




A systematic literature review of blockchain-based e-KYC systems

Md. Abdul Hannan¹ · Md. Atik Shahriar² · Md Sadek Ferdous³  ·
Mohammad Javed Morshed Chowdhury⁴ · Mohammad Shahriar Rahman⁵

Received: 2 July 2022 / Accepted: 1 April 2023 / Published online: 13 April 2023

© The Author(s), under exclusive licence to Springer-Verlag GmbH Austria, part of Springer Nature 2023

Abstract

The know your customer (KYC) guidelines in financial services require that institutions make an effort to verify the identity, suitability, and assess risks involved while maintaining a business relationship. The procedures fit within the broader scope of any financial institution's Anti-Money Laundering (AML) policy. Governments around the world have digitalised this process to make it more convenient and transparent by introducing the notion of e-KYC (Electronic KYC). e-KYC provides a flexibility for the users as they might even quickly complete the on-boarding process from the comfort of their homes. However, there are a few outstanding issues, such as the lack of any global standardisation, possible fraudulent activities during the e-KYC process and other privacy concerns, that must be addressed before its full benefits can be achieved. Recently, blockchain technology (or blockchain in short) has emerged as a foundational technology with possibilities to disrupt a wide range of application domains. Understandably, it is increasingly being investigated how blockchain can be a useful tool to address these outstanding issues. Towards this aim, a number of research works have emerged in the recent years. In this article, we present a systematic literature review (SLR) of these works using the PRISMA model in order to identify and critically analyse the existing researches at the intersection of e-KYC and blockchain. Based on our study we have also identified the limitations of the existing solutions and provided future directions.

Keywords Know your customer · KYC · e-KYC · Blockchain · PRISMA

Mathematics Subject Classification 68U99

Md. Abdul Hannan and Md. Atik Shahriar have contributed equally to this work.

✉ Md Sadek Ferdous
sadek.ferdous@bracu.ac.bd

Extended author information available on the last page of the article

1 Introduction

KYC or know your customer is the process of verifying the identity, eligibility and background verification of a customer for establishing a business relationship [1]. It is a process that all business institutions operating within the scope of existing KYC regulations must follow to analyse the suitability, and risks involved to maintain the business relationship in a continuous fashion. The KYC process, if done adequately, helps to reduce financial fraudulent activities by restricting bad actors from entering the banking system. The main objective of KYC is to prevent banks from being used for money laundering and other criminal activities and hence, the process also fits within the broader scope of any financial institution's Anti-Money Laundering (AML) policy. KYC is done by demanding valid identification documents like country-level ID cards, residential proof, income proof and so on [2].

Banks around the world have incorporated electronic KYC (e-KYC) to improve the security and customer satisfaction of the KYC process. e-KYC is the electronic version of KYC where the customer's identity and address are verified electronically through biometric or national ID authentication. The main motivation of e-KYC is to quicken up the KYC procedure for the customers as they do not need to fill up a large number of documents and FIs (Financial Institutions) do not need to store those documents for compliance. In addition, it provides additional facilities for the customers as they can open an account from the comfort of their homes which in turn might increase the number of customers for those FIs who have adopted e-KYC mechanisms. Towards this aim, Governments around the world have introduced policies to incorporate the e-KYC procedure during the customer onboarding process for FIs [3–6].

However, the traditional e-KYC approach is centralised and repetitive. There is no unification of the required documents for KYC for various banks. Users have to perform the same e-KYC procedure while creating accounts for various banks. Also, users have no control over which information they will share. As KYC data is very sensitive, it is subject to many threats from criminals. Thus, KYC systems should be properly secured from any kind of unauthorised access and denial of service attacks. In order to tackle these issues, many researchers have explored the idea of integrating blockchain/distributed ledger technology (DLT) within the e-KYC systems [7–11]. Blockchain is a distributed database formed by an immutable, cryptographically linked, and growing list of records and maintained by establishing consensus among trustless parties without the interaction of any intermediary. Because of its immutability and transparency nature, it can help to share data among various FIs in the KYC verification process in a secure and auditable way [12]. With the ever-increasing popularity of blockchain and e-KYC, there is a strong possibility that researches at the intersection of these two domains will continue to rise. However, it would be important for the researchers to have a basic understanding of the existing relevant research works along with a clear picture of their advantages and limitations as well as research gaps and potential future research directions. To lend a helping hand, we present a Systematic Literature Review (SLR) of the existing works at the intersection of Blockchain and e-KYC using the PRISMA model [13]. The main motivation of this SLR is to provide a snapshot of the existing research within this domain, identify their strengths and weaknesses and highlight possible future directions.

Towards these aims, the article is structured as follows. In Sect. 2, we present a brief summary of different concepts such as e-KYC and blockchain. Section 3 discusses the core methodologies used in the SLR while Sect. 4 elaborates on our findings. We discuss different aspects of our findings in Sect. 5 including the limitations of the SLR and future directions. Finally, we conclude in Sect. 6.

2 Background

In this section, we present a brief background on the worldwide adoption of e-KYC (Sect. 2.1), the traditional e-KYC procedure (Sect. 2.2), blockchain (Sect. 2.3), and Self-sovereign Identity (Sect. 2.4).

2.1 Worldwide e-KYC

The know your customer (KYC) procedure means making an effort to verify the identity, suitability, and risks involved with maintaining a business relationship with the customer [14]. It is done to make sure that the customers are genuinely who they claim to be.

As various financial crimes, terrorist funding, and money laundering have increased and will continue to increase worldwide, KYC has become one of the primary weapons in the fight against these types of crimes. The global anti-money laundering (AML) and countering the financing of terrorism (CFT) have increased the responsibilities of FIs all over the world [3]. International regulations guided by models like The Financial Action Task Force (FATF) are now being implemented in most countries' national laws worldwide [3]. As a result, FIs worldwide have implemented KYC procedure in their customer onboarding process as a fundamental procedure of their system. In the case of banks, it is compulsory to comply with the KYC regulations of the respective country. Failing to do so is penalised by national and international laws, which has been seen in the US and many countries in Europe, the Middle East, and the Asia Pacific in recent years [15].

At the same time, many countries are slowly shifting towards the digitalised version of KYC, named *e-KYC*, as customers' expectations of fully digital experiences have extended to every corner of the financial service domains. As a result, regulators have been slowly introducing new e-KYC guidelines to allow FIs to perform KYC checks and approve customer applications digitally over the past few years. The Aadhaar e-KYC of India is one of the pioneers of e-KYC that was launched in 2009 [16]. The Singaporean government introduced a digital personal data platform known as MyInfo in May 2016 to streamline the identity verification process during online transactions [17]. In November 2018, US agencies announced a joint declaration that promotes some banks to become more sophisticated in exploring artificial intelligence and other digital identity technologies [18]. In the same year, European Supervisory Authorities recommended maintaining a common approach for a uniform application of standards across the EU which can be possible by introducing e-KYC [3]. In February 2019, the Hong Kong Monetary Authority released a circular on "remote onboarding of

individual customers” that states that technology should be used for remote onboarding purposes that cover both identity verification and identity matching [4]. In December 2019, Bank Negara Malaysia issued draft requirements for FIs looking to implement e-KYC [5]. Similarly, the Bangladesh Financial Intelligence Unit (BFIU) issued e-KYC guidelines [6] to open accounts in the financial sector without filling up any paper-based documents. The new guidelines will be applicable for opening bank accounts, Beneficiary Owners (BO) accounts and insurance policy accounts.

In pandemic and post-pandemic times, the necessity of e-KYC has increased and will continue to increase in the near future. Thus, many regulators have already issued revised guidances on remote customer verification to help FIs to ensure business continuity and compliant client onboarding during lockdowns. For example, New Zealand reporting entities have started to accept scanned copies of documents instead of originals and to perform electronic verification to avoid physical contacts with customers [4]. The Securities and Exchange Board of India is now letting the foreign portfolio investors provide the required documents scanned [4]. The Philippine central bank has lifted the requirement of a valid ID card during client onboarding [4]. In January 2020, the Reserve Bank of India announced that it would allow video-based KYC as an option to confirm a customer’s identity [19]. The German regulator, BaFin, issued a directive that enables customer identification and verification via a live two-way video connection with a compliance professional [20]. Thus, e-KYC is being implemented worldwide using varying technologies.

2.2 Current e-KYC practice in the banking sector

Banks all around the world have started to implement e-KYC. It has made the process of opening a bank account much easier for many new bank account holders. The architecture for the current practice of e-KYC is presented in Fig. 1.

As per Fig. 1, all financial institutions rely on a central identity service provider, usually facilitated by the government, for initiating and validating the e-KYC process. There are mainly two ways: i) a new customer can complete their e-KYC process via a mobile App (the mobile app-based KYC) and in-person KYC at a bank branch [6].

- *App-based KYC* In this approach, an individual needs to put their personal information and an identification number. Then, the app will prompt for a biometric photo (selfie) which will then be uploaded to the server [6]. The identity of the individual is verified against the information stored in the central identity service provider database using the identification number.
- *In-person KYC* In this approach, individuals need to visit the bank branch. They will need to provide their personal information and identification numbers [7]. Then, they need to provide fingerprint-based biometric information at the bank. This biometric information is verified against the central identity service provider database [6]. In both cases, if the verification is successful, then bank employees will complete the account opening process for the new customer.

However, the adopted e-KYC has the following limitations:

- The current approach does not allow other organisations, particularly financial institutions such as banks, to act as KYC providers. Currently, the user registers

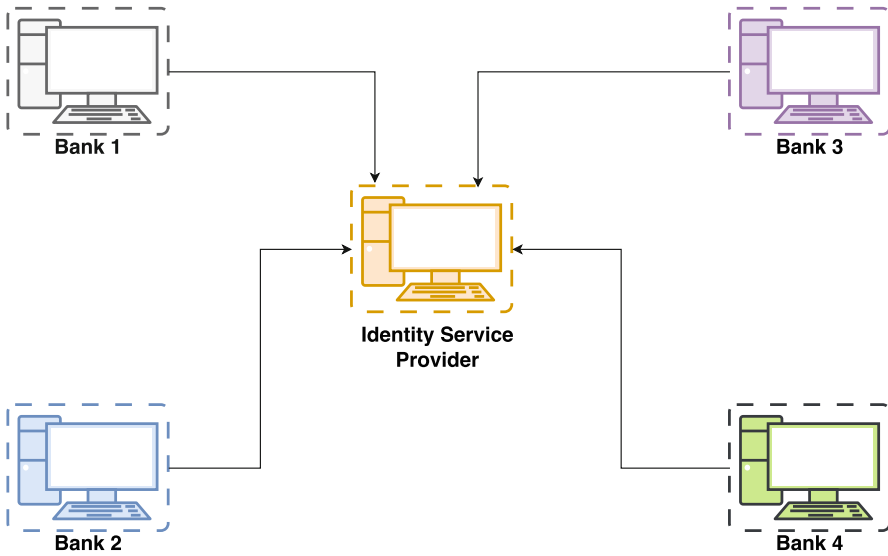


Fig. 1 Architecture of current practice of e-KYC

with the government-owned or controlled identity service provider just once and the data might not be updated for a long period, making such data not very useful [7]. Interestingly, users interact with different FIs more frequently than the identity service providers and hence, such organisations often maintain more accurate and dynamic records regarding their customers. Therefore, they could provide more updated KYC data regarding a user.

- In the current approach, there is no provision of aggregating reputation data with the KYC data regarding a user [21]. Allowing a financial institution to act as a KYC provider and to aggregate such reputation data can create a new service delivery model which can be monetised by these organisations.
- An audit trail is an indispensable component of any system which deals with critical data such as KYC. An audit trail ensures that any allowed entity involved in the system has a complete view of interactions within the system. This becomes important if there is any breach in the system by an attacker or there is any dispute [22–25]. The auditing mechanism within a centralised system might not provide a high level of security as the attacker can manipulate such audit trails [22].
- The current approach is not very privacy-friendly as the user has no control over the data that is being shared between the identity service providers and other organisations [26].
- In addition, there are no globally set standards for representing e-KYC data. The implication of this is that different countries might be following different formats to represent e-KYC data and the security assurance of exchanging such data will entirely depend on the respective implementation.

2.3 Blockchain

Blockchain has become very popular since the technology was invented and released to the world. It was first introduced by Satoshi Nakamoto as a technical paper in 2008 as the underlying technology of Bitcoin [27]. This started a revolution in the financial sector in the form of cryptocurrencies. Blockchain is a distributed, immutable, cryptographically linked, and growing list of records where consensus can be established among trustless parties without the interaction of any intermediary. The growing list of records, also known as a ledger, is distributed and stored by the nodes of a P2P network. The ledger consists of consecutive blocks chained together following a strict set of rules, and hence, the name blockchain. The blocks are created at a predefined interval in a decentralised manner by a set of rules called consensus algorithm that guarantees the immutability of data. The terms blockchain and Distributed Ledger Technology (DLT) are used interchangeably. However, there is a subtle difference. A blockchain is a particular type of ledger where data are stored in a specific format. Other types of ledgers use different data formats. A ledger can be regarded as a Distributed Ledger when it is distributed across a network.

The concept of smart contracts in blockchain was introduced in Ethereum [28]. Smart contracts, equipped with a blockchain system, enable immutable, trustless, and transparent distributed computing and autonomous code execution, which has a wide range of applications in different domains including e-KYC [29]. Smart contracts ensure that the application logic is also guarded by the tamper-proof nature of blockchain.

Even with these advantages, a major concern for blockchain is its transparency nature which allows everyone in the network to access data stored in the blockchain. Since e-KYC data are very sensitive in nature, this transparency nature of blockchain can be problematic if proper care is not given [26].

2.3.1 Blockchain and e-KYC

In order to tackle the current e-KYC issues discussed above, many researchers have explored the idea of integrating blockchain/DLT within e-KYC systems [7–11]. Blockchain can help to share data among various FIs in an immutable and distributed way which can reduce the data collection time and the overall cost involved in the process. Blockchain can also help managing an immutable audit trail of e-KYC data.

In order to ensure the confidentiality of e-KYC data, it might not be wise to store them in any public blockchain system. This is because such blockchains are fully transparent in nature and data stored in such systems are visible to everyone, thereby violating the privacy of such sensitive data. Rather, they should be stored in an off-chain database with the corresponding hashes stored on the blockchain to ensure the integrity of the e-KYC data. It might be tempting to think that utilising private blockchain systems may help tackle this problem as private blockchain systems can restrict access to such crucial data to a handful of authorised entities, however, it is to be noted that the authorised entities can access such data [30]. Therefore, the confidentiality of data must be considered with strong encryption mechanisms.

In addition to these, the smart contract can enforce distributed and autonomous code execution facility which could be useful for e-KYC applications. An agreement could be codified via a smart contract to facilitate autonomous execution of certain transactions when pre-determined conditions are satisfied. When transactions and contracts are recorded on a shared ledger, the effort of various stakeholders can be reduced [31].

2.4 Self-sovereign identity

Self-sovereign identity (SSI) has recently emerged as a new paradigm for managing digital identities [32]. It is created and controlled by the user throughout its life-cycle. SSI aims to disrupt the traditional notion of identity management which is mostly controlled by different service providers, by giving more controls to the users so that they can handle their identity life-cycle. In SSI, the user will have more control over their data for sharing with other parties using standardised security mechanisms. The Verifiable Credential (VC) is a W3C standard [33] for representing different types of credentials which have been heavily adopted within the SSI mechanism. A verifiable credential is a tamper-proof credential with cryptographically verifiable authorship. The addition of digital signatures makes verifiable credentials more tamper-proof and more trustworthy than their physical counterparts. Since there are no global standards for representing e-KYC data, a novel approach could be to represent e-KYC data with VCs and exchange them according to the SSI protocols.

3 Research methodology

In this section, we present the research methodology for the presented SLR. Particularly, we present the research questions (Sect. 3.1), discuss the search strategy (Sect. 3.2), outline the study selection (Sect. 3.3) and present the quality assessment questions (Sect. 3.4).

We have investigated the existing researches on e-KYC both based on blockchain, through this SLR. Our focus is to research existing findings and analyse them.

3.1 Research questions

PRISMA is often used in collaboration with research questions by the researchers [34, 35] to structure their surveys. Based on the existing practice, we have also opted for research questions for analysing and presenting the findings from the existing literature in this paper. We present our research questions in Table 1.

The first research question (RQ1) identifies the existing researches within the scope of this SLR. This research question will help the readers to have a brief overview of the existing works with a focus on their application domains. Research Question 2 (RQ2) analyses the architectural differences of the existing works with respect to a number of factors. RQ3 on the other hand explores different aspects related to their implementations. This will help the readers to understand how the corresponding

Table 1 Research questions

ID	Research questions
RQ1	What are the existing research works and their application domains for blockchain-based e-KYC?
RQ2	How are the works different in architectural design?
RQ3	What are the implementation considerations?
RQ4	What security and privacy aspects are considered in the research existing works?
RQ5	Are disruptive concepts such as reputation and verifiable credentials considered in existing e-KYC systems?

research works have been implemented. The motivation for RQ4 is to shed lights on different security and privacy issues in those research works. RQ5, however, requires further explanations.

As mentioned earlier, the traditional KYC process has significant issues with respect to the privacy-friendly mechanisms for sharing KYC data and the lack of any security standard. An SSI or VC-based approach would be an important mechanism to address some of the identified issues in the traditional KYC approaches. Also, generating reputation data and aggregating with e-KYC approach would facilitate novel service models, as discussed earlier. Both of them are disruptive ideas which can introduce novel service delivery models and that is why we have decided to investigate if VC and reputation have been considered as part of RQ5. They could be investigated under separate research questions, however, the amount of research works utilising any of these approaches is very small and hence, we have investigated them within one research question.

3.2 Search strategy

In order to identify relevant publications, we have searched through the electronic databases with different keywords in the primary stage. After that, we have selected papers related to our research from the first stage to the second stage. We have utilised the PRISMA framework [13] as the core methodology for record keeping and applying inclusion and exclusion rules in order to find the closely matched research publications. PRISMA stands for *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*. It is an evidence-based minimum set of items that helps authors in reporting a wide array of systematic reviews and meta-analysis. These items work as a checklist used to improve the transparency in systematic reviews.

3.2.1 Literature sources and search keywords

We have conducted searches on different digital libraries and search engines such as, Google Scholar, IEEE, Springer, ACM DL and Science Direct, using a list of search keywords. The search keywords are listed in Table 2.

Table 2 Search keywords / terms

Number	Keywords
1	Blockchain
2	KYC
3	e-KYC
4	Reputation system
5	Verifiable credentials

Table 3 Number of papers retrieved from each digital library with relation to the searched keyword combinations

Keywords	Google Scholar	Springer	ACM DL	Science-Direct
Blockchain and (KYC or e-KYC)	6510	533	1969	150
Blockchain and reputation system and (KYC or e-KYC)	1820	26	1924	3
Blockchain and verifiable credentials and (KYC or e-KYC)	914	11	1364	3
Blockchain and reputation system and verifiable credentials and (KYC or e-KYC)	899	4	1346	0

During the paper collection process, we have considered the names of publishing journal, paper titles, and publishing years. The different results have been stored in Google Docs for further use. The result of the search in different electronic databases are summarised in Table 3.

3.2.2 Search process

We have used the following approaches to identify the existing research literature.

1. *Initial Searching Phase* We have searched through every database with the keywords in different combinations. Then, the returned publications from all the sources were collected in Google Docs.
2. *Reference Searching Phase* In the next step, we have gone through the references of the publications from the previous step to identify additional relevant publications and added them to our collection of publications if anything new is found.

Figure 2 shows the search and selection process.

3.3 Study selection

A total of 17,476 candidate papers have been collected as the output from the previous steps. However, not all papers were relevant to our study. Therefore, we filtered the candidate papers in two more phases:

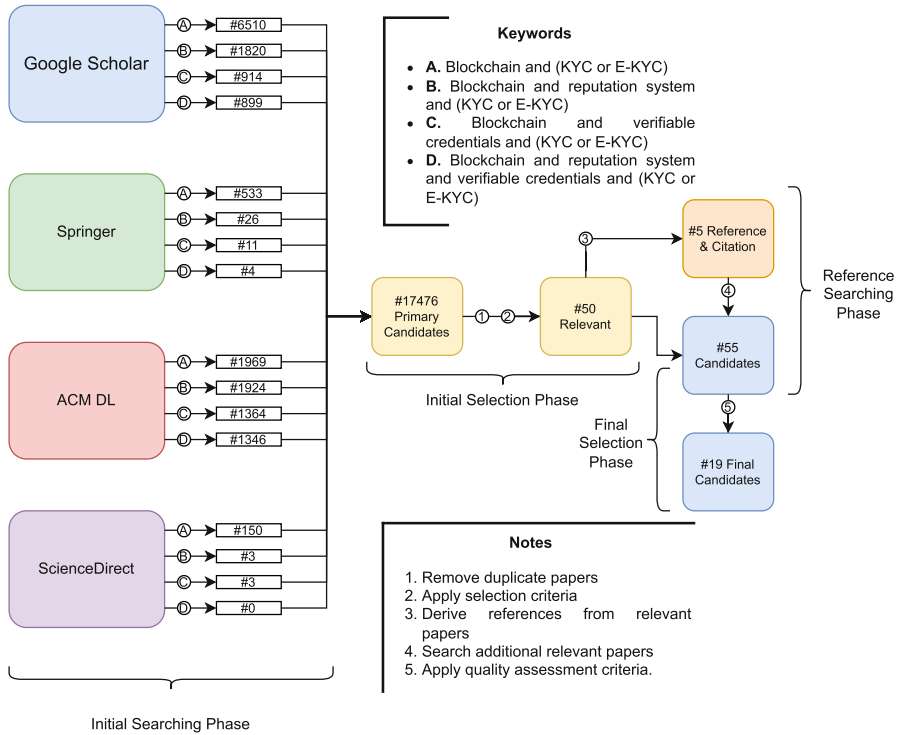


Fig. 2 Search and selection process

1. *Initial Selection Phase* We have applied some inclusion and exclusion criteria to select publications that match closely to the scope of this study. These criteria have been chosen in such a way that they will help to answer our research questions more precisely. Next, we present the utilised inclusion and exclusion criteria.

Inclusion Criteria At first, we present the inclusion criteria:

- Publications that have been written in English.
- Publications have covered blockchain and e-KYC.
- For a publication with both conference and journal versions, only the journal version is considered.

Exclusion Criteria Next, the exclusion criteria are presented:

- Publications that are partially available.
- The duplicate publications.
- Patents.
- Publications written in any other languages except English.

2. *Final Selection Phase* In this phase, we have applied our quality assessment questions (discussed next) on the publications upon which the exclusion criteria have

Table 4 Quality assessment questions

ID	Quality assessment questions
QAQ1	Is the paper peer-reviewed?
QAQ2	Has the paper described the process of collecting KYC data and converting it into e-KYC properly?
QAQ3	Are security and/or privacy issues (related to e-KYC) discussed or analyzed?
QAQ4	Does the paper have implementation?
QAQ5	Are the limitations or scope of future improvement of the proposed system discussed?

been applied. Publications found after filtering with these questions are considered for our analysis.

3.4 Quality assessment questions

Quality Assessment Questions (QAQs) have been prepared to check out the quality of the searched papers. The utilised QAQs are presented in Table 4.

If a publication is peer-reviewed, it implies that the quality of the paper is acceptable in the research community. It can be assured that some experts in this field have examined the claim and methodology presented in the paper. That is why, QAQ1 can be considered to be a good filtering mechanism.

QAQ2 is used to ensure that the respective publication is related to e-KYC, not KYC. Thus, any publications related to only KYC are discarded. On the other hand, QAQ3 ensures that security and privacy concerns have been properly addressed in the selected works.

Furthermore, the proper implementation of the proposed system in a work is a very important factor. A research work with its respective implementation showcases the applicability of the proposed approach. That is why we have considered QAQ4 an important criterion for assessing the quality of the work. However, since there a few relevant works with interesting ideas, we have also considered works that have no implementation.

Finally, the discussion about limitations and future work for a certain research work provides a way forward for the future researchers and highlight the depth of the analysis by the authors of the publications. Therefore, we have used this criteria to check the quality of the paper as QAQ5.

In order to assess the quality of each publication, the following formula has been utilised:

$$QAQ1 \wedge QAQ2 \wedge (QAQ3 \vee QAQ4 \vee QAQ5)$$

This means, to pass the quality assessment, each publication must satisfy both QAQ1 and QAQ2 and at least one of the QAQ3, QAQ4 and QAQ5.

After reviewing the candidate publications in the lens of these quality assessment questions, we have identified 19 publications as the final candidates for our analysis.

4 Analysis

In this section, we present the analysis of the 19 selected works against the research questions as presented in Table 1.

RQ1: What are the existing research works and their application domains for blockchain-based e-KYC?

Current research works related to blockchain-based e-KYC are mostly finance-oriented. However, some of them are for general purposes.

Upadhaya et al. [11] have implemented a system for e-KYC with Ethereum blockchain [36] where different banks would participate in the network. The system provides rating systems for the customers and to the banks based on their performance. Customers will save their time and hassles, and banks will save their money for KYC. The authors have a proposal to reduce the gas amount (computation fee as required in Ethereum blockchain) and increase scalability as future works.

Sinha et al. [37] have proposed a cost-efficient system with Ethereum and Inter Planetary File System (IPFS) [38] database. It has the options of uploading the ID, image, and address of a customer similar to any legacy KYC system. The blockchain system stores only the IPFS hash and the username of a customer. Public-private key pairs and wallet addresses are generated from a username. As the system stores only IPFS hashes and usernames, the system uses less gas (fees). In this way, the system is considered to be more cost-efficient.

Singhal et al. [10] have proposed a DLT-based KYC system using IPFS to store user details and documents. The proposed system has three independent components that are Document Submission, Notary Verification and Third Party Verification. The documents are first stored in IPFS, and the returned hash is then stored in the blockchain. The third party can access the document based on the IPFS hash and verify it if the IPFS hash of the document is present.

Sundareswaran et al. [39] have proposed a blockchain-based e-KYC system which is claimed to be more optimised. It uses symmetric AES for encryption and LZ compression algorithm [40] for the optimisation. The smart contract automatically validates the KYC data to be stored on Blockchain. Compressed KYC data help to lessen the number of gas fees on Ethereum, but on the contrary, increase the time required to extract data from the blockchain.

Ullah et al. [41] have proposed a blockchain-based Hyperledger Fabric network that reduces cost, speeds up transfers, secures data sharing, and brings transparency. The system has three main mechanisms: a permissioned blockchain, distributed storage database, and a REST interface. The smart contract of the permissioned blockchain offers progressive programmability to the distributed ledger, improving the efficiency of the ledger. The Hyperledger Composer is used to measure the runtime of the business network archive on the network. The customer can update the KYC data in real-time.

Bhaskaran et al. [42] have proposed a system for consent-driven and double-blind data sharing by designing and implementing a smart contract on the Hyperledger Fabric

[43] blockchain platform. There are three participants in the application: customers, service providers, and auditors or regulators. The proposed model ensures anonymous relationships among customers and service providers. In the model, the customer portal is the place to start and it is the representation of a customer. Banks having the correct private key may be a part of the blockchain network and run a smart contract code. The system architecture has three layers: UIs on top, APIs using REST protocol in the middle, and chaincode (smart contract) at the bottom.

Hanbar et al. [9] have proposed an optimised solution for e-KYC processes using blockchain where smart contracts are implemented on Hyperledger Fabric. To share KYC information between the banks with the customer's consent over blockchain, an anonymous AES Key-Sharing protocol has been proposed. The verified KYC documents are shared with other banks by taking the customer's consent. They suggested off-chain storage, IPFS, for data, and then storing the link of the related KYC data in the permissioned blockchain. They also analysed the results in terms of throughput and latency while performing the transactions to test the feasibility of the system.

Parra-Moyano et al. [7] have analysed different distributed technologies to implement the e-KYC process. The proposal proportionally distributes the costs among all the participants and the customers. Customers own their relevant information which is saved by local banks and in a permissioned database maintained by the regulator.

Thoroddsen et al. [8] have also proposed an optimised and dynamic KYC System that can be considered as an improvement to their previous paper as stated above [7]. This dynamic KYC system will have the option to update customer data and share the cost of update among all the FIs connected with that customer. For this, the authors have proposed a blockchain system based on two smart contracts, one for first-time onboarding of the customer, and the other for updating the customer data.

George et al. [44] have proposed a blockchain-based solution for unifying the KYC process of different banks. The proposed system has three roles: administrators, banks, and users. An administrator can enrol a customer in a bank, however, does not have the authority over the KYC data. Banks verify the newly submitted KYC data by the new users and enrol them in the system. Users can review only their data and let other banks access them. Every bank maintains its trust by utilising their digital signatures.

Rofiq et al. [45] have implemented a KYC mechanism for banking industry, by building a permissioned and modular system based on Hyperledger Fabric and IPFS. The system can process two types of data, credit records stored in Hyperledger Fabric, and documents stored in IPFS. By using this system, customers only upload the documents required in opening a bank account once to the system and reuse the uploaded document during the onboarding process for other banks. From the results of tests carried out on the system, they have concluded that the performance of the system depends on the specification of the computer used in the system.

Schlatt et al. [46] have identified some of the current shortcomings in the KYC process such as leaks and misuses of personal data, fear of aggregating significant power by a centralised service provider, and privacy-related problems arising from transparency and append-only structure of blockchain. They have also demonstrated a solution to tackle these shortcomings using blockchain-based SSI. In the paper, they have created a framework to utilise SSI in the KYC process following a rigorous design

science research approach. They have finally theorised on blockchain's role for SSI by deriving nascent design principles.

Sajid Amit et al. [47] have had a closer look at the Bangladeshi financial inclusion landscape. There are no clear rules and regulations regarding financial inclusion for the Financial Technology (FinTech) in Bangladesh. The authors argue that the existing KYC procedure in Bangladesh is not congenial to financial inclusion, and e-KYC can be a significant role player in the growth of the Digital Finance System (DFS).

From a report by Mohsin et al. [48], the authors find that the ongoing process of bKash, the most popular mobile banking company in Bangladesh, is time-consuming, not secure, and uncomfortable. They have suggested implementing the process in a specific area first and observing the outcome as the e-KYC process is new in Bangladesh.

Tina et al. [49] have reviewed the paper of Parra-Moyano et al. [7] and implemented it. They have identified some key disadvantages in the traditional KYC and prescribed some regulations to overcome those issues in e-KYC.

In their paper, Kulkarni et al. [50] have studied the current KYC procedures in banks worldwide. They have also highlighted the challenges and identified the steps that need to be taken to initiate and maintain an industry-wide blockchain consortium for ensuring a wide-scale e-KYC adoption.

The article by Arner et al. [51] addresses the identity management problem in financial institutions considering the e-KYC and digital identification infrastructures. The article comes up with a taxonomy of digital identities and investigates the opportunities of digital identities. The design of e-KYC infrastructure is based on the outcome of the investigation.

Malhotra et al. [52] have conducted a systematic review on blockchain-based KYC concepts and implementations. They have found that there are three kinds of research works in this domain: framework, case study, review. Framework-based works are of two kinds - storage-based and encryption-based. They have argued that a profit-based organisation may not benefit from a completely decentralised KYC system, but academic and non-profit organisations may benefit from such kinds of systems.

The work of Adel et al. [53] mainly focuses on various potential factors related to the spread of COVID-19 pandemic and government lockdowns that influenced the adoption of the e-KYC system by the banking sector of Malaysia. To their surprise, they have discovered that despite the severity of the pandemic the rate of adoption of e-KYC was very slow. However, the reason lies not in the desire of the banks, but in the lack of worldwide standards.

The identified application domains for the analysed research works are presented in Table 5.

Figure 3 presents the number of application domains as considered in the reviewed works. As per Fig. 3, FIs have been considered in 8 reviewed works whereas banks have been specifically considered in 6 works. In 7 works, other application domains such as non-financial industry, fintech, government, and even non-profit organisations, have been considered.

RQ2: How are the works different in architectural design?

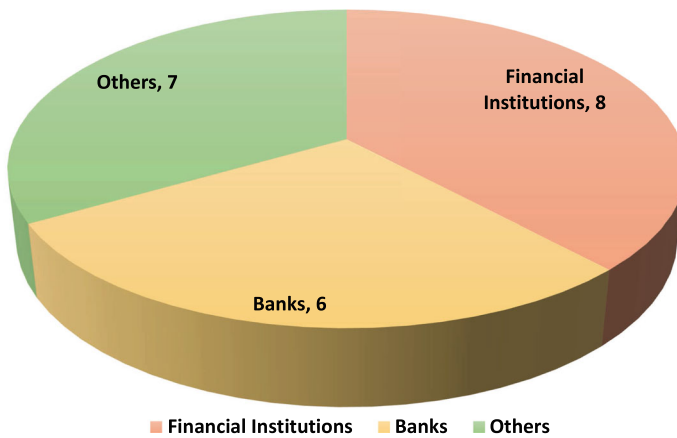
Here, we have analysed the architecture of different proposed blockchain-based e-KYC systems. Since some of the works did not provide any architectural details, they have been excluded from this analysis.

Table 5 Application domains considered in the reviewed research works (RQ1)

Title	Author	Domain
Blockchain-enabled e-KYC system	Rajyashree et al. [11]	Consortium of banks
Decentralized KYC system	Sinha et al. [37]	General industries and government sectors where e-KYC is needed
Smart KYC using blockchain and IPFS	Singhal et al. [10]	Banks and other non-banking organisations who need customer document verification
Optimised KYC blockchain system	Sundareswaran et al. [39]	Various organisations which require KYC verification of the customers
KYC optimization by blockchain-based hyperledger fabric network	Ullah et al. [41]	Financial institutions
Double-blind consent-driven data sharing on blockchain	Bhaskaran et al. [42]	Financial, healthcare and other service providers conducting KYC processes
Optimizing e-KYC process using distributed ledger technology and smart contracts	Hanbar et al. [9]	Banks
KYC optimization using distributed ledger technology	Parra-Moyano et al. [7]	Financial institutions
Optimized and dynamic KYC system based on blockchain technology	Thoroddsen et al. [8]	Financial institutions
A blockchain based solution to know your customer (KYC) dilemma	George et al. [44]	Financial institutions
Design and development of know your customer mechanism using blockchain in the process of small business loans application in Indonesia	Rofiq et al. [45]	Small businesses in Indonesia
Designing a framework for digital KYC processes built on blockchain-based self-sovereign identity	Schlatt et al. [46]	Banks
A closer look at financial inclusion in Bangladesh	Amit et al. [47]	Financial institutions in Bangladesh
e-KYC: a much-needed impetus for improving bKash's current registration method	Mohsin et al. [48]	bKash [54]
Unification of Kyc process using blockchain	Tina et al. [49]	Financial institutions
Sustainable KYC through blockchain technology in global banks	Kulkarni et al. [50]	Banks

Table 5 continued

Title	Author	Domain
The identity challenge in finance: from analogue identity to digitized identification to digital KYC utilities	Arner et al. [51]	Financial institutions
How blockchain can automate KYC: systematic review	Malhotra et al. [52]	Various institutions which require KYC verification of the customers
The attitude of potential customers toward eKYC at Malaysian Banks during the Coronavirus pandemic: perspectives of clients	Adel et al. [53]	eKYC at Malaysian Banks during the spread of COVID-19

**Fig. 3** Application domains of reviewed papers

We have explored the differences in (system) architectural design of the existing blockchain-based e-KYC systems in this section. To find out and understand the differences and their impact on the outcome of the e-KYC systems, we have used thematic analysis [55]. We have looked into the entities of the system, the interactions between them, data storage and verification mechanisms. Based on these criteria we have analysed the impact of different designs in the paper.

There are mainly three types of entities in the analysed systems: users/customers, FIs/banks and regulators. Among these entities, regulators take the roles of a relying (or trusted) third party, mostly to manage the system or to verify an identity document (only considered by George et al. [44]). However, a few works did not consider any trusted third party.

Most of the architectures have shown customers or users of the system as the sources of initial KYC. However, the system presented in [44] collected KYC information from the government database using the Aadhaar (Indian identity card) number. When KYC data are collected from a customer, it might be important to verify the authenticity of the provided information. A trust anchor (an entity who would act as the trusted

identity data provider) would be an important tool in order to verify the authenticity of information during the KYC process. Unfortunately, most of the works did not consider this crucial criterion. Only the works by Singna et al. [37] and George et al. [44] utilised identity cards as the trust anchor.

We have also explored if the research works have utilised any type of API for performing the e-KYC processes. Utilising APIs provides flexibility and extensibility. We have found that only the work of Ullah et al. [41] has used APIs to communicate between different components.

Table 6 presents a comparative analysis among different works with respect to RQ2.

RQ3: What are the implementation aspects considered in the existing research?

In this section, we explore different implementation aspects considered in the reviewed research works. Since not all reviewed works had implementation, we present here only the implemented works.

There might be a number of different types of aspects to consider during implementation. However, in this study, we are only interested in the works focused on blockchain-based e-KYC. Thus, we mostly focus on the aspects such as blockchain frameworks, off-chain/on-chain database, associated cost (if any), performance analysis and technology stacks. Reviewing these aspects will give us a clear picture of the effectiveness, feasibility and implementation cost of the systems presented in the analysed research works.

Ullah et al. [41], Bhaskaran et al. [42], Hanbar et al. [9], and Rofiq et al. [45] have implemented the system with Hyperledger Fabric, which is currently the most stable project of the Hyperledger Foundation. Generally, these works provide cheaper cost solutions in comparison to Ethereum based ones. Normally the solution platform for these type of projects are cloud and virtual machine based. The researchers have mostly analysed the performance of the works based on the Hyperledger Fabric. Hyperledger Caliper [56] is one of the most used tools for this purpose. Docker container [57] has been used in these projects as Hyperledger Fabric utilises Docker containers to deploy its network.

On the other hand, Rajyashree et al. [11], Sinha et al. [37], Sundareswaran et al. [39], Parra-Moyano et al. [8], and George et al. [44] have worked with Ethereum. Ethereum is equipped with a virtual machine, known as EVM (Ethereum Virtual Machine), which facilitates a computing platform running on top of the Ethereum blockchain. It costs significantly to carry out computations and store data in Ethereum. To mitigate these issues, either a minimum amount of data are stored on-chain (discussed next) or a private version of Ethereum is utilised. There were hardly any analysis of performance of works based on this category. Only Sundareswaran et al. [39] has provided a performance analysis.

In terms of storage, there are mainly two perspectives: on-chain and off-chain. On-chain means that data are stored within the blockchain whereas an off-chain storage implies a non-blockchain centralised or decentralised/distributed database. IPFS is the mostly used off-chain database in the reviewed works. Some authors have used local databases besides blockchain systems to speed up the processes and to extend the scalability to handle more transactions simultaneously. For example, Rofiq et al. [45], Singhal et al. [10], Hanbar et al. [9] utilised both on-chain and off-chain databases. However, some authors such as Sinha et al. [37], Parra-Moyano et al. [7], Thoroddsen et

Table 6 Different architectural aspects in the reviewed research works (RQ2)

Author	Entities	Initial KYC source	Trust anchor	Relying third party used	API used
Rajyashree et al. [11]	Customer, Bank	Customer	Not specified	No	No
Sinha et al. [37]	Customer	Customer	ID Card	Not specified	No
Ullah et al. [41]	Customer, FI, Regulator	Customer	Not specified	Yes (Regulator, role not explained)	Yes
Parra-Moyano et al. [7]	Customer, FI, Regulator	Customer	Not specified	Yes (Regulator, to manage the system)	No
Thoroddsen et al. [8]	Customer, FI, Regulator	Customer	Not specified	Yes (Regulator, to manage the system)	No
George et al. [44]	Customer, FI	Aadhaar Database	Aadhaar Database	Regulator / Government Agency	No
Schlatt et al. [46]	Customer (Holder), Verifier (FI), Issuer (FI, Government Agency)	Customer	Not clearly specified	No	No
Tina et al. [49]	Customer, FI, Regulator	Customer	Not specified	Yes (Regulator, to maintain the system)	No

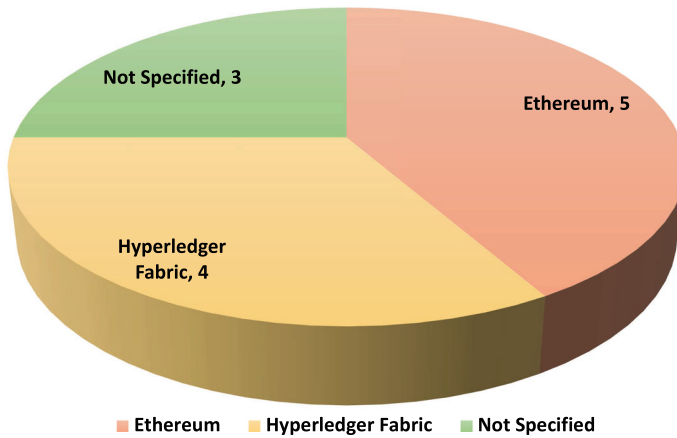


Fig. 4 Platforms of the reviewed systems

al. [8], and Schlatt et al. [46] opted for only local databases. Other works have not used any additional database except the blockchain ledger, e.g. the works by Rajyashree et al. [11], Sundareswaran et al. [39], Ullah et al. [41], Bhaskaran et al. [42], George et al. [44].

The works of Sinha et al. [37], Singhal et al. [10], Sundareswaran et al. [39], George et al. [44] have claimed to require low cost because of their cost effective storage system and compression of data. Rajyashree et al.'s work [11] is comparatively costlier. Parra-Moyano et al.'s [7] and Thoroddsen et al.'s works [8] are proportionally distributed in terms of cost. As mentioned earlier, there is no cost for transactions in Hyperledger Fabric based works.

Most of the solutions are web-based. Therefore web-based technology stacks, e.g. node.js, web3js, Postman, CryptoJS and Express.js, for blockchain and web are mostly utilised. For deploying the Hyperledger Fabric network, Docker containers have been used. On the other hand, for Ethereum either its test network or Ganache (an Ethereum simulator [58]) have been used. George et al. [44] specified the usage of Truffle (an Ethereum framework used for development with Ethereum [59]) and Metamask (a browser add on acting as a crypto-currency wallet [60]) for their implementation with Ethereum. However, such details are absent in other Ethereum-based implementations.

Different implementation aspects of the reviewed works are summarised in Table 7.

Next, we present a few figures that illustrate different implementation aspects. Figure 4 presents the utilised blockchain platforms. As evident in the figure, Ethereum has been the mostly used blockchain platform. The striking feature is that 25% of the reviewed works with implementations did not specify their blockchain platforms.

Figure 5 highlights the number of works which have discussed other aspects: performance analysis and technology stacks. Interestingly, 7 reviewed works did not carry out any performance analysis and 3 works did not mention the technology stacks used in the implementation.

Table 7 Implementation aspects in the reviewed research works (RQ3)

Author	Blockchain platform	Storage	Cost	Performance analysis	Technology stack
Rajyashree et al. [11]	Ethereum	On-chain	High cost	No	Web3.js, Node.js
Sinha et al. [37]	Ethereum	Off-chain	Medium cost	No	IPFS
Singhal et al. [10]	Not specified	On-chain and off-chain	Low cost	No	IPFS
Sundareswaran et al. [39]	Ethereum	On-chain	Low Cost	Yes	Web3.js, Node.js, CryptoJS, Lzutf8 and Nodemailer
Ullah et al. [41]	Hyperledger Fabric	On-chain	No cost	Yes (Hyperledger Composer)	AWS, Docker, Postman, Hyperledger Composer
Bhaskaran et al. [42]	Hyperledger Fabric	On-chain	No cost	Yes	A network of Docker containers, a network of VMs in a cloud
Hanbar et al. [9]	Hyperledger Fabric	On-chain and off-chain	No cost	Yes (Hyperledger Caliper)	Docker container, CouchDB container, Node.js, Express.js, MongoDB
Parra-Moyano et al. [7]	Not specified	Off-chain	Medium cost	No	Not specified
Thoroddsen et al. [8]	Ethereum	Off-chain	Medium cost	No	Not specified
George et al. [44]	Ethereum	On-chain	Low cost	No	Ganache, Truffle, Metamask, Flask, Web3.js, Google Oauth 2.0 API
Rofiq et al. [45]	Hyperledger Fabric	On-chain and off-chain	No cost	Yes (Postman, JMeter)	IPFS
Schlatt et al. [46]	Not specified	Off-chain	Low cost	No	Not specified

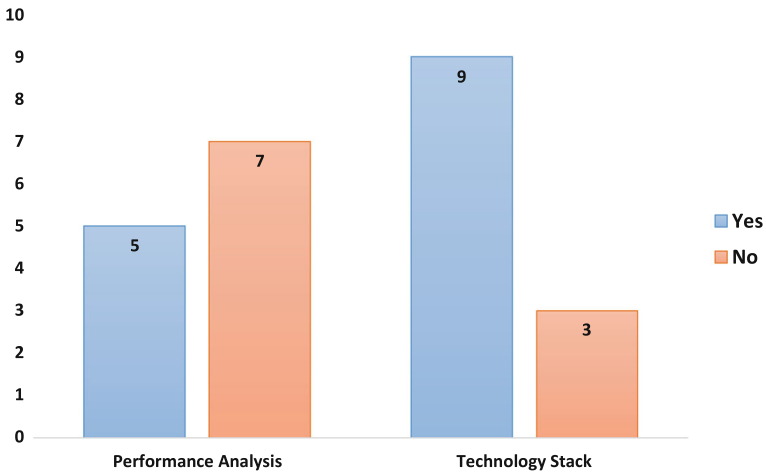


Fig. 5 Other implementation aspects considered

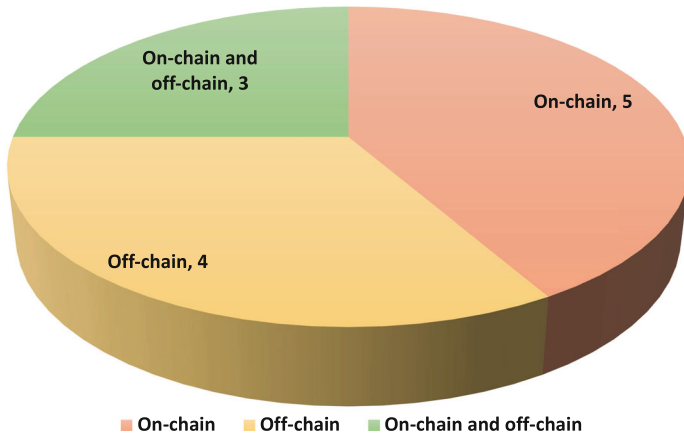


Fig. 6 Storage aspects considered

Figure 6 illustrates storage aspects considered in the reviewed works. As seen in the figure, 4 of them utilised off-chain storage, 5 used on-chain storage and the rest of 3 utilised a combination of off-chain and on-chain storage.

RQ4: What different security considerations have been considered in the existing researches?

Security and privacy are integral parts of any blockchain based research. We have reviewed all the research papers from the perspective of their different security aspects and checked different privacy concerns. The confidentiality of data utilised in the KYC system must be considered as the KYC process deals with highly sensitive data. Different encryption algorithms can be utilised to ensure the confidentiality of KYC data. On the other hand, hashing algorithms play a major role to ensure the integrity of the data. Even though authenticity is another important parameter while

considering security, we did not consider for the evaluation as e-KYC is mostly done in real-time with the presence of the user. It is also imperative to evaluate the security of any implementation with respect to a threat model and prove the implementation has mitigated the security issues by means of a formal mathematical proof or using a protocol verification tool.

Ensuring the privacy in a blockchain-based e-KYC system is another challenging part. Blockchain is a transparent system by nature. However, privacy and transparency are conflicting objectives. Thus, a blockchain-based e-KYC system must ensure that the data on blockchain is only accessible by an authorised party to ensure privacy (e.g. access control). The systems must also guarantee that users have full control on who can access their data and the unlinkability of relationships between users and organisations so that an organisation cannot build a profile for the user without their knowledge.

Rajyashree et al. [11], Sundareswaran et al. [39], Bhaskaran et al. [42] and Hanbar et al. [9] used the AES encryption algorithm within their system. In addition to AES, Sinha et al. [37] also used the DES encryption algorithm. The rest of the works did not specify the encryption algorithm that was used. On the other hand, Singhal et al. [10] and Rofiq et al. [45] did not consider any encryption algorithm in their system.

Most of the works utilised different hashing algorithms in their systems in order to ensure the integrity of the KYC data. However, the works of Singhal et al. [10], Sundareswaran et al. [39], Ullah et al. [41], Parra-Moyano et al. [7], George et al. [44] and Schlatt et al. [46] did not specify the type of hashing of algorithms utilised in their system. Rofiq et al. [45] did not consider the usage of any hashing algorithm at all in their system.

Privacy of users within the reviewed works were considered in different capacities. Bhaskaran et al. [42] considered privacy with the respect to the anonymity of relationships between users and organisations so that it becomes difficult to link users with the organisations, user consent before data sharing and via access control. Parra-Moyano et al. [7] and Thoroddsen et al. [8] mostly focused on privacy by ensuring unlinkability of users and organisations and Hanbar et al. [9] by user consent. Schlatt et al. [46] considered privacy by facilitating a full control to the users with data minimisation. Ullah et al. [41] considered privacy, however, no details were specified on how to ensure the privacy. Other authors did not consider privacy at all in their works.

Our analysis with respect to RQ4 is summarised in Table 8. A graphical summary is presented in Fig. 7 where it can be seen that only half of the works considered privacy. Similarly, 7 of the implemented works considered different security measures. One surprising factor was that none of the existing works considered any threat model. Also, all works except one did not evaluate its security. Even the one (Sundareswaran et al. [39]) that carried out security evaluation, it was quite rudimentary as there was no mathematical security proof or protocol verification by a formal verification tool.

RQ5: How do existing projects or systems implement a reputation system and use of verifiable credentials?

Reputation systems allow online users of online communities to give feedback on something, distribute and aggregate those feedback in order to build a notion of trust through that reputation scope. Such a reputation score can be a useful attribute which could help to complete the e-KYC process. For example, the information collected

Table 8 Security and privacy aspects in the reviewed research works (RQ4)

Author	Confidentiality	Integrity	Threat model	Security evaluation	Privacy
Rajyashree et al. [11]	AES encryption	SHA256 hashing algorithm	No	No	Not considered
Sinha et al. [37]	DES and AES encryption	SHA256 hashing algorithm	No	No	Not considered
Singhal et al. [10]	Not considered	Not specified	No	No	Not considered
Sundareswaran et al. [39]	AES Encryption	Not specified	No	Partial	Not considered
Ullah et al. [41]	Not specified	Not specified	No	No	Considered, but not specified
Bhaskaran et al. [42]	AES encryption	SHA-256 hashing algorithm	No	No	Considered: anonymity of relationships, consent and access control
Hanbar et al. [9]	AES encryption	SHA-256 hashing algorithm	No	No	Considered: consent
Parra-Moyano et al. [7]	Not specified	Not specified	No	No	Considered: unlikability between users and organisations
Thoroddsen et al. [8]	Not specified	SHA-256 hashing algorithm	No	No	Considered: unlikability between users and organisations
George et al. [44]	Not specified	Not specified	No	No	Not considered
Rofiq et al. [45]	Not considered	Not considered	No	No	Not considered
Schlatt et al. [46]	Not specified	Not specified	No	No	Considered: full control and data minimisation

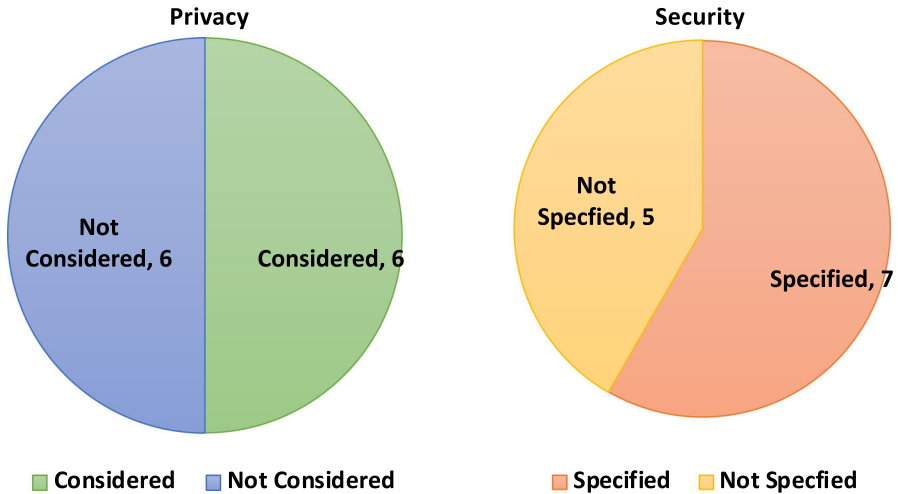


Fig. 7 Security and privacy aspects considered

Table 9 Reputation systems and VCs in reviewed research works (RQ5)

	Papers
Rating/reputation system	Rajyashree et al. [11]
Verifiable credentials	Schlatt et al. [46]

during the KYC process can be aggregated with other information while the user interacts with an FI to develop a reputation profile of users. Such a reputation profile could be shared with other FIs during the onboarding process with those FIs, thus creating a novel service delivery model. The motivations of leveraging VCs during e-KYC was discussed in Sect. 3.

In this SLR, we have reviewed if the existing works have implemented a reputation system within their system. We have also reviewed if the proposed system has used VCs. As per our analysis, only Rajyashree et al. [11] considered the possibility of a reputation score in their work while the work by Schlatt et al. [46] considered the possibility of SSI and VC within the e-KYC system. That is, reputation system and VCs are rare in practice and we have not found any work that has utilised the both.

We have summarised our analysis with respect to RQ5 in Table 9.

5 Discussion

Different authors have proposed different types of approaches for blockchain-integrated e-KYC solutions for different application domains. Even though the core e-KYC functionality remains the same, different works focus on different additional aspects as analysed in Sect. 4. Some of the reviewed works are only conceptual, however, other works have reported different levels of implementations. All these have been analysed in detail in Sect. 4. In this section, we discuss the current limitations

in the existing works (Sect. 5.1) and present a number of future research directions (Sect. 5.2).

5.1 Current limitations

From our review, we have identified a number of limitations in the reviewed works. We are presenting these limitations below.

- Very few papers clearly identify the application domains. Most of the solutions are shown to be applicable to all financial institutions in general. However, not all financial institutions are the same. They have different purposes and functionalities. As a result, the same solution should not be applicable to all. Some features can be proven to be beneficial for some FIs, and not so for others.
- Many works that have presented an implementation of the system did not carry out any performance analysis. However, it is a crucial part that a performance analysis is carried out for all implemented systems in order to test the scalability and feasibility of the system.
- One crucial aspect that was mostly overlooked in the majority of the reviewed research works is the lack of any trust anchor. As most of the works relied on customers to provide information during the KYC process, a trust anchor would be crucial to verify the authenticity of the provided information. Only two works considered this crucial factor and utilised the corresponding national identity number/card to be used as the trust anchor. However, when such a national identity card is used, the system must be carefully designed to avoid any unintentional release of sensitive data to exaggerate any privacy issue.
- It has been a surprise for us to identify that many reviewed works did not specify either any encryption and/or hashing mechanism to tackle the confidential and integrity issues. It must be noted that separate protocols must be designed in order to ensure the confidentiality or integrity of a system. The lacking of any threat model or security evaluation mean that it is difficult to judge any security claim presented in the system.
- Some of the works ignored the privacy issues altogether while their system is discussed. As e-KYC data can be very sensitive, privacy should be a major concern and specific mechanisms must be deployed to handle privacy.
- Only one paper that we have reviewed considered a reputation system in their system. However, e-KYC data can be a major source for maintaining a reputation profile of users that can largely benefit FIs to deal with their customers. This important factor was totally ignored in most of the works. However, privacy must be guaranteed while creating, maintaining and sharing such a reputation profile with other entities.
- Another limitation of reviewed works is the absence of VCs. Only Schlatt et al. [46] used VCs in their system. However, a VC-based system could streamline many security properties in a standard way and tackle a few privacy issues as it offers better user control over their private identity data.

5.2 Recommendations & future research directions

Now, we present a list of future research directions.

- Leveraging a trust anchor would be an important factor to consider during the e-KYC process as it might happen online and there is a possibility of malicious users providing falsified identity information. To mitigate this risk, an e-KYC system must integrate a trust anchor. Leveraging the national identity card within the jurisdiction of a country could act as the trust anchor for this purpose. However, this will require that the respective country has adopted such a national identity card equipped with a remote online verification mechanism. If such a scheme is not available in a country, other methods need to be devised which could be a potential research scope.
- KYC data residing within an FI could be monetised by allowing users to share their KYC data from one FI to another FI in such a way that the e-KYC process can be carried out using the data from the first FI as the source. This will open up new business models. However, as mentioned earlier, different privacy risks must be identified and mitigated while sharing such data.
- KYC data are extremely sensitive. That is why a system that is used for e-KYC must be developed following a rigorous threat model which identifies different security and privacy threats. STRIDE [61] is a well established threat model that encapsulates different security threats. However, other privacy threats must be identified as well. Once these threats are identified, proper measures must be taken to ensure that such threats are mitigated. In order to guarantee such (security) measures would in fact mitigate the threats, the security of the developed system must be proved using a mathematical proof or via a protocol verification.
- There are different privacy aspects such as anonymity and unlinkability [62] as well as explicit consent [63] and selective disclosure [64]. Adopting blockchain for e-KYC might introduce additional privacy threats because of the open and transparent nature of blockchain. In addition, the deployed system must also consider the privacy regulations imposed within the jurisdiction of the country.
- A blockchain based e-KYC system based on a standard approach could satisfy different security and privacy requirements. An SSI-based e-KYC system could be a potential candidate for this approach as proposed by Schlatt et al. [46]. However, it is to be noted that even if such an SSI-based system is developed, other recommendations need to be considered for tackling other issues as discussed above.
- Another interesting yet challenging research dimension would be to introduce a universal e-KYC scheme which would function anywhere in the world. However, integrating a trust anchor with such a universal system would be a major research challenge.

6 Conclusion

e-KYC is getting tractions as it provides a fast and convenient way to carry out the KYC process, even from the comfort of someone's home. In order to mitigate different secu-

rity issues, researchers have started exploring to introduce a blockchain-based e-KYC systems. In this article, we have conducted an SLR at the intersection of blockchain and e-KYC using the well-known PRISMA method with a number of research questions covering different aspects. Our analysis has identified that the reviewed works have different implementation perspectives. We have also analysed different implementations from the view of storage, cost, technology stack and performance analysis. A very few of existing works have utilised VCs and reputation systems. In addition, we have also identified that there are a number of serious limitations that must be addressed before we can reap the benefit of a blockchain-integrated e-KYC system. Finally, we have provided a number of recommendations and future research directions in order to address these limitations. With these contributions, we strongly believe that this SLR will be a helpful resource for any future researchers interested to work in this domain.

Acknowledgements This article is an output from an ongoing research project funded by the Ministry of Information and Communication Technology Division of the Ministry of Posts, Telecommunications and Information Technology, Bangladesh.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose. The authors have no conflicts of interest to declare that are relevant to the content of this article. All authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

References

1. What is KYC?. <https://www.swift.com/your-needs/financial-crime-cyber-security/know-your-customer-kyc/meaning-kyc>. Accessed 07 Apr 2022
2. What is KYC: steps to do KYC online: types of KYC: Paisabazaar. <https://www.paisabazaar.com/aadhar-card/what-is-kyc/>. Accessed 04 June 2022
3. Know your customer in banking. <https://www.thalesgroup.com/en/markets/digital-identity-and-security/banking-payment/issuance/id-verification/know-your-customer>. Accessed 14 Oct 2021
4. Christensen C. The four e-KYC models around the world. <https://www.regulationasia.com/the-four-e-kyc-models-around-the-world/>. Accessed 14 Oct 2021
5. Exposure draft on electronic know-your-customer (e-KYC)—Bank Negara Malaysia <https://www.bnm.gov.my/-/exposure-draft-on-electronic-know-your-customer-e-kyc-1>. Accessed 14 Oct 2021
6. Bangladesh Financial Intelligence Unit: guidelines on electronic know your customer (e-KYC). <https://www.bb.org.bd/mediaroom/circulars/aml/jan082020bfui25.pdf>. Accessed 13 Oct 2021
7. Moyano JP, Ross O (2017) Kyc optimization using distributed ledger technology. *Bus Inf Syst Eng* 59(6):411–423
8. Parra-Moyano J, Thoroddsen T, Ross O (2018) Optimized and dynamic KYC system based on blockchain technology. Available at SSRN 3248913
9. Hanbar H, Shukla V, Modi C, Vyjayanthi C (2019) Optimizing e-kyc process using distributed ledger technology and smart contracts. In: International conference on computational intelligence, security and internet of things. Springer, Singapore, pp 132–145
10. Singhal N, Sharma MK, Samant SS, Goswami P, Reddy YA (2020) Smart KYC using blockchain and IPFS. In: *Advances in cybernetics, cognition, and machine learning for communication technologies*, pp 77–84
11. Rajyashree UA, Doulani S, Pareek S (2019) Blockchain-enabled e-kyc system. *Int Res J Comput Sci VI*:137–143

12. Lootsma Y (2017) Blockchain as the newest regtech application—the opportunity to reduce the burden of KYC for financial institutions. *Bank Financ Serv Policy Rep* 36(8):16–21
13. Moher D, Altman DG, Liberati A, Tetzlaff J (2011) Prisma statement. *Epidemiology* 22(1):128
14. Know your customer. https://en.wikipedia.org/wiki/Know_your_customer. Accessed 12 June 2022
15. Jaeger J. Report: financial firms fined \$26b for AML, sanctions, KYC non-compliance since 2008. <https://www.complianceweek.com/report-financial-firms-fined-26b-for-aml-sanctions-kyc-non-compliance-since-2008/8088.article>. Accessed 14 Oct 2021
16. About Aadhaar Paperless Offline e-Kyc—Unique Identification Authority of India: Government of India. <https://uidai.gov.in/ecosystem/authentication-devices-documents/about-aadhaar-paperless-offline-e-kyc.html>. Accessed 14 Oct 2021
17. Myinfo: speed up E-KYC processes for individual users easily with data from government sources. <https://api.singpass.gov.sg/library/myinfo/business/introduction>. Accessed 14 Oct 2021
18. Board of Governors of the Federal Reserve System: joint statement on innovative efforts to combat money laundering and terrorist financing. <https://www.federalreserve.gov/newsevents/pressreleases/files/bcreg20181203a1.pdf>. Accessed 14 Oct 2021
19. RBI allows regulated entities to use video-based KYC. <https://www.regulationasia.com/rbi-allows-regulated-entities-to-use-video-based-kyc/>. Accessed 14 Oct 2021
20. Circular 3/2017 (GW)—video identification procedures. https://www.bafin.de/SharedDocs/Veroeffentlichungen/EN/Rundschreiben/2017/rs_1703_gw_videoident_en.html. Accessed 14 Oct 2021
21. Khaqqi KN, Sikorski JJ, Hadinoto K, Kraft M (2018) Incorporating seller/buyer reputation-based system in blockchain-enabled emission trading application. *Appl Energy* 209:8–19
22. Duncan RAK, Whittington M (2016) Enhancing cloud security and privacy: the power and the weakness of the audit trail. *CLOUD COMPUTING*
23. Gao L, Srivastava RP (2011) The anatomy of management fraud schemes: analyses and implications. *Indian Account Rev* 15(1):1–23
24. Khanuja HK, Adane D (2011) Database security threats and challenges in database forensic: a survey. In: *Proceedings of 2011 international conference on advancements in information technology (AIT 2011)*, pp 170–175
25. Lee KH, Zhang X, Xu D (2013) Loggc: garbage collecting audit log. In: *Proceedings of the 2013 ACM SIGSAC conference on computer & communications security*, pp 1005–1016
26. Sedlmeir J, Lautenschlager J, Fridgen G, Urbach N (2022) The transparency challenge of blockchain in organizations. *Electron Mark* 1–16
27. Nakamoto S (2019) Bitcoin: a peer-to-peer electronic cash system. Technical report, Manubot
28. Buterin V, et al (2014) A next-generation smart contract and decentralized application platform. White paper 3(37)
29. Chowdhury MJM, Ferdous MS, Biswas K, Chowdhury N, Kayes A, Alazab M, Watters P (2019) A comparative analysis of distributed ledger technology platforms. *IEEE Access* 7(1):167930–167943
30. King S, Nadal S. Ppcoin: Peer-to-peer Crypto-currency with Proof-of-stake. <https://www.swift.com/your-needs/financial-crime-cyber-security/know-your-customer-kyc/meaning-kyc>. Accessed 12 May 2022
31. El Defrawy K, Lampkins J (2014) Founding digital currency on secure computation. In: *Proceedings of the 2014 ACM SIGSAC conference on computer and communications security*, pp 1–14
32. Ferdous MS, Chowdhury F, Alassafi MO (2019) In search of self-sovereign identity leveraging blockchain technology. *IEEE Access* 7:103059–103079
33. W3.org: verifiable credentials data model 1.0. <https://www.w3.org/TR/vc-data-model/>. Accessed 13 Oct 2021
34. Sahoo P, Saraf PK, Uchil R (2022) Identification of critical success factors for leveraging industry 4.0 technology and research agenda: a systematic literature review using prisma protocol. *Asia-Pac J Bus Adm* (ahead-of-print)
35. Abelha M, Fernandes S, Mesquita D, Seabra F, Ferreira-Oliveira AT (2020) Graduate employability and competence development in higher education—a systematic literature review using prisma. *Sustainability* 12(15):5900
36. Wood G et al (2014) Ethereum: a secure decentralised generalised transaction ledger. *Ethereum Project Yellow Paper* 151(2014):1–32
37. Sinha P, Kaul A (2018) Decentralized KYC system. *Int Res J Eng Technol (IRJET)* 5(8):1209–1210
38. Benet J (2014) IpfS-content addressed, versioned, p2p file system. arXiv preprint [arXiv:1407.3561](https://arxiv.org/abs/1407.3561)

39. Sundareswaran N, Sasirekha S, Paul IJL, Balakrishnan S, Swaminathan G (2020) Optimised KYC blockchain system. In: 2020 International conference on innovative trends in information technology (ICITIIT). IEEE, pp 1–6
40. Ziv J, Lempel A (1977) A universal algorithm for sequential data compression. *IEEE Trans Inf Theory* 23(3):337–343
41. Ullah N, Al-Dhlan KA, Al-Rahmi WM (2021) Kyc optimization by blockchain based hyperledger fabric network. In: 2021 4th international conference on advanced electronic materials, computers and software engineering (AEMCSE). IEEE, pp 1294–1299
42. Bhaskaran K, Ilfrich P, Liffman D, Vecchiola C, Jayachandran P, Kumar A, Lim F, Nandakumar K, Qin Z, Ramakrishna V, et al (2018) Double-blind consent-driven data sharing on blockchain. In: 2018 IEEE international conference on cloud engineering (IC2E). IEEE, pp 385–391
43. Androulaki E, Barger A, Bortnikov V, Cachin C, Christidis K, De Caro A, Enyeart D, Ferris C, Laventman G, Manevich Y, et al (2018) Hyperledger fabric: a distributed operating system for permissioned blockchains. In: Proceedings of the thirteenth EuroSys conference, pp 1–15
44. George D, Wani A, Bhatia A (2019) A blockchain based solution to know your customer (kyc) dilemma. In: 2019 IEEE international conference on advanced networks and telecommunications systems (ANTS). IEEE, pp 1–6
45. Rofiq FA, Ekadiyanto FA (2017) Design and development of know your customer mechanism using blockchain in the process of small business loans application in Indonesia
46. Schlatt V, Sedlmeir J, Feulner S, Urbach N (2021) Designing a framework for digital kyc processes built on blockchain-based self-sovereign identity. *Inf Manag* 103553
47. Amit S (2017) A closer look at financial inclusion in Bangladesh. *CES Thought Leadership*
48. Mohsin M. e-KYC: a much-needed Impetus for Improving bKash's Current Registration Method. Internship Report. BRAC University. https://dspace.bracu.ac.bd/xmlui/bitstream/handle/10361/11041/15204075_BBA.pdf?sequence=1. Accessed 07 Feb 2022
49. Tina PQ, Shrayya A, Jain V, Kumar N (2020) Unification of kyc process using blockchain. *J Crit Rev* 7(18)
50. Kulkarni V, Singh AP et al (2019) Sustainable KYC through blockchain technology in global banks. *Annals of Dunarea de Jos University of Galati. Fascicle I. Econ Appl Inform* 25(2):34–38
51. Arner DW, Zetzsche DA, Buckley RP, Barberis JN (2019) The identity challenge in finance: from analogue identity to digitized identification to digital kyc utilities. *Eur Bus Organ Law Rev* 20(1):55–80
52. Malhotra D, Saini P, Singh AK (2021) How blockchain can automate KYC: systematic review. *Wirel Pers Commun* 1–35
53. Adel Z, Othman AHA, Bin Hasan A (2021) The attitude of potential customers toward ekyc at Malaysian banks during the coronavirus pandemic: perspectives of clients. *Rev Int Geogr Educ Online* 11(5)
54. bKashlit's that simple. <https://www.bkash.com/>. Accessed 07 Feb 2022
55. Kiger ME, Varpio L (2020) Thematic analysis of qualitative data: Ameer guide no. 131. *Med Teach* 42(8):846–854
56. Hyperledger Caliper. <https://www.hyperledger.org/use/caliper>. Accessed 18 Jan 2022
57. what is a container? <https://www.docker.com/resources/what-container>. Accessed 18 Jan 2022
58. Ganache. <https://trufflesuite.com/ganache/>. Accessed 04 Jun 2022
59. Sweet tools for smart contracts. <https://trufflesuite.com/>. Accessed 04 Jun 2022
60. The Crypto Wallet for Defi, Web3 Dapps and Nfts. <https://metamask.io/>. Accessed 04 Jun 2022
61. Shostack A (2014) Threat modeling: designing for security. Wiley, Hoboken
62. Pfitzmann A, Hansen M (2010) A terminology for talking about privacy by data minimization: anonymity, unlinkability, undetectability, unobservability, pseudonymity, and identity management. TU Dresden, Dresden
63. Lunshof JE, Chadwick R, Vorhaus DB, Church GM (2008) From genetic privacy to open consent. *Nat Rev Genet* 9(5):406–411
64. Beardsley EL (2017) Privacy: autonomy and selective disclosure. In: *Privacy & personality*, pp 56–70

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Md. Abdul Hannan¹ · Md. Atik Shahriar² · Md Sadek Ferdous³  · Mohammad Javed Morshed Chowdhury⁴ · Mohammad Shahriar Rahman⁵

Md. Abdul Hannan
hannansagar8@gmail.com

Md. Atik Shahriar
md.atikshahriar728@gmail.com

Mohammad Javed Morshed Chowdhury
m.chowdhury@latrobe.edu.au

Mohammad Shahriar Rahman
mshahriar@cse.uju.ac.bd

- ¹ Department of Computer Science and Engineering, Khulna University of Engineering and Technology, Khulna, Bangladesh
- ² Department of Computer Science and Engineering, Shahjalal University of Science and Technology, Sylhet, Bangladesh
- ³ Department of Computer Science and Engineering, BRAC University, Dhaka, Bangladesh
- ⁴ Department of Computer Science and Information Technology, La Trobe University, Melbourne, Australia
- ⁵ Department of Computer Science and Engineering, United International University, Dhaka, Bangladesh