CSE 543: Information Assurance and Security

Fall 2022, Group 1-2

Final Project Report

Securing User Identity Theft in Cloud Computing using Blockchain Technology

Dibbya Barua (**Leader**)
1227077997
dbarua2@asu.edu

Akshay Malhotra (**Deputy Leader**)
1220233742
amalho23@asu.edu

Harshita Verma
1223176266
hverma7@asu.edu

Saidubabu Mallela
1225284411
smallel2@asu.edu

Keenan Rahman
1222316745
krahman2@asu.edu

Christopher Feger
1214876221
cfeger@asu.edu

Zeal Patel
1219834291
zpatel2@asu.edu

Abstract

Identity management systems have a lot of issues and do not represent the best option when compared to other alternatives for ensuring that data is protected against misuse, theft, and illegal conduct regardless of the advancements that have taken place in this field. Blockchain technologies are being developed to mitigate a number of these hazards. In accordance with literary studies, blockchain gives individuals back control over their identities. This significantly improves user identity management.

Table of Abbreviations

IDM	Identity Management
IAM	Identity and Access Management
SSI	Self Sovereign Identity
SSO	Single SignOn
IDP	Identity Provider
DNS	Domain Name System
API	Application Programming Interface
DID	Decentralized Identifiers
SP	Service Provider
CIAM	Customer Identity and Access Management
JWT	Java Web Token
P2P	Peer to Peer
PoW	Proof of Work
PoS	Proof of Stake
IPFS	InterPlanetary File System
SDK	Software Development Kit
DIONS	Decentralized Input/Output Name Server
AES	Advanced Encryption Standard
SHA	Secure Hash Algorithms
SaaS	Software as a Service

PaaS	Platform as a Service
OPT	One Time Password
CSC	Cloud Service Consumers
CSP	Cloud Service Providers
CBA	Certificate Based Authentication
DDO	Decentralized Identitifier Descriptor Objects
IoT	Internet of Things
CMS	Certificate Management System
MFA	Multi-Factor Authentication
VPN	Virtual Private Network
PoW	Proof of Work
ISP	Internet Service Provider
DID	Decentralized Identifier

1. Introduction

Organizations should have digital identities to enable identification of individuals in an emerging digital environment that can be used with internet-based services, enabling them to communicate with one another while safeguarding their personally identifiable information. Earlier, throughout the late 1980s, people frequently utilized the email and password security mechanism to get access to networks, managing several accounts as a consequence of signing up with various Internet providers. [1] Furthermore, systems to regulate personally identifiable information are essential for the procedure of recognizing, verifying, and permitting individuals to readily access various services and infrastructure. The desire for safe access is a significant barrier to data communication technology personnel who must simultaneously provide customers' access needs and protect the privacy of their data, thus complicating the scope of the current identity and authorization management dilemma.

1.1. Motivation and Background

Identification management and access management approaches, which regulate the identification life span and offer customers access to the necessary services at the right instance and within the appropriate scenario whilst also reducing costs and energy, have been developed to meet access requirements while maintaining a certain standard of privacy and information security. [1] Regardless of the advancements that existing systems for identity management have made to the administration of authentication processes and resource access, systems still have a number of shortcomings and do not represent the best option for ensuring that data is protected against misuse, theft, and illegal conduct. Individuals believe they have almost no control over their personal information, according to assessments carried out by the European Commission. [1] Processes of identification systems which are currently in usage do not function openly and uphold granular permissions, thereby, having a direct impact on data security.

Blockchain technologies are being developed to mitigate a number of these hazards. In accordance with literary studies, this technology could significantly improve user identity management by giving individuals back control over their identities. Additionally, by combining Identity Management with Blockchain systems, it's indeed capable of holding identities decentralized, minimizing the need for a centralized authentication body and guarding against the manipulation of both the identities and even the information that is being kept. In light of all of this, various projects and applications have focused their research primarily on the development of trustworthy and consistent identity and access management systems that utilize blockchain-based technology. [1] Several of these difficulties have driven us to

examine the current blockchain-based risk mitigation and security enhancing solutions for cloud-based application security.

1.2. Goal and Scope of Study of this Project

Throughout this research, we give a detailed review of the major privacy issues surrounding user identity management systems and assess whether blockchain-based approaches help mitigate these issues and enhance data protection. We'll talk about several user authentication and authorization management solutions, such as password-based authentication, multi-factor authentication, token-based identification, and certificate-based authentication. The usefulness of blockchain for identity management will also be discussed, along with several blockchain-based methods that address user identification challenges.

This study's main objective is to educate readers about the security risks associated with user identity management platforms. We feel it is beneficial to provide a clear description of the potential threats present in user identity management systems. We think that when people have a greater awareness of cloud and blockchain technology, along with their limits, developing safe user identification systems is a feasible aim.

This study's scope of effort will include:

- Understanding the dynamics and vulnerabilities of trust relationships for both the cloud service providers and their clients.[2]
- Preserving user confidentiality and safety while insulating them from centralized third parties.[1]
- Learn about the many encryption technologies and approaches used throughout blockchain to help maintain privacy.[3]

2. Summary on the Accomplishments of the Project

Harishta Verma

- Study about the importance of Cloud Based Systems
- Understand Blockchain in Cloud Computing
- Analysis of Benefits of Integrating Blockchain in Cloud Computing
- Understand how Blockchain works

Zeal Patel

- Understanding Cloud Computing
- Study Life Cycle of Identity Management Systems
- Classification of Identity Management Systems
- Understanding Privacy Requirements of Identity Management Systems
- Study on Benefits of Cloud Computing
- Classifying various Authentication Methods

Akshay Malhotra

- Understanding classification of different types of IDM, their advantages and disadvantages.
- Studies existing IDM systems - cloud-based and blockchain-based.
- Studied problems encounters by existing, current IDM
- Analysis of Privacy concerns in an Identity Management System
- Carried out duties of a deputy-leader and ensured contant progress of project.

Christopher Feger

- Study Decentralized Systems
- Study Blockchain in Identity Management Systems

- Study Actors in Identity Management Systems

Saidubabu Mallela

- Understanding Storage of blockchain in Identity Management System
- Study Centralized Identity Management
- Study Decentralized Identity Management
- Study Using blockchain for identity and access management to avoid fraudulent activity
- Study Blockchain Structure

Keenan Rahman

- Study Blockchain's impact on identification management's security and confidentiality
- Study Blockchain Limitations
- Study Merits Of Blockchain in Identity Management
- Analysis on When to use and when not to use blockchain systems
- Study Permissioned Blockchains vs Permissionless Blockchains
- Analysis of different Identity Management blockchain Techniques
- Write Conclusion anf Recommendations part.

3. Accomplishments of Each Group Member

Dibbya Barua (Leader)

- Dibbya was responsible for understanding the team's interests as a group leader. He took account of everyone's availability and interests to ensure that all team members can meet at a convenient time.
- He contributed to project proposal by reading research papers and resources.
- He provided valuable inputs during the drafting process of weekly reports and verified the information present in them as well.

Akshay Malhotra (Deputy Leader)

- In this project, I was responsible for understanding the current state of identity management systems. This includes exploring the components compromising a digital identity and how the same is managed.
- I was accountable to read information and material that outlined different IDM architectures, and their advantages and disadvantages.
- Additionally, I was also responsible for understanding more about the vulnerabilities associated with different IDM architectures and how different variants of IDMs addressed these vulnerabilities. This includes but is not limited to privacy issues, data management issues, legality and many more.
- Moreover, I also skimmed through and briefly read about other topics related to the project. This includes popular blockchain-based IDM systems that are currently being used, their features and their components.
- As the deputy leader of this project, I was responsible for understanding the interests of all team members with regard to the project. I allocated and distributed tasks to the team members, and constantly monitored their progress to ensure that the project was progressing.
- I always offered my assistance in case any team member required additional support.
- I was also responsible for drafting and submitting the weekly report that outlined the progress of the team every week.
- In order to gather data for the same, I also organized team meetings where all team members would provide status updates and feedback.

Keenan Rahman

- I was responsible for assessing the appropriateness of using blockchain in relation to identity management systems.
- Along with this, I was responsible for understanding the current state of identity management systems using blockchain technology. This includes exploring the merits of blockchain technology in identity management systems.
- Additionally, I did an analysis of different Identity management blockchain techniques, which included studying different types of systems blockchain has to offer and what areas they target. Also, I was responsible for writing the conclusion and introduction part of this report which included the future of digital identity management, cloud computing, and blockchain systems.
- As a group member, I was responsible for developing Gantt Chart for each weekly report and helping the team approve it. I also offered my assistance in setting up the index topics for the final research report.
- As part of the research study I had focused my learning towards the Blockchain Architecture, the understanding of the Identity Management Systems. Furthermore, I did an indepth study on the topics related to the Decentralized Systems, Blockchain Techniques and the Blockchain Infrastructure.

Harishta Verma

- I was in charge of studying the limitations of blockchain technology in Identity Management Systems on cloud-based systems for this project. I began by reading a number of academic articles that explain the operation of blockchain technology and its fundamental principles, such as decentralization, immutability, smart contracts, and data integrity.
- I concentrated my study on the drawbacks of adopting blockchain technology for identity management and examined the methods used by prior systems to address these vulnerabilities.
- Throughout the process, I read about the benefits of Self-Sovereign Identity, a significant blockchain implementation that is currently in use. I studied the following papers regarding privacy aspects, principles governing blockchain system.
- In addition to this, I also authored and edited the final report, which was about 2000 words long.

Saidubabu Mallela

- For this Project I researched identity and access management issues with current cloud platforms. Evaluated, analyzed, and studied the centralized and decentralized identity management systems, as well as how a blockchain identity management system functions.

- Additionally, I read and performed in-depth studies on many identity management blockchain techniques that have been discovered.
- Moreover, I looked into a number of both permissioned and non permissioned blockchains. I studied the breadth and depth of cloud applications and developed a greater understanding of cloud applications in relation to IAM. Lastly,.
- As a group member I have given my input and review comments on the project proposal and worked on creating and evaluating weekly reports that detail the objectives and tasks for the week which in turn contributed to the addition of the final report.

Christopher Feger

- For this project I was responsible for studying the impact of blockchain on identity management, including several papers about how blockchain itself works.
- I studied three additional papers specifically about blockchain and cloud systems, such as blockchain security in decentralized systems and user identification using blockchain.
- I also wrote and proofread approximately 2000 words of the final report.

Zeal Patel

- In this project, I researched and studied Cloud Computing along with its aspects, benefits, and types. The study further led to an understanding of the various domains where cloud computing is applicable.
- Thereafter, I studied what is the Identity Management System in Cloud Computing including its working lifecycle in relation to the cloud user, the types of IDM systems in the market, and the differences between them.
- Moreover, I also studied and looked up information about the privacy friendliness standard of existing IDM systems.
- This study was further extended to understand various types of authentication methods practiced in the market at present. These include password-based authentication, token-based, multi-factor authentication and certificate-based authentication.

4. Detailed Results

4.1. Overview

4.1.1. Significance of Cloud Based System

Many cloud environments provide role assignments and identity and access management policies for end users and service accounts, as well as logging and access review auditing procedures. Implementing cloud-based identity management technologies has a number of benefits. Firstly, they assist in streamlining the various identity management operations such as access requests, password reset requests, and user provisioning.[4] This can significantly lower operational costs. In addition to this, they also give us a wide range of APIs that may be linked with user workflows, cloud services, and on-premises applications. Second, since no installation of hardware or complex infrastructure is required, cloud identity management services are significantly simpler to set up and use. Authentication, authorization, provisioning, and auditing security controls and processes can all be substantially improved by cloud-based identity management systems, which can also be incredibly helpful for regulatory audits and compliance reporting.[4]

4.1.2. How Does a Blockchain Work?

A database or ledger that is distributed by computer network nodes that stores information digitally in the format of blocks is referred to as a "blockchain." Every block does have a timestamp, a cryptographic digest of the previous one, and transaction records. A blockchain is a decentralized, distributed, digital ledger that is employed to record transactions across numerous machines in a manner that prevents the record from being updated retroactively without impacting all following blocks and necessitating network approval.

Key features of blockchain include:

Shared Ledger: A shared ledger is a system of record that is "append-only" and distributed throughout a corporate network. Transactions are recorded just once, eliminating the duplication of effort that is common on traditional business networks

Permissions: Permissions ensure the security, authenticity, and verifiability of transactions. Permissions in blockchain help adherence to data protection rules. Transactions are secure, authenticated, and verifiable due to permissions. [5]

Smart Contracts: An agreement or set of rules that regulate a business transaction is referred to as a smart contract. A smart contract is kept on the blockchain and executed automatically as part of a transaction.

Consensus: All parties must consent to the network-verified transaction in order for a transaction to be conducted on the blockchain platform. Blockchains have a variety of consensus processes, including multi-signature, PBFT, and proof of stake (practical Byzantine fault tolerance).

4.1.3. Blockchain in Cloud Applications

Public, private, and hybrid clouds are the three different forms of cloud infrastructure. The user on the leased premises owns and manages the private cloud infrastructure. [5] The cloud service provider owns the public cloud infrastructure, and they are the only ones who can manage the service remotely. Cloud computing, which can be considered to be centralized within the context of this study report, provides various advantages, including but not limited to faster speeds, lower latency, and more availability. However, many of these advantages come with serious setbacks, such as being vulnerable to security breaches and data hacks and having no control over data. Because of this, the decentralized cloud storage concept aims to prevent hacking, a single point of failure, and data loss by using a large-scale backup mechanism.[6] Users can store data anywhere on the network because the nodes are dispersed and independent of one another.

4.1.4. Advantages of Integrating Blockchain to Cloud Computing

Decentralization: Relying on a centralized server for data management is a significant issue as the entire system might be compromised by a central server failure, which could potentially result in the loss of crucial data. [7] The main server is also vulnerable to attacks. Blockchain can offer a solution to this issue since it uses a decentralized architecture that prevents a single point of failure by storing numerous copies of the same data on several computer nodes. Furthermore, copies of the data are present on multiple nodes, which prevents data loss significantly.

Increased Data Security: Storing data in the cloud can be unreliable since hackers might be able to gain access to it. The adoption of blockchain in cloud computing as a solution to this challenge has the potential to provide greater security to the entire architecture. [8]

Fault Tolerance: A network of computer servers that are robustly connected to one another via collaborative clouds might benefit from cloud replication of blockchain data to improve fault tolerance.

By doing this, the chances of a single failure brought on by the disruption of any cloud node will be reduced, enabling continuous services.

Scalability: Blockchain typically permits 10–20 transactions per second, as opposed to more than 1000 transactions per second for conventional techniques. In order to enable scalable blockchain services, it is crucial to have strong data processing services that have high transaction execution thresholds. In this area, the cloud's scaling capabilities can provide on-demand computing resources for blockchain activities. [8] Therefore, a highly scalable integrated system can be produced by combining blockchain technology and cloud computing.

4.2. Cloud Computing

4.2.1. What is Cloud Computing?

Cloud computing is a distributed system of low-cost computing units. It provides the user the benefits of computing on an external machine with a pay-for-what-you-use scheme. It is one of the fastest-growing and most in-demand technologies in the world today. Its popularity among users arises from the fact that it provides highly scalable computing units that are highly available on demand, i.e. resources get added or removed dynamically as per their use. [9]

It allows users to develop applications, store data, or access software already deployed on the cloud. It is good for innovators to focus on the development of their product instead of worrying about the logistics of gathering the resources and setting up the environment for it. The cloud is a collection of servers distributed globally. These servers can be physical or virtual computers. The web services and applications hosted on the cloud can be accessed via the internet. It supports a broad spectrum of workloads, including batch-style back-end jobs and interactive, user-facing apps. Also, it allows workloads to be swiftly delivered and expanded via rapid provisioning of virtual or physical servers. Moreover, it helps recover from many unforeseen hardware and software breakdowns. [10]

4.2.2. Services Models that a Cloud Service Provider Provides

Although there are many service models that exist in the world of cloud computing, the 4 types of service models that are extremely popular and readily available are mentioned below:

1. **Software as a Service(SaaS)** - A cloud provider maintains applications and make them accessible to target consumers through the network using the software as a service (SaaS) paradigm.. Here, the software development vendor, who can be an individual or an enterprise, hosts the application on the

third-party cloud or on its own cloud-like in the case of Google. For example, Gmail and Google Docs.[11]

2. **Platform as a Service (PaaS)** - PaaS is a platform distribution model in which a cloud provider hosts the necessary software and hardware for application development. It includes the Operating System, Middleware, and other database management tools over the cloud infrastructure.[11]

3. **Infrastructure as a Service (IaaS)** - Iaas is an infrastructure distribution model in which the cloud provides the basic infrastructure such as networking, storage, and computing resources.[12]

4. **Serverless Computing** - Serverless Computing is a service model where the cloud provider manages and scales the resources on the requirement of the code run by the developer. The term serverless is to signify the fact that the developer is free from the management of the infrastructure and the cloud takes the responsibility for it. [12]

4.2.3 Cloud Computing Architecture

The Cloud Computing architecture is based on 2 components:

1. Users - An individual or a business that uses the cloud services.

2. Cloud Provider - An enterprise that provides the above-mentioned cloud services. E.g. Amazon Web Services(AWS), Google Cloud Platform(GCP), and many more.

Some of the main cloud service providers are Amazon Web Services(AWS), Google Cloud Platform(GCP), Microsoft Azure and IBM Cloud. As of today, cloud computing is used in almost every sector, ranging from individual developer use to commercial use in the sectors of healthcare, military, e-commerce, automotive, banking and many more.

4.2.4. Advantages of Cloud Computing

There are several benefits of cloud computing for companies. Indeed, there are so many benefits that you almost have to think about moving your operations to a cloud-based platform. In spite of this, many businesses continue to adopt outdated, ineffective procedures because they are not aware of the advantages.

1. **Accessibility from any location and with any device -** Applications deployed on the cloud are accessible for any geographic location, be it a branch or office in different states or countries.

Employees are not the only ones who benefit from better accessibility; clients and customers can also log in to an account and view their information. This ensures that everyone, whether at the office or on the go, has up-to-date information.[13]

2. **Capability to remove most or all hardware** - Owning a server, network switches, backup generators, and other hardware is not necessary with cloud computing. Depending on the cloud provider you choose, they might be able to handle everything for a monthly fee. Every cloud-based platform benefits from cost reduction, which is a crucial component of any company strategy.
3. **Centralized data security -** When you utilize cloud computing, data backups are aggregated in the data centers of the cloud providers, so no specific user or team needs to maintain their own backup locally or remotely. This lessens the likelihood that data will be lost if one backup crashes or is damaged by a disaster. Cloud service providers can recover the data by using a backup copy that is continuously updated when new data is added to their cloud storage.
4. **Increased performance and availability -** Because cloud services are dispersed across various cloud facilities, they provide high availability with minimal downtime.[13] Cloud providers are responsible for upgrading cloud systems as well as resolving bugs and security flaws in cloud software, which is visible to end users.
5. **Faster Application deployment -** Unpredictable corporate operations typically require the immediate utilization of cloud computing resources. Rapidly deploying cloud apps, which are readily accessible without the need for extra hardware or for you to wait for IT specialists to set up servers, can help you boost your cloud application development.
6. **Quick business insights -** The ability to access data as soon as it is collected is a unique capability provided by cloud-based technology. This enables improved decision-making and provides insight into what your company's future may contain based on predictions made from historical data.[13]
7. **Continuity of operations -** Do you have an adequate backup plan in place in case of tragedy or unanticipated circumstances? If not, relying on cloud computing services can help your company. Cloud computing uses endless data storage capacity and technology that may be activated remotely if necessary to maintain business continuity. [13]
8. **Price efficiency and cost savings -** Although a monetary investment is necessary to implement a cloud approach, businesses save a lot over time since they no longer have to maintain pricey hardware or local data centers. [13]. Furthermore, because cloud-based solutions have no upfront expenses, firms may test them out before investing in them at their own speed.
9. **Virtualized computing -** Cloud computing is ideal for virtualized computer settings because cloud resources can be rapidly assigned to meet huge surges in demand, ensuring that downtime is never

experience again.[13] With cloud computing, your company may almost effortlessly increase its capacity to meet growing demands without expanding employees or capital costs.

10. **Environmental Friendly -** Compared to conventional IT solutions, cloud computing represents a more ecologically friendly technology. Businesses that switch to the cloud could reduce their carbon footprint and energy consumption by up to 90%. Instead of maintaining internal servers and software, businesses may utilize cloud-based services to access the same apps and data from any machine or device with an internet connection. No longer must businesses own and maintain their own IT infrastructure.[13]

4.3 User Identity Management Systems

4.3.1. Identity Management System Life Cycle

The lifecycle of identity management includes the entire lifecycle of a consumer's identity credential, from the creation of an identity, to maintaining the identity through password changes, and then, at the end, deletion of the credentials linked to the account upon deactivation of the account by the consumer.

1. **User Provisioning -** This phase is the initial phase where a cloud consumer creates an account in the Identity Management System. The user subscribes to a cloud-based service on a cloud service provider and the IDMS stores all the essential information related to the user. The IDMS, in return, provides a set of credentials used to access the resources. Along with that, a role is also assigned to the user to allow them access to different levels of services. For example, an admin has more privileges and access to many resources than a regular, genuine user.[14]

2. **Account Changes/Management -** During the course of the subscription, i.e. the user using the resources provided by the cloud service provider, many changes occur. The user might subscribe or unsubscribe to a few services/resources.[15] For example, let’s say that the role of the user (employee) changes in an organization, and the employee is promoted to manager. This would result in changes in access rights for that specific user. Thus, the IDMS has to keep all the information up-to-date. It aids in the prevention of potential conflicts or security breaches, such as unauthorized or illegal access to cloud resources.

3. **User Deprovisioning -** When a user leaves the organization or in general removes his/her account over the cloud service provider, then all the user information and the rights related to the role of the user must be deleted from the cloud. Any kind of delay in this deletion process would lead to many security risks.[16]

4.3.2. Centralized Identity Management Systems

Identity and access management are carried out in the same system. This involves a user login towards a single domain to securely connect to all applications and resources one needs in a work context. The foundation of IAM solutions is the credibility that has been built up amongst their users, networks, and providers. IAM systems are a flexible choice for organizations since they serve customers, workers, and stakeholders. Identity and access management contains user login details for the accounts with data gathered from various sources organized into groups. Similar to IAM, centralized IAM structures could also include additional features including enrollment, self-account management, permission & selection management, and common data processing. Such individual accounts that are created enable businesses and their collaborators to offer customized services to their customers.

Federalised identity and access management, for example, is just a centralized IAM tool that enables users to sign in to several entities or sectors with only one verification. When their employer permits it, individuals may use 3rd party applications such as Salesforce with just a single keystroke. To carry out their everyday activities, individuals don't need to memorize several credentials or usernames whenever they log-in in the morning, this enhances efficiency.

Customers may anticipate CIAM to provide a comparable experience. For instance, customers may quickly and easily access independently run financial services like obtaining checks, transferring funds via Zelle, or applying for a loan. When a customer updates their contact information through one program, the rest are automatically synced.

4.3.2.1. Advantages of Centralized Identity Management Systems

A smart centralised IAM solution is a game-changer both in terms of privacy and from a cost-savings perspective.

- **Speedy Deployments in Case of Crisis:** Data breaches are both embarrassing and expensive. Take into account the 2017 Equifax network attacks, that led to the company getting the UK's worst penalty. [17] Teams are given insight through a single Identity management platform, which enables users to immediately identify issues and take appropriate action, saving businesses both valuable time.
- **Homogeneous Identities Autonomous Process Management:** Bid farewell to laborious procedure for granting & rescinding access permissions on a case-by-case basis. Having consolidated identities that provide for true visibility within which individuals are entitled as for what (and the amount of

privilege), offering and not offering are all performed through one spot. Reporting on specific individuals, groups, and apps has become simpler. As a result, IT can quickly complete detailed inspections while monitoring historical records. [17]

- **Simplified Single Sign-On:** As users join to other programs, users are required to create and memorize multiple credentials. Individuals are therefore more inclined for choosing passwords which are easy to recall or even to repeat a single one throughout many systems. [17]
- **No Obstructions:** In numerous businesses, it creates a bottleneck. A single Identity management software considerably minimizes IT work and engagement, thereby decreasing obstacles, whether it is by speeding up the onboarding process for newly hired or giving individuals the ability to change their own passwords. Increase your efficiency using unified, automated authentication and authorization. [17]

4.3.2.2. Disadvantages of Centralized Identity Management Systems

- **Single Point of Failure:** Centralized identity and access management possesses drawback of introducing a single failure point. All of a participant's possessions are accessible to the hacker if their credentials are taken. But if you choose a provider which is built to stop the most cunning attacks, you could reduce your vulnerability. While assessing centralized IAM solutions, take multi-factor protection and accurate risk monitoring skills into account. Login details are secured at the point of entry, and security is handled by experts instead of the consumer. [18]

4.3.3. Decentralized Identity Management Systems

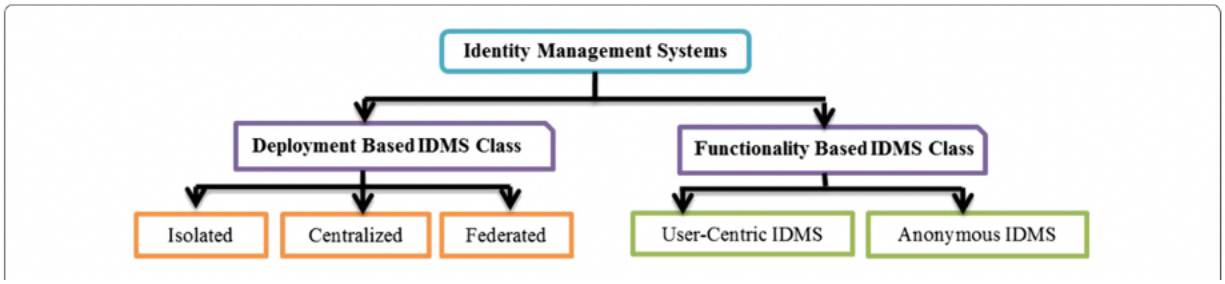
Decentralized identity management, sometimes referred to as individual identity, self-sovereign identity, or decentralized identity, gives individuals greater control. Through using dissimilar systems approach, people can keep identifier information inside a virtual wallet stored on their personal mobile. When opening an account in a bank, purchasing a vehicle or starting a fresh employment, for example, consumers could keep this knowledge updated and only disclose the information they require to businesses and other users.

Likewise towards how people store a driving permit, debit cards, and proof of insurance inside our wallets, individuals get accounts validating the identities out of a number of sources (companies, the authorities, etc.) and retain it there. Every individual could create a collection of encryption keys inside their identification wallet and decide to just reveal the minimum necessary of them.

Decentralized identification is an example of a concept that anybody might well be familiar with. In addition to confidentially proving vaccination and COVID diagnostic testing toward others, COVID Mission enables vaccine providers, businesses, and people to satisfy policies to address regarding their vaccination coverage across period via privately secure communications. People receive electronic proof of vaccinations in ShoCard online wallets that they can securely disclose to affiliated businesses like cafés and workplaces by scanning a barcode. The consequences of digital transactions might be significant. Currently, more than 6 billion smartphone subscribers are available worldwide. Masses of individuals in developing countries could rely solely on their smart phones for obtaining banking and some other services, as they've developed to be among the most potent electronic identification devices.

4.3.4. Classification of Identity Management Systems

They were classified based on their deployment architecture and functional behavior.



Cloud identity management security issues & solutions: a taxonomy; Umme Habiba; Rahat Masood; Muhammad Awais Shibli; Muaz A Niazi

4.3.4.1. Deployment-Based Classification

When classified based on the deployment strategies, there are 3 popular architectures - Isolated, Centralized, and Federated. The fundamental architecture for the storage, administration, and transfer of identity information is the main focus of this category. Depending on the requirements of CSCs and CSPs, identity information could be maintained on a single storage server or spread over numerous servers.

1. **Isolated Cloud IDMS -** Small and medium-sized businesses frequently employ this concept. A single server has served as both an identity provider (IDP) and a service provider (SP).[19] It stores identity information and various user operations. Customers authenticate themselves at the cloud service provider, which then forwards the query to the Identity Provider for completion. An authentication approval is sent back to the corresponding user. The issuing and validation of the user credentials are

not dependent on the Trusted Third Party (TTP) in this procedure. The disadvantage of this system is that it is difficult to scale up services and resources.[20]

2. **Centralized Cloud IDMS -** Unlike Isolated IDMS, Centralized IDMS has a single IDP (a Trusted third party) that issues, stores and manages the identity data of the users.[21] It manages all the information taken from the cloud service provider centrally. Following that, the CSC sends an authentication request to the CSP, the authentication request is forwarded to the appropriate IDP, and the authentication answer is returned to the CSP. The authentication response will be sent to CSC (either successful or an error message).[20] Typically, a single CSC can utilize the services of several CSPs that have access to the same IDP. CSPs and CSCs in this situation must share a trusted IDP since it is in charge of managing sensitive identifying credentials. The single point of failure is a clear disadvantage of the centralized IDMS concept.

3. **Federated Cloud IDMS -** Federated Cloud identity management system is the embodiment of the federated identity management paradigm, which allows subscribers from many organizations to utilize the same identifying information to have access to all networks within any particular trusted group of enterprises. [19] Because of the Federated Cloud Identity management solution's design agility, which naturally allows cross-domain access to its users by removing the need to create new user accounts for third parties, the IT industry has paid special attention to it. The distributed storage concept, in which identity information is held in several places, is followed by federated IDMS. [21] In this architecture, the consumer of cloud service sends a request to the cloud service provider (a federated IDMs) which authenticate the request by gathering consumer information. It then transmits the authentication request to the next IDP and retrieves the relevant attributes from its Identity data store. Finally, an authentication response is generated and returned to the requesting CSP. This method is repeated until all of the attributes required for authentication have been collected. In this, the security is enhanced as the user's information is linked over multiple IDPs on the user's authentication request.

4.3.4.2. Functional-Based Classification

This class includes the functional behavior-based IDMS named Anonymous and User-centric IDMS. These systems do not rely on the underlying architecture. For e.g. The user-Centric IDMS can have any of the following as the underlying architecture - federated or centralized identity management method for the storage of the user credentials. The main focus of this classification is the functionality of the Identity Management System.[22]

User-Centric Cloud IDMS - In this IDMS, the users are involved in every identity authentication process. In order to access any service offered by the cloud service provider, the consumer sends a credential request to the corresponding IDP which in return provides a token to complete the authentication process.[22] Here, the consumers are responsible for the storage, management and retrieval of their information related to identity. The consumer has the authority to decide with whom to share the credentials, be it the trusted entities such as the service providers, IDPs or other users. Therefore, this factor increases the privacy of the user. The disadvantage is that accessing each and every service/application requires the authentication process to be performed. Additionally, there is a huge overhead with regard to maintenance of credentials.[23]

Anonymous Cloud IDMS - An identity management system that provides anonymity as a feature is called an anonymous identity management system. An anonymous cloud identity management system can keep its entity (owner) secret from others. Anonymous identities must be strong enough to make it difficult, if not impossible, to reveal your real identity. This is because the derived data may eventually be linked to other information and republished. However, anonymous identity management also has the following drawbacks: be it a lack of trust between CSC and CSP. Also, consumers of the service use temporary identities to perform actions, which seems to defeat the purpose of logging and monitoring.[24]

4.3.5. Classification of User Identification System

4.3.5.1. Certificate-Based Authentication

In order to identify individuals, computers, or other objects before giving access to programs, networks, or other resources, certificate-based authentication (CBA) employs encrypted digital certificates. Unlike some authentication solutions that target humans, such as One-time passwords (OTP) and biometrics, certificate-based authentication can be employed for all endpoints including servers, PCs, e-passports, and literally anything that can be categorized as the Internet of Things (IoT).[25]

It is a much more secure alternative than a simple username/password authentication process. It becomes phishing attack resistant when combined with other traditional methods for a stronger user authentication process. Because the digital identity is saved with the private key on the individual's computer or smartphone and could be readily delivered as needed, the website or user may smoothly login into a variety of services without doing any additional work.[25] Client certificate-based authentication and other approaches in which the secret is never revealed to even the user are superior to password-based authentication in general. Username and password authentication rely solely on what the user knows (the

password), whereas certificate-based client authentication also relies on what the user possesses (the private key), which cannot be phished, guessed, or socially engineered.

To maintain this level of security, userse should also be vigilant and keep their private keys secure along with the security of their physical devices. The CBA architecture providers are also responsible to provide a secure path to conduct the process.[25]

4.3.5.1.1. Advantages of Certificate-Based Authentication

- Makes the authentication procedure easier.
- Reduces the use of unsafe passwords.
- Password attacks such as brute force and rainbow tables are no longer a danger.
- Resistance to phishing attacks.

4.3.5.1.2. Disadvantages of Certificate-Based Authentication

- Although setting up a digital network infrastructure for certificate-based authentication is a one-time expenditure, it is not inexpensive. It may just not be a viable alternative for many start-ups and small businesses.
- The continual maintenance of CBA, including issuance, renewal, and revocation, must always be considered. This includes, for example, operational and licensing fees for a Certificate Management System (CMS).[25]
- Digital certificates are not usually convertible between products and suppliers because they are vendor specific. This implies that certificates may be incompatible with specific devices and systems, however, this may change in the future, with Microsoft, for example, soon expanding support.

4.3.5.2. Multi-Factor Authentication

In order to access a resource like an application, online account, or VPN, the user will have to provide two or more verification factors, which is known as MFA. A effective identity and access management (IAM) policy must include MFA. MFA minimizes the probability of a successful cyberattack by demanding one or more additional verification requirements in combination to a login and password.[24] The primary benefit of MFA is that it enhances security for your business by demanding users to provide additional identification than merely a login and password. Although usernames and passwords are crucial, they are susceptible to brute-force assaults and can be stolen by other parties. Your organization's capacity to defend itself from cyber criminals will be more trusted if you require the usage of an MFA component, such as a fingerprinting or physical hardware key.

MFA operates by requesting more verification data (factors). One-time passwords are among the most frequently used MFA factors that consumers come across..[26] The 4 to 8 digit codes known as OTPs are usually sent to you via mail, Text, or a smartphone app. OTPs generate new code periodically or each time an authentication request is sent. The user is given a seed value when they first register, and a second factor—which might be anything as basic as an increasing count or a time value—is used to generate the code. With the advent of cloud computing, MFA has gained even greater significance. Companies can no longer depend on security being offered by a person physically connected to a system when they move existing systems to the cloud. In order to make sure that users accessing the systems are not hostile actors, added security precautions must be put in place. MFA can help guarantee that users are who they claim to be by requesting extra authentication factors that are more difficult to reproduce or break using brute force techniques.[26]

4.3.5.2.1 Advantages of Multi-Factor Authentication

- The obvious advantage of multi-factor authentication is that it increases the security of your firm. When you add another layer of security, you are enhancing your whole cybersecurity strategy. For example, if a cybercriminal gains access to an employee password via a brute-force attack, adding another factor will only serve to thwart them. [26]
- It is possible to deploy MFA simply by using a physical token.
- If one of your employees loses a device, which is a very real and growing possibility due to remote work, you don't have to be concerned about compromised data or access.

4.3.5.2.2 Disadvantages of Multi-Factor Authentication

- The stress of managing a second component might be felt by personnel.
- Putting multi-factor authentication into practice can be expensive and time-consuming.
- There may be some inconsistencies while adopting MFA across an organization.

4.3.5.3. Token-Based Authentication

When authenticating users, a system known as token-based authentication employs encrypted security tokens. Users can use it to log in to websites, which creates a special, encrypted authentication token. [25] Without requiring them to enter their credentials again, the token gives users momentary access to resources and pages that are secured. User identification information is securely sent between applications and websites via an authentication token. They enable businesses to enhance the way they authenticate users for these services. The three primary parts of an authentication token are the header, payload, and

signature. The signature algorithm and token type are both specified in the header. The issuer and expiration date of the token must be included in the payload. Along with other information, it also provides useful information. A message's signature attests to its authenticity and guarantees that it hasn't changed while en route.[27]

Mentioned below is an outline of the components and their steps in token-based authentication:

1. **Request:** When a user logs in to a service using their credentials, a server or protected resource generates an access request.[27]

2. **Verification:** To confirm that the authorized users have access, the server verifies the login information. This implies comparing the entered password to the supplied username. [27]

3. **Token submission:** The server generates a safe, signed access code for the user for a predetermined period of time.[27]

4. **Storage:** The user's browser receives the token back and saves it for future visits to the website. The token is encrypted and verified when the user switches to another website. On the occasion there is a match, the user may go on. [27]

5. **Expiration:** The token is only good while the user is logged in or until the service is shut off.[27]

4.3.5.3.1. Advantages of Token-Based Authentication

- Tokens are stateless: A user can validate their identity without supplying login credentials by using authentication tokens, which are created by an authentication provider. [27]
- Tokens expire: The token is erased when a user logs off of the service after concluding their browsing session. By doing this, it is made sure that user accounts are safe from hackers. [27]
- Tokens are encrypted and created by a machine: Token-based authentication uses encrypted, computer-generated codes to confirm a user's identity. Every token is specific to just a user's session and has been safeguarded by a mechanism that makes sure servers can detect and reject manipulated tokens. A far safer alternative to using passwords is encryption.
- Tokens streamline the login process: Users don't need to re-enter their account information every time they browse a website thanks to authentication tokens. This expedites the procedure and makes it easier to use, which attracts repeat visitors by prolonging their stay on websites.
- Tokens add a barrier to prevent hackers: To stop hackers from stealing user data and corporate resources, implement two-factor authentication. Passwords by themselves make it simpler for

Hackers can access user accounts while tokens enable users to validate their identity using physical tokens and smartphone applications. As a result, even if the user's login credentials are stolen, there is an additional layer of security that prevents a hacker from accessing the account.

4.3.5.3.2. Disadvantages of Token-Based Authentication

- Compromised Private Key - The fact that tokens only need one key is one of their main drawbacks. Yes, JWT only uses one key, which could have major consequences if handled improperly by a developer or administrator and endanger important data.[27]
- Data overhead - When more data is supplied to a JWT, its length grows since it has a considerably bigger overall size than a typical session token. So, adding more data to the token might result in a slower overall loading time and degradation of the user experience. [27]
- Shorter lifespan - Operating with short-lived JWT is more challenging for users. These tokens demand routine reauthorization, which can occasionally be annoying, especially for customers. [27]

4.3.5.4. Password-Based Authentication

Password-Based Authentication is the process of acquiring access to resources to which one is entitled using a set of credentials that include a username and a password. This is a widely used approach noted for its ease of use and inexpensive cost. When several users have access to resources in an organizational network, it is critical that their identities be validated before they are provided access to their entitlements. That can be accomplished with the use of a password, as it has long been one of the preferred methods of validating one's identity and relies on one's ability to verify oneself by producing the appropriate credential.[28] The complexity and secrecy of the password ensure the security of all resources in the organization. An attacker who obtains a user's password can impersonate the user and get access to sensitive resources to which the user is authorized. For the majority of resources in the organization that require a login and a password, password-based authentication is the preferable technique. The Password-Based Authentication method is popular due to its simplicity, cost-effectiveness, the convenience of use, and practicality. [28]

There are three specific types of authentication-

1. **Knowledge-based:** Also known as "something you're familiar with." Traditional passwords are included. When you, as a user, create a unique password for your account, it becomes the key to continually entering an account. It is something that only the user (ideally) is aware of. [28]

2. **Possession-based:** Also referred to as "what you have." In this scenario, a person proves themself using something that only he or she has. For example, a user might utilize a physical key card to show that they are who they say they are. [28]
3. **Inheritance-based:** Also known as "what you are." These are biometric traits used to validate an individual's identity, such as a facial scan or a fingerprint. [28]

4.3.5.4.1. Advantages of Password-Based Authentication

- Knowledge-based: Simple to manage, inexpensive
- Possession-based: Difficult to duplicate, lost is soon discovered
- Inheritance-based: Always available, difficult to steal

4.3.5.4.2. Disadvantages of Password-Based Authentication

- Knowledge-based: It is simple to leak, but it is difficult to discover when it is lost.
- Possession-based: Can be damaged; replacement is slow if lost or damaged
- Inheritance-based: When corrupted, irreplaceable, false acceptance or rejection

4.3.6. Privacy Requirements for Identity Management System

To reduce the risks associated with the unwarranted and otherwise undesired disclosure of personally identifiable information, any accessible IDM platform must protect consumer privacy. The past few years have seen a significant increase in the usage of security and visibility improvement measures within IAM systems across Europe, both in the national and European Union sectors. [29] This is because many of these regulations call for businesses to follow data reduction, data security, and, in some circumstances, record retention plans. According to the information reduction rule, private details shouldn't be given to a business associate except when it is extremely important for the deal to also be completed. The justification for the exposure of each data element that will be revealed must be provided to demonstrate such a strict need.[29] Individuals must have accessibility and the ability to modify personal details whenever it is stored by an entity. Corporations must also maintain records in a way that enables effective investigation of past purchases. In this context, "private details" describe any data that could be used to identify a specific individual, regardless of when combined with other information. There are more specific obligations as a result of the duty to limit the quantity of private information transferred among entities. An IDM platform that respects individuals' confidentiality should permit them to:

- Judiciously provide personal data to institutions as well as other users

- Create many personas or aliases.
- Connect numerous bits of private data to separate personalities.
- Analyze information that has already been made public.
- Maintain distinct personas to varied groups.
- Establish "permanent" guidelines that govern the handling, publication, and utilization of private information.
- Minimize the number of faith consumers needs to have in other people and various infrastructural parts.
- Individuals should be given the option to withdraw previously given permission and provide explicit consent before providing personal data.

It is rather challenging to do all of the aforementioned tasks in a way that is useful, such that, without placing an unnecessary strain on system administrators and individuals. To roughly meet the aforementioned conditions, employ the following metrics. They can be evaluated.

These IDM platforms' levels of confidentiality may be assessed using the next set of parameters.

1. **Trust Model -** Many IDM technologies are intended such that a remote organization designated also as an "Identity Supplier" stores and manages the individual's private data (IDP). Individuals can frequently access their data and send that to interested entities after being authenticated by an IDP. Although this paradigm has the benefit of being mobile (users may access the platform from every machine and virtually any place), it also raises serious privacy and security issues.[30] Since other entities would turn to the third authority anytime they need user data or confirmations of user authenticity, the trusted party not only acquires that lot of personal information but also details about the user's connections and interactions. Moreover, it has to be believed that indeed trusted party won't assert falsely that an individual has indeed been verified or that they have provided false details about their characteristics.[30] However, if an individual does have the option to choose which IDP to use, the privacy problems posed by using a 3rd party IDP may be slightly reduced. Passport, Microsoft's original effort to address the identification management issue via serving as that of the universal IDP, brought attention to this matter of option. The thought of trusting Microsoft with substantial quantities of private information sparked a broad and violent backlash, that took Microsoft back by surprise, as was widely documented. After all, a diverse environment of organizations ready to act as IDPs is necessary for selecting to be successful, and so this environment has still been developing.[31] Additionally, not all consumers will have access to the tools necessary to decide which IDPs to entrust with their mission-critical personal data, even if such an ecosystem does develop. An approach

that saves personal information, such as traits and certificates, on the user's machine and then discloses it directly to the individuals that need it is much more confidentiality. This tactic is less practical, though, as users would have to perform more administrative tasks and portability is no longer expected. Additionally, it highlights how important platform system security is for safeguarding user data. In the end, everything boils to faith.[29] Customers will be required to decide whether to entrust a trustworthy third party (like an IdP) or their own system when it comes to the management of their private information. Nevertheless, historical experience indicates that this is an extremely critical problem since, as demonstrated by the myriad problems involving social media sites, consumers are frequently making incorrect trust choices on the processing of their personal information.

2. **Multiple Unlinkable Identities -** Typically, a user's identity is defined as the accumulation of their personal data (characteristics, qualifications, as well as other declarations about the person). The individual can "build" one or even more individual identities by gathering pertinent details about themselves. Identification need not be continuous; for instance, one identification may include the user's actual name and address, while the other one could contain a pseudonym (such as a surname) with no address. The ability of a user to switch among settings or responsibilities is facilitated by the organizing of information into ids.[30] It ought to be mentioned that in certain literary works, a constructed persona like this is referred to as a "digital identity." Customers in complex systems were frequently confined to one identity, but customers in open platforms were typically not. If indeed the IAM system interacts among many institutions that have the ability to independently identify persons, it may be desirable for users to have the capability to be using separate identities with other institutions. To further accomplish data reduction, mechanisms must be developed to forbid cooperating organizations from linking a specific information about the user place at a single institution with the identical person's profile at the other. Allowing individuals to build and maintain many identities their own is very straightforward, but making ensuring that all these identities stay unlinkable is trickier. Consumption habits and various attributes might occasionally provide enough details to connect a particular digital credentials. The strategy must not, however, prevent people who value their anonymity from keeping credentials that are technically unlinkable.[31]

3. **Selective Disclosure -** In order to carry out selective disclosure, it must be equally possible and straightforward for a user to disclose only a portion of their identify to a given requestor. It should be possible to only show the user's age or maybe even age group (for example, 18–25) if the system has previously stored the user's date of birth, as opposed to requiring the installation of a separate identity or a drawn-out enrollment procedure. [29]

4. **Consent -** There are several uses for personal information. For instance, if a consumer wishes to purchase electronic goods delivered by email, he must supply his email address. On the other side, an online business can try to exploit a customer's email address for research or marketing. [31] The explicit consent of the user for other purposes must be supported by an I&AM system that protects privacy. Other provisions should be put in place for data retention timeframes when user input is required. If users no longer want to be contacted by the other party, the should be allowed to remove access.

5. **Privacy respecting sharing of personal information -** A "sticky policy" is similar to a consent request in that both let the user choose acceptable recipients, purposes, and retention times. The distinction is that a sticky policy is "attached" to the data, while consent only pertains to the first recipient of personal information. As a result, the policy is open and applicable to all "downstream" data processors, or anybody to whom data is disclosed during an I&AM system operation.

4.3.7. Problems Encountered by Current IDM

There are a variety of Identity Management Systems architectures that have the potential to create a huge negative impact. To start, the oldest and most traditional manner of creating, managing, and using identities linked to an individual is using paper-related documentation. An identity of an individual is documented on a paper-based physical document that is stored in a centralized location for reference. This technique of managing identities is now antiquated and outdated as it creates a host of problems related to privacy, unauthorized access, single point of failure, and management problems. Digital identity management, on the other hand, has reduced many of these problems. In a digital identity management system, there are different entities that play a unique role in authorizing and validating identity. Additionally, there are a plethora of architectures that exist in the domain of digital identity management that have their advantages and disadvantages.[32]

One of the biggest challenges in current identity management systems is establishing a trust relationship among IDPs or between IDPs and service providers. Services providers and identity providers can each have a different level of privacy protection and security mechanism, thereby making it challenging for these entities to exchange information securely, especially in accordance with the law. IDPs with higher privacy protection levels find it challenging to put their trust in other IDPs or service providers who have lower trust standards. Moreover, the lack of standardization in protection and privacy standards adds to the problem.[32]

One of the most common IDM architectures is Silo. In this architecture, the service provider is responsible for authenticating, authorizing and providing the service to the user.[32] The service provider plays the responsibility of the identity provider because it owns the namespace and authentication tokens for all of its users. As the number of users increases, the identity creation and management process can be tiresome in combination with the responsibility for providing the service. Additionally, a huge amount of diverse PII information is managed by each service provider of varying levels of security. Due to this, substantial privacy threats can arise if some service providers with lax security countermeasures reveal PII. Consumers who access services from numerous SPs experience identity overload and password fatigue. For critical, sensitive services, where password recovery must be highly secure, misplaced passwords can substantially increase the expense of providing the service.[33]

The centralized IDM model introduces the concept of a shared domain and a single IDP to reduce redundancy and privacy issues. Utilizing an ID given by a single IDP, the user can access all service providers in the same domain using the centralized architecture.[33] Despite not authenticating users when they access services, a central authority serves as an IdP and controls IDs and authentication tokens. The fundamental drawback is that establishing and maintaining a universal namespace for all users is practically and politically difficult. In terms of the global namespace of identifiers, email addresses are distinct yet unstable as individuals have many email addresses and social security numbers are inappropriate due to privacy concerns. Moreover, if an attacker gains access to a user's unique ID authentication details, they may quickly assume the user's identity and exploit it to access all domain services.

In Federated IDM, a mutual authentication, security, and protection agreement form the cornerstone, which enables user SSO to a set of service providers. These entities that form a mutual agreement and share information are referred to as “circle of trust”. Although this arrangement does improve upon multiple domains, user privacy may still be in danger. As a portion of the user's PII is shared with federated IDPs and service providers with the user's tacit consent, service providers in this architecture can track the user by mapping multiple of the user's IDs if they are malevolent. Additionally, it is challenging to discern between a genuine user and a service provider posing as a user.

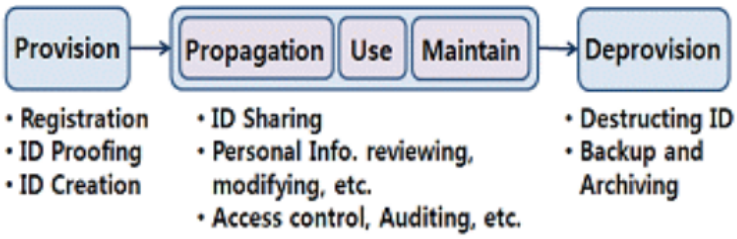
As the notion of the digital world is still evolving, a unique outlook called "self-sovereign identity" (SSI) has begun to gain hold. The objective is to boost trust and security while safeguarding user privacy and defending them against the growing domination of centralized third parties. Christopher Allen depicts a representation of self-sovereign identity that encompasses a variety of characteristics. Their characteristics revolve around the domains of existence, control, access, transparency, persistence,

operability, interoperability, consent, minimization and protection.[1] IdM's based on blockchain emulate the SSI models to a good extent, thereby giving birth to a new generation of identity systems.

4.3.8. Privacy concerns in an Identity Management System

Privacy threats in IDM systems are of the utmost importance as they can compromise the identity of an individual, thereby causing unwarranted consequences. The impact of such a disaster has always been significant. In 2017, a significant data breach involving 46.2 million mobile users occurred in Malaysia.[34]

The figure mentioned below depicts the entire ID life cycle. This life cycle involved multiple critical tasks including but not limited to propagation, use, maintenance and destruction. The lifecycle can be broadly broken down into 3 main components - ID Provision, ID Usage and Propagation, and ID destruction.[34] Each of these stages faced unique privacy-related challenges that have been mentioned below.



1. ID Provision

The process of enrolling a user who desires to utilize the IdM service and creating his or her ID from an IDP is called ID provision. This can be broadly divided into 2 different categories:

a. **User registration and ID proofing:** In order to create a digital identity, a user must register themselves with an Identity Provider. A user's identity and legal legality are verified as part of the registration procedure. The user is asked to provide PII to verify their identity.

   A significant privacy concern can arise if the IDP gathers PII more frequently than necessary or handles it carelessly. Additionally, if the IDP has a lower level of privacy protection and security mechanism then it prohibits collaboration with service providers. In order to deal with these issues, the user should be informed about the type of PII required during ID verification, its intended use, and its retention time. They should also be informed about the administrative and physical safeguards in place to protect them from potential danger.

b. **ID creation and rights assignment:** This procedure results in the creation of ID, the authentication data associated with ID, and the assignment of user permission to ID. During this process, additional PII can also be generated by the IDP to identify them. The management of these unique attributes related to the identification is important as they do not necessarily have to be stored together. However, introducing centralization with the storage of identification-specific information can cause privacy issues.

2. **ID Storage, Propagation and Maintenance**

   This life-cycle stage can be divided into 3 stages:

a. **ID Storage**: Once a digital identification is created with regard to a user, it is stored safely. Traditionally, this digital identity is maintained in a database along with its properties and attributes. If the accessibility permission to the PII information is wrongly defined, a whistleblower or collaborator of IDP may leak the user's PII. Therefore, based on their responsibilities or the IDP's policy, the accessibility permission should be restricted against the whistleblower or partner who has the right of access to the storage. Crucial data among ID-related characteristics and the PII should be encrypted before storage.

b. **ID Propagation:** This stage deals with the transfer of ID and ID-related information to other parties or service providers. This is a crucial part of the lifecycle as this helps in removing redundancy and helps reduce management costs. Moreover, this helps build security standards amongst multiple different entities. The criticality and degree of services offered by other systems and agents, as well as the extent to which PII and ID-related attributes are required, should be taken into consideration. The unauthorized release of the user's PII is likely if the excessive PII is transferred across other systems.

c. **ID Maintenance:** Once a digital identification is issued, it should be possible for the user to retrieve and modify information related to the digital identity. Storing additional PII might require additional verification, however, the IDP should be able to provide a secure mechanism for all the above-mentioned operations. It is important that only the correct individual with the right privileges is allowed access to edit the identity. The PII may be made public if the privilege is improperly assigned, thereby, raising privacy concerns.

3. **ID Destruction**

   This process deals with the deletion of ID and its related information and attributes. A major security weakness in that system could result if any privileges were still granted to the removed ID. A privacy issue might arise if the ID, characteristics, or PII of the withdrawn user is not removed, or if data

pertaining to the legitimate user is removed unintentionally or accidentally. As a result, extra verification steps or technical measures are needed to determine whether the user ID and related information were actually deleted using the right process and manner.

Mentioned below is a summarization of the various privacy risks and threats associated with each stage of the ID lifecycle.

Life Cycle	Privacy risks	Privacy threats
Provision	- Request for excessive PI - Insufficient evidence - Error in ID creation Incorrect PI	- Leakage of PI - ID creation to unintended user - Auth. Info for unintended user
Propagation, Use, Maintain	- ID/PII sharing with untrusted IdPs - Unauthorized data reviewing/modifying - Inappropriate access control /logging - Inconsistency of PII among IdPs	- ID Masquerading - High possibility of PI disclosure by unauthorized IdP or user - User tracking or profiling by ID
Deprovision	- Disclosing deleted ID,PII - Destructing valid user ID and PII - unauthorized deleting	- ID Masquerading - Loss of valid user's PI - Service inaccessible to valid user

4.4. Blockchain Technology & Identity Management Systems

4.4.1. What are Decentralized Systems

A solution to some of the problems of identity management can be found in decentralized systems. Decentralized systems spread the cost of storing and processing data across many physical computer systems (often called "nodes"), reducing costs in the process. While aggregate system costs may increase, the pooling of the resources of nodes allows large tasks that may be infeasible for a single node to do to become possible.

Decentralized data systems are one example of this pooling of resources. Isolated computer systems have limited storage space, but through the usage of thousands of nodes, a single decentralized data system could have aggregate storage space in the range of petabytes, all while only requiring single nodes to have terabytes individually. As high-density storage for a single node is expensive, this could reduce the total cost of the data storage for the system. [35]

Because of this wealth of storage space, decentralized data systems can afford duplication across nodes. This duplication helps ensure that every piece of data in the system is always available, allowing for downtime of any single node and often multiple nodes, depending on the structure of the system. This

kind of universal access and supply of information is extremely valuable in the field of identity management, as it allows for a system to verify the identity of any user at any time, possibly through its own local storage, if the data happens to reside there, or through requesting it from the network, which will contain at least one copy. This avoids current systems' problems with single points of failure. There are several types of decentralized data systems, but the one that is the focus of this paper is the blockchain.

4.4.2. What is Blockchain

A blockchain consists of a ledger, which may be either publicly or privately distributed. This ledger contains records of some sort of transaction; in many current blockchains, these transactions are of an associated monetary unit, while in an identity management system, these transactions could be akin to accesses to identification data. This ledger is decentralized, allowing anyone with access to the blockchain to view it.[35] The ledger consists of an ordered list of blocks, with the order determined by the time of submission to the blockchain. The decentralized, peer-to-peer nature of blockchain technologies allows the ledger to be copied between hosts with no major bottlenecks in the system, as no central server is involved.[1]

Blockchain blocks are linked and ordered by a cryptographic hash function. This hash function takes an arbitrarily sized input value and transforms it into a fixed size output value. Each input value has a deterministic output value. A requirement for secure hash functions is also that it be only one way, meaning that the content for an input cannot be calculated from its output. Further, while the mapping of a hash function is not one-to-one, as that would be impossible for any function with a fixed size output and a variable size input, the space and construction of the hash function should be vast yet precise enough that, for the amount of inputs required of it, it is extremely unlikely that any two given inputs compute the same two hashes. This feature allows blocks to be identified by their hashes.

General structure of blockchain blocks consists of a header that contains the hash of the previous block, along with some metadata about the data stored, including a hash of it, or, in some cases, a Merkle root, which consists of the hashes of the hashes of transactions. This previous block link through the hash in the header is what orders blocks and forces them to be immutable: if a previous block is modified, its hash is modified due to the properties of a hash function. The only way to modify a previous block would be to control a majority of the network, thus modifying the way that consensus works.

Storage of these blocks, as mentioned, is decentralized and duplicated, ensuring availability to the data from anywhere connected to the network and at any time. It is also validated in a decentralized manner, as

every duplicate contains the information required to validate it. Every node in the network validates transactions to ensure that they fit rules of transactions dependent on the blockchain itself; in cryptocurrencies, these rules may prohibit double-spending or overspending, and also ensure that a block is not being made farcical. These rules are the “consensus” of the blockchain.

Several consensus processes are in use in current publicly available blockchains. One of the most popular is proof-of-work, where “miners” race to find an acceptable solution to a hard mathematical problem. A second popular consensus process is proof-of-stake, which is based on the scarcity of a cryptocurrency.[1] Both of these processes work by making it difficult to get a block accepted, thus lowering the amount of farcical blocks that will make it through.

4.4.3. Blockchain Structure

Blockchain is now a standard protocol used during peer-to-peer (P2P) communications systems that create, authenticate, and distribute ledgers of various types of transactions. It is expressed by a predefined structure in a specific sequence (nodes).[36][1] This is relied around asymmetric encryption, cryptographic certificates, and cryptic hash methods.[36] This blockchain based system is made up from many nodes and primarily uses this described mechanisms:

- **P2P Network:** To transfer information among nodes within blockchain networks, a protected transmission mechanism is used over a Peer to peer network. Transactions are disseminated throughout all nodes in the absence of a centralized system. This architecture serves as the basis for such Blockchain's distributed features.[1]
- **Storage:** In order to maintain entire neighborhoods of replicated operations on every node, blockchain technique focuses upon state transition duplication.[1] By reducing a single failure point through decentralized storing, the blockchain based system can keep functioning even when certain system components breakdown. Large chunks, on the other hand, require a large amount of space and propagate through the system relatively slow.
- **Validation:** By using this process, repeated expenditure concerns with cryptocurrencies are avoided, as well as the integrity of blockchain is preserved.[1] For the purpose of validating transactions against a set of requirements, each nodes in the blockchain based network confirms whether transactions were genuine and haven't been expended.. After that, blocks consisting of valid records can be built.
- **Consensus:** This entails a collection of rules which syncs any nodes having knowledge of the total transaction existence and also the ledger's present condition. Many decentralized consensus techniques are proposed for such blockchain. Among the most well-known are:

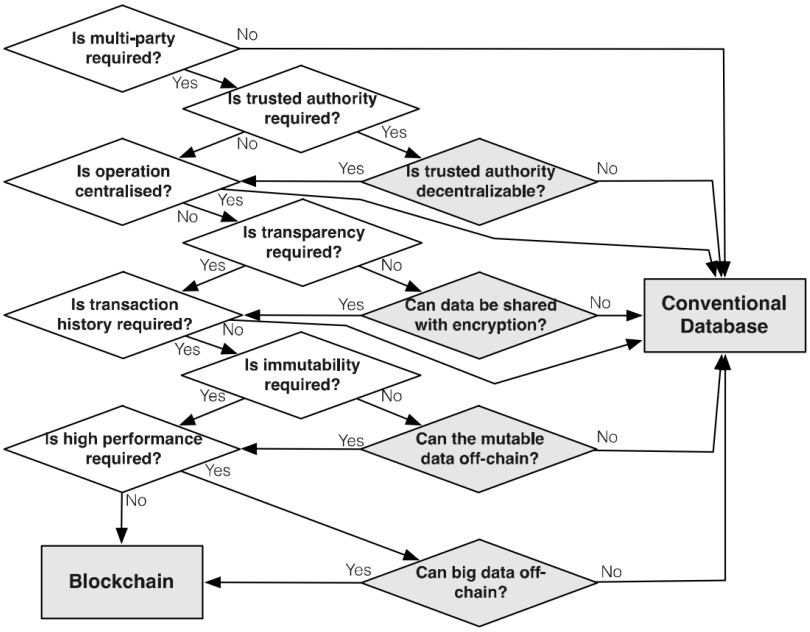
i) Proof-of-Work (PoW) This is focused on the competition among mining to discover a realistic alternative to a challenging computational problem.

ii) Proof-of-Stake (PoS) This relies mostly on commodity's shortage as for a substitute towards PoW.

- **Cryptography:** With the use of this technique, the information is offered significant security and confidentiality. Blockchain technology uses a modern cryptographic method for operations and accounts. As a result, all blocks that are created are permanent and therefore cannot be deleted or modified.

4.4.4. When should one use blockchain-based systems?

Assessing the appropriateness of using blockchain in relation to using scenario criteria is the initial stage in designing a blockchain-based system. The seven issues that must be addressed as part of the procedure to determine if blockchain technology is appropriate are known as white decision nodes. Grey decision nodes are used to indicate the sublevel questions that are generated from the basic questions.[5]



1. **MultiParty:** The very first query is if the situation involves many parties. In most cases, intermediates oversee the activities or dealings among parties.[5] Blockchain offers an independent, common infrastructure that is not dictated by any of the participating organizations. Blockchain is therefore appropriate for cases concerning numerous participants, maybe with facilitators functioning inside the existing programs.

2. **Trusted Authority:** The next query concerns whether the situation calls for reliable authorities. A trusted authority is indeed a body that has the power to carry out a specific action, change a regulation, or configure an activity.[5] Blockchain is appropriate in situations where there is no figure of authority or when the existing trusted authority may become decentralized. Blockchain is frequently referred to as a "decentralized trust" since it eliminates the necessity to rely on a single specified 3rd party to manage a transaction's log.

3. **Centralized Operation:** The implementation of process operations for smart contracts in blockchain-based platforms is more difficult than that of conventional distributed applications.[5] By using a system that uses blockchain, no one party will be in control, and each individual would be in charge of their own information and data, which also poses problems for regulation. As a result, blockchain in its present state is unsuitable for an application that necessitates centralized management.

4. **Data Transparency vs Confidentiality:** An unbiased network in which all parties may view the published information is provided by a blockchain. Prior to being stored on a blockchain, information is encrypted.[5] This might boost secrecy but also decrease performance, accessibility, or independent auditability. The primary trade-off involves the advantages of data exchange among collaborators and maintaining secrecy towards rivals when necessary.

5. **Data Integrity:** For traceability to be created, that can be utilized to trace assets across changes in possession and treatment, data integrity as in previous transactions is essential. Comparing blockchain to other persistent technologies, it could be costly to guarantee integrity.[5] The authenticity knowledge gathered by utilizing a blockchain might not be advantageous to a design with an established monitoring system.

6. **Data Immutability:** The substantial support that blockchains can offer for immutability and non-repudiation could be a considerable benefit in markets where 3rd party service providers are rarely reliable.[5] Since information on the blockchain is continuously duplicated over several places and entities, it is practically impossible to modify it; efforts to do so are denied by other stakeholders as something of an attack on system integrity. Blockchain ledgers' immutability could make them less adaptive than traditional systems managed by reputable 3rd parties who facilitate reversal.

7. **High Performance:** Although the scalability of blockchains is presently not very great, this isn't certainly an inherited constraint and could be resolved soon. Since it is unable to handle huge amounts of information moving at a fast rate, blockchain is just not suited for preserving Big Data.

Due to the extreme duplication caused by a large number of processing nodes each keeping a complete copy of the distributed ledger, this is a fundamental problem of blockchains. The existing approach is to keep the bulk of content off-chain in order to prevent data repetition among all linked peers.

4.4.5. Permissioned Blockchains vs Permissionless Blockchains

Different forms of blockchain have arisen since blockchain technology has advanced: permissionless and permissioned blockchain. The permissionless blockchain technology allows for data to still be "distributed among all users on the network, amended by miners, watched by everybody, and possessed and managed by no one," which is the best way to define it.[37] A body (person or company) may utilize its smartphones or desktop gadgets to connect to the system using a permissionless blockchain, like Bitcoin. Decentralization is really an advantage of a permissionless blockchain, which has been supported by the popularity of various well-known services, notably the cryptocurrency Bitcoins. But there are downsides. For instance, executing a lot of operations quickly on a permissionless network like Bitcoin is limited.[37] The preservation of its confidentiality is much more important, and big corporations are concerned that a distributed ledger system could jeopardize proprietary information.

The blockchain variant having constraints upon participation and regulatory control is referred to as a permissioned blockchain. In a blockchain technology like Ripple, these rights of the members are defined by the core config, which determines which individuals may acquire, and add information to the blockchain, as well as authorize the admittance of newcomers.[37] The permissioned blockchain may be regarded as somewhat distributed since various members possess varying access control permits. Alternatively, a blockchain network possesses a huge possibility to retain confidentiality and meet corporate governance objectives unlike a permissionless blockchain with proper implementation of access-control layers. On the flip side, a permissioned blockchain permits a centralized government with overriding rights, which can damage the blockchain's reputation.

Blockchains with permissions and those without permissions have different fundamental characteristics. Below, we go into further detail about the unique characteristics.

1. **Unreliability and immutability**

- In a permissionless blockchain, ledgers are permanent once they have been uploaded to the network, negating the requirement for middlemen or centralized authority.[37]

- A blockchain with permissions is not completely unreliable. A centralized organization having overriding power might undo operations. The majority of the participants could also decide to overturn the transaction history.[37]

2. Distributed Agreement and Openness

- Every connected network member retain an exact replica of the system under permissionless blockchains. The agreement is accomplished by continuously syncing the copies, ensuring that the data is accurate, clear, and always up-to-date.[37]
- A blockchain with permissions doesn't guarantee complete transparency. Not every node receives a copy of a master record of transactions. Permissioned blockchains are more acceptable in a corporate context due to the secrecy safeguards offered by such access limits.[37]

Although permissioned blockchain systems have a variety of desirable characteristics for business identity and access management (we explored why these characteristics may enhance current identity management services in preceding paragraphs), those also offer a variety of critical hurdles that the scientific world must evaluate.

4.4.6. Effectiveness of Blockchain

4.4.6.1. Decentralization

When an identity is managed by a centralized authority, it is difficult to establish trust between the user and the governing authority. No proprietary organization should centrally register and control an identity.[6] A decentralized infrastructure, such as distributed ledger technology or decentralized network technology, should be used to register and manage digital identity. Smart contracts, which are pieces of code that govern interactions between parties that are mutually untrusting, are used in blockchain-based systems. Because the code is difficult to change, this builds trust. By establishing a blockchain-based system, no one party will be in control of the system, and each user will be in charge of their own data and assets, which naturally fosters a higher level of confidence.[6]

4.4.6.2. Immutability

Once information is published on a blockchain, it cannot be changed. Each block has a hash value for both itself and the block before it. In turn, this ensures that the blocks are connected and that saved data cannot be altered. As a result, when data is controlled by more than two servers or nodes, it is practically impossible for attackers to compromise it at any moment.[6] It enables a free and transparent flow of data

and guards against data corruption because every activity taken on a digital ledger is encrypted with a specific hash code, on which the majority of nodes must agree.

4.4.6.3. Privacy

Only the minimal identification information necessary for service or verification should be asked from an identity owner, and they should be allowed to remain as anonymous as feasible. There shouldn't be any way for the identity infrastructure to connect private and biometric information to an underlying identity. Only with the owner's consent should any personally identifying information with an identity be disclosed.[6]

4.4.6.4. Integrity

Since blockchains are immutable, once a data addition or transaction has been made, they cannot be changed or deleted. Additionally, it is simple to record and keep up-to-date proof of the history of data. Data reliability and trustworthiness are referred to as "data integrity," and the immutability of a blockchain makes this guarantee possible.[6]

4.4.6.5. Trust

Blockchain is a distributed database that keeps track of all transactions going back in time.[6] Instead of being constantly verified by a centralized body, the network as a whole helps verify the identity of the individual. In this way, security is ensured without consumers having to rely on a centralized authority thanks to the strength and computational capacity of the whole network involved in the blockchain.

4.4.6.6. Cybersecurity

For any identification infrastructure, the security of identity and the communications associated with it are of utmost importance.[6] Various degrees of identity security, including cryptographically secure connections and communications, digitally signed transactions, and decentralized and encrypted storage are included.

4.4.6.7. Traceability

One of the most crucial elements of the blockchain is the shared ledger, which makes it easier to trace assets and information and builds technology-based trust between people participating in the network without a central authority. Through tracking and tracing the origin of products, blockchain-based traceability offers the potential to detect fraud or counterfeit transactions.[6]

4.4.6.8. Cost Reduction

Everyone should have access to an identity at no cost or for a very small expense; there shouldn't be any additional costs for licensing or any other type of financial transaction only to own an identity. This might not be the case for expenses associated with other resources and implementations, though. If an identity is to be made available to everyone on the earth, the cost issue is critical.[6]

4.4.7. Expense of Implementing Blockchain Solutions

The cost of setting up blockchain solutions is generally twofold: first, there is the problem of starting a network large and diverse enough such that it is able to provide meaningful privacy and integrity guarantees that do not depend on a single entity that controls a majority of the nodes in the network, and second, there is the cost of every transaction, including the original few that must be made to start the blockchain.[38]

The cost of starting a network large and diverse enough to provide privacy and integrity guarantees is somewhat inherent to starting any distributed system: it is generally necessary to start from a small seed and build upwards. What several solutions have done is integrate themselves into existing blockchain technologies through their smart-contract systems, which automatically execute code when certain conditions are met. This allows them to leverage a preexisting network over having to build one from scratch, gathering nodes from many different sources to provide the proper guarantees. This cost is thus somewhat mitigable, as long as the underlying blockchain is suitably configurable for the needs of an identity management solution.

The cost of every transaction, however, is less mitigable. Proof of work consensus systems consume huge amounts of computational resources, which in turn require huge amounts of electrical energy. This cost may not directly affect either of the user or the service, but the effect on the ecosystem through carbon emissions should be taken into account as well. Proof of stake consensus systems do lessen this somewhat, but simply the fact that the system is distributed, duplicated, and running pieces of itself all at once across a huge array of networked machines means that there’s still a large cost in this respect. More specific research would need to be done to see if the difference in environmental impact between current non-blockchain identity management solutions and solutions that do use blockchain, specifically those using proof of stake or other less-resource-heavy consensus systems, is non negligible. [38]

4.4.8. The Impact of Blockchain on Identity Management

Using blockchain in identity management has been brought up multiple times in related literature. Concerns regarding privacy, security, and socioeconomic usability (Panait) have been raised regarding this technology.[4]

Privacy and security issues regarding the usage of blockchain in identity management include two main issues, both related to verifiability: in the case of a self-sovereign identity system, there is the problem of requiring claim-verifiers, or users that endorse identity of others, much like current public-key infrastructure.[4] This requires users that wish for their identity to be verified to somehow communicate this information to other users outside of the blockchain, which requires a secure, trusted relationship between the identity issuer and the identity verifier.[4]

The second is in the case of a decentralized trusted identity system. This system assumes that there is a single service that provides the user's identity and records it on the blockchain. This service would thus be a single point of trust, which weighs on security and privacy concerns: first, it is possible that this single service could be breached and thus insert farcical data into the chain. Second, its control of enrollment of user data in the blockchain is a concern to the privacy of the person enrolling, as it requires the central service to have all of their data at once for submission.

Despite these concerns, there are several existing implementations of blockchain-based identity management. These include, in a non-exhaustive list, Sovrin, MyData, Waypoint, Bloom, BlockStack, ShoCard, uPort, I/O Digital, BlockAuth, UniquID, Jolocom, Cambridge Blockchain, KYC.LEGAL, CertCoin, and Authenteq. These solutions all have their own problems: Sovrin faces low adoption rates due to its difficulty to use, uPort's JSON structure that encodes data is visible to anyone in the network, and ShoCard requires a "bootstrapping" document that often includes more personal information than was meant to be revealed, to name a few. Similar problems exist in other solutions

4.4.9. Blockchain Identity Management Actors

When dealing with identity management in general, there are a few actors that must be considered: the user, the verifier, and the service. The user is the entity that generally has its identity managed by the system: when it requests data from the service, the service will ask the verifier to confirm that the user is who they say they are. The service, thus, is the system that wants information about the user, for whatever purpose it needs: it may be a bank establishing a user's Social Security number for opening a new account, or just an online game establishing that a phone connecting to it is not someone impersonating

the user to gain access to their purchases. Finally, the verifier is the system that tells the service upon request whether or not the user is who they say they are, according to some heuristics defined by the management system. Note that the verifier and the service can be the same entity.[32]

The user is a major crux of this system, being the entity about which data must be collected and verified for the verifier to perform its job.[32] If the service can convince either the user or the verifier (which we will assume is a trusted entity for now) to reveal more information to it than is required, that constitutes a breach of privacy, which could theoretically compromise the user's security.

If instead, we do not assume that the verifier is trusted, the verifier becomes the most important entity in the system: it has the power to lie to the service and the power to disseminate any information given to it meant to be part of its verification process. This imbalance of power is what distributed identity management systems try to solve.[32]

Specifically, when dealing with identity management using a blockchain, there are three factors again: the user, the blockchain network itself, and the service. The user is the same as before, as is the service. The blockchain network takes the place of the verifier, and due to its distributed nature, changes the weight of several of the problems mentioned above. As long as the portion of the network that is hostile does not exceed 50%, the network is trustworthy. The problem of dissemination of information is still present, possibly more so as there are more nodes that have that information.

4.4.10. Blockchain's Impact on Identity Management's Privacy and Security

Rather than examining the accuracy of the information within provided evidence, the confirming entities could utilize the blockchain's architecture to evaluate the legitimacy of the certification and certifying entity, from which they could decide to either authenticate the proof or not.[39] Whenever an identification owner presents proof of one's name and date of birth, for example, the confirmation entity would first rationalize the government's biometric data that has been accepted and certified to a qualification before deciding if they trust the government's assessment of the document's accuracy.[39] As a consequence, every evidence is validated using the verifier's evaluation of said defendant's reliability.

Leveraging blockchain technology, such as Tykn's Self-Sovereign Identification system offers, builds trust between the two parties and ensures the veracity of the data and testimony without storing any personal record of transactions. It is crucial since a decentralized system should never have contained any personal details as it is permanent, meaning that whatever is committed to it wouldn't be amended or erased.

4.4.11. Blockchain Data Preservation in Identity Management Systems

The document primarily retains references along with supporting affirmation of a participant's validated credentials. The techniques of pseudonymization as well as negative correlation can be utilized to safeguard confidentiality.[39] In order to avoid retaining actual confidential details, the ledger simply retains the following details:

- **Public Decentralized Identifiers:** This keeps access points as well as Identifying information along with verifying credentials. DIDs are the modern type of verification identifier which can only be used by the person whose identifier it really is. Decentralized identifiers are independent of centralized authority, directories, and identity providers.
- **Blueprint:** The comprehensive explanation of the credential's schema.
- **Authorization Detail Definitions:** Adminstrations issues various (usually actual) documents that serve as identity proof or eligibility, including credit and debit cards, identity documents, driving licences, and travel documents. As the name suggests, identity descriptions are simply the descriptions of the different credentials that are being recorded on ledger.
- **Revocation Registers:** Producer would be given the chance to retract the assurance. Rescinding information will be published by the supplier and will be announced in the revoking registry.
- **Data Exchange Contract Documentation:** Such authorization documents, which serve as documentation of consent, enable individuals to reach consent or get facts (simply stating that the information was being collected and checked).

4.4.12. Risks to Blockchain's Security in Identity Management

Blockchain is prone to several kinds of privacy issues. The blockchain system can be compromised, but it is vital to remember that now the technology is not yet impenetrable. Blockchain technologies have limitations such as

4.4.12.1. Phishing Attack

Phishing is the effort to obtain a participant's login information through fraud. Emails that appear to originate from a reliable source are sent to users of accounts by scammers. By means of fraudulent Web links, these emails request the user's login information. Consumers as well as the blockchain network could suffer damages if their login details or any other confidential material is compromised.

4.4.12.2. 51% Attacks

Whenever a member holds over 51% of such a system, that individual possesses a significant likelihood of altering the blockchain besides being caught because they have higher authority over the network and consequently more influence over the agreement. Smaller networks also become more vulnerable to assaults since, during the initial phases of the blockchain, an individual member might have substantially greater influence.[40]

4.4.12.3. Unintentional Centralization

It's indeed hard to prevent the weakest players from shifting commodities toward a centralized exchange network because protection on the blockchain is partially determined by the members involved. These frequently happen as a consequence of a 3rd party amassing substantial assets and preserving those on the user's behalf.[40]

4.4.12.4. Lack of Confidentiality as a result of the user's Pseudonym

It is essential to decide which data should be kept confidential and which should be publicly disclosed.[40] Additionally, it is hard to maintain complete security because it is feasible to associate some users' transactional behaviors with specific data.

4.4.12.5. Routing Attack

Blockchain systems depend on real-time, massive data exchanges. Information that is being transferred to ISPs can be intercepted by attackers. In this type of attack, blockchain users usually are blind to the danger, making everything seem to be normal. However, attackers have obtained users' personal access information behind the scenes.[40]

4.4.13. Identity Management with Blockchain: Preventing Fraud and Vandalism

Every user's digital identity authentication tokens are maintained on a biometric authentication ledger, such Tykn's SSI Digital Wallet, on his smartphone using authentication mechanism with blockchain system.

What would eventuate if a smartphone was taken or misplaced? There seem to be two stages to be undertaken, according to Sovrin. Removing the smartphone's login information is the initial step. Login details for an online identity must be utilized on a system which has been given authorization for this purpose. In the event that even a user's smartphone is stolen or broken, the member may use any

permitted machine, like a laptop, to record that now the permission for that device has now been withdrawn. This could occur right away and stop anyone from accessing the smartphone's digital identity information. Even with the passcode, identity, or smartphone, the burglar could not be able to impersonate her because the blockchain will store a secured and irreversible revoking record for that particular device. Whereas if device's authorisation is cancelled, the intruder cannot attempt to pose as the victim and make connections. The stage two stops the thief from looking into the device's present associations with those other groups. Therefore, the following action is to revoke the current relationship privileges. With the assistance of the two procedures, it's indeed difficult for such an unauthorized individual to be using online identity password to access online platforms or investigate connections, yet yet enabling individuals to keep utilizing their identities on different devices.

In several cases, an individual is required to personally visit the district or administrative agency, invalidate the identity, then begin fresh when they wished to remove a lost id card. Regardless of how much time elapsed, a fraudster could still access your data. Whenever a credit card is lost or stolen, the user should notify the institution (this requires time), and he will be unable to utilize the card till a replacement is produced and shipped to them.

4.4.14. Analysis of different Identity Management blockchain Techniques

A large number of people, businesses, and governments have begun some studies and experiments on blockchain technology since it is a new method with endless prospects. Although it could take several years for blockchains to completely develop, several applications based on the blockchain and services are now entirely doable in the coming years, & new possibilities will keep emerging as fundamental technology advances. Various identity and access management methods are devised as blockchain technology develops to strengthen the decentralized feature.[37] Blockchain options for identity and access management and verification are discussed in this chapter.

4.4.14.1. Sovrin

The decentralized, worldwide public asset for self-sovereign identification is called Sovrin Foundation. This is also the first globally accessible public asset that only supports independently validating arguments as well as self-sovereign identities.[37] Hyper-ledger Indy Project & open standards forms the foundation of a Sovrin system. Sovrin identities are written under a false name by default. Pairwise pseudonymous IDs, or a different Distributed Identifier for each connection, are the answer.[37]

4.4.14.2. ShoCard

Shocard is a commercial cellular identification system that safeguards customer privacy. Clients' identification cards are scanned and verified whenever they generate a ShoCard ID using App or SDK.[37] Shocard Identification System is based on the open blockchain architecture like BlockCypher, and any potentially vulnerable information or credentials remains off-ledger.

4.4.14.3. uPort

This self-sovereign identity scheme is reliable. This intends to serve as the decentralized site's public identification method.[42] This runs on the Ethereum blockchain that allows people to manage personal keys and information securely and also transmit and seek credentials, electronically sign transactions, and manage the credentials and information. Instances of engagements made possible by uPort involve private communications sent to other uPort users or apps in to blockchain operations like purchasing stocks on the Gnosis prediction market. The Identification and Claimant Protocols are the two protocols used by uPort.[37]

4.4.14.4. MyData

MyData is a study for managing private details that the Finnish government ordered. The idea of human-focused control, usefulness, access, and transparency is what motivates this Nordic form of self-sovereign identification.[37] Information exchange among industries including governmental, healthcare, and financial may be secured with MyData. Human-controlled access, OpenID sole control, and Oauth 2.0, that regulates web API access, are the three main components of MyData auth.

4.4.14.5. Waypoint

On the Ethereum VMs, Waypoint is really a distributed multi-factor authentication mechanism. Through the use of a Web Service, such a solution enables identification authentication to be done using blockchain. Waypoint enables applications to protect numerous modules within another item by specifying various functions, having both a mobile app and a desktop app accessible.[37]

4.4.14.6. Bloom

Bloom is an Ethereum and IPFS-based blockchain platform for access control as well as creditworthiness. This is a comprehensive system that enables owners of both conventional and digital currencies to act as financiers to customers that fail to establish a bank account or earn a credit record.[37]

4.4.14.7. BlockStack

Developers may create serverless applications using JavaScript libraries without worrying over managing architecture as it offers decentralized capabilities for name, identification, verification, and storage. Modern client/server architecture would be replaced by Blockstack, where individuals own their data, apps operate mostly on the user, while backend capability is replaced by the public Blockstack network.[37]

4.4.14.8. I/O Digital

Decentralized I/O Name Server, an enhanced blockchain, is used to handle identities, and Proof of Stake is used to protect them. Information storage is made possible through the DIONS blockchain, which also has documentation and identification storage functionality. Additionally, DIONS offers a comprehensive Aliases mechanism alongwith AES 256 cryptography communication encryption.[37] Members of the Alias platform can establish a public (unencrypted) alias, a private (encrypted) alias, or even both, giving them the option to retain important identification information as well maintain respect and govern their content.

4.4.14.9. Jolocom

It seeks to create a method for supplying a decentralized identity based on keys with a multilevel deterministic structure. This solution makes it simple to maintain several identities and maintain paired privacy in conversations that are tailored to a particular situation.[37] In addition, Jolocom enables IoT system ownership modeling for combined human and automated identification.

5. Conclusions and Recommendations

5.1. Conclusion

Cloud platforms are vulnerable to several kinds of problems from a range of sources, notably malware, bot, and DDoS assaults. The most frequent kinds of cloud breakdowns include virtual machine faults, application failures, and infrastructure issues. It's important to precisely predict or identify potential problems before implementing a successful strategy to remedy them. Therefore, solutions that could automatically monitor the anomalous behavior of the system should be created.

Throughout this paper, we addressed some key issues relating to access control and blockchain's relevance to the same. Blockchain technology is essential to tackling several problems in IDMs. It's really crucial to make such blockchain technologies interoperable among sectors since doing so will address the decentralization issue that is the core of this technology. Blockchain had also gained a great deal of traction within the identity and access management space as both a distributed and decentralized open ledger together in a peer-to-peer system. With just a typical runtime of under a second, such an infrastructure is incredibly effective.

Additionally, the top three identity management solutions that are based on blockchain technology have been covered: uPort, Sovrin, and ShoCard. There are benefits and drawbacks to each technique. The most noteworthy ones include more authority regarding identification methods based on a platform's self-sovereign identification, a decentralized identifier owing to blockchain technology, and quicker multiple entity confirmation. Moreover, there exists a glaring lack of situational knowledge when it comes to the customer journey. To ensure anonymity while using blockchain technology, it is necessary to give an even more standard view of identity and access management.

Below following is a list of prerequisites enabling cloud-based access control:

- **Effective Validation:** Multi-factor verification procedures have to be available in order for the specified personally identifiable information to be validated. Using these strong authentications are necessary for authenticating and assessing customer identities.

- **Avoidance of Information Loss:** Tracking, safeguarding, and confirming sensitive data will guarantee its integrity when it is idle, in route, and being used in the clouds.

- **Software as a Service:** To accomplish a formalized security architecture, a declaration that details which data protection services are offered and where, and also to accomplish a formalized security

architecture, a document that details all necessary components of a cloud service agreement among suppliers and users.

Deep learning may assist in the development of technologies that will improve these real-time operations and give system managers crucial data. Additional research could suggest placing the timestamp beside hashing algorithms within the Merkle tree as an enhancement. Furthermore, the distribution of different blockchains in a real blockchain network would never be uniform. Therefore, investigating how system architecture influences the efficiency of consensus mechanisms might be an interesting direction to take.

5.2. Digital Identity Future

Identification, verification, and validation are three tasks that are coordinated by digital identity systems. How or when to handle those tasks in a decentralized environment without widely accepted authority is the main issue to be solved. Our recommendation is to employ a multiple stakeholder approach to establish standard guidelines that provide a set of guidelines for engagement, as opposed to attempting to compel every user to use a particular new technology or system. Without needing approval from a certain institution or committing to the usage of a specific system, each institution will be allowed to create and employ its own solutions which might work with those created by any other institution.

Today, there are several identity management implementations and designs to carry out the three unique tasks we listed earlier: identification, validation, and verification. We propose a fourth component, auditing, which enables the system's judgments to be justified and assessed. In the end, if an identity management system enables or subjugates its consumers depends on the administration of the system, including its inherent principles and methods as well as the responsibility of the people and organizations who regulate its functioning. To prevent unexpected outcomes from the system's deployment, we contend that good governance, especially a cohesive strategy towards the technology and regulations that it includes, is crucial.

We suggest leveraging distributed ledger technology (DLT) to spread verification and certificate providers while preventing market concentration. This distributed ledger will act like a common method for certificate providers to connect with any or all the verification providers simultaneously, or the other way around. The ledger on its own would serve as a distribution mechanism for signs and expulsions; it should be distributed by users and not fall under the exclusive control of any one entity.

Future work could entail developing a POC implementation, evaluating the many execution tradeoffs pertinent to different use cases, and doing a proper study of the data security features of the system

created using this methodology. We propose that various use situations would need noticeably various design decisions.

5.3. Future Developments in Cloud Computing

One of the technology sector supporting categories with the highest growth is cloud computing. It provides enormous data throughput and retention along with adaptable, configurable computational power to accommodate various needs and supplies across the internet. Customers of the cloud dodge high initial fixed expenditures. Cloud service providers must prevent security flaws involving the customer information they store on behalf of individuals and organizations. However, before installing their system on the cloud, businesses and people must weigh the advantages and disadvantages of their operations. We discussed cloud computing and its benefits in report. We concentrated on several security issues that have to be appropriately handled and controlled, as well as the need for cloud security.

As more businesses use cloud services and invest in R&D to address its shortcomings, the future appears to be less cloudy. Being capable of working with cloud infrastructure, new security measures and guidelines must be created, and previous ones should be totally rewritten.

- **Multi-Cloud and Hybrid Solutions -** Everything that is having mixed heritage or makeup is referred to as "hybrid." This means that it refers to any item that is made up of a range of various components. The utilization of several cloud-based services is referred to by the simple phrase "multi-cloud." Businesses desire to be able to upgrade, run, and administer their applications and data securely on the clouds of their choosing without worrying about being confined. Organizations may host their local software someday, move to a cloud service the very next, and keep the option to transfer to a cloud platform in the long run with this solution.

- **Green Cloud Computing -** Data centres cost a lot to operate while having a negative influence on the environment. Energy expenses and greenhouse gas emissions are high because these storage systems require large quantities of power to operate and chill the many servers housed there. To prevent their profit margins from being significantly reduced by growing energy prices, cloud service providers must take steps.

5.4. Blockchain Systems' Future

Self-service and hardware virtualization can help blockchain continue to fulfill a variety of purposes for businesses. Blockchain technology enables businesses to increase their operations swiftly while retaining

service or product standardization and top quality. Furthermore, the robust service characteristics of block chain technology can offer businesses a favorable environment for production and service improvement. Continued development will increase the innovation and agility of businesses. The enterprise issues and challenges and prospective research directions of blockchain are:

1. Integration of blockchain with multiple industrial sectors.
2. Strategy incorporation and strategy.
3. Blockchain technology in new businesses
4. Blockchain-based and industry structure
5. C2C, B2C, and B2B interactions and engagements
6. Collaboration and cross-organizational governance
7. Low-carbon economies & blockchain

5.5. How to combat fraud and identity theft using AI

Organizations can monitor every detail at all moments by employing AI, as a computer is able to pick up on details that a human cannot. With information on the complexity of network interaction, IT firms are enabled to carry out wise administration operations while also making smarter user license decisions. When temporary privileges are allowed, role-based permission could be enhanced into a more complex solution with improved privilege access controls and a reduced risk of access control abuse. As business systems become connected, the demand for smooth, consistent, and precise data accessibility is even more crucial. Consequently, advanced authentication mechanisms using AI would be crucial, especially if data is gathered and evaluated much more quickly than by humans. AI systems would continually monitor individuals as they move around the system within the bounds of their access permissions, but they can also spot any unusual, contradictory, or inconsistent behavior. Analysts could determine whether consumers would seldom access a certain area of the network or download more data than they normally do.

5.6. Combating Identity Theft Using Machine Learning

Cyber assaults represent an increasing threat to user privacy details kept by businesses and governmental organizations every year. Every year, data breaches jeopardize the personal information of tens of

thousands of individuals and violate millions of customers' data. Monetarily, in line with data protection and credibility, as well as for people (via fraud and identity fraud), and corporations, these compromises are quite expensive. Individuals and companies must be conscious of current developments and trends, including prevalent security risks and techniques, the data categories that are heavily targeted, and the private details that are most commonly disclosed with the worst possible outcomes. Therefore, an ML-based system called PDEWS may be utilized to detect identity theft that is now taking place all over the world.

Personal Data Early Warning System (PDEWS) is an interactive portal that records and shows the present state of cyber threats and produces useful information about trends and patterns. PDEWS is an autonomous system that gathers information daily on current cybersecurity threats. PDEWS is broken down into four main stages. PDEWS first combs across daily media reports on identity fraud and identity theft and extracts the core text. The text is then put into an AWS cloud infrastructure after being formatted into the structure needed for a machine learning system. To extract pertinent risk indicators, PDEWS then uses machine learning techniques that have been developed on a proprietary corpus of identity theft articles. Lastly, PDEWS presents those patterns over an application interface along with suggestions that have been shown to have the best threat-mitigation potential. This PDEWS platform is a cutting-edge method for tracking cyber threat patterns through web articles and providing prevention and mitigation advice based on industry best practices.

6. References

1. El Haddouti, Samia, and M. Dafir Ech-Cherif El Kettani. "Analysis of identity management systems using blockchain technology." 2019 International Conference on Advanced Communication Technologies and Networking (CommNet). IEEE, 2019.

2. Yang Liu, Debiao He, Mohammad S. Obaidat, Neeraj Kumar, Muhammad Khurram Khan, Kim-Kwang Raymond Choo, Blockchain-based identity management systems: A review, Journal of Network and Computer Applications, Volume 166,2020,102731,ISSN 1084-8045

3. Keltoum Bendiab; Nicholas Kolokotronis; Stavros Shiaeles;Samia Boucherkha; “WiP: A Novel Blockchain-based Trust Model for Cloud Identity Management”, 2018 IEEE 16th Int. Conf. on Dependable, Autonomic & Secure Comp.

4. Panait Drăgnoiu, Andreea & Olimid, Ruxandra & Stefanescu, Alin. (2020). Identity Management on Blockchain - Privacy and Security Aspects. Proceedings of the Romanian Academy - Series A: Mathematics, Physics, Technical Sciences, Information Science. 21. 45-52.

5. Lo, Sin Kuang, et al. "Evaluating suitability of applying blockchain." 2017 22nd International Conference on Engineering of Complex Computer Systems (ICECCS). IEEE, 2017.

6. Naik, Nitin, and Paul Jenkins. "Governing principles of self-sovereign identity applied to blockchain enabled privacy preserving identity management systems." 2020 IEEE International Symposium on Systems Engineering (ISSE). IEEE, 2020.

7. Mohammed, Ishaq Azhar. "A systematic literature mapping on secure identity management using blockchain technology." International Journal of Innovations in Engineering Research and Technology 6.5 (2019): 86-91.

8. Agrawal, Tarun Kumar, et al. "Blockchain-based framework for supply chain traceability: A case example of textile and clothing industry." Computers & industrial engineering 154 (2021): 107130.

9. Qian, Ling, et al. "Cloud computing: An overview." IEEE international conference on cloud computing. Springer, Berlin, Heidelberg, 2009.

10. Singh, Harinder & Kumar, T.Durga & Bhasha, Shaik. (2010). CLOUD COMPUTING. International Journal of Computer and Communication Technology. 271-276. 10.47893/IJCCT.2010.1056.

11. Wesley Chai, Kathleen Casey. Software as a Service (SaaS)

12. Microsoft Azure. Serverless Computing - An Introduction to serverless technologies.

13. Oracle.com. The top 10 benefits of cloud computing topics

14. Gopalakrishnan A: Cloud computing identity management. SETLabs Briefings 2009, 7:45–54.

15. Meier JD, Farre C, Taylor J, Bansode P, Gregersen S, Sundararajan M, Boucher R: Improving Web Services Security: Scenarios and Implementation Guidance for WCF: Microsoft Developer Network; 2009

16. Slone S: Identity management. A white paper: The open group identity management work area; 2004.

17. Lim, Shu Yun, et al. "Blockchain technology the identity management and authentication service disruptor: a survey." International Journal on Advanced Science, Engineering and Information Technology 8.4-2 (2018): 1735-1745.

18. Naik, Nitin, and Paul Jenkins. "uPort open-source identity management system: An assessment of self-sovereign identity and user-centric data platform built on blockchain." 2020 IEEE International Symposium on Systems Engineering (ISSE). IEEE, 2020.

19. Alrodhan WA, Mitchell CJ: Enhancing user authentication in claim-based identity management. In CollaborativeTechnologies and Systems (CTS), 2010 International Symposium on. Piscataway, New Jersey, United States: IEEE;2010:75–83.

20. Cao Y, Yang L: A survey of identity management technology. In IEEE International Conference on Information Theory and Information Security (ICITIS): IEEE; 2010:287–293.

21. øsang A, Fabre J, Hay B, Dalziel J, Pope S: Trust requirements in identity management. In Proceedings of the 2005: Australasian workshop on Grid computing and e-research-Volume 44: Australian Computer Society, Inc.; 2005:99–108

22. Suriadi S, Foo E, Jøsang A: A user-centric federated single sign-on system. J Netw Comput Appl 2009, 32:388–401.Elsevier, 2009.

23. Conrado C, Kamperman F, Schrijen GJ, Jonker W: Privacy in an identity-based DRM system. In the Proceedings of Database and Expert Systems Applications, 2003. Proceedings. 14th International Workshop on. Piscataway, New Jersey, United States: IEEE; 2003:389–395.

24. McCallister E: Guide to Protecting the Confidentiality of Personally Identifiable Information. Collingdale, PA, United States: Diane Publishing; 2010.

25. Yubico. What is Certificate-Based Authentication? 2021.

26. Wizbuzzer. What Are the Pros and Cons of Multi-Factor Authentication? 2021.

27. Govind Malviya. What is a Token? What are its Pros and Cons? 2021. LoginRadius

28. Thesis/Dissertation: Bc. Martin Drašar, učo 98998: Password Based Authentication. 27/5/2009 16:36, Ing. Mgr. et Mgr. Zdeněk Říha, Ph.D.

29. Kormann, D.P. & Rubin, A.D (2000), Risks of the Passport Single SignOn Protocol, in: Computer Networks (pp. 51-58), Volume 33, Issues 1-6

30. Hogben, G (2007), Security Issues and Recommendations for Online Social Networks, ENISA Position Paper 1, Retrieved August 23rd, 2010

31. Privacy in Identity and Access Management Systems; 2011; Andreas Pashalidis; Chris J Mitchell

32. H. Lee, I. Jeun and H. Jung, "Criteria for Evaluating the Privacy Protection Level of Identity Management Services," 2009 Third International Conference on Emerging Security Information, Systems and Technologies, 2009, pp. 155-160, doi: 10.1109/SECURWARE.2009.

33. Audun Jøsang, Muhammed Al Zomai, and Suriadi Suriadi. 2007. Usability and privacy in identity management architectures. In Proceedings of the fifth Australasian symposium on ACSW frontiers

34. Rozanna Latiff, Jeremy Wagstaff. Malaysia investigating reported leak of 46 million mobile users' data. Media and Telecoms. Reuters. 2017

35. Nahar, Nazmun, Farah Hasin, and Kazi Abu Taher. "Application of Blockchain for the Security of Decentralized Cloud Computing." 2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD). IEEE, 2021.

36. Hammi, Mohamed Tahar, Patrick Bellot, and Ahmed Serhrouchni. "BCTrust: A decentralized authentication blockchain-based mechanism." 2018 IEEE wireless communications and networking conference (WCNC). IEEE, 2018.

37. Lim, Shu Yun & Fotsing, Pascal & Almasri, Abdullah & Musa, Omar & Mat Kiah, Miss Laiha & Ang, Tan & Ismail, Reza. (2018). Blockchain Technology the Identity Management and Authentication Service Disruptor: A Survey. International Journal on Advanced Science, Engineering and Information Technology. 8. 1735. 10.18517/ijaseit.8.4-2.6838.

38. Naik, Nitin, and Paul Jenkins. "uPort open-source identity management system: An assessment of self-sovereign identity and user-centric data platform built on blockchain." 2020 IEEE International Symposium on Systems Engineering (ISSE). IEEE, 2020.

39. Davies, A. (2021). 6 benefits of blockchain identity management

40. Nawari O. Nawari, Shriraam Ravindran, Blockchain and the built environment: Potentials and limitations, Journal of Building Engineering, Volume 25, 2019, 100832, ISSN 2352-7102.