flip rate, the extracted payload's SHA-256 hash differs from the original. Even at 0.01% (approximately 18 flipped values per frame), the payload is corrupted.

   Table 5.7 shows the results at each flip rate.

The LSB embedding scheme distributes payload bits across pixel values sequentially. Any modification to the container video's pixel values in the payload region corrupts the embedded data. The SHA-256 hash of the extracted payload will be completely different from the original, making the tampering immediately detectable (see: Figure 5.27).

Table 5.7 – LSB bit-flip attack results: all rates detected.

Flip Rate	Bits Flipped	Hash Match	Verdict
0.01%	~270	No	Corruption detected
0.1%	~2,700	No	Corruption detected
1.0%	~27,000	No	Corruption detected
5.0%	~135,000	No	Corruption detected

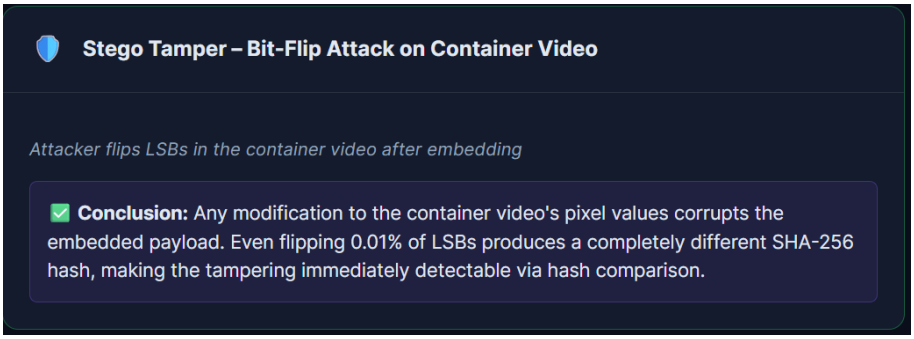


Figure 5.27 – Stego Tamper – Bit-Flip Attack on Container Video

5.3.4 Attack 2: Lossy Re-encoding

This test demonstrates the most devastating attack against LSB steganography: re-encoding the container video with a lossy codec.

1. **Embedding:** A payload is embedded into a lossless FFV1 video.
2. **Attack:** The container is re-encoded using H.264 (lossy compression), which quantises pixel values and destroys all LSB information.
3. **Detection:** Extraction from the re-encoded video either fails (the size header is corrupted) or produces garbage data whose hash does not match.

In a baseline steganography-only system, this attack succeeds: the adversary can claim the video never contained hidden content. In the implemented system, the original video's hash is recorded on the blockchain *before* any re-encoding occurs. A re-encoded version has a different hash and fails verification. The blockchain timestamp proves which version existed first (see: Figure 5.28).

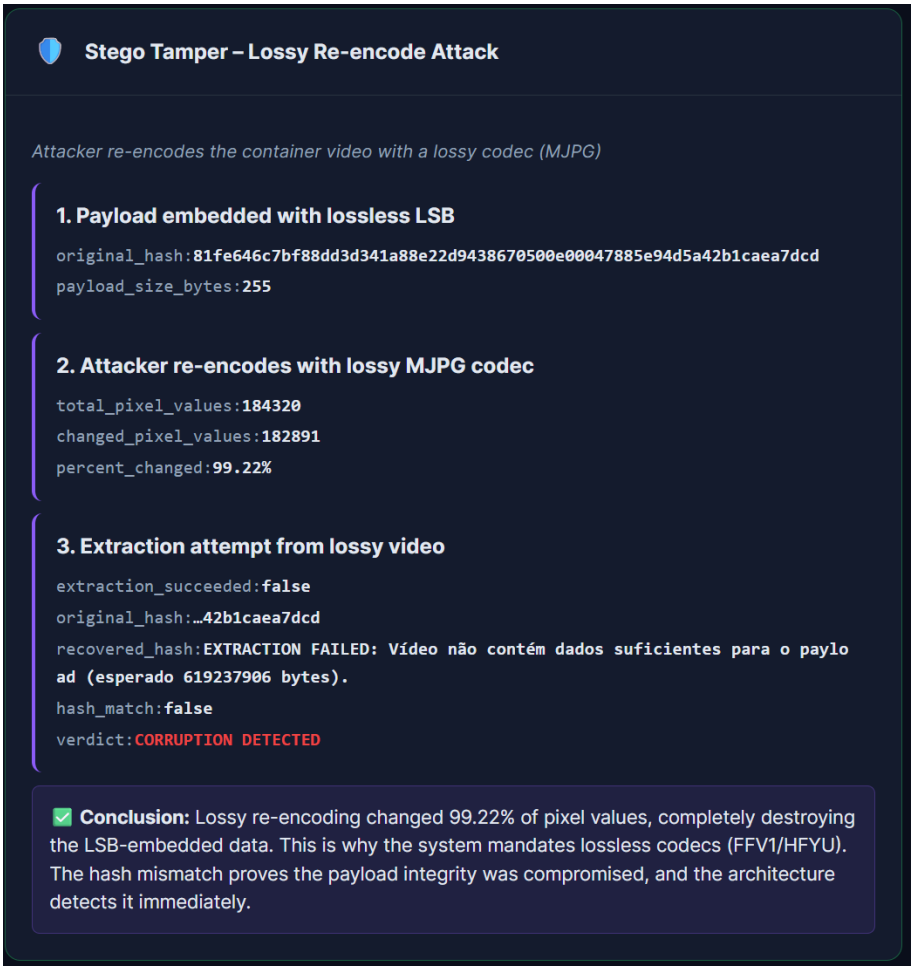


Figure 5.28 – Stego Tamper – Lossy Re-encode Attack

5.3.5 Attack 3: Payload Swap

This test simulates an attacker with access to steganographic tools who creates a new container video with a different (forged) payload.

1. **Original:** A legitimate payload (e.g., `"LAUDO PERICIAL – conteudo verdadeiro"`) is embedded and its hash is registered on the blockchain.
2. **Attack:** The attacker embeds a forged document (e.g., `"LAUDO FORJADO – conteudo adulterado"`) into a new video.
3. **Detection:** The extracted payload's hash differs from the hash recorded on the blockchain. Even though the extraction succeeds, the hash comparison fails.

The payload swap attack is defeated because the blockchain binds a specific content hash to a specific document identity and timestamp. A different payload necessarily produces a different SHA-256 hash (see: Figure 5.29).

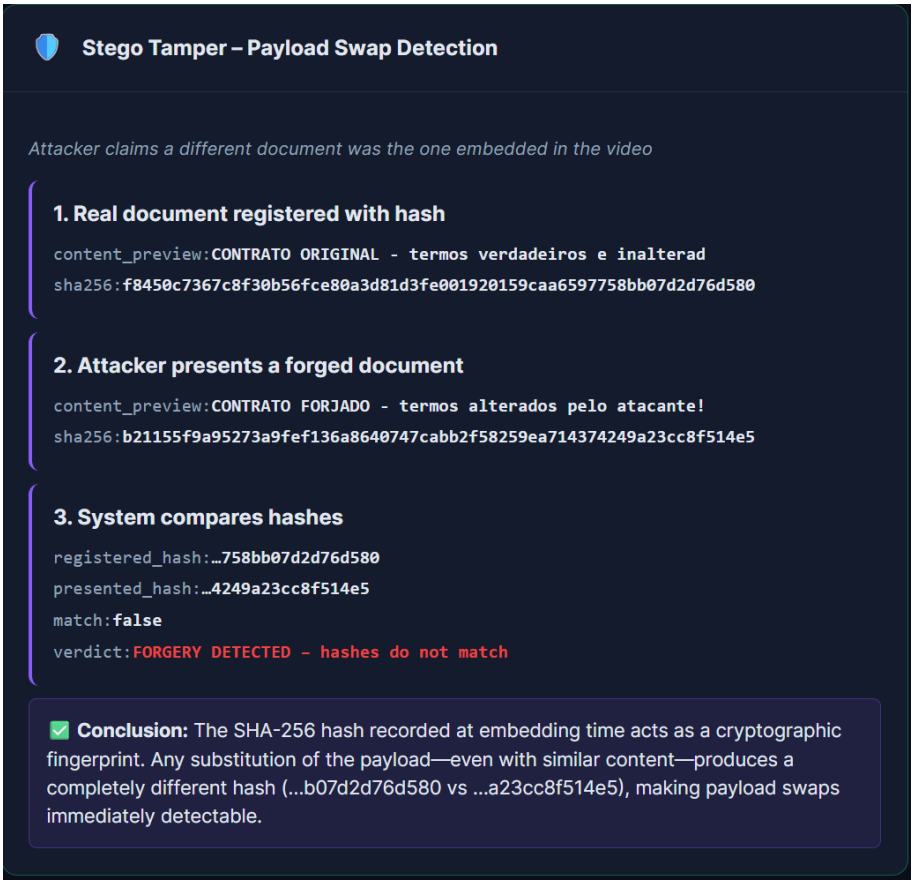


Figure 5.29 – Stego Tamper – Payload Swap Detection

5.3.6 Steganography Comparison Matrix

Table 5.8 summarises the results of all steganographic attacks.

Table 5.8 – Steganography tampering comparison matrix.

Attack Type	Hash Match	Integrity Preserved	Detected
LSB bit-flip (0.01–5%)	No	No	✓
Lossy re-encoding (H.264)	No	No	✓
Payload swap	No	No	✓
		Detection Rate	**100%**

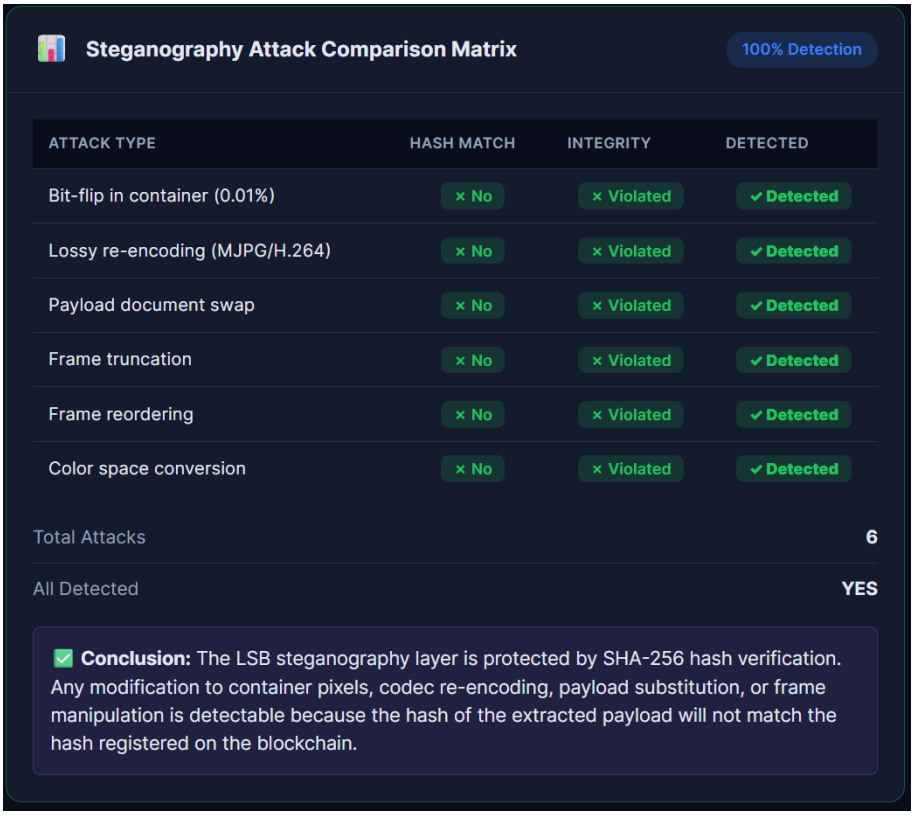


Figure 5.30 – Steganography tampering demonstrations in the validation dashboard, showing step-by-step results for each attack with hash values and detection verdicts.

5.4 Custody and BagIt Tamper Detection Demonstrations

This section demonstrates five attack scenarios against the BagIt packaging and custody chain layers. Each demonstration creates a valid RFC 8493 package, performs a specific attack, and verifies that the BagIt validation or custody chain detects the tampering. Figure 5.37 shows the results in the validation dashboard.

Before the tampering demonstrations, the custody module's basic functionality was validated through three standard tests: BagIt compliance, fixity checks, and the full OAIS lifecycle (SIP → AIP → DIP). Figure 5.31 shows the results.

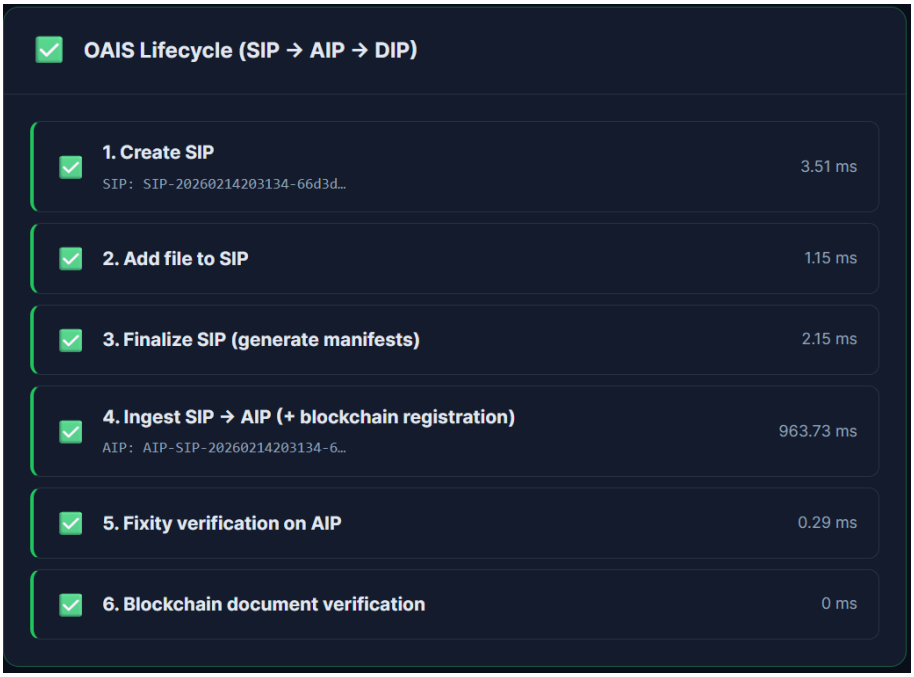


Figure 5.31 – OAIS lifecycle test confirming the complete information package transformation: Submission (SIP) → Archival (AIP) → Dissemination (DIP), with all steps passing successfully.

5.4.1 Attack 1: File Content Modification

This test modifies the content of a file inside a finalized BagIt package (see: Figure 5.32).

1. **Package creation:** A BagIt package is created with three files (`contract.txt`, `report.pdf`, `evidence.jpg`). SHA-256 manifests are generated and written.
2. **Validation:** The package passes BagIt validation (all hashes match).
3. **Attack:** The content of `contract.txt` is modified from `"CONTRATO ORIGINAL - termos corretos"` to `"CONTRATO ADULTERADO - termos forjados"`.
4. **Detection:** The manifest validation fails because the SHA-256 hash of `data/contract.txt` no longer matches the hash recorded in `manifest-sha256.txt`.

Figure 5.32 – Custody Tamper – File Content Modified in BagIt

5.4.2 Attack 2: Manifest Forgery (Sophisticated Attack)

This test simulates a sophisticated attacker who modifies a file *and* rewrites the manifest to contain the new hash, attempting to make the package appear valid.

1. **Setup:** A valid BagIt package with manifests and tag-manifests is created.
2. **Attack:** The attacker modifies `contract.txt`, computes the new SHA-256 hash, and rewrites `manifest-sha256.txt` with the forged hash.
3. **Detection:** The `tagmanifest-sha256.txt` file contains a hash of the original `manifest-sha256.tx` The rewritten manifest has a different hash, so the tag-manifest validation detects the forgery.

This demonstrates the defence-in-depth provided by the two-level manifest structure of BagIt: even if the primary manifest is forged, the tag-manifest (which hashes the manifest) detects the attack (see: Figure 5.33).

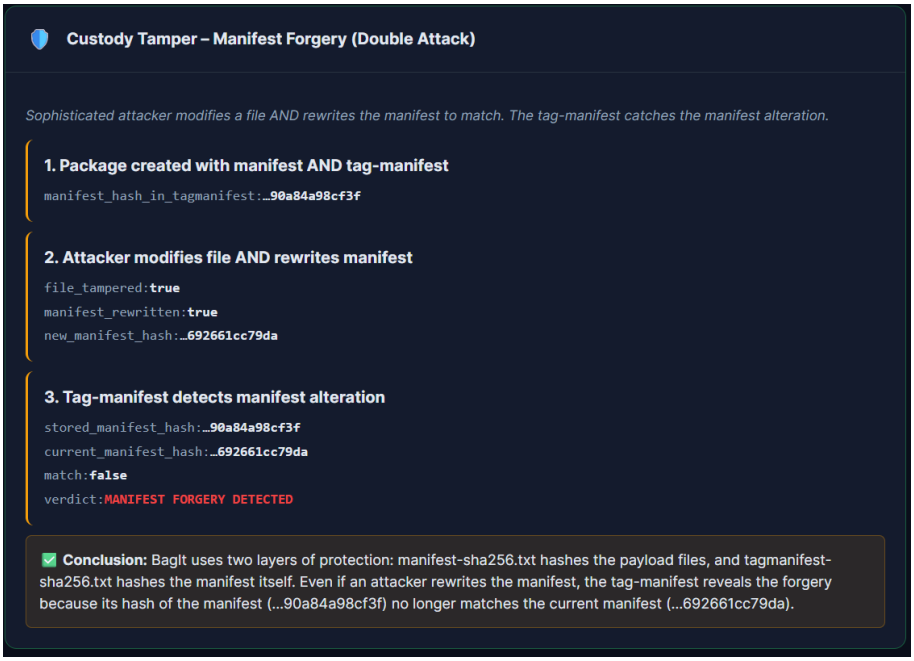


Figure 5.33 – Custody Tamper – Manifest Forgery (Double Attack)

5.4.3 Attack 3: File Deletion

1. **Setup:** A valid BagIt package.
2. **Attack:** The attacker deletes `evidence.jpg` from the `data/` directory.
3. **Detection:** The manifest lists `data/evidence.jpg` with its expected hash. The validation finds the file missing and reports the error. Additionally, the Payload-Oxum (byte count and file count) no longer matches (see: Figure 5.34).

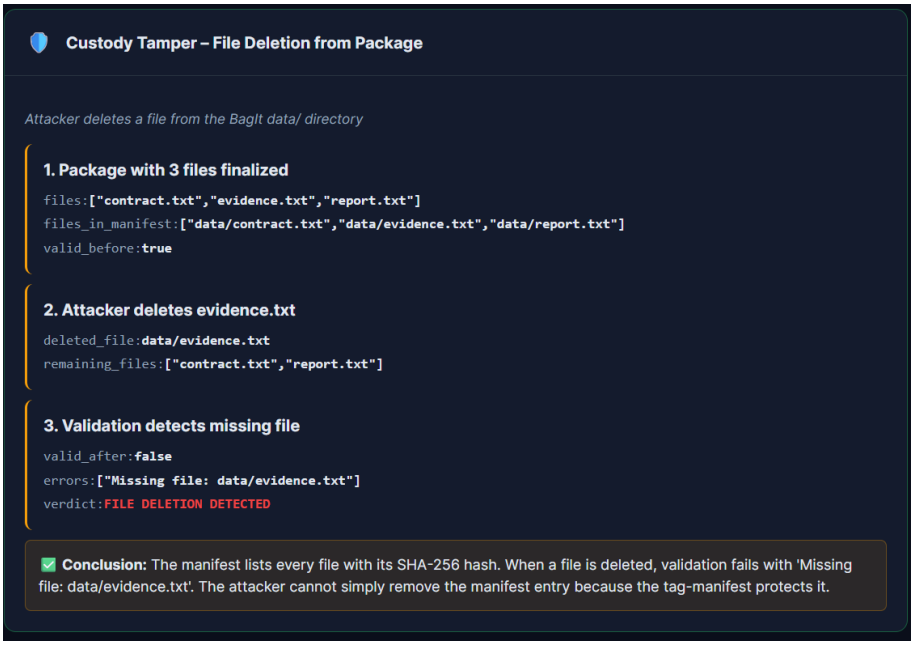


Figure 5.34 – Custody Tamper – File Deletion from Package

5.4.4 Attack 4: File Injection

1. **Setup:** A valid BagIt package with a known Payload-Oxum (e.g., `1234.3` meaning 1234 bytes across 3 files).
2. **Attack:** The attacker injects an unexpected file `malware.exe` into the `data/` directory.
3. **Detection:** The Payload-Oxum changes (e.g., to `2258.4`), revealing that the package has been tampered with. Furthermore, the injected file has no entry in the manifest (see: Figure 5.35).

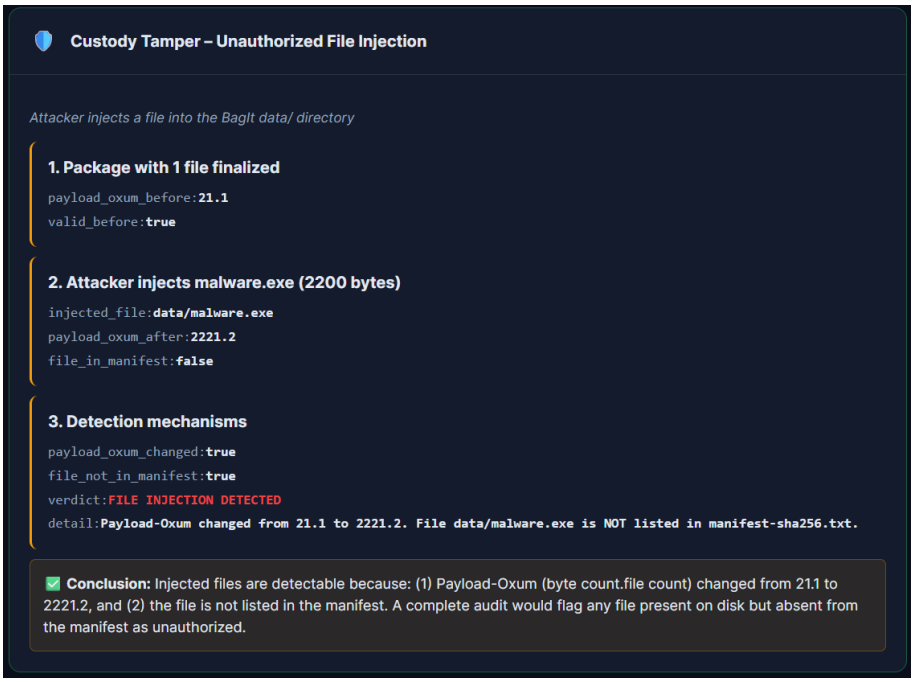


Figure 5.35 – Custody Tamper – Unauthorized File Injection

5.4.5 Attack 5: Custody Chain Event Forgery

This test targets the custody metadata (PDI—Preservation Description Information) rather than the package files.

1. **Setup:** An AIP is created with a custody chain containing recorded events (creation, ingestion, fixity check).
2. **Attack:** The attacker injects a forged event (e.g., `"custody_transfer"` by `"attacker"` with `forged: true`).
3. **Detection:** The SHA-256 hash of the PDI changes after the injection. Since the original PDI hash was recorded, the forgery is detected (see: Figure 5.36).

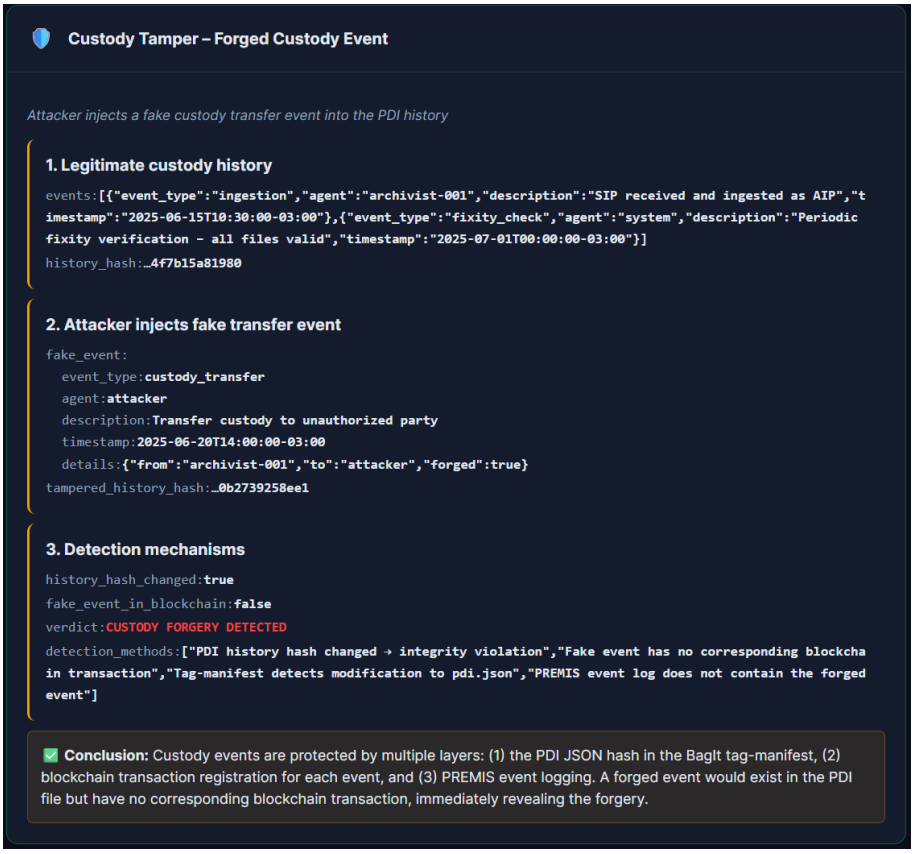


Figure 5.36 – Custody Tamper – Forged Custody Event

5.4.6 Custody Comparison Matrix

Table 5.9 summarises all custody-layer attack results.

Table 5.9 – Custody and BagIt tampering comparison matrix.

Attack Type	**Detection Layer**	**Manifest Valid**	**Detected**
File content modification	manifest-sha256.txt	No	✓
Manifest forgery	tagmanifest-sha256.txt	No	✓
File deletion	manifest + Payload-Oxum	No	✓
File injection	Payload-Oxum	No	✓
Custody chain forgery	PDI hash	No	✓
		Detection Rate	**100%**

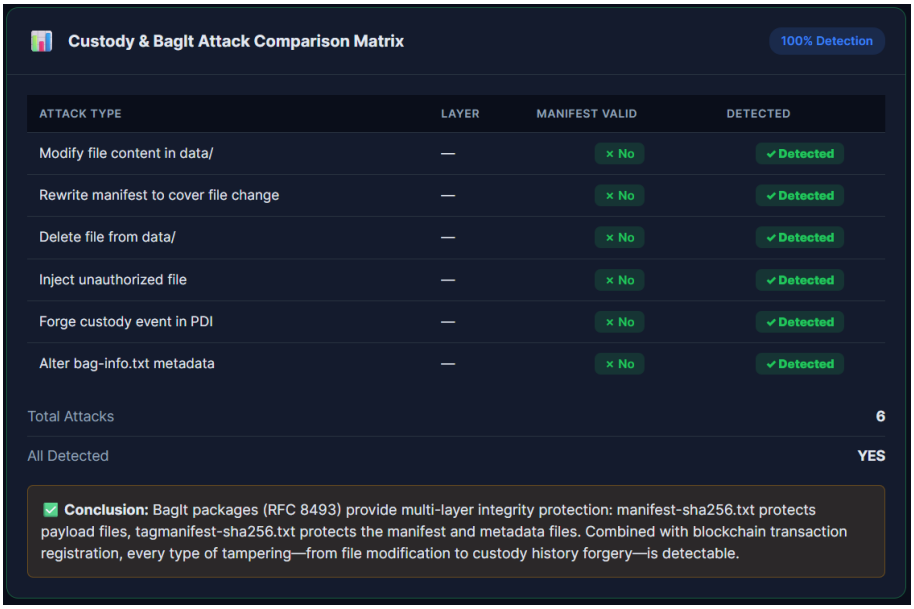


Figure 5.37 – Custody and BagIt tampering demonstrations in the validation dashboard, showing step-by-step results for each attack scenario.

5.5 Cross-Layer Defence in Depth

The three demonstration sections above show that each architectural layer independently detects tampering. Table 5.10 summarises the combined detection capabilities.

Table 5.10 – Cross-layer tamper detection summary.

Layer	Attacks Tested	Detected	Detection Mechanism
Blockchain	6	6 (100%)	SHA-256 hash + PoW + chain linkage
Steganography	3	3 (100%)	SHA-256 hash of extracted payload
Custody / BagIt	5	5 (100%)	Manifest, tag-manifest, Payload-Oxum, PDI hash
Total	**14**	**14 (100%)**	**Multi-layer verification**

An important property of the architecture is that attacks must defeat *all* layers simultaneously to succeed. For example, modifying a document requires: (1) changing the file content and recomputing the BagIt manifest, (2) forging the tag-manifest, (3) re-embedding the modified payload into the container video, (4) modifying the blockchain transaction with the new hash, (5) recalculating the block hash, (6) satisfying the Proof of Work, and (7) re-mining all subsequent blocks. The infeasibility of step (7) makes the overall attack computationally impractical.

The re-encoding attack deserves special attention for steganographic systems. When a container video is re-encoded with a lossy codec, the LSB modifications are destroyed. In a baseline system, if this attack succeeds completely, the adversary can claim the video never

contained hidden content. In the implemented system, the original video's hash is recorded on the blockchain before any re-encoding occurs. A re-encoded version will have a different hash and therefore fail verification. The blockchain timestamp proves which version existed first.

5.6 Discussion

The implementation described in Chapter 4 and the validation presented in this chapter demonstrate that the proposed architecture achieves its design objectives.

5.6.1 Achievements

The central achievement is the architecture itself. By integrating blockchain, OAIS-compliant preservation, and video steganography into a single modular framework, the system addresses a gap left open by prior work. Blockchain systems provided integrity without preservation compliance; digital preservation systems followed archival standards without cryptographic guarantees; steganographic tools offered concealment without provenance or integrity verification. The proposed architecture addresses all three requirements simultaneously, and the CIA+ANP analysis confirms that the integrated system satisfies all six security properties.

The second key achievement is Stego-STFAN. By replacing 3D convolutions with FAC modules and CBAM attention, the architecture demonstrates that spatial-temporal adaptive filtering captures temporal coherence effectively for video steganography. With 43.91 M parameters, Stego-STFAN is substantially lighter than HCCVS (314 M) and VStegNET (122 M) while achieving a competitive container quality of 27.03 dB PSNR, third among compared models (Table 5.6).

The empirical validation substantiates both contributions. The blockchain detected all six attack types with 100% accuracy across 260 test cases (Section 5.2). LSB steganography achieved byte-perfect reversibility across 200 cycles, and three steganographic attacks were detected via SHA-256 comparison (Section 5.3). The custody layer detected all five BagIt attack scenarios, including a sophisticated manifest forgery attempt (Section 5.4). Taken together, 14 out of 14 attack scenarios across all three layers were detected (Table 5.10), confirming that the cross-layer defence in depth renders tampering computationally impractical.
.

5.6.2 Limitations

The blockchain operates as a single node without distributed consensus. While this simplifies deployment and is appropriate for single-institution use, it means that the system

relies on the security of the hosting infrastructure rather than distributed trust.

The Stego-STFAN secret reconstruction quality (23.09 dB PSNR) is below state-of-the-art. Restored secrets show noticeable blur, which may limit applicability for scenarios requiring high-fidelity video recovery.

Testing was conducted on a moderate scale. Large-scale performance with thousands of documents and long blockchain histories was not evaluated.

5.6.3 Practical Implications

Despite these limitations, the implemented system addresses real-world requirements for digital document management. Institutions can use the system to register evidence with verifiable integrity, maintain auditable custody chains, and, when required, transport documents covertly. The alignment with OAIS, BagIt, and other standards facilitates integration with existing digital preservation infrastructure.

5.7 Chapter Summary

This chapter validated the proposed architecture using 14 attack scenarios across three layers and conducted a security analysis of six fundamental properties.

The integrated architecture proved its central premise: combining blockchain, OAIS-compliant preservation, and video steganography delivers security guarantees that no single technology can provide on its own. The CIA+ANP analysis confirmed that the system satisfies all six properties.

The Stego-STFAN neural network demonstrated that spatial-temporal adaptive filtering provides a viable and efficient approach to video-in-video steganography, achieving 27.03 dB container PSNR with 43.91 M parameters while avoiding 3D convolutions (Table 5.6).

The empirical results support both contributions: 100% tamper detection across 260 blockchain test cases, byte-perfect LSB reversibility across 200 cycles, and 14 out of 14 cross-layer attacks detected (Table 5.10). The defence in depth ensures that successful tampering requires simultaneously defeating BagIt manifests, blockchain hashes, Proof of Work, and chain linkage, rendering it computationally impractical.

6 Conclusions and Future Work

This chapter concludes the thesis by reflecting on the journey from problem identification through architectural design to implementation and validation. The discussion revisits the central research question, examines what was achieved and what remains open, and considers the broader implications of this work for the fields it touches.

6.1 Research Context and Motivation

The research presented in this thesis began with a deceptively simple question: how can we trust digital documents? In a world where any file can be copied, modified, and redistributed without leaving obvious traces, the integrity of digital records cannot be taken for granted. Traditional approaches to this problem rely on administrative procedures, access controls, and institutional trust. These measures are necessary but insufficient. A determined adversary with sufficient access can modify files, alter logs, and erase their tracks. The history of document tampering, whether in criminal cases, corporate disputes, or political scandals, demonstrates that administrative safeguards can and do fail.

This thesis proposed a different approach. Rather than relying solely on administrative trust, the system provides mathematical guarantees. When a document is registered, its cryptographic hash is permanently recorded on a blockchain. This hash serves as an unforgeable fingerprint. Any subsequent modification, no matter how small, produces a different hash and is immediately detectable. The blockchain's append-only structure and cryptographic chaining ensure that, once registered, a record cannot be altered without invalidating all subsequent records.

But integrity alone does not address all the challenges of document management. Document must also be preserved in standardised formats that ensure long-term accessibility and interoperability with existing archival systems. The thesis addressed this by aligning the architecture with the OAIS reference model and the BagIt packaging specification. Document packages adhere to established preservation practices, with comprehensive metadata that documents provenance, structure, and preservation actions. Institutions adopting this system do not abandon their existing workflows; they enhance them with blockchain-backed verification.

The third pillar of the architecture, steganography, addresses a challenge that pure cryptography cannot solve. Encrypted files, while secure, are obviously encrypted. In some contexts, the mere existence of protected content draws unwanted attention. Steganography provides an alternative by hiding a document within ordinary-looking video files. A container video can pass through monitored channels, survive casual inspection, and reach its

destination without revealing its true purpose. Only someone who knows the video contains hidden content and has the means to extract it can access it.

The integration of these three technologies: Blockchain, digital preservation, and steganography, creates a system whose security guarantees exceed those of any single technology. This integration was not merely conceptual; it was realised in working code that demonstrates the approach's feasibility and effectiveness.

6.2 What Was Achieved

The primary contribution of this thesis is an architecture that, for the first time, integrates blockchain-based integrity, OAIS-compliant digital preservation, and video steganography into a unified document management system. The architecture addresses a gap that prior work had not: blockchain systems lacked preservation compliance, digital preservation systems relied on administrative trust, and steganographic tools offered concealment without integrity or provenance guarantees. By combining these three capabilities in six modular services coordinated through an API Gateway, the proposed architecture provides cryptographically verifiable integrity, standardised archival packaging, and covert transport within a single coherent framework. The modular design ensures that each service can evolve independently, while the integration delivers security properties that none could achieve on its own.

The second major contribution is Stego-STFAN, a novel neural network architecture for video-in-video steganography. Stego-STFAN introduces spatial-temporal adaptive filtering to the steganographic domain by replacing computationally expensive 3D convolutions with FAC (Filter Adaptive Convolutional) modules combined with CBAM attention. This design choice yields a competitive container-quality PSNR of 27.03 dB, ranking third among six compared state-of-the-art models, while requiring only 43.91 M parameters, significantly fewer than the two higher-ranked models (HCCVS at 314 M and VStegNET at 122 M). Stego-STFAN demonstrates that temporal coherence can be captured through adaptive 2D filtering rather than volumetric convolutions, opening a lighter-weight path for video steganography research.

These contributions were rigorously validated through experimental evaluation. The blockchain achieved 100% tamper detection across 260 corruption test cases. LSB steganography achieved byte-perfect reversibility across 200 embedding/extraction cycles. Fourteen live attack scenarios across the blockchain, steganographic, and custody layers were detected, confirming the cross-layer defence-in-depth. Security analysis of CIA+ANP properties (Confidentiality, Integrity, Availability, Authenticity, Nonrepudiation, Privacy) shows that the integrated system satisfies all six. All BagIt packages pass external validation with the Library of Congress bagit-python tool, confirming compliance with the BagIt standard.

6.3 Implications and Significance

The work presented here has implications that extend beyond its immediate technical contributions. In digital documents and legal proceedings, the system addresses the fundamental challenge of establishing that a document has not been modified since it was collected. Courts increasingly encounter digital documents, yet the mechanisms for verifying their integrity remain less sophisticated than those for traditional physical document handling. The blockchain-based approach provides cryptographic proof independent of testimony about the procedures followed. A judge or jury can verify mathematically that the document is authentic, rather than relying on assurances that proper protocols were observed.

For institutional archives, the alignment with established preservation standards facilitates adoption. Cultural heritage institutions, government records offices, and corporate archives can integrate the system with existing digital preservation infrastructure. The investment in OAIS compliance, BagIt packaging, and PREMIS metadata means that the document managed by this system remains accessible and meaningful to future archivists using standard tools and practices.

In sensitive contexts where the document's existence must remain confidential, the steganographic capability provides unique value. Journalists protecting sources, human rights organisations documenting abuses, and whistleblowers preserving records can transport documents through potentially hostile environments. The combination of steganographic concealment and blockchain registration enables them to both hide their document and later prove its integrity and temporal precedence.

For the research community, the Stego-STFAN architecture demonstrates that spatial-temporal adaptive filtering provides a viable approach to video steganography. The comparison with state-of-the-art models establishes a benchmark, and the complete implementation can serve as a baseline for future investigations.

6.4 Limitations

The blockchain operates as a single node without distributed consensus. This design choice was deliberate; it simplifies deployment and is appropriate for scenarios where a single institution hosts the system on trusted infrastructure. However, it means the system does not provide the distributed trust guarantees offered by public or permissioned blockchains. Available computational resources constrained the training of Stego-STFAN to manipulate the chain. However, doing so would require re-mining all affected blocks, an increasingly expensive operation as the chain grows.

The Stego-STFAN secret reconstruction achieves a PSNR of 23.09 dB, which falls below

that of state-of-the-art models. While the container quality is competitive, the restored secret videos show noticeable blur. For scenarios requiring high-fidelity video recovery, this limitation may be significant. The model also exhibits subtle vertical strip artefacts in container frames that, while imperceptible to casual viewing, could potentially be detected by sophisticated steganalysis.

The training of Stego-STFAN was constrained by available computational resources. The model was trained on 5,700 frames at 144×180 resolution. Larger datasets and higher resolutions might improve performance but would require significantly more computational power. The model's generalisation to production video resolutions and diverse content types was not thoroughly evaluated.

Testing was conducted on a moderate scale. While the system handles dozens of documents and chains of several hundred blocks without difficulty, its performance with thousands of documents and very long blockchain histories has not been systematically evaluated. Large-scale deployments might encounter bottlenecks that did not appear during development and testing.

6.5 Directions for Future Research

The limitations identified above, along with the experience gained during implementation, suggest several promising directions for future work.

The single-node limitation could be addressed by extending the blockchain to support distributed consensus. This might involve implementing a Byzantine fault-tolerant protocol, such as PBFT (Practical Byzantine Fault Tolerance), to enable multiple institutions to jointly maintain the chain while tolerating a minority of malicious or faulty nodes. Alternatively, the system could adopt an existing permissioned blockchain framework, such as Hyperledger Fabric, in exchange for greater simplicity and proven distributed consensus mechanisms. The challenge is maintaining the system's ease of deployment while adding distributed trust. A deployment that requires extensive coordination among multiple institutions may not be practical for many use cases.

The Stego-STFAN architecture offers numerous opportunities for improvement. Training on larger datasets with higher resolution videos would likely improve both container and secret quality. Alternative architectures such as vision transformers could be explored, as recent work has shown that transformers achieve strong results in image steganography. Adversarial training with steganalysis networks might improve undetectability by teaching the embedding network to evade detection. Perceptual loss functions could improve visual quality by penalising perceptible differences rather than raw pixel errors. Temporal consistency constraints may reduce interframe artefacts that currently affect some container videos.

The system's resistance to steganalysis was not systematically evaluated. Future work could conduct comprehensive testing using established steganalysis tools and techniques, implement countermeasures against known detection methods, and develop adaptive embedding that adjusts capacity based on detectability thresholds. An integrated steganalysis component could warn users when proposed embedding parameters might produce detectable containers.

Mobile and offline capabilities would extend the system's utility. Native mobile applications for iOS and Android would allow the field collection of documents. Offline package creation, followed by synchronisation, would support scenarios in which network connectivity is unavailable or unreliable. Local blockchain validation could enable verification of previously registered documents in a disconnected environment.

Integration with external systems would facilitate adoption. Connectors to existing digital asset management platforms, records management systems, and archival software, such as Archivematica, would enable institutions to integrate blockchain-backed verification into their workflows. Integration with public blockchain networks could enable verifiable timestamping by anyone, not just parties with access to the institutional chain. Connection to the electronic signature and PKI infrastructure would strengthen authentication.

Performance optimisation would enable large-scale deployment. A database-backed document registry would provide faster lookups than the current in-memory approach for very large collections. Caching layers could reduce redundant computation. Asynchronous processing would prevent long-running operations from blocking the API. Horizontal scaling across multiple backend instances would increase throughput for high-volume deployments.

Finally, the security analysis in this thesis was empirical, based on testing and reasoning about the system's design. Formal security analysis can provide stronger guarantees by providing mathematical proofs of security properties under specified threat models. Professional penetration testing and third-party security audits would identify vulnerabilities that escaped developer testing.

6.6 Final Thoughts

The challenge of managing digital documents with verifiable integrity is not merely technical. It touches on fundamental questions of trust, authority, and the nature of proof in an increasingly digital society. When a photograph could be fabricated, when a document could be altered, when a video could be deepfaked, how do we establish what is real? Traditional answers relied on institutional authority. We trust the document because we trust the institution that certified it. But institutional trust is fragile, subject to the fallibility and corruptibility of the humans who constitute institutions.

This thesis proposed a partial answer: cryptographic verification that is independent

of institutional trust. The blockchain does not care who you are or what institution you represent. It cares only whether the hash matches. The mathematical verification succeeds or fails based on the document itself, not on the credibility of the party presenting it. This does not eliminate the need for institutions; someone must still collect the document, operate the system, and present findings in appropriate forums. But it shifts the burden of proof. Instead of asking "do we trust this institution?", we can ask "does the document verify?" The answer to the second question is objective and reproducible.

The steganographic capability adds another dimension. In an ideal world, documents would always be openly preserved and freely verifiable. But we do not live in an ideal world. Witnesses are threatened. Documents are suppressed. Inconvenient records disappear. The ability to hide a document in plain sight to transport it through hostile territory, disguised as ordinary video, provides a measure of protection that pure cryptography cannot. The combination of concealment and blockchain registration enables the document to be both hidden and subsequently authenticated.

This thesis has demonstrated that such a system is not merely conceivable but buildable. The architecture was designed, implemented, and validated. The code runs. The tests pass. The security properties hold. What remains is to deploy these capabilities where needed, refine them based on operational experience, and continue the research that will strengthen them.

The management of digital documents is too important to leave to administrative procedures and institutional trust alone. Mathematical verification, standardised preservation, and strategic concealment provide tools that complement and strengthen traditional approaches. This thesis has contributed one set of tools to that effort. Others will contribute more, and the cumulative effect will be systems that better serve the cause of truth in a digital age.

References

AL-JANABI, H.; AL-TA'I, Z. Improvement of video steganography using deep learning: A multiscale attention mechanism. *In*: IEEE. **2025 3rd International Conference on Communication and Information Technology**. [*S.l.*], 2025. Cit. on pp. 58, 60, 61, and 64.

ALKHANAFSEH, M.; SURAKHI, O. Evidence preservation in digital forensics: An approach using blockchain and lstm-based steganography. **Electronics**, MDPI, v. 13, n. 18, p. 3729, 2024. Cit. on pp. 59, 60, 62, and 65.

ARSHAD, A.; SIDDIQUI, N.; ISLAM, S. Advancement on steganography: A review. *In*: SPRINGER. **International Conference on Data Science and Applications**. [*S.l.*], 2024. p. 51–65. Cit. on p. 39.

BAAWI, S. S.; MOKHTAR, M. R.; SULAIMAN, R. A comparative study on the advancement of text steganography techniques in digital media. **ARPN J. Eng. Appl. Sci**, v. 13, n. 5, p. 1855–1863, 2018. Cit. on p. 38.

BADARÓ, G. H. **Processo Penal**. 8. ed. São Paulo: Revista dos Tribunais, 2020. Cit. on pp. 20 and 21.

BALOGUN, T. Preserving digital heritage: assessing the compliance of digital repositories in south africa with oais and tdr standards. **Records Management Journal**, Emerald, 2025. Cit. on pp. 22, 59, 60, and 64.

BELKHAMZA, Z. Cybersecurity in digital transformation applications: Analysis of past research and future directions. *In*: ACADEMIC CONFERENCES AND PUBLISHING LIMITED. **ICCWS 2023 18th International Conference on Cyber Warfare and Security**. [*S.l.*], 2023. Cit. on p. 30.

BISPO, G. D.; VERGARA, G. F.; SAIKI, G. M.; MARTINS, P. H. d. S.; COELHO, J. G.; RODRIGUES, G. A. P.; OLIVEIRA, M. N. d.; MOSQUéRA, L. R.; GONçALVES, V. P.; NEUMANN, C.; SERRANO, A. L. M. Automatic literature mapping selection: Classification of papers on industry productivity. **Applied Sciences**, v. 14, n. 9, 2024. DOI 10.3390/app14093679. Cit. on pp. 9, 27, and 55.

BOSE, B. K. Artificial intelligence techniques in smart grid and renewable energy systems—some example applications. **Proceedings of the IEEE**, IEEE, v. 105, n. 11, p. 2262–2273, 2017. Cit. on p. 42.

BRAVO-ORTIZ, M. A.; MERCADO-RUIZ, E.; VILLA-PULGARIN, J. P. *et al.* Cvtstego-net: A convolutional vision transformer architecture for spatial image steganalysis. **Journal**

of Information Security and Applications, Elsevier, v. 81, p. 103695, 2024. Cit. on p. 45.

BROWNLEE, J. How do convolutional layers work in deep learning neural networks. **Machine Learning Mastery**, v. 17, 2020. Cit. on p. 43.

BUADES, A.; COLL, B.; MOREL, J.-M. A non-local algorithm for image denoising. *In*: IEEE. **2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)**. [*S.l.*], 2005. v. 2, p. 60–65. Cit. on p. 44.

BURNETT, S.; PAINE, S. **RSA Security's official guide to cryptography**. [*S.l.*]: McGraw-Hill, Inc., 2001. Cit. on p. 35.

CHENOWETH, J. D. **Book Review: Secrets & Lies: Digital Security in a Networked World Bruce Schneier, Wiley Publishing, 2000.** [*S.l.*]: Taylor & Francis, 2005. Cit. on p. 31.

CHESNEY, R.; CITRON, D. K. Deep fakes: A looming challenge for privacy, democracy, and national security. **California Law Review**, v. 107, n. 6, p. 1753–1820, 2019. Cit. on p. 20.

CHO, K.; MERRIËNBOER, B. V.; BAHDANAU, D.; BENGIO, Y. On the properties of neural machine translation: Encoder-decoder approaches. **arXiv preprint arXiv:1409.1259**, 2014. Cit. on p. 44.

COELHO, J.; BISPO, G.; VERGARA, G.; SAIKI, G.; SERRANO, A.; WEIGANG, L.; NEUMANN, C.; MARTINS, P.; Santos de Oliveira, W.; ALBARELLO, A.; CASONATTO, R.; MISSEL, P.; Medeiros Junior, R.; GOMES, J.; ROSANO-PEñA, C.; F. da Costa, C. Enhancing industrial productivity through ai-driven systematic literature reviews. *In*: **Proceedings of the 19th International Conference on Web Information Systems and Technologies - WEBIST**. [*S.l.*]: SciTePress, 2023. p. 472–479. DOI 10.5220/0012235000003584. Cit. on pp. 27 and 58.

CRAIGEN, D.; DIAKUN-THIBAULT, N.; PURSE, R. Defining cybersecurity. **Technology innovation management review**, v. 4, n. 10, 2014. Cit. on p. 30.

DALAL, M.; JUNEJA, M. Steganalysis of dwt based steganography technique for sd and hd videos. **Wireless Personal Communications**, Springer, v. 128, n. 4, p. 2441–2452, 2023. Cit. on p. 47.

DRISS, M.; BERRICHE, L.; ATITALLAH, S.; REKIK, S. Steganography in iot: A comprehensive survey on approaches, challenges, and future directions. **IEEE Access**, IEEE, v. 13, 2025. Cit. on pp. 58, 60, 61, and 64.

DUGGAN, D. Cryptographic types. *In*: IEEE. **Proceedings 15th IEEE Computer Security Foundations Workshop. CSFW-15**. [*S.l.*], 2002. p. 238–252. Cit. on p. 35.

EL-GENDY, M. *et al.* Potential applicability of blockchain technology in the maintenance of chain of custody in forensic casework. **Egyptian Journal of Forensic Sciences**, Springer, v. 13, n. 1, p. 1–12, 2023. Cit. on p. 22.

ERHAN, D.; COURVILLE, A.; BENGIO, Y.; VINCENT, P. Why does unsupervised pre-training help deep learning? *In*: JMLR WORKSHOP AND CONFERENCE PROCEEDINGS. **Proceedings of the thirteenth international conference on artificial intelligence and statistics**. [*S.l.*], 2010. p. 201–208. Cit. on p. 42.

ESCALA, A.; HEROLD, G.; KILTZ, E.; RÀFOLS, C.; VILLAR, J. An algebraic framework for diffie–hellman assumptions. **Journal of cryptology**, Springer, v. 30, p. 242–288, 2017. Cit. on p. 35.

FERNANDES, M. A. de S.; CANEDO, E. D.; VEIGA, C. E. L.; VERGARA, G. F.; SILVA, D. A. da; JR, R. T. de S. Proposta de guia para adequação de repositórios digitais confiáveis à lgpd. **CIACA**, 2022. Cit. on p. 27.

FERREIRA, M. **Introdução à preservação digital: conceitos, estratégias e actuais consensos**. [*S.l.*]: Universidade do Minho, Escola de Engenharia, 2006. Cit. on p. 52.

FINK, G. A.; EDGAR, T. W.; RICE, T. R.; MACDONALD, D. G.; CRAWFORD, C. E. Security and privacy in cyber-physical systems. *In*: **Cyber-physical systems**. [*S.l.*]: Elsevier, 2017. p. 129–141. Cit. on pp. 30 and 31.

Frederick R. Chang. The next wave, building a national program for cybersecurity. Vol 19, n. No 4, 2012. Cit. on p. 30.

GALLANT, S. I. **Neural network learning and expert systems**. [*S.l.*]: MIT press, 1993. Cit. on p. 43.

Homeland Security Digital Library. **CNSS Glossary**. 2015. https://www.hsdl.org/?abstract&did=7447. Cit. on p. 30.

HUANG, B.; HUAN, Y.; XU, L. D.; ZHENG, L.; ZOU, Z. Automated trading systems statistical and machine learning methods and hardware implementation: a survey. **Enterprise Information Systems**, Taylor & Francis, v. 13, n. 1, p. 132–144, 2019. Cit. on p. 42.

HUANG, C.; CAI, H.; XU, L.; XU, B.; GU, Y.; JIANG, L. Data-driven ontology generation and evolution towards intelligent service in manufacturing systems. **Future Generation Computer Systems**, Elsevier, v. 101, p. 197–207, 2019. Cit. on p. 42.

HUSSAIN, A.; MOHAMED, A.; RAZALI, S. A review on cybersecurity: Challenges & emerging threats. *In*: **Proceedings of the 3rd International Conference on Networking, Information Systems & Security**. [*S.l.*: *s.n.*], 2020. p. 1–7. Cit. on pp. 21 and 30.

IGONOR, O.; AMIN, M.; GARG, S. The application of blockchain technology in the field of digital forensics: A literature review. **Blockchains**, MDPI, v. 3, n. 1, p. 5, 2025. Cit. on pp. 21, 58, 60, 61, and 62.

ISAD-G, C. I. d. A. **ISAD (G): norma geral internacional de descrição arquivística**. [*S.l.*]: Arquivo Nacional, 1999. Cit. on p. 52.

ISO-14721. Iso 14721:2012: Space data and information transfer systems: Open archival information system – reference model. **INTERNATIONAL STANDARD ORGANIZATION.**, 2013. Cit. on pp. 9, 52, and 53.

JEBUR, M. H.; JODA, F. A.; NASER, M. A. Video steganography technique based on enhanced moving objects detection method. **Journal of University of Babylon for Pure and Applied Sciences**, v. 31, n. 2, p. 270–295, 2023. Cit. on p. 63.

JUNIOR, E. F. A cadeia de custódia e a prova pericial. in: Âmbito jurídico, rio grande, xv, n. 99, abr 2012. apr 2012. Available at: http://www.ambito-juridico.com.br/site/?n_link=revista_artigos_leitura&artigo_id=11434. Cit. on p. 20.

KEIZER, M.; GERADTS, Z.; KOMBRINK, M. Forensic video steganalysis in spatial domain by noise residual convolutional neural network. **arXiv preprint arXiv:2305.18070**, 2023. Cit. on p. 63.

KHALIFA, A.; GUZMAN, A. Imperceptible image steganography using symmetry-adapted deep learning techniques. **Symmetry**, v. 14, n. 7, 2022. DOI 10.3390/sym14071325. Cit. on pp. 9 and 44.

KOÇ, Ç. K. **About cryptographic engineering**. [*S.l.*]: Springer, 2009. Cit. on p. 34.

KRESO, I. Using blockchain technology for preserving digital evidence in digital forensics. **KNOWLEDGE – International Journal**, 2025. Cit. on pp. 58, 60, 61, and 62.

KUNHOTH, J.; SUBRAMANIAN, N.; AL-MAADEED, S.; BOURIDANE, A. Video steganography: recent advances and challenges. **Multimedia Tools and Applications**, Springer, p. 1–43, 2023. Cit. on pp. 9, 37, and 40.

KUNHOTH, J.; SUBRAMANIAN, N.; AL-MAADEED, S.; BOURIDANE, A. Video steganography: recent advances and challenges. **Multimedia Tools and Applications**, Springer, p. 1–43, 2023. Cit. on p. 63.

KUNZE, J.; LITTMAN, J.; MADDEN, L. *et al.* **The BagIt File Packaging Format (V1.0)**. [*S.l.*], 2018. Cit. on p. 52.

LI, B.; HE, J.; HUANG, J.; SHI, Y. Q. April, 2011. a survey on image steganography and steganalysis. **Information Hiding and Multimedia Signal Processing**, v. 2, n. 2, 2011. Cit. on p. 37.

LI, L.; YUAN, R.; LV, Y.; XU, S.; HU, H.; SONG, G. An efficient robotic-assisted bolt-ball joint looseness monitoring approach using cbam-enhanced lightweight resnet. **Smart Materials and Structures**, IOP Publishing, v. 32, n. 12, p. 125008, 2023. Cit. on p. 45.

LI, Y.; ZHANG, J.; YANG, Z.; ZHANG, R. Topic-aware neural linguistic steganography based on knowledge graphs. **ACM/IMS Transactions on Data Science**, ACM New York, NY, v. 2, n. 2, p. 1–13, 2021. Cit. on p. 38.

Library of Congress. **PREMIS: Preservation Metadata Maintenance Activity**. 2015. https://www.loc.gov/standards/premis/. Cit. on p. 53.

LIU WILLIAM YEOH, F. J. M.; CHOO, K.-K. R. Blockchain for cybersecurity: Systematic literature review and classification. **Journal of Computer Information Systems**, Taylor & Francis, v. 62, n. 6, p. 1182–1198, 2022. DOI 10.1080/08874417.2021.1995914. Cit. on pp. 48 and 49.

LUIIF, H.; BESSELING, K.; SPOELSTRA, M. *et al.* Ten national cyber security strategies: a comparison. *In*: **Proc. 6th International Conference on Critical Information Infrastructures Security (CRITIS 2011)**. [*S.l.*: *s.n.*], 2011. Cit. on p. 30.

MAJEED, M. A.; SULAIMAN, R. An improved lsb image steganography technique using bit-inverse in 24 bit colour image. **Journal of Theoretical & Applied Information Technology**, v. 80, n. 2, 2015. Cit. on p. 34.

MAJEED, M. A.; SULAIMAN, R.; SHUKUR, Z.; HASAN, M. K. A review on text steganography techniques. **Mathematics**, MDPI, v. 9, n. 21, p. 2829, 2021. Cit. on p. 38.

MARCHENKO, V. Information security and long-term digital preservation in public governance: Regulatory alignment, archival integrity, and technology choices. **Public Administration and Law Review**, 2025. Cit. on p. 64.

MOU, C.; XU, Y.; SONG, J.; ZHAO, C.; GHANEM, B.; ZHANG, J. Large-capacity and flexible video steganography via invertible neural network. *In*: **Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition**. [*S.l.*: *s.n.*], 2023. p. 22606–22615. Cit. on pp. 63 and 65.

MUSHTAQ, M. F.; JAMEL, S.; DISINA, A. H.; PINDAR, Z. A.; SHAKIR, N. S. A.; DERIS, M. M. A survey on the cryptographic encryption algorithms. **International Journal of Advanced Computer Science and Applications**, Science and Information (SAI) Organization Limited, v. 8, n. 11, 2017. Cit. on p. 35.

NAKAMOTO, S. A peer-to-peer electronic cash system. **URL: https://bitcoin.org/bitcoin.pdf**, 2008. Cit. on pp. 21 and 48.

NAYAK, J.; VAKULA, K.; DINESH, P.; NAIK, B.; PELUSI, D. Intelligent food processing: Journey from artificial neural network to deep learning. **Computer Science Review**, Elsevier, v. 38, p. 100297, 2020. Cit. on p. 43.

NILSSON, N. J. **Principles of artificial intelligence**. [*S.l.*]: Springer Science & Business Media, 1982. Cit. on p. 42.

OWASP Foundation. **OWASP Top Ten 2021**. 2021. https://owasp.org/Top10/. Cit. on p. 32.

PEFFERS, K.; TUUNANEN, T.; ROTHENBERGER, M. A.; CHATTERJEE, S. A design science research methodology for information systems research. **Journal of management information systems**, Taylor & Francis, v. 24, n. 3, p. 45–77, 2007. Cit. on p. 25.

PONNUSAMY, V.; MANICKAM, N. Blockchain technology for evidence integrity. *In*: **Forensic Intelligence and Deep Learning Approaches**. [*S.l.*]: IGI Global, 2025. Cit. on p. 22.

RAJKOMAR, A.; OREN, E.; CHEN, K. *et al.* Scalable and accurate deep learning with electronic health records. **NPJ digital medicine**, Nature Publishing Group UK London, v. 1, n. 1, p. 18, 2018. Cit. on p. 42.

REINSEL, D.; GANTZ, J.; RYDNING, J. **The Digitization of the World: From Edge to Core**. [*S.l.*]: International Data Corporation (IDC), 2018. IDC White Paper US44413318. Cit. on p. 20.

RLG-OCLC. **Trusted Digital Repositories: Attributes and Responsibilities–An RLG-OCLC Report**. [*S.l.*]: Research Libraries Group available at: www. rlg. org/longterm/repositories. pdf, 2002. Cit. on p. 51.

ROCHA, C. L. Repositórios para a preservação de documentos arquivísticos digitais. **Acervo**, v. 28, n. 2 jul-dez, p. 180–191, 2015. Cit. on p. 52.

RODRIGUES, G. A. P.; SERRANO, A. L. M.; VERGARA, G. F.; ALBUQUERQUE, R. d. O.; NZE, G. D. A. Impact, compliance, and countermeasures in relation to data breaches in publicly traded us companies. **Future Internet**, MDPI, v. 16, n. 6, p. 201, 2024. Cit. on pp. 21, 22, and 27.

SABEENA, M.; ABRAHAM, L. Convolutional block attention based network for copy-move image forgery detection. **Multimedia Tools and Applications**, Springer, v. 83, n. 1, p. 2383–2405, 2024. Cit. on p. 45.

SANJALAWE, Y.; AL-E'MARI, S.; FRAIHAT, S.; ABUALHAJ, M. *et al.* A deep learning-driven multi-layered steganographic approach for enhanced data security. **Scientific Reports**, Nature Publishing Group, v. 15, 2025. Cit. on pp. 58, 60, 61, and 64.

SANTOS, H. M. dos; FLORES, D. Modelo lógico da informação no open archival information system: uma reflexão arquivística sobre o pacote de informação para arquivamento. 2020. Cit. on p. 50.

SASI, S. B.; SIVANANDAM, N. A survey on cryptography using optimization algorithms in wsns. **Indian Journal of Science and Technology**, Indian Society for Education and Environment, v. 8, n. 3, p. 216, 2015. Cit. on pp. 9, 35, and 36.

SHI, Z.; HUANG, Y.; HE, Q.; XU, L.; LIU, S.; QIN, L.; JIA, Z.; LI, J.; HUANG, H.; ZHAO, L. Msminer—a developing platform for olap. **Decision Support Systems**, Elsevier, v. 42, n. 4, p. 2016–2028, 2007. Cit. on p. 42.

SINGH, A. K.; DAVE, M.; MOHAN, A. Hybrid technique for robust and imperceptible multiple watermarking using medical images. **Multimedia Tools and Applications**, Springer, v. 75, p. 8381–8401, 2016. Cit. on p. 47.

SMITH, J. *et al.* Multiparty trust levels in evidence management: Ensuring tamper-proof chain of custody in blockchain. **Information and Automation Sciences**, Sciendo, 2024. Cit. on p. 62.

SMITH, M.; BRONNER, W.; SHIMOMURA, E.; LEVINE, B.; FROEDE, R. Quality assurance in drug testing laboratories. **Clinics in laboratory medicine**, v. 10, n. 3, p. 503–516, 1990. Cit. on p. 20.

SRIKUMAR, R.; MALARVIZHI, C. Strong encryption using steganography and digital watermarking. *In*: **Proceedings of the 22nd Picture Coding Symposium, Seoul, Korea**. [*S.l.*: *s.n.*], 2001. p. 25–27. Cit. on p. 34.

STALLINGS, W. **Cryptography and network security: principles and practice**. [*S.l.*]: Pearson Upper Saddle River, 2017. Cit. on pp. 32 and 36.

SURESH, M.; BINOY, T. A comprehensive review based on video steganography for secure data transmission. **Multimedia Tools and Applications**, Springer, 2026. Cit. on p. 64.

SURESH, M.; SAM, I. S. Optimal wavelet transform using oppositional grey wolf optimization for video steganography. **Multimedia Tools and Applications**, Springer, v. 79, n. 37-38, p. 27023–27037, 2020. Cit. on p. 47.

TELLI, M.; OTHMANI, M.; LTIFI, H. A new approach to video steganography models with 3d deep cnn autoencoders. **Multimedia Tools and Applications**, Springer, p. 1–17, 2023. Cit. on pp. 63 and 65.

THAMES, L.; SCHAEFER, D. Cybersecurity for industry 4.0 and advanced manufacturing environments with ensemble intelligence. *In*: **Cybersecurity for Industry 4.0: Analysis for Design and Manufacturing**. [*S.l.*]: Springer, 2017. p. 243–265. Cit. on p. 31.

TORRES, J. A. S.; SANTOS, P. H. D.; SILVA, D. A. D.; VEIGA, C. E. L.; MEDEIROS, M. B.; VERGARA, G. F.; MENDONÇA, F. L. L.; JÚNIOR, R. T. D. S. Using spatial data and cluster analysis to automatically detect non-trivial relationships between environmental transgressors. *In*: IEEE. **2022 IEEE International Conference on Data Mining Workshops (ICDMW)**. [*S.l.*], 2022. p. 98–104. Cit. on p. 27.

UNDERWOOD, S. Blockchain beyond bitcoin. **Communications of the ACM**, ACM New York, NY, USA, v. 59, n. 11, p. 15–17, 2016. Cit. on pp. 9, 48, and 76.

United States District Court, Southern District of New York. Zubulake v. ubs warburg llc: The duty to preserve and produce electronic evidence. **229 F.R.D. 422**, 2004. Cit. on p. 20.

VARADARAJAN, M.; RAJKUMAR, N.; MOHANRAJ, A. Safeguarding digital archives with advanced strategies. *In*: **Blockchain-Based Digital Preservation**. [*S.l.*]: IGI Global, 2025. Cit. on pp. 58, 60, and 64.

VARGHESE, M. F. A clear-cut journey over image steganography techniques. **NAIVIGYAN**, p. 1, 2021. Cit. on p. 39.

VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, Ł.; POLOSUKHIN, I. Attention is all you need. **Advances in neural information processing systems**, v. 30, 2017. Cit. on pp. 44 and 45.

VERGARA, G. F.; GIACOMELLI, P.; SERRANO, A. L. M.; MENDONÇA, F. L. L. d.; RODRIGUES, G. A. P.; BISPO, G. D.; GONÇALVES, V. P.; ALBUQUERQUE, R. d. O.; JÚNIOR, R. T. d. S. Stego-stfan: A novel neural network for video steganography. **Computers**, MDPI, v. 13, n. 7, p. 180, 2024. Cit. on p. 27.

VERGARA, R. F.; SANTOS, P. H. D.; VERGARA, G. F.; MENDONÇA, F. L. L.; VEIGA, C. E. L.; PRACIANO, B. J. G.; SILVA, D. A. D.; JÚNIOR, R. T. D. S. A study of automatic speech recognition in portuguese by the brazilian general attorney of the union. *In*: IEEE. **2022 IEEE International Conference on Data Mining Workshops (ICDMW)**. [*S.l.*], 2022. p. 226–231. Cit. on p. 27.

WANG, X.; GIRSHICK, R.; GUPTA, A.; HE, K. Non-local neural networks. *In*: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [*S.l.*: *s.n.*], 2018. p. 7794–7803. Cit. on p. 45.

WERTHMULLER, N. **Blockchain applied to digital archives**. Dissertation (Master of Science) — HES-SO, 2025. Cit. on pp. 58, 59, 60, and 64.

WOO, S.; PARK, J.; LEE, J.-Y.; KWEON, I. S. Cbam: Convolutional block attention module. *In*: **Proceedings of the European conference on computer vision (ECCV)**. [*S.l.*: *s.n.*], 2018. p. 3–19. Cit. on p. 45.

WU, N.; YANG, Z.; YANG, Y.; LI, L.; SHANG, P.; MA, W.; LIU, Z. Stbs-stega: Coverless text steganography based on state transition-binary sequence. **International Journal of Distributed Sensor Networks**, SAGE Publications Sage UK: London, England, v. 16, n. 3, p. 1550147720914257, 2020. Cit. on p. 38.

XU, S.; LI, Z.; ZHANG, Z.; LIU, J. An end-to-end robust video steganography model based on a multi-scale neural network. **Electronics**, MDPI, v. 11, n. 24, p. 4102, 2022. Cit. on pp. 64 and 65.

YAN, W.; SHEN, J.; CAO, Z.; DONG, X. Blockchain based digital evidence chain of custody. *In*: **Proceedings of the 2020 The 2nd International Conference on Blockchain Technology**. [*S.l.*: *s.n.*], 2020. p. 19–23. Cit. on pp. 62 and 65.

ZAREMBA, W.; SUTSKEVER, I.; VINYALS, O. Recurrent neural network regularization. **arXiv preprint arXiv:1409.2329**, 2014. Cit. on p. 44.

ZEAR, A.; SINGH, A. K.; KUMAR, P. A proposed secure multiple watermarking technique based on dwt, dct and svd for application in medicine. **Multimedia tools and applications**, Springer, v. 77, p. 4863–4882, 2018. Cit. on p. 47.

ZHANG, S.; WU, Y.; CHE, T.; LIN, Z.; MEMISEVIC, R.; SALAKHUTDINOV, R. R.; BENGIO, Y. Architectural complexity measures of recurrent neural networks. **Advances in neural information processing systems**, v. 29, 2016. Cit. on p. 44.

ZHOU, S.; ZHANG, J.; PAN, J.; XIE, H.; ZUO, W.; REN, J. Spatio-temporal filter adaptive network for video deblurring. *In*: **Proceedings of the IEEE/CVF international conference on computer vision**. [*S.l.*: *s.n.*], 2019. p. 2482–2491. Cit. on p. 64.

Appendix A – API REST Specification

This appendix documents the complete REST API exposed by the system's API Gateway. All endpoints accept and return JSON unless otherwise noted.

A.1 Blockchain Endpoints

Method	Endpoint	Description
GET	/api/blockchain/status	Returns chain status including length, difficulty, and pending transactions
GET	/api/blockchain/blocks	Returns list of all blocks with pagination support
GET	/api/blockchain/blocks/{index}	Returns specific block by index
GET	/api/blockchain/documents	Returns list of all registered documents
GET	/api/blockchain/documents/{id}	Returns specific document by ID
POST	/api/blockchain/register	Registers new document on blockchain
POST	/api/blockchain/verify	Verifies document integrity against blockchain
POST	/api/blockchain/validate	Validates entire chain integrity

A.2 SIP Management Endpoints

Method	Endpoint	Description
POST	/api/sip/create	Creates new Submission Information Package
POST	/api/sip/{id}/upload	Uploads file to existing SIP
PUT	/api/sip/{id}/metadata	Sets or updates SIP metadata
POST	/api/sip/{id}/steganography	Applies steganographic embedding to SIP files
POST	/api/sip/{id}/finalize	Finalizes SIP for ingestion
GET	/api/sip/{id}	Returns SIP details and status

A.3 AIP Management Endpoints

Method	Endpoint	Description
POST	/api/aip/ingest/{sip_id}	Ingests finalized SIP, creating AIP
GET	/api/aip/{id}	Returns AIP details including preservation metadata
POST	/api/aip/{id}/fixity	Performs fixity check on AIP
GET	/api/aip/{id}/premis	Returns PREMIS events for AIP
GET	/api/aip/{id}/mets	Returns METS document for AIP

A.4 DIP Management Endpoints

Method	Endpoint	Description
POST	/api/dip/generate/{aip_id}	Generates DIP from AIP
GET	/api/dip/{id}	Returns DIP details
GET	/api/dip/{id}/download	Downloads DIP as ZIP archive
GET	/api/dip/{id}/files/{filename}	Downloads specific file from DIP

A.5 Custody Endpoints

Method	Endpoint	Description
GET	/api/custody/{doc_id}/chain	Returns complete custody chain for document
GET	/api/custody/{doc_id}/history	Returns custody history with details
POST	/api/custody/transfer	Records custody transfer event
GET	/api/custody/events	Returns all custody events with filtering
GET	/api/custody/statistics	Returns custody statistics

A.6 Administration Endpoints

Method	Endpoint	Description
GET	/api/admin/status	Returns system status for all services
GET	/api/admin/storage	Returns storage information
POST	/api/admin/cleanup	Performs data cleanup (requires confirmation)
GET	/api/admin/version	Returns system version information

Appendix B – BagIt Package Structure

This appendix illustrates the structure of information packages as implemented in the system, following RFC 8493 BagIt specification.

B.1 Submission Information Package (SIP)

```
SIP_2024-01-15_abc123/
|-- bagit.txt
|-- bag-info.txt
|-- manifest-sha256.txt
|-- manifest-sha512.txt
|-- tagmanifest-sha256.txt
|-- tagmanifest-sha512.txt
|-- data/
|   |-- document_001.pdf
|   |-- document_002.jpg
|   `-- document_003.mp4
`-- metadata/
    `-- dublin_core.xml
```

B.2 Archival Information Package (AIP)

```
AIP_2024-01-15_def456/
|-- bagit.txt
|-- bag-info.txt
|-- manifest-sha256.txt
|-- manifest-sha512.txt
|-- tagmanifest-sha256.txt
|-- tagmanifest-sha512.txt
|-- data/
|   |-- document_001.pdf
|   |-- document_002.jpg
|   `-- document_003.mp4
`-- metadata/
    |-- dublin_core.xml
    |-- isadg.xml
    |-- premis.xml
    |-- mets.xml
    `-- pdi.xml
```

B.3 Example bag-info.txt

```
Source-Organization: Digital Custody System
Contact-Name: System Administrator
Contact-Email: admin@example.org
```

```
External-Description: document package for case 2024-001
Bagging-Date: 2024-01-15
External-Identifier: AIP_2024-01-15_def456
Bag-Size: 15.7 MB
Payload-Oxum: 16478208.3
```

B.4 Example manifest-sha256.txt

```
a7f5c8d9e2b1a3f4c6d8e0a2b4c6d8e0f2a4b6c8d0e2f4a6  data/document_001.pdf
b8e6d7c5a4f3e2d1c0b9a8f7e6d5c4b3a2f1e0d9c8b7a6f5  data/document_002.jpg
c9f7e8d6b5a4c3f2e1d0c9b8a7f6e5d4c3b2a1f0e9d8c7b6  data/document_003.mp4
```

Appendix C – Stego-STFAN Architecture Details

This appendix provides detailed specifications of the Stego-STFAN neural network architecture.

C.1 Network Configuration

Parameter	Value
Total Parameters	43.91 M
Input Resolution	144×180 pixels
Color Space	YUV
Wavelet Transform	Discrete Wavelet Transform (DWT)
Embedding Subbands	LH, HL
Base Channels	64

C.2 Training Configuration

Parameter	Value
Framework	PyTorch 2.0
GPU	NVIDIA L4 (22 GB VRAM)
Batch Size	4
Optimizer	Adam
Learning Rate	0.0001
Training Epochs	100
Dataset Size	5,700 frames
Train/Validation Split	80% / 20%
Datasets	UCF-101, Weizmann Action

C.3 Loss Function

The training loss function combines three components:

$$\mathcal{L}_{total} = \lambda_c \mathcal{L}_{container} + \lambda_s \mathcal{L}_{secret} + \lambda_b \mathcal{L}_{balance} \tag{C.1}$$

where:

$$\mathcal{L}_{container} = \text{MSE}(I_{cover}, I_{container}) \tag{C.2}$$

$$\mathcal{L}_{secret} = \text{MSE}(I_{secret}, I_{restored}) \tag{C.3}$$

$$\mathcal{L}_{balance} = |\mathcal{L}_{container} - \mathcal{L}_{secret}| \tag{C.4}$$

The balance weights used were $\lambda_c = 1.0$, $\lambda_s = 1.0$, and $\lambda_b = 0.5$.

C.4 Attention Mechanisms

Non-Local Self-Attention (NLSA): Captures long-range dependencies within individual frames by computing attention weights between all spatial positions.

Non-Local Co-Attention (NLCA): Identifies correlations between cover and secret frames, determining optimal embedding locations.

CBAM (Convolutional Block Attention Module): Applies sequential channel and spatial attention to refine feature representations.

C.5 Filter Adaptive Convolutional (FAC) Layer

The FAC layer generates dynamic convolutional filters conditioned on input features:

$$F_{out} = \text{FAC}(F_{in}, F_{cond}) = F_{in} * \varphi(F_{cond}) \tag{C.5}$$

where φ is a filter-generating network that produces spatially-varying kernels based on the conditioning features F_{cond}.

Appendix D – System Deployment Configuration

This appendix documents the deployment configuration for the complete system.

D.1 Docker Compose Configuration

```
version: '3.8'

services:
  api_gateway:
    build:
      context: .
      dockerfile: Dockerfile.api
    ports:
      - "8000:8000"
    volumes:
      - ./data:/app/data
      - ./stego_files:/app/stego_files
    environment:
      - BLOCKCHAIN_DIFFICULTY=4
      - IPFS_ENABLED=true
      - IPFS_HOST=ipfs
      - IPFS_PORT=5001
    depends_on:
      - ipfs

  frontend:
    build:
      context: .
      dockerfile: Dockerfile.frontend
    ports:
      - "4200:80"
    depends_on:
      - api_gateway

  ipfs:
    image: ipfs/kubo:latest
    ports:
      - "5001:5001"
      - "8080:8080"
    volumes:
      - ipfs_data:/data/ipfs

volumes:
  ipfs_data:
```

D.2 Environment Variables

Variable	Default	Description
BLOCKCHAIN_DIFFICULTY	4	Number of leading zeros required in block hash
BLOCKCHAIN_DATA_PATH	./data/blockchain	Path to blockchain persistence file
CUSTODY_DATA_PATH	./data/custody	Path to custody packages
IPFS_ENABLED	true	Enable/disable IPFS integration
IPFS_HOST	localhost	IPFS daemon hostname
IPFS_PORT	5001	IPFS API port
STEGO_CODEC	FFV1	Lossless codec for steganography
LOG_LEVEL	INFO	Logging verbosity

D.3 System Requirements

Minimum Requirements:

- CPU: 4 cores
- RAM: 8 GB
- Storage: 50 GB SSD
- Docker 20.10+
- Docker Compose 2.0+

Recommended for Stego-STFAN:

- GPU: NVIDIA with 8+ GB VRAM
- CUDA 11.0+
- cuDNN 8.0+

Appendix E – Source Code Organization

This appendix documents the organization of the system's source code.

E.1 Backend Structure

```
/
|-- api_gateway/
|   |-- __init__.py
|   |-- main.py               # FastAPI application entry point
|   `-- dependencies.py       # Dependency injection
|
|-- blockchain_module/
|   |-- api/
|   |   `-- routes.py          # Blockchain REST endpoints
|   |-- models/
|   |   |-- block.py           # Block dataclass and mining
|   |   `-- transaction.py    # Transaction types
|   |-- services/
|   |   `-- blockchain_service.py  # Chain management
|   `-- config/
|       `-- settings.py       # Blockchain configuration
|
|-- chain_of_custody_module/
|   |-- api/
|   |   `-- routes.py          # Custody REST endpoints
|   |-- models/
|   |   |-- sip.py             # Submission Information Package
|   |   |-- aip.py             # Archival Information Package
|   |   |-- dip.py             # Dissemination Information Package
|   |   `-- bag.py             # BagIt container
|   |-- services/
|   |   |-- bagit_service.py   # BagIt operations
|   |   `-- stego_lsb.py       # LSB steganography
|   `-- schemas/
|       `-- isad_g.py          # ISAD(G) metadata schema
|
|-- common_utils/
|   |-- __init__.py
|   `-- timezone.py           # Timezone utilities
|
|-- auth_module/
|   |-- api/
|   |   `-- routes.py          # Authentication endpoints
|   |-- models/
|   |   `-- user.py            # User model
|   `-- services/
|       `-- auth_service.py    # Authentication logic
|
`-- data/
```

```
    `-- blockchain_data/
        `-- chain.json         # Persisted blockchain
```

E.2 Frontend Structure

```
frontend-angular/
|-- src/
|   |-- app/
|   |   |-- app.routes.ts            # Route definitions
|   |   |-- app.config.ts            # Application config
|   |   |
|   |   |-- core/
|   |   |   `-- services/
|   |   |       |-- api.service.ts        # HTTP client
|   |   |       |-- blockchain.service.ts
|   |   |       |-- custody.service.ts
|   |   |       `-- status.service.ts
|   |   |
|   |   |-- features/
|   |   |   |-- dashboard/
|   |   |   |   `-- dashboard.component.ts
|   |   |   |-- admin/
|   |   |   |   `-- admin.component.ts
|   |   |   |-- blockchain/
|   |   |   |   `-- blockchain.component.ts
|   |   |   |-- upload/
|   |   |   |   `-- upload.component.ts
|   |   |   `-- custody/
|   |   |       `-- custody.component.ts
|   |   |
|   |   `-- layout/
|   |       `-- shell/
|   |           `-- shell.component.ts    # Main layout
|   |
|   `-- assets/
|       `-- images/
|
`-- angular.json
```


UnB