# Blockchain Models that Enable Users to Retain Data Ownership:

# A Systematic Literature Review

Garreth Curran

M.Sc. in Management of Information Systems
2019

# Blockchain Models that Enable Users to Retain Data Ownership: A Systematic Literature Review

## Abstract

The disruptive technology the blockchain is in its' technological infancy and new implementations are proposed daily. Enabling individuals to control data about themselves is a challenge that may benefit from a system that integrates with blockchain technology to create personal data stores and personal data management tools. This dissertation conducts a systematic literature review into models that propose leveraging blockchain technology that provides individuals with a personal data store, to achieve a well-founded understanding of the research literature available. The review applies the eight-step guide as set out by Okoli and Schabram (2010). The review is conducted with rigour, and every step is reproducible as per the goals of systematic literature reviews. This dissertation aims to create a solid starting point for academics interested in further research on the topic of individuals managing and owning their personal data.

## Acknowledgements

I would like to thank my supervisors Dr Aideen Keaney and Diana Wilson for their time, direction and experience. Without them, this dissertation would still be at the idea stage.

I would like to thank my friends, colleagues and fellow students for their encouragement and guidance.

I would like to thank my employer Infosys Limited, especially my team for enabling me to achieve my academic goals.

Finally, I would like to thank my wife Helen and my son Cuán for their continuous support, patience and encouragement during my academic journey.

## Declaration

I declare that the work described in this dissertation is, except where otherwise stated, entirely my own work, and has not been submitted as an exercise for a degree at this or any other university. I further declare that this research has been carried out in full compliance with the ethical research requirements of the School of Computer Science and Statistics.

Signed: Garreth Curran
25 April 2019

## Permission to lend and/or copy

I agree that the School of Computer Science and Statistics, Trinity College may lend or copy this dissertation upon request.

Signed: Garreth Curran
25 April 2019

## Table of Contents

## List of Figures

## List of Tables

## List of Abbreviations

| | |
|---|---|
| AmI | Ambient Intelligence |
| ANAC | AutoNomy-based Access Structure |
| BC-PDS | BlockChain-based Personal Data Store |
| CJEU | Court of Justice of the European Union |
| DIM | Decentralised Identity Manager |
| ECtHR | European Court of Human Rights |
| GSS | General Secret Sharing Scheme |
| IPFS | InterPlanetary File System |
| JSON-LD | JavaScript Object Notation for Linked Data |
| LISTA | Library, Information Science & Technology Abstracts database in EBSCOhost repository |
| openPDS | Open Personal Data Store framework |
| PDM | Personal Data Management |
| PDS | Personal Data Store |
| QA | Quality Appraisal |
| Rpf | Reputation Fluctuation |
| TSS | Threshold Secret Sharing |
| TURS | Tsinghua University User Reputation System |
| UINP | Unitary Interchain Network Protocol |

# 1. Introduction

While there may be limited trust with the initial service or internet-connected device manufacturer that users provide data to (Golbeck, 2009; Christidis and Devetsikiotis, 2016; Porambage et al., 2016), the users continue to permit the collection of personal data and may not be aware as to how this data will be used or that it may transfer to third parties (Moor, 1997). The lack of transparency is also an issue (Christidis and Devetsikiotis, 2016), and can lead to significant data breaches such as the recent Cambridge Analytica scandal on Facebook where 50 million user profiles had personal data harvested (Cadwalladr and Graham-Harrison, 2018). Therefore, in the age of information, where people share personal data with trusted and untrusted third-party services and internet-connected devices, the protection of this information is essential.

The two highest courts in Europe, the European Court of Human Rights and the Court of Justice of the European Union "treat data protection as an expression of the right to privacy" (Kokott and Sobotta, 2013). The European Convention on Human Rights and the Charter of Fundamental Rights have specific privacy clauses. Article 8 of the European Convention on Human Rights, and Article 7 of the Charter of Fundamental Rights "provide that everyone has the right to respect for his or her private and family life, home, and communications" (Kokott and Sobotta, 2013). Privacy is not a simple concept and differs depending on a number of factors including an individuals' culture (Moor, 1997; Moore (1984) as cited in Acquisti, Brandimarte and Loewenstein, 2015), behavioural intentions (Stewart and Segars, 2002), and confusion caused by services privacy settings (Acquisti, Brandimarte and Loewenstein, 2015). Therefore, as an expression of privacy, information privacy is also complicated.

Blockchain technology, also referred to as 'the blockchain' and 'blockchain' is a disruptive technology, and the current rapid uptake within corporations will displace some traditional systems and technologies (Tapscott, 2018). A blockchain is a peer to peer network of distributed ledgers used for recording transactions efficiently, verifiably and permanently between two parties that may be unknown to each other (Iansiti and Lakhani, 2017). Therefore, blockchain is distributed with no single databases that could be hacked, publicly viewable by anyone, and it is encrypted for privacy (Tapscott, 2018). In 1994, Nick Szabo proposed a new technological concept called "smart contracts" (Szabo, 1994). Smart contracts are self-executing computer-based transactions. The creation of blockchain technology in 2009 with the launch of Bitcoin (Barber et al. 2012), led to the development of the Ethereum blockchain in 2015 (Tual, 2015) enabled the implementation of Szabo's concept.

Through the use of smart contracts on blockchain technology, it may be possible to protect the privacy of individuals data by implementing a personal data store (PDS) combined with applications that provide personal data management (PDM). The PDS and PDM application may leverage the pervasive properties offered by distributed technology such as the blockchain (Alessi et al., 2018) and provide users control of what personal information they share, when they share it, and for how long when interacting with services and connected devices. The rules surrounding the sharing of data may utilise the smart contracts as theorised by Szabo (1994).

Due to the limited trust while interacting with services and devices (Golbeck, 2009; Porambage et al., 2016), the attitude and behaviours of people (Moore (1984), as cited in Acquisti, Brandimarte and Loewenstein (2015)), the example of harvesting personal data by Cambridge Analytica (Cadwalladr and Graham-Harrison, 2018), combined the suggestion that blockchain may assist in the creation of personal data stores and management solutions (Alessi et al., 2018), curiosity was piqued in these subjects. It was of personal interest and considered a benefit to other academics, to identify potential blockchain technology models that aim resolve existing failures that may lead to the theft or misuse of personal information caused by some interactions with internet connected services and devices. As a result, in preparing for this dissertation, no evidence was found to demonstrate that a systematic literature review has ever been conducted to identify proposed models that suggest blockchain technology can provide a complete solution or contribute to a solution that ensures the secure transfer of private information between users and third parties.

## 1.1. The Contribution of this Dissertation

Information Privacy is a complicated issue as there are various values, beliefs, cultures and even confusion surrounding the relationship people have with it (Moore, 1984, as cited in Acquisti, Brandimarte and Loewenstein 2015; Moor, 1997; Post, 2001). Privacy is subjective and can depend on the context. Defining privacy is thus difficult, the definition of information privacy used in this dissertation is that proposed by Clarke (1997), "Information Privacy is the interest an individual has in controlling, or at least significantly influencing, the handling of data about themselves" is used to normalise information privacy.

A systematic literature review is a "form of secondary study" of individual "primary studies" (Ryan, 2010, pp 1). Okoli and Schabram (2010, pp 1) put forward that a systematic literature review "constitutes an original and valuable work of research in and of itself", and "rather than providing a base for the researcher's own endeavours, it creates a solid starting point for all other members of the academic community interested in a particular topic." As a result, they developed an eight-step guide to conducting systematic literature reviews for information system research (figure 1). They based the guide on Kitchenham's and Charters three phases of a systematic literature review combined with their adaptation of Finks (2005) definition of a research literature review; "a systematic, explicit, [comprehensive, (2007, p. 17)] and reproducible method for identifying, evaluating, and synthesizing the existing body of completed and recorded work produced by researchers, scholars, and practitioners" (Okoli and Schabram, 2010, pp 4). The eight-steps are applied in this dissertation and described in chapter 2.

As Clarke's (1997) definition demonstrates, it is important that individuals have control or influence over the use of their data and with blockchain being a new field of research (Zhao, Fan and Yan, 2016), this dissertation has identified a gap in the existing research literature. To date, no systematic literature review has been conducted on research papers that propose models that incorporate blockchain technology to enable individuals to manage their personal information and the transfer of the data to third parties in a secure method. Therefore, this dissertation intends to fill the research gap and create a "solid starting point" as referred to by Okoli and Schabram (2010, pp 4) for use by other academics interested in the provision of personal privacy models that use or incorporate blockchain technology.

## 1.2. Research Questions

The research objective of conducting a rigorous systematic literature review as per section 2.1.1. Purpose of the Systematic Literature Reviewed translates into the following research question:
**RQ:** What generic models are proposed to leverage blockchain technology to provide individuals with a personal data store?

## 1.3. Timeframe

This process of conducting the necessary research, the systematic literature review and the writing of this dissertation took place between December 2018 and April 2019. Chapter 1, Introduction was completed in January 2019 with periodical reviews after that. The Planning phase in chapter 2 occurred from December 2018 to February 2019, the Selection phase in chapter 3 during March 2019, and the Extraction and Execution phases in chapter 3 and chapter 4 were carried out between March and April 2019.

## 1.4. Structure

This dissertation is organised as follows:

- **Chapter 1 - Introduction:** This chapter provides the background and context on which the systematic literature review is based and the research questions that arose from the systematic literature review planning stage (section 2.1.). Also provides is the contribution that this dissertation provides to academia.
- **Chapter 2 - Methodology and Fieldwork:** This chapter outlines the methodology applied to identify the Data Privacy Management models that will be subject to review intros dissertation. It provides the purpose of the systematic literature review, the protocol used for the literature search, and the selection and the screening process used to select relevant research literature.
- **Chapter 3 - Extraction:** This chapter defines the quality appraisal process to ensure that only higher quality research literature is selected. The extraction of pertinent data is conducted on all the remaining high-quality papers.
- **Chapter 4 - Execution:** This chapter analyses and evaluates each model outputted during the systematic review process through the application of specific criteria i. e. adherence to Clarke's information privacy definition, the security of data, and if it is trustless.
- **Chapter 5 - Conclusion:** The conclusion of the dissertation is contained in this chapter. The conclusion summarises the research conducted, offers recommendations for further research, and highlights the limitations of the research methodology,

## 2. Methodology and Fieldwork

The research methodology approach in this dissertation is a systematic literature review (also referred to as a 'systematic review'). Systematic literature reviews were introduced in the 1970s and "examined the effectiveness of a healthcare intervention" (Ridley, 2012, pp 188-189). In the 1990s, these systematic literature reviews became more common in general medical research, and standards have been well documented (Babar and Zhang, 2009). They have since extended into other fields of research (Ridley, 2012). The expansion arose for the "recognition of the importance of evidence based practice to inform policy decisions and professional practice" (Ridley, 2012, pp 189). Systematic reviews strive "to comprehensively identify, appraise, and synthesize all the relevant studies on a given topic" (Petticrew and Roberts, 2006, pp 19).

Okoli and Schabram (2010, pp 2) identified three common types of literature reviews; "theoretical background", "literature review" and a "stand-alone literature review". The theoretical background literature review is a section within a journal artifice that sets the context and provides a theoretical foundation to aid the focusing of the research question. The "literature review" is identified as a graduate thesis chapter, and a "stand-alone literature review" is a journal article that reviews literature in a given area without collating or analysing any primary data (Okoli and Schabram, 2010, pp 2). However, they also specify that if "a stand-alone literature review is conducted using a systematic, rigorous standard", it is known as a systematic literature review (Okoli and Schabram, 2010, pp 2). A systematic literature review is "a means of identifying, evaluating and interpreting all available research relevant to a particular research question, or topic area, or phenomenon of interest. Individual studies contributing to a systematic review are called primary studies; a systematic review is a form of secondary study" (Kitchenham and Charters, 2007, pp 3). However, primary studies are used as source material (Ryan, 2010). Okoli and Schabram (2010, pp 4) suggest that a systematic literature review should be stand-alone and carried out with various degrees of rigour "ranging from little more than an annotated bibliography to scientifically rigorous syntheses of a body of primary research". They also affirm that a systematic literature review is "an original and valuable work of research", and that they assist other academics was they create "a solid starting point" for "other members of the academic community interested in a particular topic" (Okoli and Schabram, 2010, pp 1).

Fink (2005; 2014) refers to a systematic literature review as a “research literature review” and used four keywords to define it. A research review must be “systematic” by following a methodological approach, “explicit” in explaining all processes and procedures used while conducting the review, “comprehensive” in scope in order to include all "the existing body of completed and recorded work produced by researchers, scholars, and practitioners", and “reproducible” by anyone who wishes to conduct the same review (Fink, 2005; 2014, pp 3).

Systematic reviews are required to understand large bodies of information (Ridley, 2012), to highlight spurious claims, create certainty by asking questions, eliminate intentional and unintentional bias in traditional reviews and to “help organize and prioritize the most relevant information” (Petticrew and Roberts, 2006, pp 9).

Okoli and Schabram (2010, pp 4) succinctly adapted Fink’s (2005) definition and described a systematic literature review as “a systematic, explicit, [comprehensive, (2007, p. 17)] and reproducible method for identifying, evaluating, and synthesizing the existing body of completed and recorded work produced by researchers, scholars, and practitioners.”

The three distinct systematic literature review phases as identified in the guidelines created by Kitchenham and Charters (2007, pp 6) are “planning the review”, "conducting the review” and “reporting the review”. Coalescing their definition of systematic reviews, and expanding on Kitchenham’s and Charter’s three phases, Okoli and Schabram (2010, pp 6-7) created an eight-step guide (figure 1) tailored to conducting a systematic literature review specifically for use in information systems research as follows:

1. *Purpose of the literature review:* Ascertain and identify the purpose and aim of the review. The importance of “being clear about the purpose” (Okoli and Schabram, 2010, pp 14) is important to the systematic literature review.
2. *Protocol and training:* Outline a precise procedure for conducting the review when multiple people are involved in conducting the research required for the review. Also included in this step is the production of a training document for each author to use and hence ensuring consistency.
3. *Searching for the literature:* An explicit description of the literature search strategy must be provided, explained and justified to assure that the search is comprehensive.
4. *Practical screen:* List all included and excluded research and the practical reasons for not including specific studies. A justification for the comprehensiveness of the review must be provided taking into consideration the “practical exclusion criteria”.

5. *Quality appraisal:* This screening for exclusion requires the reviewer to state the judging criteria for excluding low-quality studies explicitly, conversely included studies should be scored for quality.
6. *Data extraction:* After identifying all included studies, pertinent information must be extracted from them.
7. *Synthesis of studies:* Analyse the included studies using qualitative, quantitative or both qualitative and quantitative techniques.
8. *Writing the review:* The systematic review process should be explicitly reported in detail to enable the exact reproduction of the outcomes if the process is followed.

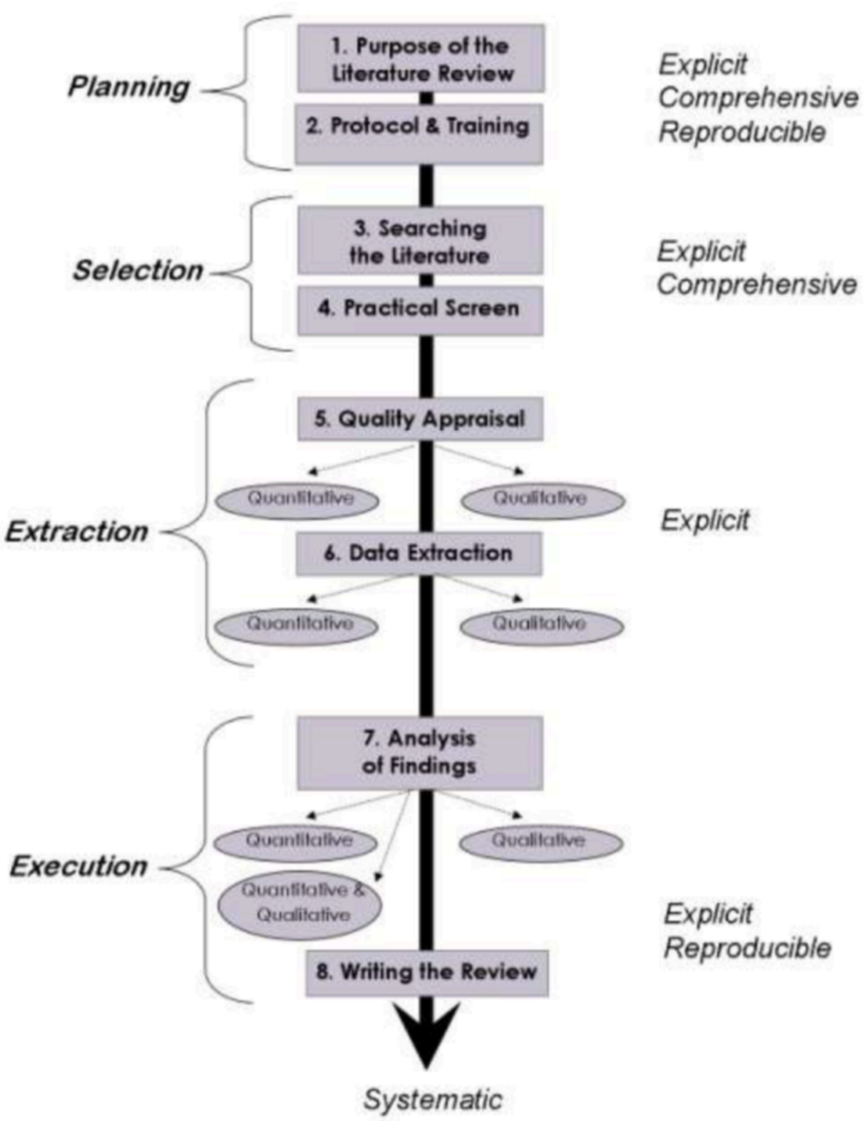

*Figure 1: A Systematic Eight-Step Guide to Literature Review Development (Okoli and Schabram, 2010)*

Systematic literature review - “a systematic, explicit, [comprehensive, (2007, p. 17)] and reproducible method for identifying, evaluating, and synthesizing the existing body of completed and recorded work produced by researchers, scholars, and practitioners.”

(Okoli and Schabram, 2010, pp 4)

*Figure 2: Systematic Literature Review Definition*

## 2.1. Planning

### *2.1.1. Purpose of the Systematic Literature Review*

Kokott and Sobotta (2013) investigated the “distinction between privacy and data protection in the jurisprudence” of the Europes two highest courts, the Court of Justice of the European Union (CJEU) and the European Court of Human Rights (ECtHR). They identified that the European Convention on Human Rights and the Charter of Fundamental Rights have provisions for privacy, Article 8 of the ECtHR European Convention on Human Rights and Article 7 of the CJEU Charter of Fundamental Rights “provide that everyone has the right to respect for his or her private and family life, home, and communications” (Kokott and Sobotta, 2013). They concluded that while not identical, privacy and the protection of personal data are closely connected as per ECtHR and CJEU jurisprudence. Also identified was the significant overlap of rights for both privacy and data protection and “the requirements that personal data must be processed fairly and for a specified purpose cover many instances where an interference with privacy would have to be justified” (Kokott and Sobotta, 2013).

Post (2001, pp 2087) cautions that privacy “is a value so complex, so entangled in competing and contradictory dimensions, so engorged with various and distinct meanings” and wonders if it is possible to address it effectively. He postulates that privacy has three concepts; the concepts connect privacy to the “creation of knowledge”, “dignity” and “freedom”. As a result, the value placed on privacy incorporating personal data is inconsistent amongst the world's population. There are two schools of thought on the value of privacy. On the one hand, defending privacy is required as it is critical to life, while on the other hand, it is also difficult to justify as it is "matter of individual preference" and "culturally relative” (Moor, 1997, pp 28). Therefore, to understand the importance of information privacy, it is paramount to have some knowledge of human attitudes and behaviours.

The risk of harm befalling individuals due to an invasion of privacy impelled Moor (1997) to take an ethical theory approach to privacy. In an information age where systems continuously collect, store, move and massage private data without the individuals' knowledge is known as "greased data” (Moor, 1997, pp 27). An example of greasing data is a supermarket loyalty card. During the purchase process, the loyalty card gets scanned, and the associated customer profile has detailed data added to it. The customer is often oblivious to the fact that the data moves, is stored indefinitely, and assists with such things as inventory management and targeted advertising.

There are two values for privacy and questions where privacy can be grounded inferred Moor (1997). The first is instrumental, and this indicates that it will result in a positive outcome for society through the security element. The second is intrinsic, and therefore it is valued by individuals for its own sake. Using these values emphasise that privacy can be grounded with either. However, Moor proposes that privacy is not a core value for society as it is not prevalent in all cultures. Although, security is a core value within all cultures and privacy is just one element used when expressing security. Therefore, Moor suggests that the traditional understanding of privacy as an instrumental or intrinsic value is flawed as it guides towards the instrumental value for “summum bonum” (Moor, 1997, pp 30). Moor (1997, pp 30) stated: “Privacy, as an expression of security, is a critical, interlocking member in our systems of values in our increasingly computerized culture”. Therefore, as a core value, the understanding that privacy as the expression of security for society, and privacy as a core value for an individual proves that privacy can be both instrumental and intrinsic in the information age.

Various papers highlight different aspects of the multifaceted relationship between humans and privacy. A study by Stewart and Segars (2002) used the ‘concern for information privacy instrument’ as put forward by Smith, Milberg and Burke (1996) to demonstrate that concern for information privacy is at the centre of the relationship between computer anxiety and behavioural intentions. Moore (1984), as cited in Acquisti, Brandimarte and Loewenstein (2015) suggest that there are significant cultural differences in behaviour and norms within public and private domains.

Acquisti, Brandimarte and Loewenstein (2015, pp 8) suggest that people face a “dilemma of what to share and what to keep private” due to the complexities of information systems and the available privacy settings. This quandary reinforces that trust is a critical issue as users interact with services and create more data (Golbeck, 2009). Golbeck (2009, pp 2) continues to state that there is an increase in risk due to dealing with “unreliable parties” and that is a “matter of opinion and perspective”. There is a lack of trust in the protection of information privacy when people interact with information systems (Golbeck, 2009).

As indicated by Acquisti, Brandimarte and Loewenstein (2015), the cultural differences, along with the inconsistent and intricate systems people interact with, means that many individuals require assistance in navigating through data privacy minefields. The difficulty individuals face is compounded with the distrust people have in services that they interact and share information with (Golbeck, 2009).

Clarke (1999, pp 60) affirms that people often think of privacy as “a moral right or a legal right” and “Information privacy refers to the claims of individuals that data about themselves should generally not be available to other individuals and organizations”. Clarke (1997) provided the following definition that will be used as a basis of this dissertation: “Information Privacy is the interest an individual has in controlling or at least significantly influencing, the handling of data about themselves”. The right of individuals to controlling their information provides the reasoning behind this thesis and leads into the investigation of potential solution opportunities provided by the blockchain, a technology in its infancy.

As an emerging technology, the digital framework known as the Blockchain has the means to revolutionise many fields (Zhao, Fan and Yan, 2016) including data privacy through methods of interaction with governments, companies, devices, and other individuals. The blockchain is considered a disruptive technology and should be considered as a component in information technology architecture that can operate as a software connector located between other systems (Xu et al., 2016). The current rapid uptake within corporations will displace some traditional systems and technologies.

“Blockchain is an incorruptible digital ledger of economic transactions that can be programmed to record not just financial transactions but virtually everything of value. Each unit of value is represented by a transaction recorded in a blockchain, which leverages the resources of a large peer-to-peer network to verify and approve each transaction” (Tapscott, 2018). To further this definition, Iansiti and Lakhani (2017), describe the blockchain as a public or private, peer to peer network of distributed ledgers that records transactions between two parties in an efficient, verifiable and permanent method.

The blockchain is constructed from a network of peers/nodes. Each node contains a full copy of the ledger, and once a transaction is verified and validated on one node, it propagates to all remaining nodes. Blockchain technology is in its' technological infancy; therefore it is still in the early adoption stage. Due to the decentralised nature where every node has a complete copy of the ledger, and the consensus protocol, it is an indestructible and immutable system for transferring value that is extremely difficult to alter. It requires 51% of the nodes to be compromised to enable the exploitation of a blockchain through a double spend fraud (Bae and Lim, 2018; Zhu et al., 2019). The value contained in blockchains can be any format of digital data, or monetary such as digital currencies and payment systems, e.g. Bitcoin.

Swan (2015) suggests that it is possible to segregate blockchain applications into three technological iterations. Blockchain v1.0 is for the most basic blockchain implementations used to operate cryptocurrencies such as Bitcoin and is already in a state of maturity (Zhao, Fan and Yan, 2016). The blockchain is relatively dumb, does not execute any procedures, and only provides recordkeeping functionality in a decentralised ledger. Blockchain v2.0 permits smart contracts to coexist with cryptocurrency, and once the smart contract is created, it will operate and execute specified terms when programmed conditions are satisfied without any human intervention required. Iteration v3.0 of blockchain technology is for applications that go further than financial and have not yet been conceived or executed.

Szabo (1994) initially conceptualised smart contracts in a short, unpublished manuscript titled “Smart contracts”. He describes smart contracts as “a computerized transaction protocol that executes the terms of a contract.” The premise is that intelligent decision-making contracts could be run on computer systems and require virtually no human interaction, therefore reducing errors and the need for trusted intermediaries. The purpose of smart contracts “are to satisfy common contractual conditions, (such as payment terms, liens, confidentiality, and even enforcement), minimize exceptions both malicious and accidental, and minimize the need for trusted intermediaries. Related economic goals include lowering fraud loss, arbitration and enforcement costs, and other transaction costs” (Szabo, 1994).

In the economics dictionary reference book, “The new Palgrave: allocation, information and markets”, Eatwell et al. (1989) as cited in Szabo (1994), state that related economic goals of “lowering fraud loss, arbitration and enforcement costs, and other transaction costs” would be created. Szabo proposed multiple protocols and subprotocols within the subject of ‘Digital Cash’ and suggested that emerging technologies during the 1990s could be used to facilitate these protocols. Szabo highlights an example where smart contracts can provide the transparency that is hidden from customers where personal data is collected during a point of sale transaction in a shop.

Fourteen years after Szabo proposed smart contracts, on 31 October 2008, an unknown person or group of people using the pseudo-name Satoshi Nakamoto published a paper detailing a proposed electronic currency called Bitcoin. Bitcoin relied on “the block chain”, a new technology that Nakamoto proposed (Nakamoto, 2008). Bitcoin and blockchain technology became a reality on 3 January 2009 when the genesis block was created (Barber et al., 2012) with the message “The Times 03/Jan/2009 Chancellor on brink of second bailout for banks” along with 50 Bitcoins (Ali, 2014, pp 267). The technological breakthrough made by Nakamoto was the creation of the blockchain framework, and it was the first step to the realisation of Szabo's smart contracts.

The innovative Ethereum blockchain was launched on 30 July 2015 (Tual, 2015) by Vitalik Buterin. Buterin’s white paper stated that “Ethereum intends to provide is a blockchain with a built-in fully fledged Turing-complete programming language that can be used to create ‘contracts’” (Buterin, 2014; pp 1). These contracts ”can be used to encode arbitrary state transition functions” (Buterin, 2014; pp 1) and so allows users to create systems described in the white paper, along with others that have yet to be envisaged. These systems would be created “simply by writing up the logic in a few lines of code” (Buterin, 2014, pp 1). Ethereum permits the executing of code written in a language called Solidity to successfully implement Szabo’s smart contact vision.

It was only through the development of blockchain v2.0 that smart contracts on the Ethereum blockchain became a reality. The Ethereum blockchain provides the technology that enabled the implementation of self-executable contracts. The third iteration of the blockchain, v3.0, is reserved for applications “beyond currency, finance, and markets” that have not yet been thought of (Swan, 2015). Smart contracts are a core element enabling advanced applications through interactions with and between blockchains. Enhancing data privacy through the use of blockchain technology is not possible without smart contracts.

Due to escalating global concerns surrounding data privacy and data theft Zyskind, Nathan and Pentland published a short paper in 2015. The article is titled “Decentralizing Privacy: Using Blockchain to Protect Personal Data” and focuses on the data privacy of individuals (Zyskind, Nathan and Pentland, 2015). It provided an exciting overview of a potential personal information privacy management solution using blockchain technology combined with “off-blockchain storage” to “construct a personal data management platform” (Zyskind, Nathan and Pentland, 2015, pp 180). The proposed system permits access control and data storage and retrieval. The authors have provided protocols and procedures for programming the system. During the process of designing the platform, Zyskind, Nathan and Pentland (2015) identified three problems faced by such systems:

- the importance of users owning and controlling their data;
- transparency of how the data is used and who has access to it and the reasons why, and;
- “Fine-grained Access Control” (Zyskind, Nathan and Pentland, 2015, pp 181) as on any platform, while using online services users are required to indefinitely agree to a set of permissions when signing up for an online service, and the provider may change the terms of service.

These three problems overlap with the privacy issues mentioned earlier in this section of this dissertation.

A Gartner forecast estimated that 7 billion Internet of Things devices would be in use by the consumer segment in 2018 (Gartner Inc., 2017). These internet-connected devices are designed to enhance the lives of individuals. However, there is a risk to individuals as the integrity of data transmitted is at risk from hackers, viruses, and malicious software (Jing et al., 2014; Yang et al., 2017).

A proposal by Robles et al. (2018) suggests the use of blockchain technologies for the management of private data in ambient intelligence (AmI) environments. Vasilakos and Pedrycz (2006) as cited in Cook, Augusto and Jakkula (2009), describe an AmI environment as one that immerses people in networks of intelligent devices that through the use of sensors can determine their state, anticipate change, and adapt to a person's needs. Many Internet of Things devices are AmI systems; therefore by default information privacy becomes a concern. If implemented, the proposal may provide an effective data management solution on blockchain technology. However, due to the storing of data on the blockchain, it is publicly viewable. The privacy issue needs resolving for all blockchain solutions attempting to store private data.

“Blockchain Challenges and Opportunities: A Survey” by Zheng et al. (2016) details the operation of blockchain technology. The authors combined research that they reviewed and proposed new blockchain applications that would avail of ideas they garnered. While noting enterprise applications, also referenced is the Internet of Things concerning safety and privacy. Importantly for individuals privacy, Zheng et al. (2016) reference a paper by Hardjono and Smith (2016) that postulates blockchain can be used to help a device "prove its manufacturing provenance without authentication of third party and it is allowed to register anonymously" (Zheng et al., 2016, pp 364). With this, it is possible that only the device manufacturers can install software updates and so nullify the opportunity for the installation of hacked or infected software that may comprise information privacy. Zheng et al. (2016) demonstrated that privacy protection using the blockchain to authorise the access to, and transfer of personal data, between an individual and companies such as Facebook would impact individuals directly.

The Internet of Things provides excellent opportunities to improve human life and with the advent of AmI technology new privacy and authentication issues for connected devices arise (Yang et al., 2017). These new technologies compound privacy issues created by existing technology and services, and as Clarke (1999) deduced, there are trust issues with the handling of private data and individuals want more control and influence on how their information is used and processed. Protagonists believe that control can be given back to the individual through data privacy management systems, that make use of the technological advancements blockchain provides (Alessi et al., 2018; Cai, Yuan and Wang, 2017; Yan, Gan and Riad, 2017; Zyskind, Nathan and Pentland, 2015). Therefore, a systematic literature review is of great benefit to academia as it provides a starting point in future research.

### *2.1.2. Research Question*

The research objective of conducting a rigorous systematic literature review translates into the following research question:

**RQ:** What generic models are proposed to leverage blockchain technology to provide individuals with a personal data store?

## 2.2. *Protocol*

### *2.2.1. Search Process*

The search process consists of a search of electronic databases for journal papers and conference proceedings dated from 1 August 2015 to 24 March 2019. The start month of August 2015 was chosen as the concept for the Ethereum blockchain which enabled the use of smart contracts on blockchain technology (Buterin, 2014) was launched on 30 July 2015 (Tual, 2015). However, as some of the electronic repositories and databases do not permit searching by exact date, the initial search which was conducted on 24 March 2019 used a year range of 2015-2019. The Practical Screen process (section 2.3.) excluded any papers published before 1 August 2015. The electronic repositories and databases searched are as per table 1 below.

*Table 1: Search Process Electronic Repository/Database Name*

| Electronic Repository/Database Name |
|---|
| EBSCOhost - Academic Search Complete and LISTA (Library, Information Science & Technology Abstracts) |
| IEEE Explore Digital Library |
| SAGE Journals |
| Science Direct |
| Scopus |
| Springer Link |

#### *2.2.1.1. Search Structure*

Keywords were selected and arranged into boolean commands (table 2) for use when searching the repositories and databases. The search commands used for each source repository are contained in table 3.

*Table 2: Search Structure Keywords*

| **Includes** | **AND** | **AND** | **AND** | **AND** | **AND** |
|---|---|---|---|---|---|
| | | **OR** | **OR** | **OR** | **OR** |
| Blockchain | Data | Privacy | Store | Personal | Model |
| | | Private | Storage | Personalised | System |
| | | | | Personalized | Framework |
| | | | | | Scheme |

#### *2.2.1.2. Limiters Applied for Inclusion*

Limiters were applied to the electronic repository searches in order to obtain more relevant results.

- Publication Date: 2015 to 2019
- Language: English
- Full text is available in the specified repositories and databases

#### *2.2.1.3. Limiters for Exclusion*

Limiters were applied to the electronic repository to eliminate papers that were deemed unsuitable for selection.

- Language: Non-English
- Full text is not available in the specified repositories and databases
- Paper is a review of another research paper
- Paper is a review of conference proceedings
- Contains the following additional keywords: medical, health, healthcare, education, automotive, government, analytics

*Table 3: Search Commands Used*

| Electronic Repository/ Database Name | Search Commands Used |
|---|---|
| EBSCOhost - Academic Search Complete and LISTA (Library, Information Science & Technology Abstracts) | blockchain AND data AND (personal OR personalised OR personalized) AND ( privacy OR private ) AND ( model OR system OR framework OR scheme ) AND ( store OR storage ) |
| IEEE Explore Digital Library | (("blockchain" AND "data") AND ("personal" OR "personalised" OR "personalized") AND ("privacy" OR "private") AND ("model" OR "system" OR "framework" OR "scheme") AND ("store" OR "storage")) |
| SAGE Journals | [Title blockchain OR Abstract blockchain] AND [Title data OR Abstract data] AND [[Title personal] OR [Title personalised] OR [Title personalized] OR [Abstract personal] OR [Abstract personalised] OR [Abstract personalized]] AND [[Title model] OR [Title system] OR [Title framework] OR [Title scheme] OR [Abstract model] OR [Abstract system] OR [Abstract framework] OR [Abstract scheme]] AND [[Title store] OR [Title storage] OR [Abstract store] OR [Abstract storage]] AND [[Abstract private] OR [Abstract privacy] OR [Title private] OR [Title privacy]] |
| Science Direct | tak: ((blockchain AND data AND (privacy OR private) AND (personal OR personalised OR personalized) AND (framework OR model OR system) AND (store OR storage)) |
| Scopus | (TITLE-ABS(blockchain) AND TITLE-ABS(data) AND TITLE-ABS(personal OR personalised OR personalized) AND TITLE-ABS(privacy OR private) AND TITLE-ABS(framework OR model OR system OR scheme) AND TITLE-ABS(store OR storage)) AND ( LIMIT-TO ( DOCTYPE,"cp" ) OR LIMIT-TO ( DOCTYPE,"ar" ) ) |
| SpringerLink | ("blockchain" AND "data" AND ("personal" OR "personalised" OR "personalized") AND ("private" OR "privacy") AND ("framework" OR "model" OR "system" OR "scheme") AND ("store" OR "storage") |

#### *2.2.1.4. Search Results*

The search commands (table 3) were executed and 323 papers were returned as in table 4. A large number of results can be accounted for by the lack of specification that SpringerLink permits when searching for articles. It is not as specific as the other repositories and databases. It is not possible to be as accurate with the sections of the document to be searched. Therefore, the search required some refining.

*Table 4: Search Result Totals*

| Repository/Database | Total Number of Papers Found Matching |
|---|---:|
| EBSCOhost | 4 |
| IEEE Explore Digital Library | 14 |
| SAGE | 0 |
| ScienceDirect | 0 |
| SCOPUS | 10 |
| SpringerLink | 295 |
| Total Papers Found | 323 |

#### *2.2.1.5. Refining the Search*

As this dissertation is investigating generic models that can apply to a multitude of data sharing scenarios, the electronic repositories that had previously returned results as in table 5 were searched again.

*Table 5: Refined Electronic Repository/Database Names*

| Electronic Repository/Database Name |
|---|
| EBSCOhost - Academic Search Complete and LISTA (Library, Information Science & Technology Abstracts) |
| IEEE Explore Digital Library |
| Scopus |
| Springer Link |

The new search command changed the ('store' OR 'storage') section to ('store' OR 'storage' OR 'management') as it is considered that if storing is required then management is also required, and different authors may reference a variant of 'store' or the word 'management'. Limiters were then added to eliminate research papers targeting specific industries or areas of interest as these do not provide for the generic model requirement. The results must not include the following additional keywords (table 6):

- medical
- health
- healthcare
- education
- automotive
- government
- analytics
- bitcoin transaction

*Table 6: Refined Search Structure Keywords*

| **Includes** | **AND** | **AND** | **AND** | **AND** | **AND** | **NOT** |
|---|---|---|---|---|---|---|
| | | **OR** | **OR** | **OR** | **OR** | |
| Blockchain | Data | Privacy | Management | Personal | Framework | Analytics |
| | | Private | Store | Personalised | Model | Automotive |
| | | | Storage | Personalized | Scheme | Bitcoin Transaction |
| | | | | | System | Education |
| | | | | | | Government |
| | | | | | | Health |
| | | | | | | Healthcare |
| | | | | | | Medical |

*Table 7: Refined Search Commands Used*

| Electronic Repository/ Database Name | Search Commands Used |
|---|---|
| EBSCOhost - Academic Search Complete and LISTA (Library, Information Science & Technology Abstracts) | blockchain AND data AND (personal OR personalised OR personalized) AND ( privacy OR private ) AND ( model OR system OR framework OR scheme ) AND ( store OR storage OR management) AND NOT medical AND NOT health AND NOT healthcare AND NOT education AND NOT automotive AND NOT government AND NOT analytics AND NOT bitcoin transaction |
| IEEE Explore Digital Library | ((("blockchain" AND "data") AND ("personal" OR "personalised" OR "personalized") AND ("privacy" OR "private") AND ("model" OR "system" OR "framework" OR "scheme") AND ("store" OR "storage" OR "management") AND NOT ("medical" OR "health" OR "healthcare" OR "education" OR "automotive" OR "government" OR “analytics” OR "bitcoin transaction"))) |
| Scopus | ( TITLE-ABS ( blockchain ) AND TITLE-ABS ( data ) AND TITLE-ABS ( personal OR personalised OR personalized ) AND TITLE-ABS ( privacy OR private ) AND TITLE-ABS ( framework OR model OR system OR scheme ) AND TITLE-ABS ( store OR storage OR management) AND NOT TITLE-ABS ( "medical" OR "health" OR "healthcare" OR "education" OR "automotive" OR "government" OR "analytics" OR “bitcoin transaction”) ) AND ( LIMIT-TO ( LANGUAGE , "English" ) ) |
| SpringerLink | ("blockchain" AND "data" AND ("personal" OR "personalised" OR "personalized") AND ("private" OR "privacy") AND ("framework" OR "model" OR "system" OR "scheme") AND ("store" OR "storage" OR "management") NOT ("medical" OR "health" OR "healthcare" OR "education" OR "automotive" OR "government" OR "analytics" OR "bitcoin transaction")) |

The refined research reduced the number of unique papers to 81 (table 8). SpringerLink continued to return a large number of results. However, all 81 articles passed onto the next phase, the Practical Screen where papers that are not pertinent will face elimination. The papers that passed this phase and continued to be considered are listed with the results of the Practical Screen in Appendix 1.

*Table 8: Refined Search Result Totals*

| **Repository/Database** | **Total Number of Papers Found Matching** |
|---|---|
| EBSCOhost | 0 |
| IEEE Explore Digital Library | 17 |
| SCOPUS | 10 |
| SpringerLink | 62 |
| **Total Papers Found** | **89** |
| Duplicated papers | 8 |
| **Total Unique Papers** | **81** |

### *2.3. Training*

As this dissertation is a solo effort, it was deemed that the training document was not required. The protocol describes the in-depth procedure used to conduct this systematic literature review.

### *2.4. Practical Screen*

The Practical Screening incorporated the critical evaluation of the abstract contents of all unique papers identified during the search process (table 8) to select all papers that are relevant to the research questions (section 2.1.2.), and therefore meet the following criteria:

- Published between 1 August 2015 and 24 March 2019
- Refers to personal data storage or management
- Propose a model, system, framework or scheme
- Blockchain technology is part of or all of the proposed solution
- Does not target a specific industry

The results of the Practical Screen process are contained in Appendix 1., and identify where papers reviewed pass or fail. The Practical Screen Results table is sorted by 'Repository' and then the 'Paper Title' columns. 81 papers were involved in the Practical Screen process. 69 papers were awarded a fail, and 12 papers passed.

Table 9 (see below) lists all the papers that passed the Practical Screen stage and is sorted alphabetically by the Paper Title. Each paper is assigned a number to be used for reference throughout the remainder of the systematic literature review.

*Table 9: List of Papers that Passed the Practical Screen*

| Paper No. | Paper Title | Authors |
|---|---|---|
| 1 | A Novel Sustainable Interchain Network Framework for Blockchain | Q. Yang, H. Guo, V. Zhu, X. Fan, X. Cui, X. Kong, B.K. Bobby |
| 2 | An Identity Management System Based on Blockchain | Y. Liu; Z. Zhao; G. Guo; X. Wang; Z. Tan; |
| 3 | An Online Identity and Smart Contract Management System | A. Yasin; L. Liu |
| 4 | BC-PDS: Protecting Privacy and Self-Sovereignty through BlockChains for OpenPDS | Z. Yan; G. Gan; K. Riad |
| 5 | Blockchain-based Trusted Computing in Social Network | F. Dongqi; L. Fang |
| 6 | Decentralizing Privacy: Using Blockchain to Protect Personal Data | G. Zyskind; O. Nathan; A. Pentland |
| 7 | DStore: A Distributed Cloud Storage System Based on Smart Contracts and Blockchain | J. Xue; C. Xu; Y. Zhang; L. Bai |
| 8 | Make Users Own Their Data: A Decentralized Personal Data Store Prototype Based on Ethereum and IPFS | M. Alessi; A. Camillo; E. Giangreco; M. Matera; S. Pino; D. |
| 9 | Mapping Requirements Specifications into a Formalized Blockchain-Enabled Authentication Protocol for Secured Personal Identity Assurance | B. Leiding; A. Norta |

| Paper No. | Paper Title | Authors |
|---|---|---|
| **10** | Peer to Peer for Privacy and Decentralization in the Internet of Things | M. Conoscenti; A. Vetrò; J. C. De Martin |
| **11** | Scalable and Privacy-Preserving Data Sharing Based on Blockchain | B. Zheng, L. Zhu, M. Shen, F. Gao, C. Zhang, Y. Li, J. Yang |
| **12** | Towards Trustworthy and Private Keyword Search in Encrypted Decentralized Storage | C. Cai; X. Yuan; C. Wang |

# 3. Extraction

It must be acknowledged that in theory, there is a risk of data-extraction bias introduced at this stage due to “the reviewer’s awareness of the study authors or the journal, or their disciplinary background, or by awareness of other aspects of the study being reviewed.” (Petticrew and Roberts, 2006, pp 155). Methods to ensure that this does not occur requires obscuring some of the literature details or using at least two reviewers (Petticrew and Roberts, 2006). As this dissertation has one author, it does not have the luxury of guaranteeing that bias will not occur; however, a consistent approach to extracting data is used to attempt to limit any bias and criteria is explicitly stated.

## 3.1. Quality Appraisal

This step of the process incorporates examining the papers and screening for exclusion. Stricter criteria are applied when appraising for quality to assist in the selection of articles. The articles that pass the quality appraisal will “continue to be considered for the literature review” (Okoli and Schabram, 2010, pp 25). Papers are scored on quality to ensure that low-quality studies do not proceed past this point. It also ensures that papers selected are relevant to the research goals of the dissertation.

Eleven quality appraisal (QA) questions were identified to asses the quality of the fifteen papers that passed the Practical Screen stage. Some papers may have passed the Practical Screen as their abstract did not make it evident that the paper was not aligned with the goals of this systematic literature review. Therefore, question one is used to eliminate the papers that are not relevant, and the paper will not continue the QA process. The questions used for the Quality Appraisal are as follows:

**QA1.** Does the paper propose a solution incorporating blockchain for use by individuals to manage personal data?

- Y (yes). The paper explicitly proposes a solution that aligns with the goals of the systematic literature review.
- P (partly). The paper partially proposes a solution that aligns with the goals of the systematic literature review.
- N (no). The paper does not propose a solution that aligns with the goals of the systematic literature review, and this paper is automatically eliminated.

**QA2.** Does the paper identify the problem that it aims to solve?

- Y (yes). The problem to be solved is explicitly described.
- P (partly). The problem to be solved is implicitly described.
- N (no). The problem to be solved cannot be readily inferred.

**QA3.** Is the structure of the paper provided?

- Y (yes). The structure of the paper is explicitly described.
- P (partly). The structure of the paper is implicitly described.
- N (no). The structure of the paper is not provided.

**QA4.** Does the paper contain a general overview of the proposed solution?

- Y (yes). A general overview of the proposed solution explicitly described.
- P (partly). A general overview of the proposed solution is partially described.
- N (no). A general overview of the proposed solution is not provided.

**QA5.** Does the paper outline the high-level system design?

- Y (yes). The high-level system design is explicitly described.
- P (partly). The high-level system design is implicitly described.
- N (no). The high-level system design is not provided.

**QA6.** Does the paper provide a detailed system design?

- Y (yes). The detailed system design is explicitly described.
- P (partly). The detailed system design is implicitly described.
- N (no). The detailed system design is not provided.

**QA7.** Has an evaluation test case or proof of concept been provided in the paper?

- Y (yes). An evaluation test case or proof of concept is explicitly described.
- P (partly). An evaluation test case or proof of concept is implicitly described.
- N (no). An evaluation test case or proof of concept is not provided.

**QA8.** Has a production implementation process described in the paper?

- Y (yes). A production implementation process is explicitly described.
- P (partly). A production implementation process is implicitly described.
- N (no). A production implementation process is not provided.

**QA9.** Does the paper identify any limitations with the proposed system?

- Y (yes). System limitations are explicitly described.
- P (partly). System limitations are implicitly described.
- N (no). System limitations are not provided.

**QA10.** Does the paper identify future expansions of the solution or discuss further research?

- Y (yes). Future expansions are explicitly identified, or further research is explicitly discussed.
- P (partly). Future expansions are implicitly identified, or further research is implicitly discussed.
- N (no). Future expansions are not identified, and there is no discussion of further research.

**QA11.** Is the proposal contained in the paper credible?

- Y (yes). The proposal contained in the paper is credible.
- P (partly). The proposal contained may have some merit.
- N (no). The proposal contained in the paper is not credible.

The quality appraisal process was conducted twice in order to minimise the risk of performance and attrition bias occurring due to having only one person conducting the appraisal. The first pass focused on all papers and the rationale for the results are contained in table 10; Quality Appraisal Rationale First Pass. The second pass focused on the papers that returned ’N’ for question QA1. The results of the second pass reinforce the results of the first pass and therefore are contained in Appendix 2. - Quality Appraisal Rationale: Second Pass.

*Table 10: Quality Appraisal Rationale First Pass*

| **Paper No.** 1 | | |
|---|---|---|
| **Paper Title:** A Novel Sustainable Interchain Network Framework for Blockchain | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | N | This paper is not explicitly providing a solution for individuals to manage their data. The solution can be used for the transfer of any data type between two or more blockchains. |

| **Paper No.** 2 | | |
|---|---|---|
| **Paper Title:** An Identity Management System Based on Blockchain | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | N | Proposed is an identity authentication and reputation management |

| **Paper No.** 3 | | |
|---|---|---|
| **Paper Title:** An Online Identity and Smart Contract Management System | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | N | The proposal is for the Tsinghua University User Reputation System (TURS) that can be used to identify individuals in various fields. It uses online behaviour to identify and rate individuals. This data is collected from social/online media, manually entered and browser history. It is not aimed at the individual managing their data. |

| **Paper No.** 4 | | |
|---|---|---|
| **Paper Title:** BC-PDS: Protecting Privacy and Self-Sovereignty through BlockChains for OpenPDS | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | Y | The proposal is specifically for personal metadata, an area frequently overlooked in the search for privacy solutions. |
| QA2 | Y | Personal metadata is constantly being collected and includes identity, location details and other information and could become a corporate asset. Individuals do not have access to this metadata, and so there is an issue with ownership and privacy. Any anonymised stored metadata unique to an individual stored can be used to re-identify that person. The paper aims to find more successful techniques that enhance personal meta-data privacy against re-identification through building on the non-blockchain OpenPDS/SafeAnswers framework. |
| QA3 | Y | The structure of the paper is included in the last paragraph of the Introduction under the heading “Organization”. |
| QA4 | Y | The introduction in section I includes a very high-level overview under the heading "Our Contribution”. Section II provides more clarity where and existing work is discussed. |
| QA5 | Y | Section IV covers the high-level system design. |

| **Paper No.** 4 | | |
|---|---|---|
| **Paper Title:** BC-PDS: Protecting Privacy and Self-Sovereignty through BlockChains for OpenPDS | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA6 | P | Section IV delves into detail at times. |
| QA7 | Y | A case study is provided. |
| QA8 | N | No evidence is provided for an implementation process in the paper. |
| QA9 | N | No limitations are identified in the paper. |
| QA10 | N | No future expansion or research are discussed. |
| QA11 | Y | The proposal is credible. It introduces a blockchain layer onto the existing OpenPDS/SafeAnswers framework. OpenPDS/SafeAnswers enables users to collect and store personal data, and also the ability to bestow fine-grained access to personal data. The addition of blockchain also aids to protect their privacy. This proposal would make it more secure and impossible to re-identify an individual through their metadata. |

| **Paper No.** 5 | | |
|---|---|---|
| **Paper Title:** Blockchain-based Trusted Computing in Social Network | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | N | This paper proposes an encryption algorithm for the model proposed in the paper “Decentralizing Privacy: Using Blockchain to Protect Personal Data” authored by Zyskind, Nathan and Pentland (2015). It does not propose a system. |

| **Paper No.** 6 | | |
|---|---|---|
| **Paper Title:** Decentralizing Privacy: Using Blockchain to Protect Personal Data | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | Y | A combined on-blockchain and off-blockchain solution are proposed in this paper. |
| QA2 | Y | Section II discusses the issue of privacy that the authors aim to solve. |
| QA3 | Y | The structure of the paper is included in the last paragraph of the Introduction under the heading “Organization”. |
| QA4 | Y | Section III provides an overview of the solution. |

| **Paper No.** 6 | | |
|---|---|---|
| **Paper Title:** Decentralizing Privacy: Using Blockchain to Protect Personal Data | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA5 | P | The overview provides some high-level design detail. |
| QA6 | Y | Section IV identifies in detail the underlying protocols used in the system for creating identities, permission checks, access control and storing or loading data. |
| QA7 | N | No evaluation or test case provided within the paper. |
| QA8 | N | No evidence is provided for an implementation process in the paper. |
| QA9 | N | No limitations are identified in the paper. |
| QA10 | Y | The authors discuss two extensions to their system in detail. They also provide an example flowchart of secure computation using multiple cryptographic protocols to enable the hiding of raw data from the service requesting information and the computations included. Also suggested is rewarding the behaviour of nodes and trusting some more than others. |
| QA11 | Y | This proposal is credible and as a result, has been cited many times in blockchain and data privacy research papers and used as a base for paper number 5. However, if a future expansion of rewarding nodes for good behaviour is developed, it will make hacking the network easier as the node gets more voting rights during the consensus process it could reach 51% and write entries to the blockchain. |

| **Paper No.** 7 | | |
|---|---|---|
| **Paper Title:** DStore: A Distributed Cloud Storage System Based on Smart Contracts and Blockchain | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | N | It is a solution that leases unused cloud space to data owners that wish to store data distributed on blockchains in the cloud. |

| **Paper No.** 8 | | |
|---|---|---|
| **Paper Title:** Make Users Own Their Data: A Decentralized Personal Data Store Prototype Based on Ethereum and IPFS | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | Y | A personal data store and personal data management are proposed. |
| QA2 | Y | The introduction highlights the issue of third-party service's collecting, storing, moving data internally and transferring personal data to unknown third-parties without the individuals' knowledge. The paper aims to resolve this through Personal Data Management and give control of personal data back to the individual. |
| QA3 | Y | The structure of the paper is included in section I. Introduction. |
| QA4 | Y | An overview is provided in section I. Introduction. |
| QA5 | Y | Section III, The Proposed Solution provides a high-level system design. |
| QA6 | N | There is a lack of technical detail of how the specific components connect. |
| QA7 | Y | A proof of concept was created, tested in a laboratory setting, and documented in section IV. Validation. |
| QA8 | N | No evidence is provided for an implementation process in the paper. |
| QA9 | Y | A potential external limitation has been identified. The European General Data Protection Regulation provides the right for an individual to have all personal data stored about them erased. At the time of writing the paper, it was not technologically possible to delete records contained on a blockchain. |
| QA10 | Y | It is proposed to implement the solution external to the Servify ecosystem that was used for the proof of concept. Also, mentioned as future features are is the ability to manage the users Ethereum identity, and the creation of a safe address for users, to enable users to expose different levels of personal data to services. |
| QA11 | Y | The proof of concept demonstrates that the proposal is credible and has potential applications. |

| **Paper No.** 9 | | |
|---|---|---|
| **Paper Title:** Mapping Requirements Specifications into a Formalized Blockchain-Enabled Authentication Protocol for Secured Personal Identity Assurance | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | N | Proposed is a potential replacement for public key infrastructure. |

| **Paper No.** 10 | | |
|---|---|---|
| **Paper Title:** Peer to Peer for Privacy and Decentralization in the Internet of Things | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | N | The paper proposes a research idea. No research has been conducted on the proposed idea. |

| **Paper No.** 11 | | |
|---|---|---|
| **Paper Title:** Scalable and Privacy-Preserving Data Sharing Based on Blockchain | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | N | A multiparty, multi-layered system that uses the cloud, on-blockchain and off-blockchain (local storage) is proposed. Multiparty is where the data of multiple users is shared among all users. Each user has part of the data. This proposal is not designed for individuals. |

| **Paper No.** 12 | | |
|---|---|---|
| **Paper Title:** Towards Trustworthy and Private Keyword Search in Encrypted Decentralized Storage | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA1 | Y | Proposed is a secure decentralised personal data management system that enables keyword searches. |
| QA2 | Y | No current encrypted decentralised storage solutions support secure keyword searching. |
| QA3 | Y | The structure of the paper is included in the last paragraph of section I Introduction. |
| QA4 | Y | The introduction in section I provides an overview of issues faced and how to deal with them. |
| QA5 | Y | Section IV System Model provides a high-level design. |

| **Paper No.** 12 | | |
|---|---|---|
| **Paper Title:** Towards Trustworthy and Private Keyword Search in Encrypted Decentralized Storage | | |
| **Question** | **Grade** | **Rationale for Grade** |
| QA6 | Y | Section V System Design provides a detailed system design for all modules |
| QA7 | P | Section VI performs an analysis of security, performance and cost. It is not a full proof of concept or end to end test case. |
| QA8 | N | No evidence is provided for an implementation process in the paper. |
| QA9 | Y | The scalability of blockchain technology is mentioned as a limitation but is currently being researched by others. |
| QA10 | N | No future expansion or research are discussed. |
| QA11 | Y | This proposal is credible as it uses existing technology and adding a layer of searchable symmetric encryption. |

Each paper is awarded a score per appraisal question and is calculated with 1 point for 0 points for N (no), Y (yes), and 0.5 points for P (partially) as per the following table.

*Table 11: Quality Appraisal Scoring*

| Paper No. | QA 1 | QA 2 | QA 3 | QA 4 | QA 5 | QA 6 | QA 7 | QA 8 | QA 9 | QA 10 | QA 11 | Score |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | N | - | - | - | - | - | - | - | - | - | - | **0** |
| **2** | N | - | - | - | - | - | - | - | - | - | - | **0** |
| **3** | N | - | - | - | - | - | - | - | - | - | - | **0** |
| **4** | Y | Y | Y | Y | Y | P | Y | N | N | N | Y | **7.5** |
| **5** | N | - | - | - | - | - | - | - | - | - | - | **0** |
| **6** | Y | Y | Y | Y | P | Y | N | N | N | Y | Y | **7.5** |
| **7** | N | - | - | - | - | - | - | - | - | - | - | **0** |
| **8** | Y | Y | Y | Y | Y | N | Y | N | Y | Y | Y | **10** |
| **9** | N | - | - | - | - | - | - | - | - | - | - | **0** |
| **10** | N | - | - | - | - | - | - | - | - | - | - | **0** |
| **11** | N | - | - | - | - | - | - | - | - | - | - | **0** |
| **12** | Y | Y | Y | Y | Y | Y | P | N | Y | N | Y | **8.5** |

Due to the specific nature of this systematic literature review, it was decided that papers would require a minimum score of 65% (7.15 out of the maximum of 11) to continue to be considered. Papers numbers 4, 6, 8 and 12 were deemed to be of high quality having scored a minimum of 7.5, and these papers continue to be considered for the review.

Paper numbers 1, 2, 3, 5, 7, 9, 10 and 11 are eliminated from the process as they identified as not aligning with the goals of this systematic literature review. Through an in-depth appraisal of the papers mentioned above, they were awarded a grade of ’N’ (no) to question 1: ‘Does the paper propose a solution incorporating blockchain for use by individuals to manage personal data?’ Therefore, each of the papers as mentioned above scored 0 points.

*Table 12: Papers that Continue to be Considered*

| Paper No. | Paper Title | Authors |
|---|---|---|
| **4** | BC-PDS: Protecting Privacy and Self-Sovereignty through BlockChains for OpenPDS | Z. Yan; G. Gan; K. Riad |
| **6** | Decentralizing Privacy: Using Blockchain to Protect Personal Data | G. Zyskind; O. Nathan; A. Pentland |
| **8** | Make Users Own Their Data: A Decentralized Personal Data Store Prototype Based on Ethereum and IPFS | M. Alessi; A. Camillo; E. Giangreco; M. Matera; S. Pino; D. Storelli |
| **12** | Towards Trustworthy and Private Keyword Search in Encrypted Decentralized Storage | C. Cai; X. Yuan; C. Wang |

## 3.2. Data Extraction

A data extraction form was developed to assist with the organisation and analysation of the data contained in the six papers. The form gathers the data as recorded in the paper. The extracted data is required to address the research questions of this systematic literature review.

Items I01 to I07 identify general information about the papers including the paper number (as assigned during the Quality Appraisal in section 3.1.), paper title, author(s), country, citation, country, and the publication details. Items I07 to I10 contain specific data gathered from reading the papers including the abstract, the aim of the study, the system design, and the future extensions for the proposal. The data extraction forms are amalgamated in table 13.

*Table 13: Amalgamated Data Extraction Forms*

| Item | Data Item | Description |
|---|---|---|
| **I01** | Paper No. | 4 |
| **I02** | Title | BC-PDS: Protecting Privacy and Self-Sovereignty through BlockChains for OpenPDS |
| **I03** | Authors | Z. Yan; G. Gan; K. Riad |
| **I04** | Country | China |
| **I05** | Citation | (Yan, Gan and Riad, 2017) |
| **I06** | Publication | 2017 IEEE Symposium on Service-Oriented System Engineering |
| **I07** | Type | Conference Paper |
| **I08** | Abstract | In the Big Data era, personal metadata may will become a new type of corporate asset, however there have already been a growing public concern about user's privacy mined from metadata. In this paper we address the problem of implementing the self-sovereignty of personal metadata on the existing OpenPDS/SafeAnswers framework according to the Windhover Principle. In order to do that, we propose a new framework, called BlocakChain-based Personal Data Store (BC-PDS), to realize two basic properties: notary and autonomy. This framework, firstly introduces the BlockChain as a notary, into OpenPDS/SafeAnswers for secure storage of personal meta-data instead of the original database. Next, we present an AutoNomy based Access Control (ANAC) to improve the SafeAnswers module, where ANAC is a new mechanism that enforces access based on the relationship among all authorized users and metadata's owner. In addition, we also propose General Access Structure (GAS) and threshold secret sharing scheme in BlockChain as an implementation method for our BC-PDS framework. |
| **I09** | Aim of Paper | A system for self-ownership of personal metadata on the OpenPDS/ SafeAnswers model using the Windhover Principle. A new framework called BlockChain-based OpenPDS (BC-PDS) is presented to enable notary and autonomy, where the Blockchain, acts as a notary for the secure storage of personal metadata. An AutoNomy-based Access Structure is proposed to improve the SafeAnswers module. |

| Item | Data Item | Description |
|---|---|---|
| **I10** | System Design | BlockChain based Personal Data Store (BC-PDS) is a framework built on the OpenPDS/SafeAnswers model. This platform permits the collection and storing of data, while also giving fine-grained access control to protect user data as per the Windhover principles. To do this, personal metadata requires robust verification and enforcement. The following properties are the focus of the design:<br>• Notary: This is blockchain based and removes the need for a central authority for certifying the authenticity of added and altered data. It is used for proof of existence, integrity and validity. It prevents tampering and counterfeiting of stored personal metadata.<br>• AutoNomy: AutoNomy-based Access Structure (ANAC) is a new module that provides the user full access to their data, and enables them to grant and remove access to data depending on the relationships among all authorised users. This permits one-to-one and one-to-many data sharing.<br><br>Cryptography Used:<br>• Threshold Secret Sharing (TSS)<br>o The access structure in the BC-PDS enforces the use of TSS which is based on Shamir's Secret Sharing algorithm. It breaks the metadata amongst all authorised users. Due to secret sharing, all parties with access must cooperate in order to see the full set of metadata.<br>• General Secret Sharing Scheme (GSS)<br>o The GSS is a set of recovery algorithms that combines multiple TSS's into a tree using 'AND' and 'OR' logical functions.<br>Without the owners Personal Data Store providing authorisation via GSS, the notary cannot access the data as requested by any other party and so the self-sovereignty of personal metadata is provided. |
| **I10** | Future Extensions | None specified. |

| Item | Data Item | Description |
|---|---|---|
| **I01** | Paper No. | 6 |
| **I02** | Title | Decentralizing Privacy: Using Blockchain to Protect Personal Data |

| Item | Data Item | Description |
|---|---|---|
| **I03** | Authors | G. Zyskind; O. Nathan; A. Pentland |
| **I04** | Country | United States of America |
| **I05** | Citation | (Zyskind, Nathan and Pentland, 2015) |
| **I06** | Publication | 2015 IEEE Security and Privacy Workshops |
| **I07** | Type | Workshop Paper |
| **I08** | Abstract | The recent increase in reported incidents of surveillance and security breaches compromising users' privacy call into question the current model, in which third-parties collect and control massive amounts of personal data. Bitcoin has demonstrated in the financial space that trusted, auditable computing is possible using a decentralized network of peers accompanied by a public ledger. In this paper, we describe a decentralized personal data management system that ensures users own and control their data. We implement a protocol that turns a block chain into an automated access-control manager that does not require trust in a third party. Unlike Bitcoin, transactions in our system are not strictly financial -- they are used to carry instructions, such as storing, querying and sharing data. Finally, we discuss possible future extensions to block chains that could harness them into a well-rounded solution for trusted computing problems in society. |
| **I09** | Aim of Paper | To solve the data privacy issues of data ownership, data transparency and audibility, and enable individuals to manage the data through fine-grained access control with revocation rights. |
| **I10** | System Design | A hybrid blockchain and off-blockchain storage solution that enables blockchain to act as an access control manager is proposed. Two transactions would be permitted; $T_{access}$ for the access control management, and $T_{data}$ for the storage and retrieval of data. These transactions contain a pointer to an off-blockchain key-value store.<br>The user downloads a mobile application that uses the proposed platform and signs up to create a new user profile. The profile is sent with associated permissions to the third-party service using $T_{access}$. Any additional data is sent using the $T_{data}$ transaction to a blockchain with an accompanying encryption key. The user and third-party |

| Item | Data Item | Description |
|---|---|---|
| | | use $T_{data}$ for queries, and digital signatures then provide authentication.<br>As blockchain records are immutable, they cannot be modified, and a new $T_{access}$ transaction is required to revoke the third-party access to data.<br>The data is stored in the cloud, and it is hashed and randomised using Kademilia.<br>Network Protocol:<br>• Compound identity: shared identity but owned by one party.<br>• Blockchain memory: use the latest entry to allow updates, deletions and inserts.<br>• Policy: permissions granted to a service.<br>• Auxiliary functions: verify permissions.<br>Access Control Protocol:<br>• $T_{access}$ allows setting and changing of permissions.<br>Storing and Loading Data Protocol:<br>• $T_{data}$ transaction results in the sending of off-blockchain read and write messages to the data store. |
| **I10** | Future Extensions | • Storage to Processing: Implement processing to ensure that the third-party service does not have access to raw data. Hiding the raw data may be achieved using Shamir's Secret Sharing cryptography algorithm and the Oblivious Transfer cryptography protocol.<br>• Trust and Decision Making in Blockchains: Enable node behaviour monitoring and reward good behaviour with more trust. |

| Item | Data Item | Description |
|---|---|---|
| **I01** | Paper No. | 8 |
| **I02** | Title | Make Users Own Their Data: A Decentralized Personal Data Store Prototype Based on Ethereum and IPFS |

| Item | Data Item | Description |
|---|---|---|
| **I03** | Authors | M. Alessi; A. Camillo; E. Giangreco; M. Matera; S. Pino; D. Storelli |
| **I04** | Country | Italy |
| **I05** | Citation | (Alessi et al., 2018) |
| **I06** | Publication | 3rd International Conference on Smart and Sustainable Technologies, SpliTech 2018 |
| **I07** | Type | Conference Paper |
| **I08** | Abstract | In the times we are living, data protection infringements, at local, national or international level, are a daily occurrence, highlighting how important is the problem of users' awareness and “consent” about what data should or not be shared. A vast number of service providers strives to have access to users' personal data. While users may be aware of sharing their data with services they receive, they may be still unaware if their data is passing in others' hands and unknown third parties. But the sharing of personal data remains unavoidable, in this always connected digital era, contextualized services are not only fancy desires, they could save money, time, and even lives. The problem becomes even more complicate if we try to consider the devices around us: how to share devices we own, so that we can receive pervasive services, based on our contexts and device functionalities. The European Authority has provided regulations about personal data protection, but there are still significant differences in the ways each EU member state would implement the protection of privacy and personal data in national laws, policies, and practices. The tool that should empower users with the personal data protection has to face two problems: data privacy and control. Due to the lack of central authorities, blockchain based technologies would seem fit for the challenge, but such solutions are not fully exploited. One possible reason could be that distributed architectures alone do not achieve privacy of data. In this paper we tackle the challenge of a novel Personal Data Store, by making use of a distributed architecture, based on the Ethereum framework, together with an ontology to model user profile and data/device sharing towards services. Such solution, The Decentralized Identity Manager, solves personal data protection by offering a |

| Item | Data Item | Description |
|---|---|---|
| | | unique endpoint, without any central authority, where users can manage their data/device access, their privacy levels, and grant or deny sharing consent, every time services ask for personal data. |
| **I08** | Aim of Paper | The paper provides a prototype of a personal data management application that permits users to control personal data and who can access the data. This included considering personal devices as personal data. The proposed solution takes advantage of distributed technology in order to eliminate central authorities. Therefore, it would be trustless and leverage the pervasive properties of distributed technology |
| **I09** | System Design | A Decentralised Identity Manager (DIM) is created on a desktop computer and smartphone (application). The DIM is responsible for managing the user's identity and profile. It also provides static and dynamic user profile data for contextualisation of pervasive services once the user has provided authorisation. Using the DIM, individuals can create, alter and delete profile data. Personal data and devices are considered as profile data. Interlinked data is encoded using the human-readable JSON-LD (JavaScript Object Notation for Linked Data).<br>The user profile uses schema.org's standard schemas to describe the user context. The user profile consists of:<br>• Person entity schema describes the user and all its' attributes.<br>• Product entity schema is a representation of user-owned devices.<br>• Place entity schema is a representation of places.<br>• Action entity is a representation of actions performed by a user.<br>• Attributes of the user profile can be public or private depending on user privacy settings.<br>• Shared attributes indicate attributes that users wish to share with third-party services. |

| Item | Data Item | Description |
|---|---|---|
| | | Technological Requirements:<br>• Ethereum blockchain<br>• InterPlanetary File System (IPFS). IPFS is a protocol and peer-to-peer network for decentralised storage and file sharing. When a file gets added to the IPFS network, it is hashed uniquely. Each subsequent file change results in a new hash.<br><br>Deployment of Prototype on a Smartphone:<br>• Ethereum Blockchain contains:<br>o Username - autogenerated on the first use of the DIM<br>o Ethereum address<br>o Hash of the user profile - used as a key to connect to the IPFS and access content<br>• IPFS<br>o IPFS is used due to the cost (in Ethereum cryptocurrency) of recording transactions of the Ethereum blockchain.<br>o Stores decentralised files on the Android platform using the IpfsDroid library.<br>o Lightwallet is an Android application for sending transactions to the remote Ethereum node and create an Ethereum account.<br>• The application permits the deletion of the user profile, adding and editing data, sharing attributes with third-party services, integrates with social media and setting attributes to public or private. |
| **I10** | Future Extensions | • Develop a large scale prototype that integrates with more services.<br>• Transition from the Ethereum test bed to the production Ethereum environment.<br>• Enable users to manage their Ethereum identity from within the smartphone application.<br>• Enable users to have a safe list of addresses connected with their distributed profile in order to insert access rights and provide different layers of data to services.<br>• Implement the KSI Technology Stack standard to plug cryptography vulnerability of the Ethereum blockchain dependence on RSA standards which can be exposed by quantum computers.<br>• Address General Data Protection Regulation compatibility issues. |

| Item | Data Item | Description |
|---|---|---|
| **I01** | Paper No. | 12 |
| **I02** | Title | Towards Trustworthy and Private Keyword Search in Encrypted Decentralized Storage |
| **I03** | Authors | C. Cai; X. Yuan; C. Wang |
| **I04** | Country | China |
| **I05** | Citation | (Cai, Yuan and Wang, 2017) |
| **I06** | Publication | IEEE ICC 2017 Communication and Information Systems Security Symposium |
| **I07** | Type | Conference Paper |
| **I08** | Abstract | Emerging decentralized storage services such as Storj and Filecoin show promise as a new paradigm for data outsourcing. These services tie cryptocurrency to personal storage resources and leverage blockchain technology to ensure data integrity in distributed networks. Compared to current cloud storage, they are expected to be more scalable, cost effective, and secure. In addition to the features above, strong guarantees of data privacy are seriously desired due to today's prevalent data leak and abuse incidents. However, simply using end-to-end encryption limits the search capability and thus will degrade the user experience. In this paper, we propose an encrypted decentralized storage architecture that can support trustworthy and private keyword search functions. We start from searchable encryption to achieve search on encrypted data. Yet, only adopting this primitive is not sufficient to address particular threats in our target decentralized service model. Service peers would maliciously return incorrect results, while user peers would fraudulently refuse to pay service fees. To resolve those threats, we devise specific secure data addition and keyword search protocols to enable client-side verifiability and blockchain based fair judgments on the search results. For practical considerations, we integrate an efficient dynamic searchable encryption scheme to our protocols as an instantiation to lower the blockchain overhead. Our security and performance analysis indicates the advance of the proposed architecture. |

| Item | Data Item | Description |
|---|---|---|
| **I09** | Aim of Paper | To provide a decentralised storage architecture that enables an encrypted keyword search function and incentivises use. |
| **I10** | System Design | Two types of peers are utilised in this storage service, a client peer (data owners) and a storage peer. Client peers provide the files and indexes to storage peers that can conduct keyword searches and return the search results for which they receive payment. Peers can switch roles depending on the contract. The blockchain is a global ledger for fair judgements, i.e. consensus.<br><br>The system has four functions:<br>• System Setup<br>○ Negotiates and establishes smart contracts with one or many storage peers to perform the search and storage services. This also specified the storage rental connections and length of time storage will be used.<br>○ In the client peer, private keys are created along with a client checklist to store set hashes an empty search history, and a (multi)set hashing function.<br>○ The storage peer initialises the encrypted file index, keyword search index, posting list of search results, and a digest index that stores set hashes.<br>○ All peers create public/secret key pairs.<br>• File Addition<br>○ When the client peer sends a new file out for storage, the file is parsed into keywords.<br>○ The client peer builds an Add token consisting of the file ID, a cryptographic file digest for integrity checking, a hash of the file ID, a list of previously searched keywords, the encrypted file, and a file index. This token is signed using a standard RSA signature scheme.<br>○ The token and signature are sent to the storage peer, and the client pier checklist is updated with the set hash of the file.<br>○ The storage peer updates the file index, keyword search index and digest index. |

| Item | Data Item | Description |
|---|---|---|
| | | o The storage peer creates the Add transaction consisting of the set hash of the file, the storage location, and the signature of the Add token and writes it to the blockchain.<br>• Keyword Search<br>o The client peer creates a Search token of the pseudo-random function of the private key and updates the search history. The token is signed using the standard RSA signature scheme and sent to the storage peer. The storage peer verifies the authenticity of the token using the client peers public key.<br>o If the token is valid, and the keyword has not been used previously, the file index is scanned and a posting list is created using a set hash function of encrypted files. If the keyword has not been searched for previously, the set of results files,<br>o The search index is updated with the newly created set hash.<br>o The set hash is then compared to the client peer checklist.<br>• Fair Judgement<br>o The storage rental service permits both types of peer to publish a fair judgement request within the smart contract.<br>o Both peers must furnish data to support their behaviour.<br>o Other peers then recreate the event and it is compared to the previous transaction recorded on the blockchain.<br><br>Rewards Offered:<br>Storage peers are incentivised to store files and conduct keyword searches. Five reward functions can be coded into the smart contracts for use in the storage rental service:<br>• Deposit - transfer cryptocurrency as part of the smart contract.<br>• Add Charge - sends a defined amount of cryptocurrency for each file added.<br>• Search Charge - sends a defined amount of cryptocurrency for each keyword search.<br>• Judge - a fee for conducting a fair judgment operation.<br>• Finalise - when the duration of the file storage is complete, and the smart contract has returned the files. |

| Item | Data Item | Description |
|---|---|---|
| | | Blockchain Optimisation:<br>As records cannot be removed from a blockchain, the search is an overhead. A global time variable is used to identify a smart contracts' status. As a result, a search will start from the earliest valid record using this variable and ignore any older expired records. |
| **I11** | Future Extensions | None specified. |

# 4. Execution

## 4.1. Analysis and Findings

The systematic literature review was conducted using the eight-step guide developed for Information System by Okoli and Schabram (2010). The process consisted of: identifying the purpose of the review and generating the research question, developing a protocol, searching for the literature, conducting a screen of the search results to identify relevant papers, appraising the quality of the papers to be selected and eliminating those of low quality, extracting pertinent data and synthesising the same data. The out from this rigorous undertaking was four papers. The final step of Okoli's Schabram's guide is this section, writing the review.

A cursory search of the databases used in this systematic literature review (listed in section 2.2.1.) for the word 'blockchain' returns only 8628 results (Appendix 3). Surprisingly, only 4 (0.046%) of these papers where identified as relevant to this systematic literature review during the rigorous systematic process conducted for this dissertation. The results highlight a distinct lack of research into personal data management that incorporates blockchain technology and an opportunity for further research.

There is currently a significant issue concerning the security of personal data due to theft and harvesting. Who can access the individual's private data and how to prevent malicious parties from accessing the data is a challenge. Zyskind, Nathan and Pentland published the paper "Decentralizing Privacy: Using Blockchain to Protect Personal Data" in 2015. The paper was the first to propose the use of blockchain technology for individuals to overcome data privacy concerns and can be identified as the source for the idea behind this dissertation. Their paper focuses on providing a solution incorporating blockchain technology for individuals to manage personal data. Zyskind, Nathan and Pentland (2015, pp 180) proposed a hybrid model that combined blockchain with "off-blockchain storage" to "construct a personal data management platform".

In their paper "Make users own their data: a decentralized personal data store prototype based on Ethereum and IPFS", Alessi et al. (2018) propose a model that includes a decentralised identity manager (DIM) that is within desktop computer software or a smartphone application. The DIM is a user-controlled and responsible for managing a user's identity and profile, and for enabling the contextualisation of data depending on the service that the data with which the data is shared.

Yan, Gan and Riad (2017) suggest BlockChain based Personal Data Store (BC-PDS), a framework built on the existing OpenPDS/SafeAnswers model. The OpenPDS/SafeAnswers platform is off-blockchain and permits the collection and storing of data, while also giving fine-grained access control to protect user data as per the Windhover principles. The addition proposed to the OpenPDS/SafeAnswers platform is the storing of metadata files on a blockchain, and an off-blockchain access management module.

The paper “Towards Trustworthy and Private Keyword Search in Encrypted Decentralized Storage” authored by Cai, Yuan and Wang (2017) proposed another hybrid solution. They introduce the concept of client and storage peers. The client peers are the device owned by the user, while the storage peers are hosts that are unknown to the client peer.

### *4.1.1. Information Privacy*

Clarke (1997) defined information privacy as “the interest an individual has in controlling, or at least significantly influencing, the handling of data about themselves”. Clarke (1999, pp 60) also put forward that “Information privacy refers to the claims of individuals that data about themselves should generally not be available to other individuals and organizations”. Therefore, to dissect Clarke’s views, three aspects should be considered; data ownership, data transparency, and access control.

It is interesting that whilst designing their platform, Zyskind, Nathan and Pentland (2015) identified three very similar issues faced by personal data management systems; 1) the importance of users owning and controlling their data; 2) transparency of how the data is used and who has access to it and the reasons why, and; 3) the need for fine-grained access control. The paper's findings appear to agree with Clarke's definition.

The framework Zyskind, Nathan and Pentland (2015) proposed in this paper uses an off-blockchain personal data store. When a third-party requests access to data, the individual can, if they want to share the data with a service, issue a transaction $T_{access}$ and a new shared compound identity is then created. The $T_{access}$ transaction writes a record to the blockchain ledger that contains a pointer for the data that remains in the off-blockchain personal data store. The third-party accesses the data using the compound identity. As the individual has full control over the compound identity and the third-party only have access rights assigned, there is data transparency and access control. The individual can issue a new $T_{access}$ to rescind access to the data, and therefore they retain data ownership. While this appears to meet the three properties garnished from Clarke's definition, the issue of ensuring that the third-party deletes data that it has accessed is still present.

Alessi et al. (2018) have identified the lack of awareness of whom users share data with and what happens to the data once it is forwarded to a service. The lack of awareness could be due to a lack of understanding of the service's privacy settings (Acquisti, Brandimarte and Loewenstein, 2015) or misrepresentation of behavioural intentions (Stewart and Segars, 2002). To alleviate information privacy issues, the prototype offered by Alessi et al. (2018) involves hosting a Decentralised Identity Manager (DIM) on a local desktop computer or smartphone. The DIM is an application and is used to manage an individuals identity and profile. The identity and profile data get stored in the DIM and access are provided to datasets by uploading hashed profile flies to the InterPlanetary File System (IPFS), a decentralised peer-to-peer storage network. Using the DIM, the user can create a context relevant dataset to share with each service, and there is transparency as to who can view the data. Access to this data can also be removed. Clarke's definition criteria have been met, but there are risks involved. Alessi et al. (2018) inform that profiles can be deleted from the application. While this may be true, the profile will still be stored on an IPFS node as deletion is not a supported system function (IPFS, 2019). The data will be ever present on the IPFS.

Yan, Gan and Riad cite an increased public concern with regard to a user's privacy in a world of intelligence surveillance, big data and the ability to re-identify individuals from their metadata as the reason for their system. The openPDS (Open Personal Data Store) aspect of the BC-PDS (BlockChain based Personal Data Store) model suggested by Yan, Gan and Riad (2017), provides users with the ability to collect and store personal metadata. Individuals retain ownership of the metadata. The AutoNomy-based Access Structure module gives the user full access to their metadata and enables them to grant and remove access to data in a fine-grained manner through the distribution of pointers to where the data is stored. Therefore, there is transparency of what information third-parties can access. The individual's ownership of the metadata is also guaranteed using the General Secret Sharing Scheme for retrieval of data as covered in section 4.1.2. The three aspects of Clarke's definition are met with this solution. However, while the keys protect the metadata, and access can be revoked, the third-party may have stored a copy locally, and there is no method to guarantee that the third-party will delete it.

Cai, Yuan and Wang (2017) take a different approach. They were not as interested in preserving data privacy as others have already developed personal data management systems that already do it. Their concern was the application of searchable symmetric encryption to a personal data solution that utilises blockchain. Their system uses two types of peers. The first is a client peer and second is known as the storage peer. The user operates the client peer and has control and ownership of the personal data. To share the data the client peer negotiates a smart contract on the blockchain that contains the rental connection details and duration for the data to be stored. The user also controls access as they have the private key required to access the data file. Data transparency is present as the user knows whom they provide with keys.

### *4.1.2. Security of Data*

Encryption is applied to blockchain technology by default and through the consensus protocol blockchain are averse to malicious behaviour. However, when a blockchain is fledgeling or not used much, it is susceptible to a 51% attack where nefarious agents control the majority of nodes on the network (Bae and Lim, 2018; Zhu et al., 2019). Each of the papers will be looked int eh following subsections to see what the authors suggest as additional security measures.

With the proposed Zyskind, Nathan and Pentland (2015) hybrid model, data is partially secure. All transactions that incorporate the public blockchain only contain pointers to where the data is and they are encrypted, digitally signed, and protected by a compound identity that only authorised participants can access. The off-blockchain personal data store is key-value and implemented using Kademlia. Kademlia is a distributed hash table store that operates on nodes similar to blockchain technology. However, Kademlia is susceptible to Eclipse, Sybil, Churn, adversarial routing, denial of service and data storage attacks (Baumgart and Mies, 2007).

Alessi et al. (2018) do not provide specifics on security or encryption protocols in their papers. However, it does state that the user can set attributes within the data they share to public viewable or private. The context-dependent profile files that are shared with specified services are stored in a decentralised manner on the IPFS. They are uniquely hashed as a form of encryption and when updated a new hash is created. The Ethereum blockchain acts as a trustless intermediary. The records added to the Ethereum blockchain contain the hash of the user profile that acts as a pointer to where the file is stored on the IPFS.

Storing the data on the data on a blockchain rather than the original database that is incorporated into openPDS/SafeAnswers ensures that the data is encrypted in the BC-PDS framework devised by Yan, Gan and Riad (2017). The BC-PDS access structure enforces the use of the Threshold Secret Sharing (TSS) mechanism. TSS is based on Shamir's Secret Sharing algorithm, an established and effective means of protecting data through sharing parts of the data with multiple parties without any of them having full access. The retrieval of metadata uses a General Secret Sharing Scheme (GSS). The GSS is a collection of recovery algorithms that combines multiple TSS mechanisms into a tree using 'AND' and 'OR' logical functions. The metadata is shared amongst all authorised users, and with the secret sharing, all parties with access must cooperate in order to see the full metadata set. The use of GSS adds a layer of complexity and therefore ensures that the data is adequately protected.

The solution put forward by Cai, Yuan and Wang (2017) has a high level of security inbuilt. Each step in the process including the keyword search, creation of the smart contract, transferring, storing and retrieving the data file are encrypted and have validation checks included. The blockchain uses asymmetric cryptography to encrypt the smart contract and data file. There is a risk that a 51% attack (Bae and Lim, 2018; Zhu et al., 2019) on the blockchain until the node network becomes large enough through system use to make an attack too expensive to carry out.

### *4.1.3. Trustless*

The limited trust users have with service providers, and the manufacturers of internet-connected devices were identified in several studies (Golbeck, 2009; Christidis and Devetsikiotis, 2016; Porambage et al., 2016). There is also an issue with the lack of transparency (Christidis and Devetsikiotis, 2016) and this may lead to data theft or harvesting. Therefore, the inclusion of blockchain technology in any solution proposing to provide data privacy for individuals should remove the requirement of a central authority that stores the data, thus reducing the exposure to data theft and harvesting. The removal of the central authority results in trustless data sharing in a secure manner where the user does not necessarily know the party with whom they are sharing personal data. Thus, the implementation of the blockchain solution for each of the papers identified must be considered.

Zyskind's, Nathan's and Pentland's (2015) blockchain implementation and the storing of the data in a decentralised manner ensures that the proposed framework is trustless. Data is transferred between the individual, and potentially unknown third-party does not require a central authority that stores the personal data in one location or to verify the identity of either party to the transaction.

In the design by Alessi et al. (2015), all transactions to add and update the IFPS are conducted through the Ethereum blockchain. Using the Ethereum blockchain contains to store the username and hash of the profile as a pointer to its' location on the IPFS ensures that no central authority is required to store user data and distribute it.

Blockchain is used by Yan, Gan and Riad (2017) as a notary and this ensures that the solution is trustless. It does not require a central authority to verify the authenticity of data added or altered as the notary service provides this.

Cai, Yuan and Wang (2017) use blockchain smart contracts and record the data on the blockchain to guarantee that this system is trustless. In addition, a fair judgement module has been incorporated that can be triggered within a smart contract if either the client or storage peer suspects the other of malicious behaviour. Other nodes on the network can then arbitrate through the blockchain consensus protocol. The arbitration creates a double trustless environment that does not require a central authority and the risk of the central authority being compromised.

### *4.1.4. Challenges*

A study in 2014 estimated that the energy consumption for the Bitcoin network is A study in 2014 estimated that the energy consumption for the Bitcoin network is comparable to the energy consumption of Ireland (O'Dwyer and Malone, 2014).From the fourth quarter in 2014 to the first quarter of 2019, the size of the Bitcoin network has increased 7.5 times (Statista, 2019). Generating enough power to operate the network causes environmental damage.Introducing new blockchains or extending existing blockchains to operate personal data management solutions may be unsustainable. Factoring the cost in cryptocurrencies, it may not make economic sense.

Another challenge that to mention is that blockchain currently faces scalability issues and is unsuitable for large applications (Khan and Salah, 2018; Reyna et al., 2018). These issues include “delayed transaction confirmation, data retention, and communication failures” (Nofer et al., 2017).

## 5. Conclusion

The process of conducting a systematic literature review with rigour is viewed as “an original and valuable work of research in and of itself” (Okoli and Schabram, 2010, pp 1). An in-depth understanding of the research question is required, and the different boolean command structure nuances for performing comprehensive searches of each literature databases must be known. The process was a struggle at times due to a lack of exemplars in the field of information systems. With this systematic literature review being the first of its’ kind applied to blockchain technology, new ground was continually broken. The qualitative nature of the research papers, along with the complex cryptographic equations included within was a challenge. However, it is proposed that this systematic literature review can be viewed as “a solid starting point” for use by “other members of the academic community” (Okoli and Schabram, 2010, pp 1) as it provides the structure and repeatable steps necessary to obtain precise results.

From the beginning of the blockchain technology framework when Satoshi Nakamoto catapulted the Bitcoin blockchain and the Bitcoin currency into the world, few realised the potential of the technology. It was not until Vitalik Buterin created the Ethereum blockchain that the potential was clear. Since the launch of Ethereum, innovators have found extensive uses for smart contracts and in doing so have demonstrated that future opportunities are virtually limitless and enabled the research that was assessed for this dissertation.

Conducting the systematic literature review highlighted that to date, providing individuals with the ability to manage the personal data they own has been largely ignored by researchers when investigating novel applications that incorporate blockchain technology. Some papers did touch on the topic briefly, but it was an indirect result of examining how the Internet of Things would be affected by the implementation of blockchain technology. This systematic literature review considered the four most reliable papers in the final results.

While the four papers did propose models that met the rigour of the systematic literature review, none proposed a solution to the problem created when personal data is shared. Once personal data is exposed to a third-party service, there are no guaranteed means of retracting the data when the individual no longer wishes to share it. The services may store the data on their local databases and even distribute it unknowingly to other parties without the individual's consent. Preventing malicious behaviour of once approved services may be a challenge too far that may never be met.

## 5.1. Limitations of this Systematic Literature Review

The privacy of personal data is a highly researched area within information systems with many research papers and books available. Conducting this systematic literature review has identified that even though blockchain technology celebrated its' tenth anniversary in January 2019, and that it is starting to become a more mainstream technology, there is a distinct lack of research available. Through a cursory search (see Appendix 3 for results) only 8298 papers (not accounting for duplicates) were identified as referring to 'blockchain'. Through the rigorous systematic process conducted in this paper, only four papers were identified to be relevant to this systematic literature review.

Data-extraction bias has been identified as a limitation of this systematic literature review. Every effort was made to ensure that potential data-extraction bias did not occur by adhering to a robust and rigorous systematic approach as described in Section 2. However, the risk must be highlighted as only one person is defining the search, researching, assessing and appraising the papers involved. If this systematic literature review is used as “a solid starting point” by “other members of the academic community” (Okoli and Schabram, 2010, pp 1) in the future, this risk could be negated by building a team to conduct the systematic literature review.

Only academic publications in databases that were accessible to the author were considered for this dissertation. It is credible that other databases may contain other relevant publications. It is also possible that corporate entities are conducting research that would be relevant to this study and as they have not been published in academic journals, they are not included.

## 5.2. Further Research

In answering the research question, this systematic literature review is the first of its' kind and “constitutes an original and valuable work of research in and of itself”, and it “creates a solid starting point for all other members of the academic community” that are interested in this topic (Okoli and Schabram, 2010, pp 1). This starting point opens up opportunities for expanding on the work of this dissertation. It would be beneficial to academia to follow the process as defined in chapters 2, 3 and 4 and applying it to corporate sources, and other academic databases that were inaccessible for this dissertation.

As the rigorous process of conducting a systematic literature review identified only four papers that propose models leveraging blockchain to provide personal data stores, further research in this field provides an opportunity for researchers.

## References

Acquisti, A., Brandimarte, L. and Loewenstein, G. (2015) ‘Privacy and human behavior in the age of information’, *Science*, 347(6221), pp. 509–514. doi: 10.1126/science.aaa1465.

Alessi, M., Camillo, A., Giangreco, E., Matera, M., Pino, S. and Storelli, D. (2018). Make Users Own Their Data: A Decentralized Personal Data Store Prototype Based on Ethereum and IPFS. In: *2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech)*. [online] Split: FESB, University of Split, pp.1-7. Available at: https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=8430097 [Accessed 3 Mar. 2019].

Ali, R. (2014). Innovations in payment technologies and the emergence of digital currencies. *Bank of England Quarterly Bulletin*, (3rd Quarter), p.267.

Babar, M. and Zhang, H. (2009). Systematic Literature Reviews in Software Engineering: Preliminary Results from Interviews with Researchers." In: *3rd International Symposium on Empirical Software Engineering and Measurement*. [online] IEEE. Available at: https://ieeexplore.ieee.org/document/5314235 [Accessed 12 Mar. 2019].

Bae, J. and Lim, H. (2018). Random Mining Group Selection to Prevent 51% Attacks on Bitcoin. In: *2018 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops*. [online] Luxembourg City: IEEE Communications Society. Available at: https://ieeexplore-ieee-org.elib.tcd.ie/stamp/stamp.jsp?tp=&arnumber=8416225 [Accessed 4 Mar. 2019].

Barber S., Boyen X., Shi E., Uzun E. (2012). *Bitter to Better — How to Make Bitcoin a Better Currency. In:* Keromytis A.D. (eds) Financial Cryptography and Data Security. FC 2012. Lecture Notes in Computer Science, vol 7397. Springer, Berlin, Heidelberg.

Baumgart, I. and Mies, S. (2007). S/Kademlia: A practicable approach towards secure key-based routing. *2007 International Conference on Parallel and Distributed Systems*.

Buterin, V. (2014). *A NEXT GENERATION SMART CONTRACT & DECENTRALIZED APPLICATION PLATFORM*. [ebook] ethereum.org. Available at: https://www.weusecoins.com/assets/pdf/library/Ethereum_white_paper-a_next_generation_smart_contract_and_decentralized_application_platform-vitalik-buterin.pdf [Accessed 27 Dec 2018].

Cadwalladr, C. and Graham-Harrison, E. (2018). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach. *The Guardian*. [online] Available at: https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election [Accessed 14 Mar. 2019].

Cai, C., Yuan, X. and Wang, C. (2017). Towards trustworthy and private keyword search in encrypted decentralized storage. *2017 IEEE International Conference on Communications (ICC)*.

Christidis, K. and Devetsikiotis, M. (2016). Blockchains and Smart Contracts for the Internet of Things. *IEEE Access*, [online] 4, pp.2292-2303. Available at: https://ieeexplore.ieee.org/abstract/document/7467408/ [Accessed 30 Dec 2018].

Clarke, R. (1997). *Introduction to Dataveillance and Information Privacy, and Definitions of Terms*. [online] Rogerclarke.com. Available at: http://www.rogerclarke.com/DV/Intro.html [Accessed 2 Mar. 2019].

Clarke, R. (1999). Internet privacy concerns confirm the case for intervention. *Communications of the ACM*, 42(2), pp.60-67.

Cook, D., Augusto, J. and Jakkula, V. (2009). Ambient intelligence: Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5(4), pp.277-298.

DuckDuckGo.com (2017). *A Study on Private Browsing: Consumer Usage, Knowledge, and Thoughts*. [online] DuckDuckGo.com, p.7. Available at: https://duckduckgo.com/download/Private_Browsing.pdf [Accessed 13 Apr. 2019].

Fink, A. (2005). *Conducting research literature reviews*. 2nd ed. Thousand Oaks.: Sage Publications.

Fink, A. (2014). *Conducting research literature reviews*. 4th ed. Thousand Oaks: SAGE Publications, pp.3, 14.

Gartner, Inc. (2017). Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016. [online] Available at: https://www.gartner.com/newsroom/id/3598917 [Accessed 17 Jan 2019].

Golbeck, J. (2009). Computing with social trust. London: Springer, pp.1-7.

Hardjono, T. and Smith, N. (2016). Cloud-Based Commissioning of Constrained Devices using Permissioned Blockchains. *Proceedings of the 2nd ACM International Workshop on IoT Privacy, Trust, and Security - IoTPTS '16*.

Iansiti, M. and Lakhani, K. (2018). *The Truth About Blockchain*. [online] Harvard Business Review. Available at: https://hbr.org/2017/01/the-truth-about-blockchain [Accessed 30 Jan 2019].

IPFS. (2019). *IPFS is the Distributed Web*. [online] Available at: https://ipfs.io [Accessed 21 Apr. 2019].

Jing, Q., Vasilakos, A., Wan, J., Lu, J. and Qiu, D. (2014). Security of the Internet of Things: perspectives and challenges. *Wireless Networks*, 20(8), pp.2481-2501.

Khan, M. and Salah, K. (2018). IoT security: Review, blockchain solutions, and open challenges. *Future Generation Computer Systems*, 82, pp.395-411.

Kitchenham, B. and Charters, S. (2007). *Guidelines for performing Systematic Literature Reviews in Software Engineering (Version 2.3)*. EBSE Technical Report EBSE-2007-01. [online] Keele University and Durham University. Available at: http://community.dur.ac.uk/ebse/guidelines.php [Accessed 9 Mar. 2019].

Kokott, J. and Sobotta, C. (2013). The distinction between privacy and data protection in the jurisprudence of the CJEU and the ECtHR. *International Data Privacy Law*, [online] 3(4), pp.222-228. Available at: https://academic.oup.com/idpl/article/3/4/222/727206 [Accessed 15 Mar. 2019].

Moor, J. (1997). Towards a theory of privacy in the information age. *ACM SIGCAS Computers and Society*, 27(3), pp.27-32.

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. [online] Bitcoin.org. Available at: https://bitcoin.org/bitcoin.pdf [Accessed 27 Dec 2018].

Nofer, M., Gomber, P., Hinz, O. and Schiereck, D. (2017). Blockchain. *Business & Information Systems Engineering*, 59(3), pp.183-187.

O'Dwyer, K. and Malone, D. (2014). Bitcoin Mining and its Energy Footprint. 25th IET Irish Signals & Systems Conference 2014 and 2014 China-Ireland International Conference on Information and Communities Technologies (ISSC 2014/CIICT 2014). [online] Available at: http://mural.maynoothuniversity.ie/6009/1/DM-Bitcoin.pdf [Accessed 23 Apr 2019].

Okoli, C. and Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *SSRN Electronic Journal*.

Oxford English Dictionary. (2019). In: *Oxford English Dictionary*. [online] Oxford: Oxford University Press. Available at: https://en.oxforddictionaries.com/definition/reputation [Accessed 13 Apr. 2019].

Petticrew, M. and Roberts, H. (2006). *Systematic Reviews in the Social Sciences*. 1st ed. Malden (Mass.): Blackwell Publishing.

Porambage, P., Ylianttila, M., Schmitt, C., Kumar, P., Gurtov, A. and Vasilakos, A. (2016). The Quest for Privacy in the Internet of Things. *IEEE Cloud Computing*, 3(2), pp.36-45.

Post, R. (2001). Three Concepts of Privacy. *Georgetown Law Journal*, [online] 89(6), pp. 2087-2098. Available at: https://heinonline.org/HOL/P?h=hein.journals/glj89&i=2109 [Accessed 16 Feb. 2019].

Reyna, A., Martín, C., Chen, J., Soler, E. and Díaz, M. (2018). On blockchain and its integration with IoT. Challenges and opportunities. *Future Generation Computer Systems*, 88, pp.173-190.

Ridley, D. (2012). *The Literature Review*. 2nd ed. London: SAGE, pp.188-201.

Robles, T., Bordel, B., Alcarria, R., Sánchez-de-Rivera, D. (2018) ‘Blockchain Technologies for Private Data Management in AmI Environments’, *Proceedings*, 2(19), 1230, available: http://dx.doi.org/10.3390/proceedings2191230.

Ryan, G. (2010). *Guidance notes on planning a systematic review*. [ebook] Galway: James Hardiman Library, NUI Galway, p.1. Available at: https://www.tcd.ie/library/support/subjects/psychology/Guidance%20on%20planning%20a%20systematic%20review%20(2).pdf [Accessed 28 Mar. 2019].

Statista. (2019). *Bitcoin blockchain size 2010-2019 | Statistic*. [online] Available at: https://www.statista.com/statistics/647523/worldwide-bitcoin-blockchain-size/ [Accessed 22 Apr. 2019].

Smith, H., Milberg, S. and Burke, S. (1996). Information Privacy: Measuring Individuals' Concerns about Organizational Practices. *MIS Quarterly*, 20(2), p.167.

Stewart, K. and Segars, A. (2002). An Empirical Examination of the Concern for Information Privacy Instrument. *Information Systems Research*, 13(1), pp.36-49.

Swan, M. (2015). *Blockchain: Blueprint for a New Economy*. 1st ed. Sebastopol: O'Reilly & Associates, pp.1, 9, 27.

Szabo, N., 1994. Smart contracts. *Unpublished manuscript*.

Tapscott, D. (2018). *Blockchain Revolution The Internet of Value*. [online] Insightinvestment.com. Available at: https://www.insightinvestment.com/globalassets/documents/recent-thinking/uk-blockchain-revolution.pdf [Accessed 30 Dec 2018].

Techfak.uni-bielefeld.de. (n.d.). *Petri Nets*. [online] Available at: https://www.techfak.uni-bielefeld.de/~mchen/BioPNML/Intro/pnfaq.html [Accessed 14 Apr. 2019].

Tual, S. (2015). Ethereum Launches. [Blog] *Ethereum Blog*. Available at: https://blog.ethereum.org/2015/07/30/ethereum-launches/ [Accessed 23 Feb. 2019].

Yan, Z., Gan, G. and Riad, K. (2017). BC-PDS: Protecting Privacy and Self-Sovereignty through BlockChains for OpenPDS. *2017 IEEE Symposium on Service-Oriented System Engineering (SOSE)*.

Yang, Y., Wu, L., Yin, G., Li, L. and Zhao, H. (2017). A Survey on Security and Privacy Issues in Internet-of-Things. *IEEE Internet of Things Journal*, 4(5), pp.1250-1258.

Zhao, J., Fan, S. and Yan, J. (2016). Overview of business innovations and research opportunities in blockchain and introduction to the special issue. *Financial Innovation*, 2(1).

Zheng, Z., Xie, S., Dai, H., Chen, X. and Wang, H. (2016). *Blockchain Challenges and Opportunities: A survey*. [online] Inderscience Enterprises Ltd. Available at: https://www.researchgate.net/publication/319058582_Blockchain_Challenges_and_Opportunities_A_Survey [Accessed 20 Jan 2019].

Zhu, L., Wu, Y., Gai, K. and Choo, K. (2019). Controllable and trustworthy blockchain-based cloud data management. *Future Generation Computer Systems*, 91, pp.527-535.

Zyskind, G., Nathan, O. and Pentland, A. (2015). Decentralizing Privacy: Using Blockchain to Protect Personal Data. *2015 IEEE Security and Privacy Workshops*. [online] Available at: https://ieeexplore.ieee.org/abstract/document/7163223/ [Accessed 9 Feb 2019].

# Appendices

## Appendix 1 - All Papers for Practical Screen Review

This table contains all tables that were subjected to the Practical Screen. The highlighted rows contain the papers that passed the Practical Screen on all criteria.

*Table 14: Practical Screen Results*

| Paper Title | Authors | Between 1 Aug 2015 and 24 Mar 2019 | PDS/ PDM | Model / System / Framework / Scheme | Uses Blockchain | Does not Target Specific Industry |
|---|---|---|---|---|---|---|
| **IEEE Digital Explore** | | | | | | |
| An Identity Management System Based on Blockchain | Y. Liu; Z. Zhao; G. Guo; X. Wang; Z. Tan; S. Wang | PASS | PASS | PASS | PASS | PASS |
| An Online Identity and Smart Contract Management System | A. Yasin; L. Liu | PASS | PASS | PASS | PASS | PASS |
| BC-PDS: Protecting Privacy and Self-Sovereignty through BlockChains for OpenPDS | Z. Yan; G. Gan; K. Riad | PASS | PASS | PASS | PASS | PASS |
| Blockchain Privacy-Preservation in Intelligent Transportation Systems | L. Hîrtan; C. Dobre | PASS | PASS | PASS | PASS | **FAIL** |
| Blockchain-based Trusted Computing in Social Network | F. Dongqi; L. Fang | PASS | PASS | PASS | PASS | PASS |
| Decentralizing Privacy: Using Blockchain to Protect Personal Data | G. Zyskind; O. Nathan; A. Pentland | PASS | PASS | PASS | PASS | PASS |
| Design of privacy-preserving mobile Bitcoin client based on γ-deniability enabled bloom filter | K. Kanemura; K. Toyoda; T. Ohtsuki | PASS | PASS | PASS | PASS | **FAIL** |
| Make Users Own Their Data: A Decentralized Personal Data Store Prototype Based on Ethereum and IPFS | M. Alessi; A. Camillo; E. Giangreco; M. Matera; S. Pino; D. Storelli | PASS | PASS | PASS | PASS | PASS |
| Peer to Peer for Privacy and Decentralization in the Internet of Things | M. Conoscenti; A. Vetrò; J. C. De Martin | PASS | PASS | PASS | PASS | PASS |

| Paper Title | Authors | Between 1 Aug 2015 and 24 Mar 2019 | PDS/ PDM | Model / System / Framework / Scheme | Uses Blockchain | Does not Target Specific Industry |
|---|---|---|---|---|---|---|
| Poster: Towards Fully Distributed User Authentication with Blockchain | L. Zhang; H. Li; L. Sun; Z. Shi; Y. He | PASS | **FAIL** | PASS | PASS | PASS |
| Risk Management to Cryptocurrency Exchange and Investors Guidelines to Prevent Potential Threats | C. Y. Kim; K. Lee | PASS | **FAIL** | PASS | PASS | **FAIL** |
| The effect of a blockchain-supported, privacy-preserving system on disclosure of personal data | R. M. Frey; P. Bühler; A. Gerdes; T. Hardjono; K. L. Fuchs; A. Ilic | PASS | PASS | **FAIL** | PASS | **FAIL** |
| Towards a Secure and GDPR-Compliant Fog-to-Cloud Platform | S. Crompton; J. Jensen | PASS | **FAIL** | PASS | PASS | PASS |
| Towards Trustworthy and Private Keyword Search in Encrypted Decentralized Storage | C. Cai; X. Yuan; C. Wang | PASS | PASS | PASS | PASS | PASS |
| Transaction Immutability and Reputation Traceability: Blockchain as a Platform for Access Controlled IoT and Human Interactivity | D. W. Kravitz | PASS | **FAIL** | PASS | PASS | PASS |
| Transforming Face-to-Face Identity Proofing into Anonymous Digital Identity Using the Bitcoin Blockchain | D. Augot; H. Chabanne; O. Clémot; W. George | PASS | **FAIL** | PASS | PASS | PASS |
| WiP: A Novel Blockchain-Based Trust Model for Cloud Identity Management | K. Bendiab; N. Kolokotronis; S. Shiaeles; S. Boucherkha | PASS | **FAIL** | PASS | PASS | PASS |
| **Scopus** | | | | | | |
| A Novel Sustainable Interchain Network Framework for Blockchain | Q. Yang, H. Guo, V. Zhu, X. Fan, X. Cui, X. Kong, B.K. Bobby | PASS | PASS | PASS | PASS | PASS |
| The blockchain as a backbone of GDPR compliant frameworks | Hristov P., Dimitrov W. | PASS | **FAIL** | **FAIL** | PASS | PASS |
| **SpringerLink** | | | | | | |

| Paper Title | Authors | Between 1 Aug 2015 and 24 Mar 2019 | PDS/ PDM | Model / System / Framework / Scheme | Uses Blockchain | Does not Target Specific Industry |
|---|---|---|---|---|---|---|
| A Blockchain Implementation of an Attendance Management System | Jingyao TuZhenhua DuanCong TianNan ZhangYing Wu | PASS | **FAIL** | PASS | PASS | PASS |
| A Blockchain-Assisted Hash-Based Signature Scheme | Ahto BuldasRisto LaanojaAhto Truu | PASS | **FAIL** | PASS | PASS | **FAIL** |
| A Business-Oriented Schema for Blockchain Network Operation | Sheng HeChunxiao XingLiang-Jie Zhang | PASS | **FAIL** | PASS | PASS | **FAIL** |
| A Distributed Digital Asset-Trading Platform Based on Permissioned Blockchains | Rong WangWei-Tek TsaiJuan HeCan LiuEnyan Deng | PASS | **FAIL** | PASS | PASS | PASS |
| A Dynamic Scalable Blockchain Based Communication Architecture for IoT | Han QiuMeikang QiuGerard MemmiZhong MingMeiqin Liu | PASS | **FAIL** | PASS | PASS | PASS |
| A novel secure relay selection strategy for energy-harvesting-enabled Internet of things | Yan HuoMi XuXin FanTao Jing | PASS | **FAIL** | PASS | PASS | PASS |
| A Privacy-Preserving Networked Hospitality Service with the Bitcoin Blockchain | Hengyu ZhouYukun NiuJianqing LiuChi ZhangLingbo WeiYuguang Fang | PASS | **FAIL** | PASS | PASS | **FAIL** |
| A Secure and Targeted Mobile Coupon Delivery Scheme Using Blockchain | Yingjie GuXiaolin GuiPan XuRuowei GuiYingliang ZhaoWenjie Liu | PASS | **FAIL** | PASS | PASS | PASS |
| A Secure Provenance Scheme for Detecting Consecutive Colluding Users in Distributed Networks | Idrees AhmedAbid KhanAdeel AnjumMansoor AhmedMuhammad Asif Habib | PASS | **FAIL** | PASS | PASS | PASS |

| Paper Title | Authors | Between 1 Aug 2015 and 24 Mar 2019 | PDS/ PDM | Model / System / Framework / Scheme | Uses Blockchain | Does not Target Specific Industry |
|---|---|---|---|---|---|---|
| A Server-Assisted Hash-Based Signature Scheme | Ahto BuldasRisto LaanojaAhto Truu | PASS | **FAIL** | PASS | **FAIL** | PASS |
| A Vision for Trust, Security and Privacy of Blockchain | Wenshi Wang | PASS | **FAIL** | **FAIL** | PASS | PASS |
| An Associated Deletion Scheme for Multi-copy in Cloud Storage | DulinZhiwei ZhangShichong TanJianfeng WangXiaoling Tao | PASS | **FAIL** | PASS | **FAIL** | PASS |
| An Immunity-Based Security Threat Detection System for Cyberspace Digital Virtual Assets | Ping LinTao LiXiaojie LiuHui ZhaoJin YangFangdong Zhu | PASS | **FAIL** | PASS | **FAIL** | PASS |
| Blockchain Securities, Insolvency Law and the Sandbox Approach | Renato Mangano | PASS | **FAIL** | **FAIL** | PASS | **FAIL** |
| Blockchain-Based Fair Certified Notifications | Macià Mut-PuigserverM. Magdalena Payeras-CapellàMiquel A. Cabot-Nadal | PASS | **FAIL** | PASS | PASS | PASS |
| Blockchain-Based Privacy Preserving Deep Learning | Xudong ZhuHui LiYang Yu | PASS | **FAIL** | PASS | PASS | PASS |
| Blockchain-Based Solution for Proof of Delivery of Physical Assets | Haya R. HasanKhaled Salah | PASS | **FAIL** | PASS | PASS | PASS |
| Bootstrapping the Blockchain, with Applications to Consensus and Fast PKI Setup | Juan A. GarayAggelos KiayiasNikos LeonardosGiorgos Panagiotakos | PASS | **FAIL** | **FAIL** | PASS | PASS |
| Confidential and efficient asset proof for bitcoin exchanges | Maya MohanM K Kavitha DeviV Jeevan Prakash | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Data Acquisition and Analysis of Smart Campus Based on Wireless Sensor | Li Luo | PASS | **FAIL** | PASS | **FAIL** | **FAIL** |

| Paper Title | Authors | Between 1 Aug 2015 and 24 Mar 2019 | PDS/ PDM | Model / System / Framework / Scheme | Uses Blockchain | Does not Target Specific Industry |
|---|---|---|---|---|---|---|
| Decentralized Blacklistable Anonymous Credentials with Reputation | Rupeng YangMan Ho AuQiuliang XuZuoxia Yu | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Decentralized Voting: A Self-tallying Voting System Using a Smart Contract on the Ethereum Blockchain | Xuechao YangXun YiSurya NepalFengling Han | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Defend the Clique-based Attack for Data Privacy | Meng HanDongjing MiaoJinbao WangLiyuan Liu | PASS | **FAIL** | PASS | **FAIL** | **FAIL** |
| Designing Proof of Human-Work Puzzles for Cryptocurrency and Beyond | Jeremiah BlockiHong-Sheng Zhou | PASS | **FAIL** | **FAIL** | PASS | **FAIL** |
| Development of Means for the Formation of a Corporate Distributed Register (Blockchain) | A. Yu. Shcherbakov | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Distributed Random Process for a Large-Scale Peer-to-Peer Lottery | Stéphane GrumbachRobert Riemann | PASS | **FAIL** | PASS | PASS | **FAIL** |
| DLoc: Distributed Auditing for Data Location Compliance in Cloud | Mojtaba EskandariBruno CrispoAnderson Santana de Oliveira | PASS | PASS | PASS | **FAIL** | PASS |
| DStore: A Distributed Cloud Storage System Based on Smart Contracts and Blockchain | J. Xue; C. Xu; Y. Zhang; L. Bai | PASS | PASS | PASS | PASS | PASS |
| Embedding the MRC and SC Schemes into Trust Management Algorithm Applied to IoT Security Protection | Joy Iong-Zong Chen | PASS | **FAIL** | **FAIL** | **FAIL** | **FAIL** |
| Framework for Collaborative Software Testing Efforts Between Cross-Functional Teams Aiming at High Quality End Product | Prabal MahantaGeorg Bischoff | PASS | **FAIL** | PASS | **FAIL** | **FAIL** |

| Paper Title | Authors | Between 1 Aug 2015 and 24 Mar 2019 | PDS/ PDM | Model / System / Framework / Scheme | Uses Blockchain | Does not Target Specific Industry |
|---|---|---|---|---|---|---|
| Fully Distributed Indexing over a Distributed Hash Table | Simon DésaulniersAdrien BéraudAlexandre Blondin MasséNicolas Reynaud | PASS | **FAIL** | **FAIL** | **FAIL** | **FAIL** |
| Holistic Tracking of Products on the Blockchain Using NFC and Verified Users | Vanesco A. J. BoehmJong KimJames Won-Ki Hong | PASS | **FAIL** | PASS | PASS | PASS |
| Interacting with the Internet of Things Using Smart Contracts and Blockchain Technologies | Nikos FotiouVasilios A. SirisGeorge C. Polyzos | PASS | **FAIL** | PASS | PASS | PASS |
| IPFS-Blockchain-Based Authenticity of Online Publications | Nishara NizamuddinHaya R. HasanKhaled Salah | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Lifelogging Protection Scheme for Internet-Based Personal Assistants | David Pàmies-EstremsNesrine KaanicheMaryline LaurentJordi Castellà-RocaJoaquin Garcia-Alfaro | PASS | PASS | PASS | **FAIL** | PASS |
| Mapping Requirements Specifications into a Formalized Blockchain-Enabled Authentication Protocol for Secured Personal Identity Assurance | B. Leiding; A. Norta | PASS | PASS | PASS | PASS | PASS |
| Novel architectures and security solutions of programmable software-defined networking: a comprehensive survey | Shen WangJun WuWu YangLong-hua Guo | PASS | **FAIL** | **FAIL** | **FAIL** | **FAIL** |
| On and Off-Blockchain Enforcement of Smart Contracts | Carlos Molina-JimenezEllis SolaimanIoannis SfyrakisIrene NgJon Crowcroft | PASS | **FAIL** | PASS | PASS | PASS |

| Paper Title | Authors | Between 1 Aug 2015 and 24 Mar 2019 | PDS/ PDM | Model / System / Framework / Scheme | Uses Blockchain | Does not Target Specific Industry |
|---|---|---|---|---|---|---|
| Privacy Dashcam – Towards Lawful Use of Dashcams Through Enforcement of External Anonymization | Paul WagnerPascal BirnstillErik KrempelSebastian BretthauerJürgen Beyerer | PASS | PASS | PASS | **FAIL** | **FAIL** |
| Privacy-Preserving Public Auditing for Non-manager Group Shared Data | Longxia HuangGongxuan ZhangAnmin Fu | PASS | **FAIL** | PASS | PASS | PASS |
| Privacy-Preserving Trade Chain Detection | Stefan WüllerMalte BreuerUlrike MeyerSusanne Wetzel | PASS | **FAIL** | PASS | **FAIL** | **FAIL** |
| Research on Cross-Chain Technology Based on Sidechain and Hash-Locking | Liping DengHuan ChenJing ZengLiang-Jie Zhang | PASS | **FAIL** | PASS | PASS | PASS |
| Scalable and Privacy-Preserving Data Sharing Based on Blockchain | Bao-Kun ZhengLie-Huang ZhuMeng ShenFeng GaoChuan ZhangYan-Dong LiJing Yang | PASS | PASS | PASS | PASS | PASS |
| Security Risk Management in the Aviation Turnaround Sector | Raimundas MatulevičiusAlex NortaChibozur UdokwuRein Nõukas | PASS | **FAIL** | **FAIL** | **FAIL** | **FAIL** |
| SIoTFog: Byzantine-resilient IoT fog networking | Jian-wen XuKaoru OtaMian-xiong DongAn-feng LiuQiang Li | PASS | **FAIL** | PASS | **FAIL** | **FAIL** |
| Smart Grid Power Trading Based on Consortium Blockchain in Internet of Things | Dong ZhengKaixin DengYinghui ZhangJiangfan ZhaoXiaokun ZhengXinwei Ma | PASS | **FAIL** | PASS | **FAIL** | **FAIL** |

| Paper Title | Authors | Between 1 Aug 2015 and 24 Mar 2019 | PDS/ PDM | Model / System / Framework / Scheme | Uses Blockchain | Does not Target Specific Industry |
|---|---|---|---|---|---|---|
| Smart Papers: Dynamic Publications on the Blockchain | Michał R. HoffmanLuis-Daniel IbáñezHuw FryerElena Simperl | PASS | **FAIL** | PASS | PASS | **FAIL** |
| SRS-LM: differentially private publication for infinite streaming data | Hao WangKaiju Li | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Strain: A Secure Auction for Blockchains | Erik-Oliver BlassFlorian Kerschbaum | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Strong anonymous mobile payment against curious third-party provider | Chenglong CaoXiaoling Zhu | PASS | PASS | PASS | PASS | **FAIL** |
| Survey and Analysis of Cryptographic Techniques for Privacy Protection in Recommender Systems | Taiwo Blessing OgunseyiCheng Yang | PASS | PASS | **FAIL** | **FAIL** | **FAIL** |
| TARE: Topology Adaptive Re-kEying scheme for secure group communication in IoT networks | Anshul AnandMauro ContiPallavi KaliyarChhagan Lal | PASS | **FAIL** | PASS | **FAIL** | **FAIL** |
| The Bitcoin Backbone Protocol with Chains of Variable Difficulty | Juan GarayAggelos KiayiasNikos Leonardos | PASS | **FAIL** | **FAIL** | PASS | **FAIL** |
| Towards a Blockchain-Based SD-IoV for Applications Authentication and Trust Management | Léo MendibooureMohamed Aymen ChaloufFrancine Krief | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Towards Efficient and Secure Encrypted Databases: Extending Message-Locked Encryption in Three-Party Model | Yuuji FurutaNaoto YanaiMasashi KarasakiKatsuhiko EguchiYasunori IshiharaToru Fujiwara | PASS | PASS | PASS | **FAIL** | PASS |
| Towards the Blockchain Technology for Ensuring the Integrity of Data Storage and Transmission | Michał PawlakJakub GuziurAneta Poniszewska-Marańda | PASS | **FAIL** | **FAIL** | PASS | PASS |

| Paper Title | Authors | Between 1 Aug 2015 and 24 Mar 2019 | PDS/ PDM | Model / System / Framework / Scheme | Uses Blockchain | Does not Target Specific Industry |
|---|---|---|---|---|---|---|
| Towards the Blockchain Technology for System Voting Process | Michał PawlakJakub GuziurAneta Poniszewska-Marańda | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Transparent Personal Data Processing: The Road Ahead | Piero BonattiSabrina KirraneAxel PolleresRigo Wenning | PASS | PASS | **FAIL** | **FAIL** | PASS |
| Using Blockchains to Strengthen the Security of Internet of Things | Charalampos S. Kouzinopoulos Georgios SpathoulasKonstantinos M. GiannoutakisKonstantinos VotisPankaj PandeyDimitrios TzovarasSokratis K. KatsikasAnastasija CollenNiels A. Nijdam | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Using Economic Risk to Model Miner Hash Rate Allocation in Cryptocurrencies | George BissiasBrian N. LevineDavid Thibodeau | PASS | **FAIL** | PASS | PASS | **FAIL** |
| Verifiable Delay Functions | Dan BonehJoseph BonneauBenedikt BünzBen Fisch | PASS | **FAIL** | **FAIL** | PASS | PASS |
| Watermarking Public-Key Cryptographic Functionalities and Implementations | Foteini BaldimtsiAggelos KiayiasKaterina Samari | PASS | **FAIL** | PASS | **FAIL** | PASS |

## Appendix 2 - *Quality Appraisal Rationale: Second Pass*

A full secondary Quality Appraisal was conducted on the papers eliminated from the first pass to ensure that the papers were correctly eliminated.

*Table 15: Quality Appraisal Rationale Second Pass*

| **Paper No.** 1 | | |
|---|---|---|
| **Paper Title:** A Novel Sustainable Interchain Network Framework for Blockchain | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Does the paper propose a solution incorporating blockchain for use by individuals to manage personal data? | N | This paper is not explicitly providing a solution for individuals to manage their data. The solution can be used for the transfer of any data type between two or more blockchains. |
| Does the paper identify the problem that it aims to solve? | Y | Quote: "In this paper, we provide the design of unitary, a novel blockchain sustainable interchain network framework that connects all possible blockchain networks in the future through the decentralisation." |
| Is the structure of the paper provided? | Y | Yes, at the end of chapter 1. |
| Does the paper contain a general overview of the proposed solution? | Y | Yes, chapters 3 and 4 provide an overview of the Unitary interchain network protocol (UINP). |
| Does the paper outline a high-level system design? | Y | Yes, chapters 3 and 4 provide a high-level system design for the Unitary interchain network protocol (UINP). |
| Does the paper provide a detailed system design? | N | No detailed system design provided in the paper. |
| Has an evaluation test case or proof of concept been provided in the paper? | P | Chapters 2 and 5 mention a real-world application that uses UINP. However, no references are provided. An internet search reveals that UINP has a website http://uinp.io. However, no evidence of a real-world application was returned in the search results. |
| Has a production implementation process described in the paper? | N | No evidence is provided for an implementation process in the paper. |
| Does the paper identify any limitations with the proposed system? | N | No limitations are provided in the paper. |
| Does the paper identify future expansions of the solution or discuss further research? | P | A brief conclusion suggests that UINP can be used for data management and with other applications that have yet to be developed. It lacks specifics. It lacks specifics. |

| **Paper No.** 1 | | |
|---|---|---|
| **Paper Title:** A Novel Sustainable Interchain Network Framework for Blockchain | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Is the proposal contained in the paper credible? | Y | Blockchain technology requires software inter-connectors in order to communicate with other blockchains. There is no valid reason why this blockchain solution would not be plausible. |

| **Paper No.** 2 | | |
|---|---|---|
| **Paper Title:** An Identity Management System Based on Blockchain | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Does the paper propose a solution incorporating blockchain for use by individuals to manage personal data? | N | Proposed is an identity authentication and reputation management system. |
| Does the paper identify the problem that it aims to solve? | Y | Highlighted is the issue of personal information being “misused or leaked and financial assets been hacked” while it is stored on central servers. |
| Is the structure of the paper provided? | N | No structure is provided in this paper. |
| Does the paper contain a general overview of the proposed solution? | Y | A detailed overview of the various concepts involved is provided in section III. |
| Does the paper outline a high-level system design? | Y | This is provided in conjunction with the detailed design. |
| Does the paper provide a detailed system design? | Y | Section IV contains detailed computations and protocols. |
| Has an evaluation test case or proof of concept been provided in the paper? | Y | Section V conducts experiments on Identity Authentication, Identity Modification, Reputation Fluctuation (Rpf) Validation, Reputation Task, and Incentive Task. |
| Has a production implementation process described in the paper? | P | The authors plan to evaluate and improve the proposed system by conducting experiments on a large scale and using real data in Ethereum blockchains. |
| Does the paper identify any limitations with the proposed system? | N | No limitations are identified in the paper. |
| Does the paper identify future expansions of the solution or discuss further research? | Y | The authors plan to investigate more complex reputation voting. |
| Is the proposal contained in the paper credible? | Y | The proposal is credible. |

| **Paper No.** 3 | | |
|---|---|---|
| **Paper Title:** An Online Identity and Smart Contract Management System | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Does the paper propose a solution incorporating blockchain for use by individuals to manage personal data? | N | The proposal is for the Tsinghua University User Reputation System (TURS) that can be used to identify individuals in various fields. It uses online behaviour to identify and rate individuals. This data is collected from social/online media, manually entered and browser history. It is not aimed at the individual managing their data. |
| Does the paper identify the problem that it aims to solve? | Y | Section I, the introduction discusses the need for users to identify clients or other people of interest based on their reputation. |
| Is the structure of the paper provided? | N | No structure is provided in this paper. |
| Does the paper contain a general overview of the proposed solution? | Y | Section VI, VII and VIII provide an overview of the solution. |
| Does the paper outline a high-level system design? | Y | Section VIII provides diagram of how the system would work and some reputation calculations. |
| Does the paper provide a detailed system design? | N | The paper does not provide any detailed designs of how it will integrate with social/online media to collate personal information. |
| Has an evaluation test case or proof of concept been provided in the paper? | N | No evaluation or test case provided within the paper. |
| Has a production implementation process described in the paper? | N | No evidence is provided for an implementation process in the paper. |
| Does the paper identify any limitations with the proposed system? | N | No limitations are identified in the paper. |
| Does the paper identify future expansions of the solution or discuss further research? | P | Future research is highlighted. The authors propose that other researchers can optimise the online reputation, personality rating and professional rating lists. |

| **Paper No.** 3 | | |
|---|---|---|
| **Paper Title:** An Online Identity and Smart Contract Management System | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Is the proposal contained in the paper credible? | N | • Users could create a fake online presence to be used by the TURS system in order to keep their real personality separate, and so it does not harm their reputation score.<br>• The use of browsing history may fail as "32.9 ±2.3%" of people using computers and "31.5 ±3.9%" of mobile users surf the internet daily in private mode (DuckDuckGo.com, 2017, pp 7).<br>• The reputation attributes are a weakness as they do not abide by standard definitions of reputation e.g. "The beliefs or opinions that are generally held about someone or something" (Oxford English Dictionary, 2019). Examples that are not in sync with the dictionary meaning of reputation are: "No of views per month" and "No of tweets re-shared". |

| **Paper No.** 5 | | |
|---|---|---|
| **Paper Title:** Blockchain-based Trusted Computing in Social Network | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Does the paper propose a solution incorporating blockchain for use by individuals to manage personal data? | N | This paper proposes "a better encryption algorithm" for the model proposed in the paper "Decentralizing Privacy: Using Blockchain to Protect Personal Data" authored by Zyskind, Nathan and Pentland (2015). It does not propose a system. |
| Does the paper identify the problem that it aims to solve? | Y | Section IV. Improvement of the Platform discusses the replacement of the dynamic method of measuring the trust of a node and rewarding nodes that behave as proposed by Zyskind, Nathan and Pentland (2015). The replacement is a proof-of-credibility score that is calculated on the number of contracts the node is part of, combined with a proof-of-state score. |
| Is the structure of the paper provided? | Y | The structure of the paper is included in the last paragraph of section I; Introduction. |
| Does the paper contain a general overview of the proposed solution? | N | No overview is provided. |
| Does the paper outline a high-level system design? | P | Section IV provides a diagram as an indicator of a high-level system design. |
| Does the paper provide a detailed system design? | N | The paper does not provide any detailed designs. |

| **Paper No.** 5 | | |
|---|---|---|
| **Paper Title:** Blockchain-based Trusted Computing in Social Network | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Has an evaluation test case or proof of concept been provided in the paper? | Y | An Attack Situation Analysis is provided in section V. |
| Has a production implementation process described in the paper? | Y | No evidence is provided for an implementation process in the paper. |
| Does the paper identify any limitations with the proposed system? | Y | No limitations are identified in the paper. |
| Does the paper identify future expansions of the solution or discuss further research? | P | The paper proposes that more in-depth research and simulations should be concisered. |
| Is the proposal contained in the paper credible? | Y | The proposal is credible and does enhance the security of data transfer as it creates a method of avoiding what is known as a 51% attack, where a node has 51% or more of the computing power and therefore can provide consensus for a nefarious transaction on the blockchain. |

| **Paper No.** 7 | | |
|---|---|---|
| **Paper Title:** DStore: A Distributed Cloud Storage System Based on Smart Contracts and Blockchain | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Does the paper propose a solution incorporating blockchain for use by individuals to manage personal data? | N | It is a solution that leases unused cloud space to data owners that wish to store data distributed on blockchains in the cloud. |
| Does the paper identify the problem that it aims to solve? | Y | This is covered in detail in section 1 Introduction. |
| Is the structure of the paper provided? | Y | The structure of the paper is included in the last paragraph of section 1 Introduction. |
| Does the paper contain a general overview of the proposed solution? | Y | An overview is provided in section 1; Introduction. |
| Does the paper outline a high-level system design? | Y | A system model diagram with explanation is provided in section 3.1. |
| Does the paper provide a detailed system design? | Y | Algorithms for setup, genblock, subscribe, store, gencontract and audit are defined along with protocols required in section 4. |
| Has an evaluation test case or proof of concept been provided in the paper? | Y | In section 5, security is analysed and performance is evaluated to demonstrate the feasibility. |

**Paper No.** 7

**Paper Title:** DStore: A Distributed Cloud Storage System Based on Smart Contracts and Blockchain

| Assessment Question | Grade | Rationale for Grade |
|---|---|---|
| Has a production implementation process described in the paper? | N | No evidence is provided for an implementation process in the paper. |
| Does the paper identify any limitations with the proposed system? | N | No limitations are identified in the paper. |
| Does the paper identify future expansions of the solution or discuss further research? | Y | Expansion into the fields of data privacy and digital asset management are mentioned in section 6. |
| Is the proposal contained in the paper credible? | Y | There is much excess storage that remains unused in the cloud. It may be attractive to smaller companies in particular to lease out this space in order to reduce costs. |

**Paper No.** 9

**Paper Title:** Mapping Requirements Specifications into a Formalized Blockchain-Enabled Authentication Protocol for Secured Personal Identity Assurance

| Assessment Question | Grade | Rationale for Grade |
|---|---|---|
| Does the paper propose a solution incorporating blockchain for use by individuals to manage personal data? | N | Proposed is a potential replacement for public key infrastructure. |
| Does the paper identify the problem that it aims to solve? | Y | Risks associated with insufficient specifications, design flaws, privacy and security are risks to users when creating new security protocols. The authors believe that Authcoin can eliminate these risks. |
| Is the structure of the paper provided? | Y | The last paragraph of section 1 Introduction provides an outline of the paper structure. |
| Does the paper contain a general overview of the proposed solution? | Y | Section 2 provides an overview of Authcoin and how it works. |
| Does the paper outline a high-level system design? | Y | Section 3 and 5 provides model of the Authcoin and associated protocols. |
| Does the paper provide a detailed system design? | P | Some of section 5 delves into extra detail. |
| Has an evaluation test case or proof of concept been provided in the paper? | Y | Section 6 evaluates the coloured Petri net model. Petri nets are a “collection of directed arcs connecting places and transitions” that may contain hold authentication tokens (techfak.uni-bielefeld.de, n.d.). Coloured Petri nets are more complex versions of standard Petri nets. |
| Has a production implementation process described in the paper? | N | No evidence is provided for an implementation process in the paper. |

| **Paper No.** 9 | | |
|---|---|---|
| **Paper Title:** Mapping Requirements Specifications into a Formalized Blockchain-Enabled Authentication Protocol for Secured Personal Identity Assurance | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Does the paper identify any limitations with the proposed system? | N | No limitations are identified in the paper. |
| Does the paper identify future expansions of the solution or discuss further research? | Y | Section 7 mentions investigating blockchain specific implementation for better validation and authentication request creation process into the Authcoin model. |
| Is the proposal contained in the paper credible? | Y | The proposal is appears to be credible but a lack of implementation since the publishing of the paper may prove otherwise. |

| **Paper No.** 10 | | |
|---|---|---|
| **Paper Title:** Peer to Peer for Privacy and Decentralization in the Internet of Things | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Does the paper propose a solution incorporating blockchain for use by individuals to manage personal data? | N | The paper proposes a research idea. No research has been conducted on the proposed idea. |
| Does the paper identify the problem that it aims to solve? | Y | It identifies centralised storage of personal data as the issue. It is easier for the data to be hacked or appropriated by state actors. |
| Is the structure of the paper provided? | N | No structure is provided in this paper. |
| Does the paper contain a general overview of the proposed solution? | N | As the paper is a research idea it does not propose a researched solution. |
| Does the paper outline a high-level system design? | N | As the paper is a research idea it does not propose a researched solution. |
| Does the paper provide a detailed system design? | N | As the paper is a research idea it does not propose a researched solution. |
| Has an evaluation test case or proof of concept been provided in the paper? | N | As the paper is a research idea it does not propose a researched solution. |
| Has a production implementation process described in the paper? | N | As the paper is a research idea it does not propose a researched solution. |
| Does the paper identify any limitations with the proposed system? | Y | Scalability using blockchain technology is identified a barrier |
| Does the paper identify future expansions of the solution or discuss further research? | N | As the paper is a research idea it does not propose a researched solution. |

| **Paper No.** 10 | | |
|---|---|---|
| **Paper Title:** Peer to Peer for Privacy and Decentralization in the Internet of Things | | |
| **Assessment Question** | **Grade** | **Rationale for Grade** |
| Is the proposal contained in the paper credible? | N | As the paper is a research idea it does not propose a researched solution. |

| **Paper No.** 11 | | |
|---|---|---|
| **Paper Title:** Scalable and Privacy-Preserving Data Sharing Based on Blockchain | | |
| **Question** | **Grade** | **Rationale for Grade** |
| Does the paper propose a solution incorporating blockchain for use by individuals to manage personal data? | N | A multiparty, multi-layered system that uses the cloud, on-blockchain and off-blockchain (local storage) is proposed. Multiparty is where the data of multiple users is shared among all users. Each user has part of the data. This proposal is not designed for individuals. |
| Does the paper identify the problem that it aims to solve? | Y | The issue of storing private data in the cloud results in people losing control of their data. |
| Is the structure of the paper provided? | Y | The last paragraph of section 1 Introduction provides an outline of the paper structure. |
| Does the paper contain a general overview of the proposed solution? | Y | An overview is provided in section 1 Introduction. |
| Does the paper outline a high-level system design? | Y | Section 3 provides a high-level design and model diagram. |
| Does the paper provide a detailed system design? | Y | Section 4 identifies the protocols and explores the computations required. |
| Has an evaluation test case or proof of concept been provided in the paper? | Y | Experiments were conducted in a cloud storage environment. Section 5 details a privacy analysis. Section 6 contains analysis's of performance, efficiency (key generation, encryption, decryption), and the confirmation time and concurrent transactions relationships. |
| Has a production implementation process described in the paper? | N | No evidence is provided for an implementation process in the paper. |
| Does the paper identify any limitations with the proposed system? | N | No limitations are identified in the paper. |
| Does the paper identify future expansions of the solution or discuss further research? | P | The authors mention that they will focus on improving the system efficiency, and explore what other fields where the application may be used. |
| Is the proposal contained in the paper credible? | Y | The use of the Paillier cryptosystem is a proven method of securing data, and in this case, it enables one-to-one and one-to-many data sharing. |

## Appendix 3 - *Cursory Search of Databases for 'Blockchain*'

The results of a cursory search for papers published in journals and conference proceedings in the English language that reference 'blockchain'.

*Table 16: Cursory Search of Databases for 'Blockchain'*

| Repository/Database | Total Number of Papers Found Matching |
|---|---:|
| EBSCOhost | 1260 |
| IEEE Explore Digital Library | 1713 |
| SAGE | 133 |
| ScienceDirect | 1233 |
| SCOPUS | 2858 |
| SpringerLink | 1071 |
| **Total Papers Found** | **8268** |