National College of Ireland

A Hybrid Blockchain Solution for Privacy-Preserving Civil Status Data Verification Using Zero-Knowledge Proofs, AES-RSA Encryption, IPFS, and Cloud Services

MSc Research Project
Cloud Computing

Junior Khan Azamah
Student ID: x23110970

School of Computing
National College of Ireland

Supervisor: Prof. Sean Heeney

National College of Ireland
Project Submission Sheet
School of Computing

National College of Ireland

Student Name:	Junior Khan Azamah
Student ID:	x23110970
Programme:	Cloud Computing
Year:	2024
Module:	MSc Research Project
Supervisor:	Prof. Sean Heeney
Submission Due Date:	12/12/2024
Project Title:	A Hybrid Blockchain Solution for Privacy-Preserving Civil Status Data Verification Using Zero-Knowledge Proofs, AES-RSA Encryption, IPFS, and Cloud Services
Word Count:	6388
Page Count:	19

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	12th December 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

A Hybrid Blockchain Solution for Privacy-Preserving Civil Status Data Verification Using Zero-Knowledge Proofs, AES-RSA Encryption, IPFS, and Cloud Services

Junior Khan Azamah
x23110970

Abstract

Civil status documents, such as birth certificates, are the basis for important legal operations, but their vulnerability to forgery remains an ongoing challenge. Previous approaches to address this have focused on ensuring the integrity of the document itself, rather than the data it contains. To address this, this study proposes a hybrid blockchain-based solution that integrates zero-knowledge proofs, a hybrid encryption model with AES and RSA encryption, IPFS, and cloud services to provide a secure, scalable, and privacy-preserving solution for verifying civil status data. And by conducting extensive experiments to evaluate the performance and effectiveness of the proposed system, the study validates the feasibility of the proposed architecture by refining the system for real-world implementation, and lay the groundwork for further research in this area.

Index Terms–Civil status, AES-RSA Encryption, IPFS, Zero-Knowledge Proofs(ZKP), Cloud, Data verification, Ethereum, Polygon zkEVM

1 Introduction

Civil status records, such as birth and marriage certificates, are essential for a variety of social and legal purposes, and form the backbone of identity verification and legal entitlements, both domestically and internationally. These records are critical for governments to maintain accurate population data, administer social services, and enforce laws, and internationally, they are often required for immigration, visa applications, and other cross-border transactions. Although countries have diverse implementations of their civil registry systems, these records have traditionally been managed through centralized registries. However, such systems are prone to a range of issues, including security breaches, poor organization and inadequate measures to ensure the integrity of the documents making the officially issued documents highly susceptible to forgery. This has created a lucrative black market for forged documents , which are often used for international purposes, such as visa applications, posing significant challenges for foreign authorities who lack the means to adequately verify the authenticity of the information contained within these documents."

The International Commission on Civil Status (ICCS), an intergovernmental body of EU and non-EU member states dedicated to ensuring the reliability and integrity

of civil status records, conducted several comprehensive studies on civil status fraud in its member states through surveys and analysis of data collected from member states (Macniven; 2012).

The study concluded that civil status fraud is steadily increasing. Member States are still frequently confronted with foreign civil status documents that do not accurately reflect the true circumstances, even when the documents appear to be legitimate, complete with official stamps, signatures, and holograms, the information contained within them is often inaccurate and/or fraudulent. In its study, the ICCS identified two main types of fraud: (1) false statements made to civil registrars at the time of creating and registering a new civil status event, such as births and marriages and (2) the intentional submission to foreign states government bodies such as embassies, forged or altered civil documents which were issued by other countries. The latter was found to be the most prevalent type of fraud, with birth certificates being the the most commonly falsified, where key information such as birth and identity details, was often altered, primarily driven by motives like immigration benefits, inheritance claims and criminal activities. Since the information contained in birth certificates serve as the foundation for other crucial documents, such as passports, residence permits and other forms of identification Loon (2024) , such alteration have more far reaching consequences.

Civil status fraud remains a key concern for ICCS. Although ICCS has explored various approaches to address these challenges, they continue to fall short. The proposed "ICCS Platform for Electronic Transmission of Documents and Civil Status Data" aims to provide Member States with an additional tool to combat fraud, but this platform is still under development. Current solutions remain inadequate as the centralised system suffers from slow and unreliable verification processes and is vulnerable to data breaches and cyber-attacks. Another growing issue with one of their proposed solutions; consular cooperation, is the "paradoxical effect of secure true-but-false documents" (Macniven; 2012). This occurs when a state verifies a document as authentic because it appears physically secure and difficult to falsify, yet it still conveys incorrect information about a person's civil status. This leads to subsequent mis-identifications when the document is shared with other states

There is a critical need to protect the accuracy and authenticity of civil status data itself, rather than focusing solely on the security of the physical documents. To address this need and overcome current limitations, a fundamentally different approach is necessary. I propose a novel blockchain-based system specifically designed for secure, privacy-preserving, real-time verification, authentication, and efficient management of civil status data.

Recent research surveys have already highlighted the potential of blockchain technology to revolutionise public service delivery by increasing efficiency, transparency, and security (Aliti et al.; 2022; Alam et al.; 2021). While blockchain has shown promise in various public sectors such as e-voting, land registration, identity management, and supply chain and logistics [1] with an example of Estonia adopting blockchain technology on a national scale and integrating the technology into critical sectors such as healthcare, law, security, and commerce, revolutionizing data management and public service delivery [2]. However, the application of this technology to securing civil status data presents a unique opportunity to address critical challenges such as the aforementioned document forgery.

The proposed system prioritises privacy through the use of Zero-Knowledge Proofs

[1]IBM. Blockchain for Government: https://www.ibm.com/blockchain/industries/government

[2]BSV blockchain: https://www.bsvblockchain.org/news/6-countries-using-blockchain-right-now

(ZKPs), enhances security by employing advanced cryptographic techniques in a hybrid approach that combines Advanced Encryption Standard (AES) for symmetric encryption and Rivest-Shamir-Adleman (RSA) for asymmetric encryption to create a more resilient encryption framework, and achieves scalability by leveraging the Ethereum blockchain platform. To handle large storage needs, the system uses the InterPlanetary File System (IPFS) for decentralised storage of civil status data, ensuring data availability, integrity and resistance to tampering. It also incorporates existing cloud technologies, such as DynamoDB and other AWS services, to support real-time access and retrieval of civil records, providing a balance between decentralised resilience and the performance benefits of centralised storage.

Objective:

- Design a blockchain-based civil status system, to combat fraud and forgery in civil status documents, by securing the integrity of the underlying civil status data.
- Base the system around privacy-preserving techniques, with a key focus on security and scalability by leveraging appropriate blockchain and cloud technologies
- Enable consular cooperation by allowing one country to request and verify civil data issued by another country securely.
- Evaluate the proposed system's performance based on the chosen technologies, assessing factors such as efficiency, scalability, security, and reliability.

Research Question: This research is guided by the following question:

Can a hybrid blockchain-based system that integrates zero-knowledge proofs, AES and RSA encryption, IPFS for decentralised storage, and cloud services such as DynamoDB provide a secure, scalable, and privacy-preserving solution for verifying civil status data while addressing performance and security challenges?

Document Structure: The rest of the document is structured as follows:

1. **Literature Review -** This section lays the foundation for this research by critically reviewing similar work and situating it within the academic literature of previously published studies.
2. **Methodology -** This section describes the steps followed in the research, the materials and equipment used, and the methods employed to gather samples.
3. **Design Specification -** This section identifies and presents the techniques, workflows and architecture that underlie the implementation of the proposed system.
4. **Implementation -** This section describes the implementation of the proposed system, including the tools and programming languages used.
5. **Evaluation -** This section provides an analysis of the experiments conducted in this research, including the results and a discussion of their implications.
6. **Conclusion Future Work -** This section discusses the implications of the research and provides directions for future research.

2 Literature Review

2.1 Combating Fraudulent Civil Status Documents

Prior to the ICCS proposals, registrars in different countries used a variety of methods to verify documents. In some countries, verification was limited to assessing the physical form of submitted documents and confirming the authenticity of signatures. In others, registrars were authorised to verify the existence and content of foreign documents by contacting diplomatic or consular representatives or by contacting foreign authorities directly. However, these methods had significant limitations: (1) the verification process by contacting consular representatives was slow, and (2) they focused primarily on the security of the physical document rather than ensuring the accuracy of the civil status data itself.

Current ICCS efforts to combat fraud is focused on two main areas: "Fraud by electronic means, civil status and identification" and "Consular cooperation in the verification of foreign documents" (Macniven; 2012). To address the first area, ICCS proposed "ICCS Platform for Electronic Transmission of Documents and Civil Status Data" a platform for the electronic exchange of civil data and delivering civil status documents between member states as a countermeasure to electronic fraud affecting the security of civil documents. And for the next area, they based their study on the idea that a document verified by one ICCS member state can be considered valid in a case where another state wants to verify the same document. However, the experiments revealed different practices and levels of verification between states due to the lack of a legal framework for consular cooperation (Macniven; 2012). Additionally, as previously mentioned the issue of "secure true but false documents" still persists, even more so with consular cooperation, where documents may appear secure yet still convey incorrect information about a person's civil status data data.

Interpol, one of the organizations that has played a pivotal role in combating fraudulent civil status documents by providing specialized tools and databases accessible to law enforcement and governmental agencies for detecting document fraud. Tools such as the Frontex INTERPOL Electronic Library Document System (FIELDS) provide visual information on key markers indicating forged documents, while EdisonTD serves as a reference tool containing images of genuine identity documents. Interpol also offers specialized databases like DISCS, which stores records of civil status certificates, and SLTD, a database of lost, stolen, or invalid travel and identity documents, populated jointly by partner organizations. [3]. Although Interpol's solutions are effective in identifying most counterfeit civil documents, their success is often limited by the failure of officials to check documents against these databases, as Interpol Secretary General Jürgen Stock points out: "The failure to to screen identity documents against INTERPOL databases is a security gap." While these tools help to detect fraudulent documents, they do not address the need to protect the integrity of the underlying civil status data itself.

2.2 Blockchain for Securing Data Contained in Documents

A number of studies have explored the potential of blockchain technology as a tool to enhance the verification, authentication and security of both digital and physical certi-

[3]Interpol: https://www.interpol.int/en/Crimes/Counterfeit-currency-and-security-documents/Identity-and-travel-document-fraud

ficates, such as academic and land title certificates. In this section, I review the progress made in this area, highlighting the innovative blockchain-based approaches used in previous research and their applicability in ensuring the integrity and reliability of the data contained in civil status documents.

(Zainuddin and Choo; 2022; Leka and Selimi; 2021) both explored and proposed Ethereum blockchain-based applications for the validation and verification of academic certificates. While their approaches demonstrated the potential to improve the reliability of academic certificates, each had shortcomings in their implementation. Both studies proposed the use of IPFS; (Zainuddin and Choo; 2022) used IPFS as a decentralised storage solution, storing the certificates on IPFS and maintaining a pointer to the files on smart contracts. This approach, similar to the method proposed by (Nyaletey et al.; 2019), has notable advantages. However, (Zainuddin and Choo; 2022)'s implementation has two significant drawbacks. First, they did not include a separate backend for storing the IPFS hash and metadata, which would have greatly improved lookup speed, search and other functionalities, and ultimately reduced gas costs. Second, while they proposed to encrypt data before storing it on the IPFS, they required users to manually enter encryption/decryption keys for each document during the upload/verification process.

This approach, also used by (Leka and Selimi; 2021), presents several challenges: managing and securely storing the encryption keys is complex, and relying on users to manually enter the keys adds unnecessary friction and potential for error. To address these shortcomings, my proposed solution adopts a hybrid encryption approach, where civil data rather than the PDF documents themselves are encrypted using symmetric encryption, and the keys themselves are secured using asymmetric encryption and securely managed and stored using the AWS Key Management Service (KMS) to provide a centralised key management solution. I also propose an off-chain backend based on AWS cloud services, implemented to serve as a reference store of civil metadata, IPFS and transaction hashes, this approach will enable fast access and lookup functionality, not only does it reduce gas costs when doing on-chain lookup, it also provides a way to automate the retrieval of data in IPFS and subsequently its transaction hash.

While this research primarily focuses on securing the integrity of the data within civil documents, it is important to recognize that many countries still rely solely on physical documents for critical processes, such as visa applications. Addressing this situation in a digital-based system requires a thoughtful approach. Mthethwa et al. (2018) proposed a blockchain-based solution incorporating Optical Character Recognition (OCR), cryptographic hashing and digital signatures. While their approach shows potential, I argue that OCR introduces unnecessary complexity and an additional point of failure. It assumes that physical documents will always be in optimal condition for accurate OCR readings, which is not realistic in practical scenarios. For example, a visa officer often only needs to review key details, such as a name or date of birth, rather than the entire document. With this in mind, I propose an architecture leveraging Zero-Knowledge Proofs (ZKPs) to enable selective verification of specific attributes. By generating ZKPs for critical details (e.g. name,date of birth, place of issue), verifiers can authenticate these attributes without needing access to the full document, they only need to enter the information they see for these attributes on the civil document. This approach eliminates the complexities and uncertainties of OCR, reduces unnecessary data processing, and enhances data transmission efficiency.

2.3 Enhancing Security and Privacy with Advanced Cryptographic Techniques

Building on the approach proposed by ICCS for consular cooperation between member states as a method of verifying civil data Macniven (2012), I propose the use of Zero-Knowledge Proofs (ZKPs) as a tool to ensure privacy when implementing consular cooperation within my proposed system. ZKPs are cryptographic techniques that allow one party (the prover) to prove to another party (the verifier) that a certain statement is true or that it possesses certain knowledge, without revealing the underlying secret or any additional information about the statement itself Goldwasser et al. (1985). In essence, a state could prove the integrity of civil data to another state without having to reveal or transmit additional sensitive information about it;Notably, several governments, such as Estonia, have already integrated ZKPs into their digital identity infrastructure to enhance security and privacy in their operations [4]. This approach not only minimises network bandwidth requirements, but also mitigates the risks associated with data breaches during transmission.

2.4 Optimizing Performance of Blockchain Systems

The AWS backend is also designed with scalabiliy in mind, taking advantage of serveless technlogies like lambda, dynamoDB and other things to ensure autoscaling to handle concurrent operations and more

Given the potentially vast amount of civil records and the concurrent verifications that the system may need to handle, scalability and performance are crucial factors to consider. For a blockchain-based application, the choice of blockchain platform significantly impacts scalability. Various solutions have been proposed to enhance blockchain scalability, such as increasing block size, sharding the network, and improving consensus algorithms, as proposed by (Kim; 2022). However, these approaches often involve trade-offs between decentralization and security, a challenge commonly referred to as the blockchain trilemma (Pierro and Tonelli; 2022). Ethereum, in particular, has faced several criticisms for its scalability issues [5]. In contrast, (Pierro and Tonelli; 2022) analyzed Solana, a blockchain designed to address the blockchain trilemma through its Proof of History (PoH) consensus mechanism, and found it to be more scalable and significantly more cost-efficient than Ethereum.

Despite these findings, I argue that Ethereum remains the better choice for the system I propose. Its scalability challenges have been significantly mitigated by advances such as Ethereum 2.0 and the adoption of layer 2 solutions [6]. More importantly, Ethereum is a well-established platform that has been extensively tested and proven capable of supporting large-scale decentralised applications, and its widespread adoption and large developer community provide a level of reliability that newer platforms such as Solana, despite their promise, have yet to match.

[4]Forbes: https://www.forbes.com/councils/forbesbusinesscouncil/2023/07/25/zero-knowledge-proof-a-revolutionary-leap-in-data-protection/

[5]Lite finance: https://www.litefinance.org/blog/for-beginners/how-to-trade-crypto/solana-vs-ethereum/

[6]: https://medium.com/coinmonks/ethereum-scalability-challenges-and-innovative-solutions-0730a3153

2.5 Research Niche

A critical review of the literature reveals significant gaps in current methods for securing and ensuring the integrity of civil data, as well as shortcomings in existing blockchain-based approaches for validating the authenticity of civil documents.It provides a basis for justifying my proposed solution, highlighting the need for an architecture that is not only highly scalable, but also optimised to preserve and secure the privacy of civil data.

The expected contribution of this research is a decentralised blockchain-based system, using Ethereum to ensure scalability without compromising security or decentralisation.The proposed system will use a hybrid cryptographic system and Zero-Knowledge Proofs (ZKPs) to verify the integrity of civil documents without exposing the entire dataset. This approach, while creating the possibility for consular cooperation, also effectively addresses issues of fraud and forgery in civil status documents by securing the data within the documents themselves. In addition, it provides secure, real-time verification, authentication and efficient management capabilities while preserving privacy and ensuring interoperability across multiple systems and platforms.

3 Methodology

3.1 Steps followed in the research

This research was conducted through a series of well-defined methodological steps:

- Literature Review: Began by conducting a comprehensive review of existing methods for combating fraud in civil status documents and blockchain-based solutions for verifying data integrity within documents in order to identify the gaps in existing solutions.
- Design Specification: Designed the architecture of the two-tier blockchain-based solution, including the encryption system and the process workflows
- Implementation: Developed the system components, the ethereum smart contracts, encryption mechanisms, and the backend infrastructure. Integrated IPFS for decentralized storage and AWS services for managing encryption keys and user data.
- Simulation: I simulated the various workflows using sample civil data; uploading and verifying a civil record, and while doing this I measured the systems performance based on metrics such as transaction gas prices, encryption/decryption times, and AWS service costs.
- Evaluation: Based on the simulations i compared transaction costs (gas fees) with that of other blockchain platforms, such as Hyperledger Fabric to justify the choice of blockchain platform and conducted performance comparisons between IPFS and traditional storage methods.

3.2 Materials used in the research

- Blockchain Platform:
  - Ethereum blockchain (testnets and local development environment)

  - Binance Smart Chain testnet
  - Polygon zkEVM testnet
- Cryptographic Tools:
  - SnarkJS for Zero-Knowledge Proofs (ZKPs) implementation
  - CryptoJS
- Backend Infrastructure:
  - AWS Lambda for serverless compute
  - DynamoDB for database storage
  - AWS KMS (Key Management Service) for encryption key management
  - AWS Amplify for frontend deployment and integration with backend services
  - Amazon Cognito for user authentication and management
  - AWS IAM (Identity and Access Management) for access control and permissions
- Development Tools:
  - Solidity for smart contract development
  - Truffle and Hardhat for testing and deployment
  - Ganache for local blockchain simulation
  - AWS CDK (Cloud Development Kit) for infrastructure as code
  - TypeScript for backend and frontend development
  - ReactJS for building the user interface
  - Pinata for working with IPFS decentralized storage
  - MetaMask for Ethereum wallet integration and user authentication
  - VsCode IDE
  - Remix IDE

3.3 How samples were gathered

The sample data used to simulate the processes of the proposed system consisted of mock civil status records, specifically birth certificate data, stored in a CSV file; the dataset included fields such as name, date of birth and place of issue. a small set of records was initially created manually to ensure that the data was both realistic and diverse, and accurately reflected real-world scenarios. A script was then used to automatically generate thousands of additional records based on the structure and diversity of the manually created samples.

3.4 How measurements were made

In this study, experiments were conducted to simulate and stress test various likely scenarios for the system. This included comparing the encryption times of the symmetric encryption algorithms AES, DES, 3DES and RC4 with a range of key sizes, as well as measuring the transaction costs of deploying and interacting with the proposed smart contracts on the Ethereum, Binance Smart Chain and Polygon zkEVM testnets.

To compare the encryption times for civil records, all algorithms were tested on the same dataset using a TypeScript script that ran the simulations. To stress test the smart contracts on each network, a custom script written in Solana using the Remix IDE was used to run on each test blockchain network. Each test was repeated five times to ensure reliability.

3.5 Statistical techniques used

Descriptive statistical analysis was the main technique used in this study. The data collected from the experiments were summarised into simpler and more compact forms, after which data visualisation tools in Google Sheets, such as graphs and tables, were used to more easily analyse and interpret the data

4 Design Specification

4.0.1 System architecture

The architecture of the proposed system is illustrated in Figure 1. It follows a two-tier structure, comprising the blockchain tier and the backend tier. The blockchain tier is based on smart conotracts built on the Ethereum blockchain a user would have to be authenticated through MetaMask inorder to invoke these smart contracts. For secure decentralized storage, the system incorporates IPFS to securely store and retrieve encrypted civil data.

On the other hand, the backend tier is based on AWS serverless technologies. AWS Amplify is used together with Amazon Cognito and IAM for user authentication and authorization, while AWS Amplify also serves the frontend. The backend APIS are implemented with API gateway and links requests to AWS Lambda, which handles backend logic,and DynamoDB as a serveless database solution for storing IPFS hash keys, transaction hashes and civil metadata of processed civil documents. AWS KMS is used to securely manage the symmetric encryption keys used to encrypt civil data prior to storage on IPFS.

4.0.2 Encryption architecture

The proposed system uses a hybrid encryption approach, which is illustrated in Figure 2. First, an AES encryption key is generated in the back-end to encrypt the civil data. AWS KMS is used to securely store and manage this AES key. The process starts with AWS KMS sending an asymmetric public key, which is used to encrypt the AES key after which the encrypted AES key is then securely returned to AWS KMS for storage.

When a user logs into the system and needs access to encrypt or decrypt civilian data, a request is made to AWS KMS to retrieve the decrypted AES key, ensuring both secure key management and seamless encryption and decryption processes.

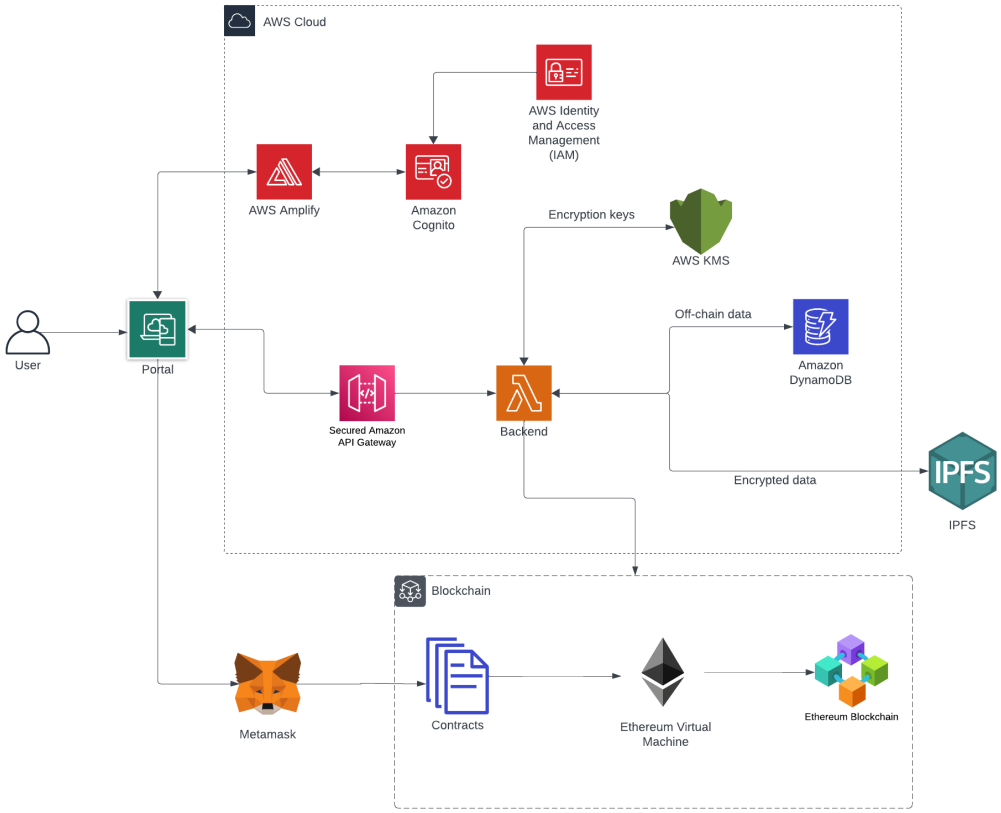

Figure 1: System architecture

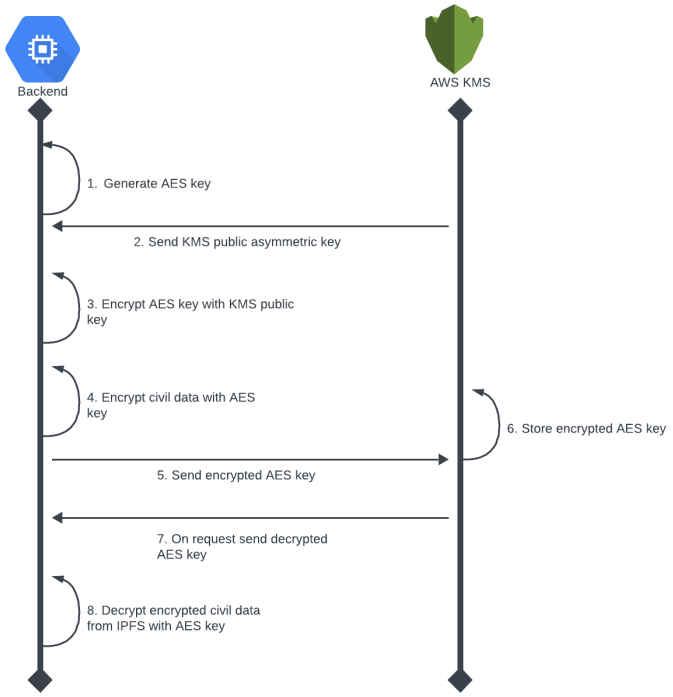

Figure 2: Encryption architecture

4.1 System Workflow

4.1.1 Data upload

Figure 3 illustrates the workflow for uploading civil data. First, the civil data is encrypted using AES encryption and the encrypted data is stored in IPFS. IPFS then returns a document hash, which serves as a unique reference to the content stored on IPFS.

Next, the IPFS hash of the uploaded civil data, along with its metadata, is encrypted and written to the blockchain using smart contracts. Finally, the blockchain transaction hash, the IPFS hash and the civil metadata are stored in DynamoDB to enable efficient search and lookup capabilities.

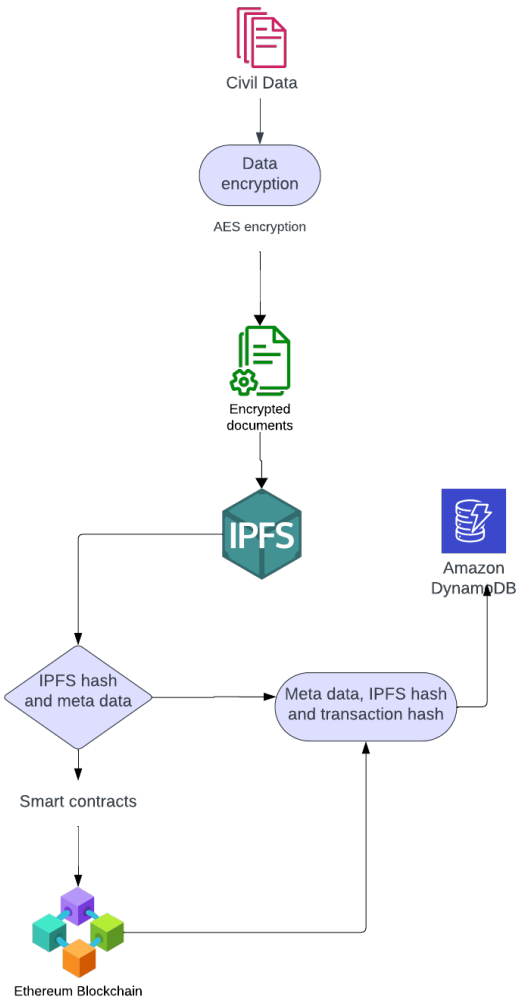

Figure 3: Uploading civil documents

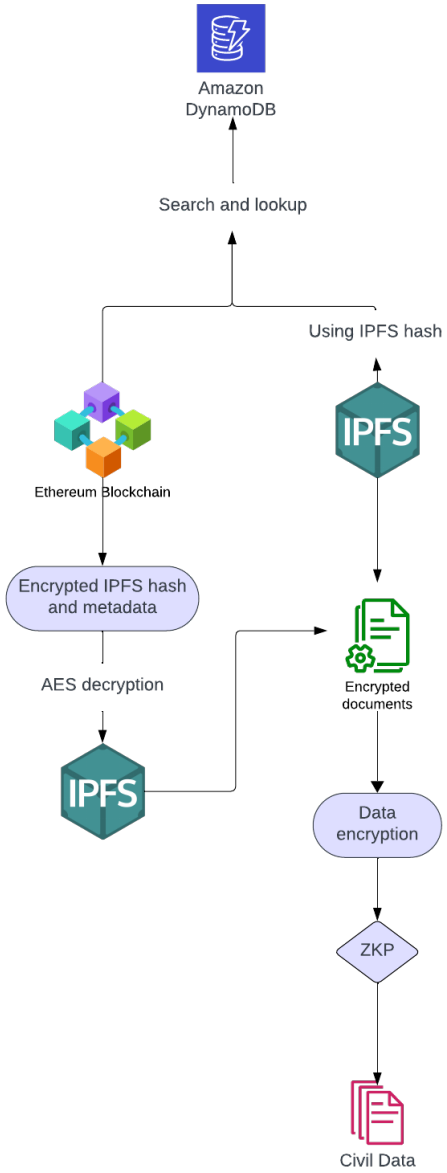

Figure 4: Verifying documents

4.1.2 Document verification

Figure 4 illustrates the workflow for verifying a civil records. First, specific civil data or a person's data can be retrieved by performing a lookup in the off-chain database, DynamoDB, to obtain the metadata, IPFS hash and/or transaction hash. A registrar can use the IPFS hash to directly retrieve the encrypted civil data from IPFS.

To verify the integrity of the data, the system uses the transaction hash to locate the blockchain block where the data was originally written. The system then decrypts the blockchain data to retrieve the stored IPFS hash. The first step in the verification process is to compare the IPFS hash that was stored off-chain with that retrieved from the transaction block, this simple step can highlight tampered data without having to perform any more computations.

Next, the encrypted civil data is retrieved from IPFS and decrypted using the AES encryption key securely obtained from AWS KMS. Finally, zero-knowledge proofs (ZKPs) are performed to verify the authenticity and integrity of the data without revealing the entire dataset.

4.2 Smart Contracts Design

To minimise blockchain interaction costs, I limited the use of smart contracts to a single CivilMetadata smart contract, designed and developed in Solidity using the Sepolia testnet for Ethereum.

The smart contract is designed to securely store and retrieve civil document metadata. It uses a structure to encapsulate two fields: ipfsHash, which stores the IPFS Content Identifier (CID) pointing to the off-chain document content, and encryptedMetadata, which contains the civil metadata that has been encrypted for privacy purposes.A mapping is then used to associate a unique document ID with the corresponding civil metadata. Several functions are implemented to store and retrieve the data based on the document ID.

- **storeDocument**: This function stores civil metadata and updates the corresponding entry in the ‘documents‘ mapping with the provided details.
- **getDocument**: Retrieves stored metadata by looking up the ‘documents‘ mapping for the given document ID specified in its parameters and returns the IPFS hash and the encrypted metadata associated with that specified document ID.

5 Implementation

As previously discussed, the proposed system follows a two-tier design structure: the backend tier and the blockchain tier. In this section, I will explain the implementation of each tier extending to include the tools and languages used to produce the outputs

5.1 Cloud Tier

5.1.1 Infrastructure as Service

I took an Infrastructure as a Service (IaaS) approach to provisioning and deploying the backend infrastructure programmatically using the AWS Cloud Development Kit (CDK)

and Typescript. Here's how it works: when the AWS CDK code is run, it generates an AWS CloudFormation template based on the resources defined. CloudFormation then takes this template and deploys the specified resources to the AWS environment.

5.1.2 Microservices

The backend of the system is powered by serverless microservices using AWS Lambda, enabling the execution of discrete, event-driven functions that are securely triggered using secure Amazon API Gateway endpoints.The API Gateway endpoints are secure and require JSON Web Tokens (JWT) obtained from Amazon Cognito upon successful user authentication. Each AWS Lambda microservice is designed to perform a specific task, such as storing civil status records in AWS DynamoDB and querying a stored record based on its record ID.

5.1.3 Database

The DynamoDB schema is designed to ensure fast and efficient searches: each record is uniquely identified by an ID that serves as the primary key, with a Metadata field that stores detailed information about the individual, including their name, age, and gender, within a nested map structure, and an IPFS field that stores the link to the record's corresponding document stored in the IPFS network. And for decentralized storage with IPFS, I used Pinata to manage and pin files to the IPFS network.

5.1.4 Encryption

As mentioned above, the proposed system follows a hybrid encryption architecture. One part of the system was implemented using AWS KMS to encrypt the symmetric encryption keys, while the other part used AES encryption to encrypt the civil record, implemented in code using the CryptoJS library.

5.2 Blockchain Tier

The smart contracts were written in Solidity, a programming language used to develop smart contracts on EVM-based blockchains. To minimise local configuration time and complexity, I used a cloud-based development environment, Remix IDE for both writing and deploying the contracts and Metamask to sign the transactions.

The smart contract, is designed to store and retrieve the encrypted civil metadata and corresponding IPFS CID on the blockchain. The contract uses a mapping to associate each document ID with its corresponding metadata.

The contract provides two main functions: storeDocument, which allows the storage of a document's IPFS hash and encrypted metadata on the blockchain, using a unique document ID. And getDocument, which retrieves the stored metadata based on the provided document ID. An event, DocumentStored, is emitted whenever a document is stored, allowing external applications to track when new data is added to the blockchain. This design ensures that civil document metadata is securely stored and easily retrievable on the blockchain, providing transparency and immutability.

Metric	# Sepolia	# BNB	# Polygon zkEVM
Transaction Fee (Euros)	4.42	1.12	0.035
Average Gas Price (Gwei)	2.178	3	0.0173
Average Burnt Fee	1.38	0.11	0
Transaction Cost (Gas)	537,778	539,590	537,634

Figure 5: Civil Smart Contract Deployment to Various Testnets

6 Evaluation

In this section I evaluate the effectiveness of the proposed system. A series of experiments were conducted to analyse the performance and efficiency of the system. These experiments included a cost analysis, which compared the cost of deploying the proposed smart contract on the system's blockchain environment, Ethereum using Sepolia ETH with alternative platforms ; Polygon zkEVM and BNB Smart Chain, as well as the cost of interacting with the smart contract. And in a second experiment, the encryption architecture used in the system, which uses a hybrid approach of AES-RSA where AES is used to encrypt the civil data. Experiments were conducted to evaluate and compare AES, a symmetric encryption algorithm, with other symmetric encryption algorithms to assess their relative security and performance in encrypting the same dateset of civil records.

6.1 Experiment / Cost Analysis of Deploying and Interacting with the Smart Contract Across Different Blockchains: Sepolia, Polygon zkEVM and Binance Smart Chain

For the experiment, I chose blockchains based on the Ethereum Virtual Machine (EVM). The main reason for this choice was the ease of deploying Solidity-based smart contracts on EVM-compatible chains, so the smart contract was developed once and could be deployed on each chain without any additional configuration. In addition, I used testnets rather than mainnets for testing and development, as testnets provide free and easily accessible tokens. The specific chains chosen for the experiments were Sepolia (Ethereum testnet), Polygon zkEVM and Binance Smart Chain (BNB). The smart contract was deployed on each of these chains and data from each deployment was recorded (see figure 5). To make the results more relatable to real-world terms, the transaction costs were converted from their respective cryptocurrencies into euros. As seen in 5 Sepolia ETH had the highest transaction fee of €4.42 for deploying the smart contract, while BNB had a much lower fee of €1.12. However, BNB's fee was still higher than that of Polygon zkEVM, which had the lowest transaction fee of €0.035.

In the second experiment, I built a simulation using Solidity to simulate the storage of multiple civil status data. The simulation was run to store 100 documents and then scaled to multiples of that to observe the increase in transaction costs per number of interactions. This experiment was particularly important because if a government were to implement such a system, it would need to store data for every citizen in the country, which means a high volume of interactions. The smart contracts needed to be stress-tested to see the costs associated with a high volume of interactions. As shown in Figure 6, the results were consistent with the previous comparison of smart contract deployments: Sepolia ETH had the highest transaction costs per interaction, while Polygon zkEVM had

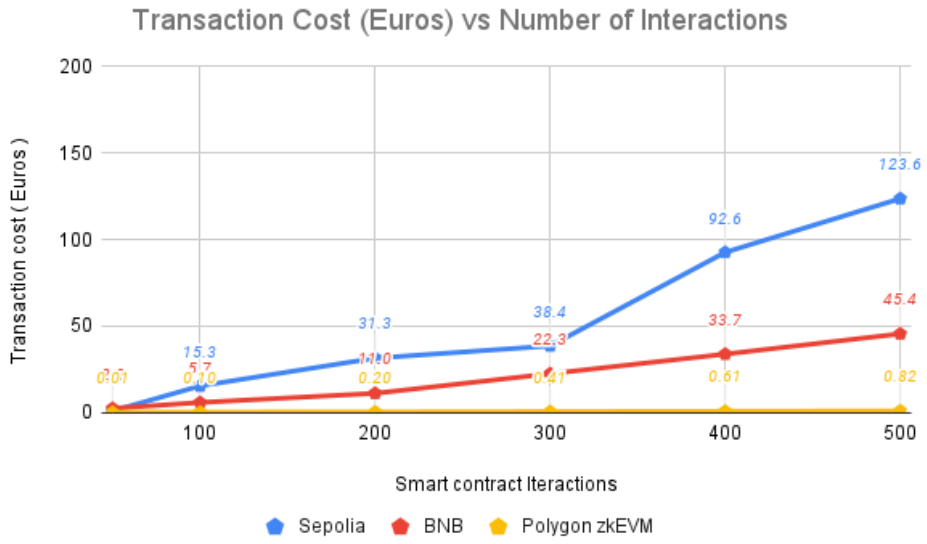

Figure 6: Transaction Cost (Euros) vs Number of Interactions

significantly lower costs, making it more efficient for handling large data stores.

These differences in transaction fees are due to the design and scalability handling of each blockchain. Although BNB and Sepolia ETH are Layer 1 blockchains, BNB incorporates performance optimisations that contribute to its lower transaction fees compared to Sepolia ETH, such as its Proof-of-Staked-Authority (PoSA) consensus mechanism, reduced block time and support for larger block sizes. [7], On the other hand, Polygon zkEVM, a Layer 2 scaling solution for Ethereum, enhances throughput and reduces transaction costs by using zk-SNARKs (Zero-Knowledge Succinct Non-Interactive Arguments of Knowledge). This technology allows Polygon zkEVM to bundle hundreds of transactions into a single proof, which is then verified on the Ethereum mainnet [8], Another key standout feature, as seen in Figure 5, is that Polygon zkEVM had no burnt fee, unlike Sepolia and BNB, which had burnt fees of €1.38 and €0.11, respectively. These features of Polygon zkEVM result in significantly lower transaction fees compared to both BNB and Sepolia ETH. It is also important to note that this experiment only focused on write operations, such as recording civil records on the blockchain, as there were no fees or transaction costs associated with reading from the chain.

6.2 Experiment / Comparing Civil Record Encryption Times for Various Symmetric Encryption Algorithms

The purpose of this experiment was to compare the encryption times for encrypting civil records using different symmetric encryption algorithms. The aim was to be able to justify the choice of encryption algorithm initially selected for the proposed system. As mentioned above, I proposed a hybrid encryption method based on AES-RSA encryption technologies for encrypting civil records and managing the encryption key with AWS KMS, where the symmetric part of the encryption is handled by the AES encryption algorithm, which is used to encrypt each civil status record before uploading to IPFS, as

[7]Medium: https://getblock.medium.com/bnb-smart-chain-bsc-vs-ethereum-whats-the-difference-a9f44b5
[8]Dune: https://docs.dune.com/data-catalog/evm/polygon-zkEVM/overview

Records	AES-128	AES-256	DES-56	3DES-168	RC4-128	RC4-256
3000	0.31	0.3	1.09	2.67	0.24	0.24
6000	0.55	0.56	2.03	5.66	0.49	0.48
9000	0.87	0.92	3.14	8.22	0.77	0.76
12000	1.18	1.22	4.14	11.84	0.99	0.98
15000	1.69	1.45	5.14	13.78	1.29	1.27
18000	1.71	1.77	6.68	16.72	2.26	2.03
21000	2.16	2.34	7.6	20.03	1.9	1.95
25000	2.73	2.61	9.65	24.75	2.41	2.26
30000	3	3.06	11.33	28.53	2.5	3.09

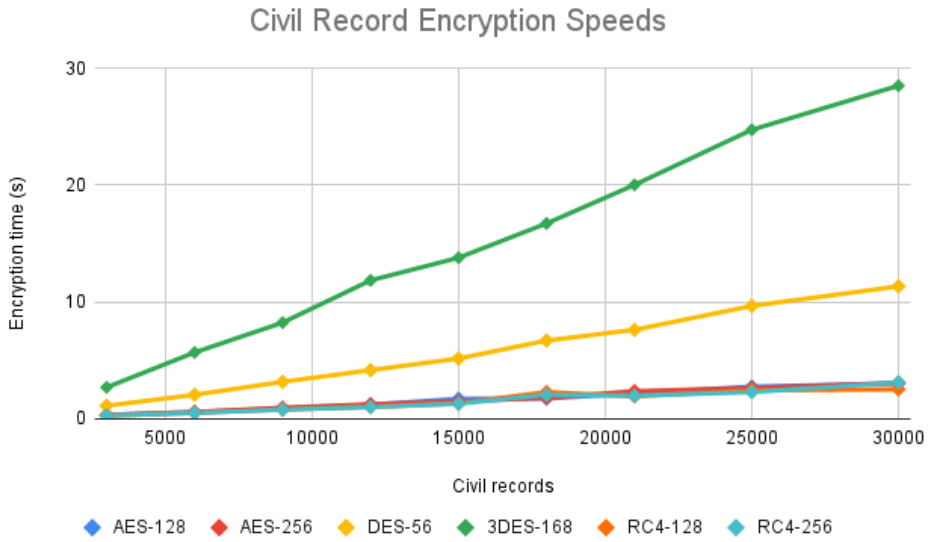

Figure 7: Comparisons of Civil Record Encryption Times

well as encrypt the metadata to be stored on-chain, while the asymmetric part, RSA, is only used in AWS KMS to encrypt the symmetric encryption key.

As the RSA operations would not occur frequently and don't significantly affect the overall performance of the system, running simulations to benchmark these operations is not a major concern. In contrast, AES encryption will be applied to potentially millions of civil records, so it is important to benchmark and compare the encryption times for encrypting the civil records using different symmetric encryption algorithms. For the experiment, I focused on the following encryption algorithms; AES, Data Encryption Standard(DES), Tripple DES and Rivest Cipher(RCA), and also compared them using different encryption key lengths, for those that had support for different key lengths, DES notably only supported 56 bit keys, so its key length wasn't varied in the experiment. Figure 7 shows the results of the experiment

As can be seen in Figure 7, the standout in terms of encryption times for all record sizes was RC4-128, with AES-128 in second place, showing very comparable encryption times. Although RC4 was the fastest, it is an outdated algorithm with well-documented limitations, such as susceptibility to biases and statistical attacks [9]. On the other hand, although AES was slightly slower, it is highly secure.

[9]Tutorials point: https://www.tutorialspoint.com/difference-between-aes-and-rc4

6.3 Discussion

The results of the experiments were quite insightful, and laid the groundwork for future improvements to the proposed system design. The most significant finding, in terms of scalability, came from the experiment of deploying the smart contracts on multiple blockchain networks and analysing the interactions. This experiment showed that the Polygon zkEVM was the most cost-effective per interaction, in contrast to the blockchain network I had originally proposed for the system, Ethereum. The difference was so significant that it clearly demonstrated that in order to achieve better blockchain economics, Polygon zkEVM is the network to use. Given that the system has the potential to store millions of civil record metadata on the blockchain, the cost of each of these interactions is critical.

Secondly, the experiment comparing civil record encryption times justified the use of AES encryption in the proposed system design to encrypt the civil records. For this experiment, I evaluated the performance of the encryption algorithms available in the crypto.js library. For a clearer and unbiased comparison, I need to compare with other algorithms not available in this library.

7 Conclusion and Future Work

This research aimed to design a hybrid blockchain-based system that incorporates AES and RSA encryption algorithms, ZKPs, IPFS, and cloud services in a way that provides a secure, scalable, and privacy-preserving solution for securing and validating the data contained within civil status documents. The proposed system addressed several key objectives. First, the hybrid encryption approach, where the AES encryption algorithm was used to encrypt the civil records prior to storage, while the symmetric encryption keys were encrypted with the RSA encryption algorithm and managed in AWS KMS, proved to be highly efficient, as demonstrated by the experiments conducted. Furthermore, using AWS KMS to manage the encryption keys was a more effective approach compared to traditional methods. Second, using IPFS to store the civil records in an immutable manner and storing the metadata and IPFS hash on the blockchain proved to be very effective in ensuring the immutability and validity of the records, as the data cannot be altered. Third, the decision to use cloud technologies such as DynamoDB to store and index civil records greatly improved the speed and efficiency of searches. This approach is more effective than indexing the blockchain itself, which would incur additional financial costs.

The research also revealed several flaws in the design of the system. Firstly, Polygon zkEVM proved to be far superior to the proposed blockchain technology, Ethereum, in terms of cost per interaction with blockchain smart contracts. Although there were no transaction fees for reading smart contracts, there were significant fees for writing civil records. However, overall, the experiments showed that Polygon zkEVM had significantly lower write interaction fees due to its layer 2 design and improved scalability. Future research should consider this technology, Polygon zkEVM, as well as experimenting with non-EVM chains such as Solana.

Secondly, the design did not include a means of restricting access to the smart contracts once they were deployed. Even though the data in the contracts was encrypted, it would be beneficial to explore an approach that restricts access to the contracts, even after deployment, through authentication mechanisms. And thirdly, this study did not

fully explore or experiment with Zero-Knowledge Proofs (ZKPs) as a tool for consular cooperation in the verification of civil status data. Future research should explore this further and consider whether the additional computational overhead is justified.

In conclusion, this research successfully establishes the foundations on which future work can be built. The plan for future work is to refine the proposed design based on the lessons learned from the experiments, with the aim of developing a fully operational end-to-end civil status verification and authentication platform that provides real-world functionality.

References

Alam, S., Shuaib, M., Khan, W. Z., Garg, S., Kaddoum, G., Hossain, M. S. and Zikria, Y. B. (2021). Blockchain-based initiatives: Current state and challenges, *Computer Networks* **198**: 108395.
URL: *https://www.sciencedirect.com/science/article/pii/S138912862100373X*

Aliti, A., Leka, E., Luma, A. and Trpkovska, M. A. (2022). A systematic literature review on using blockchain technology in public administration, *2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO)*, pp. 1031–1036.

Goldwasser, S., Micali, S. and Rackoff, C. (1985). The knowledge complexity of interactive proof-systems, *Providing sound foundations for cryptography: On the work of shafi goldwasser and silvio micali*, pp. 203–225.

Kim, S.-K. (2022). Blockchain smart contract to prevent forgery of degree certificates: Artificial intelligence consensus algorithm, *Electronics* **11**: 2112.

Leka, E. and Selimi, B. (2021). Development and evaluation of blockchain based secure application for verification and validation of academic certificates, *Annals of Emerging Technologies in Computing (AETiC)* **5**(2): 22–36.

Loon, H. V. (2024). Requiem or transformation? perspectives for the ciec/iccs and its work, *Journal or Publication Name* .
URL: *https://www.hansvanloon-pil.nl/assets/ciec-iccs-requiem-or-transformation.pdf*

Macniven, D. (2012). Fraud in respect of civil status.
URL: *https://www.europarl.europa.eu/RegData/etudes/note/join/2012/462499/IPOL-JURI$_N$T(2012)462499$_E$N.pdf*

Mthethwa, S., Dlamini, N. and Barbour, G. (2018). Proposing a blockchain-based solution to verify the integrity of hardcopy documents, *2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC)*, pp. 1–5.

Nyaletey, E., Parizi, R. M., Zhang, Q. and Choo, K.-K. R. (2019). Blockipfs - blockchain-enabled interplanetary file system for forensic and trusted data traceability, *2019 IEEE International Conference on Blockchain (Blockchain)*, pp. 18–25.

Pierro, G. A. and Tonelli, R. (2022). Can solana be the solution to the blockchain scalability problem?, *2022 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pp. 1219–1226.

Zainuddin, M. D. R. and Choo, K. Y. (2022). Design a document verification system based on blockchain technology, *Multimedia University Engineering Conference (MECON 2022)*, Atlantis Press, pp. 229–244.