DESIGNING INCENTIVES ENABLED DECENTRALIZED USER DATA SHARING FRAMEWORK

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
In Partial Fulfillment of the Requirements
For the Degree of Doctor of Philosophy
In the Department of Computer Science
University of Saskatchewan
Saskatoon

By

Ajay Kumar Shrestha

PERMISSION TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of the material in this thesis in whole or part should be addressed to:

Head of the Department of Computer Science
176 Thorvaldson Building
110 Science Place
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5C9
Canada

Or

Dean
College of Graduate and Postdoctoral Studies
116 Thorvaldson Building
110 Science Place
University of Saskatchewan
Saskatoon, Saskatchewan S7N 5C9
Canada

ABSTRACT

Data sharing practices are much needed to strike a balance between user privacy, user experience, and profit. Different parties collect user data, for example, companies offering apps, social networking sites, and others, whose primary motive is an enhanced business model while giving optimal services to the end-users. However, the collection of user data is associated with serious privacy and security issues. The sharing platform also needs an effective incentive mechanism to realize transparent access to the user data while distributing fair incentives. The emerging literature on the topic includes decentralized data sharing approaches. However, there has been no universal method to track who shared what, to whom, when, for what purpose and under what condition in a verifiable manner until recently, when the distributed ledger technologies emerged to become the most effective means for designing a decentralized peer-to-peer network. This Ph.D. research includes an engineering approach for specifying the operations for designing incentives and user-controlled data-sharing platforms. The thesis presents a series of empirical studies and proposes novel blockchains- and smart contracts-based DUDS (Decentralized User Data Sharing) framework conceptualizing user-controlled data sharing practices. The DUDS framework supports immutability, authenticity, enhanced security, trusted records and is a promising means to share user data in various domains, including among researchers, customer data in e-commerce, tourism applications, etc. The DUDS framework is evaluated via performance analyses and user studies. The extended Technology Acceptance Model and a Trust-Privacy-Security Model are used to evaluate the usability of the DUDS framework. The evaluation allows uncovering the role of different factors affecting user intention to adopt data-sharing platforms. The results of the evaluation point to guidelines and methods for embedding privacy, user transparency, control, and incentives from the start in the design of a data-sharing framework to provide a platform that users can trust to protect their data while allowing them to control it and share it in the ways they want.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisor Dr. Julita Vassileva for all the support, guidance, encouragement, and motivation during my Ph.D. journey. I am grateful that I had the chance to work under your supervision. Thank you so much for your mentorship, and for all the resources you provided me with.

I would also like to thank my advisory committee members: Dr. Ralph Deters, Dr. Gordon McCalla, Dr. Rajesh Karki (Cognate), Dr. Chanchal Roy and Dr. Jim Greer for your support, constructive feedback, and invaluable suggestions. My thanks also go to Dr. Stephen Marsh (External examiner), Dr. Ian McQuillan (Chair of Advisory Committee) and Dr. Jaswant Singh (Dean's Designated Chair).

I am deeply thankful to all the members of the Multi-User Adaptive Distributed Mobile and Ubiquitous Computing (MADMUC) Lab and Computer Science Department, who supported me in one way or the other throughout this journey. Thank you for creating a home away from home.

Finally, I would like to thank my parents and my loving wife, Sandhya for their continuous support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AP	Attitudinal Privacy
ATS	Attitude Towards System
AVE	Average Variance Extracted
BBS	Blockchain-Based System
BP-GC	Behavioral Privacy-General Caution
BP-TP	Behavioral Privacy-Technical Protection
CR	Composite Reliability
DHT	Distributed Hash Table
DLT	Distributed Ledger Technology
DPoS	Delegated Proof of Stake
DSR	Design Science Research
DSS	Data Sharing System
DUDS	Decentralized User Data Sharing
ERC	Ethereum Request for Comments
ETL	Extract, Transform and Load
EVM	Ethereum Virtual Machine
FBA	Federated Byzantine Agreement
FTP	File Transfer Protocol
GDPR	General Data Protection Regulation
HIPPA	Health Insurance Portability and Accountability Act
IBFT	Istanbul Byzantine Fault Tolerance
IOT	Internet of Things
IPFS	Inter-Planetary File System
IRB	Institutional Review Board
ITS	Intelligent Tutoring Systems
ITU	Intention to Use
JSON	JavaScript Object Notation
MVP	Minimum Viable Products
OECD	Organization for Economic Cooperation and Development
OPC	Office of the Privacy Commissioner of Canada
P	Privacy
PBFT	Practical Byzantine Fault Tolerance
PEnj	Perceived Enjoyment
PEOU	Perceived Ease of Use
PII	Personally Identifiable Information
PIPEDA	Personal Information Protection and Electronic Documents Act
PLS-SEM	Partial Least Square Structural Equation Modeling
PoET	Prof of Elapsed Time

PoW	Proof of Work
PoS	Proof of Stake
PU	Perceived Usefulness
QOS	Quality of System
rho_A	Dillon-Goldstein's rho
S	Security
SCS	Shopping Cart System
SMC	Squared Multiple Correlation
SNOMED-CT	Systematized Nomenclature of Medicine - Clinical Terms
T	Trust
TAM	Technology Acceptance Model
TRA	Theory of Reasoned Action
UCC	User- Created Content
UGC	User-Generated Content
UM	User Model
VAF	Variance Accounted For
XML	Extensible Markup Language
β or Std β	Direct Effect (Path Coefficient)

PUBLICATIONS/ARTICLES RELATED TO THE PH.D. THESIS

- **Shrestha, A. K.,** & Vassileva, J. (2016). Towards decentralized data storage in general cloud platform for meta-products. Proceedings of the International Conference on Big Data and Advanced Wireless Technologies - BDAW '16, 1–7. https://doi.org/10.1145/3010089.3016029.
- **Shrestha, A. K.**, Deters, R., & Vassileva, J. (2017). User-Controlled Privacy-Preserving User Profile Data Sharing based on Blockchain. In Future Technologies Conference (FTC) (pp. 31–40). https://saiconference.com/Downloads/FTC2017/Proceedings/3_Paper_127-User-Controlled_Privacy-Preserving_User_Profile_Data_Sharing.pdf. Conference Date: 2017/11, Vancouver, Canada.
- **Shrestha, A. K.,** & Vassileva, J. (2018). Blockchain-Based Research Data Sharing Framework for Incentivizing the Data Owners. In S. Chen, H. Wang, & L.-J. Zhang (Eds.), Blockchain -- ICBC 2018 (pp. 259–266). USA: Springer International Publishing. https://doi.org/10.1007/978-3-319-94478-4_19. Conference Date: 25-30 June 2018, Seattle, USA.
- **Shrestha, A. K.,** & Vassileva, J. (2018). Bitcoin Blockchain Transaction Visualization. IEEE. 2018 International Conference on Cloud Computing, Big Data and Blockchain, ICCBB 2018, pp. 1-6. https://doi.org/10.1109/ICCBB.2018.8756455. Conference Date: 15-17 November 2018, Fuzhou, China.
- **Shrestha, A. K.** (2019). Transparency and privacy: Empowering people through blockchain. Article published in The Conversation: http://theconversation.com/transparency-and-privacy-empowering-people-through-blockchain-104887.
- **Shrestha, A. K.**, & Vassileva, J. (2019). User acceptance of usable blockchain-based research data sharing system: An extended TAM-based study. Proceedings - 1st IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications, TPS-ISA 2019, pp. 203–208. https://doi.org/10.1109/TPS-ISA48467.2019.00033. Conference Date: 12-14 December 2019, Los Angeles, California, USA.
- **Shrestha, A. K.**, & Vassileva, J. (2019). User Data Sharing Frameworks: A Blockchain-Based Incentive Solution. 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2019, pp. 360-366. https://doi.org/10.1109/IEMCON.2019.8936137. Conference Date: 17-19 October 2019, UBC, Vancouver, Canada.
- **Shrestha, A. K.**, Vassileva, J., & Deters, R. (2020). A Blockchain Platform for User Data Sharing Ensuring User Control and Incentives. Frontiers in Blockchain, 3, 48. https://doi.org/10.3389/fbloc.2020.497985.

- **Shrestha, A. K.**, Joshi, S., & Vassileva, J. (2020). Customer Data Sharing Platform: A Blockchain-Based Shopping Cart. *IEEE International Conference on Blockchain and Cryptocurrency, ICBC 2020*. https://doi.org/10.1109/ICBC48266.2020.9169421. Demo, Conference Date: 3-6 May 2020, Toronto, Canada.
- **Shrestha, A. K.**, Vassileva, J., Joshi, S., & Just, J. (2021). Augmenting the technology acceptance model with trust model for the initial adoption of a blockchain-based system. PeerJ Computer Science, 7, e502. https://doi.org/10.7717/peerj-cs.502.

1 INTRODUCTION

1.1 Introduction and Motivation

The last decade has seen a revolutionary advancement in technological innovation and related research activity in collaborative approaches to sharing users' data among enterprises of similar interests (Shrestha & Vassileva, 2016). Numerous state-of-the-art incentive systems incorporate pricing mechanisms or reputation components to incentivize cooperation in networks (Xu & Van Der Schaar, 2014). The business models of most internet companies rely heavily on the personalization of their services and offers, based on data about individual users and their behaviors. User data are a valuable commodity shared among enterprises, allowing them to target users and customize or personalize offers to make them more effective. Similarly, in the scientific research domain, data sharing practices are much needed to maximize the collective knowledge gain from individual research efforts. Sharing research data among researchers accelerates discovery by reducing duplicative trials. In medicine and healthcare, both personalized patient care and medical research can benefit from sharing patient data and data from clinical trials (Lo & DeMets, 2016). Much of the data are contributed voluntarily by the user; others are obtained by the system from observation of user activities or inferred through advanced analysis of volunteered or observed data (Poslad, 2009).

In various domains, such as tourism, e-commerce, news aggregators, etc., data analytics and personalization enhance the users' interaction with the system and the overall quality of services offered to the users (Shrestha & Vassileva, 2019a). The applications that aim for personalization need to gather information about their users to adapt their functionality to the specific users' requirements (Proctor & Vu, 2002). The user information drawn in this way forms a user model. The process of user modeling requires collecting user data and making inferences from this data by finding patterns and similarities across the many users of a service or by abstracting user features and building user profiles from the history of the interaction of a user, which is a slow process. To speed it up, applications can share relevant data about the same user, leading to the need to share user interaction data and user profiles. There are various problems and trade-offs related to user

privacy, control over data, preserving the richness of data, as well as fairness, which have been addressed to a different extent by different existing architectures and methods.

The internet from its inception aimed to facilitate user data sharing, which was enabled through centralized (e.g., FTP and cloud file management systems) or decentralized (e.g., email) services. With the development of the social web or Web 2.0 (O'Reilly, 2005), it became very easy for users to share creative products on social sites (user-generated content on YouTube[1], Wikipedia[2], blogs, and microblogging tools like Twitter[3] and Facebook[4]).

Over the last 30 years, research in the user modeling community has developed architectures and ontologies to support personalization across applications – user modeling servers. Some of the early examples of user modeling can be seen in the late 1980s, for example, Kobsa and Wahlster (1989) and Finin (1989), derived from the need to offer better support for human-computer collaboration with a focus on human emulation (Fischer, 2001). The human emulation approach was initially used to provide better human-computer collaboration by recognizing the computer as something that has "human-like abilities." Flexible user models in dialog systems or conversational agents (Kobsa & Wahlster, 1989) could gradually develop background knowledge about users during an online interaction with the systems. The research trend then shifted towards the complementing approach (e.g., expert support system) due to the limited success of the emulating approach (Bobrow, 1991). In the meantime, a similar shift in the research direction also took place in the Intelligent Tutoring Systems (ITS) community, as described in Kay (2001). The complementing approach considers that computers are not human and aims to fill the gap between computers and humans by offering a human-centred design of interactions (Suchman, 1987). The complementing approach became more popular with commercial applications such as Lumiere prototypes (e.g., the Clippie), used for the Office Assistant in the Microsoft Office 97 suite (Horvitz et al., 1998). Later, technological advancement made personalization feasible by correlating the data of many users to find patterns of behaviors and generate recommendations for users based on the behaviors of other users. Since then, the major emphasis in user modeling has focused on

[1] https://www.youtube.com/
[2] https://www.wikipedia.org/
[3] https://twitter.com/
[4] https://www.facebook.com/

providing personalized services to users beyond the "one-size-fits-all" application design paradigm (Kobsa, 2001).

In parallel, the development of cloud-computing technologies has allowed users to share their data conveniently across devices and with each other. For example, federated learning[5] allows data mining scattered in distributed locations. It creates a collaborative learning platform for devices from a shared prediction model by allowing only the update to be uploaded to the model in the cloud while keeping the training data on the user's device and decoupling the ability to do machine learning from the need to store the data in the cloud.

However, there is no universal method to track who shared what, with whom, when, by what means and for what purposes in a verifiable fashion. There are intense debates on how credit should be awarded to the data owner for sharing their data. Privacy is another critical issue that needs to be addressed when sharing user data. The users' data are collected by different parties whose primary motive is an enhanced model while enabling the maximal research knowledge, scientific and commercial benefits, and giving their end-users or customers the best services. Those parties should take responsibility for protecting the personally identifiable information (PII) of their users. Data analytics significantly improve the quality of services, but they depend on collecting, sharing, and mining users' data. Moreover, the ownership of the data is transferred from a user to the enterprise and then into the entire network.

The existing data privacy and protection laws (e.g., PIPEDA in Canada, GDPR in the EU, Privacy Act in the USA, etc.) require informed consent from users and specify that enterprises must limit their collection to what is necessary for the identified purposes. To support regulatory compliance initiatives, every concerned enterprise and institution employs data governance procedures and implements policies to ensure data integrity, security, usability and availability. Poor data management can result in questions being raised on the data trustworthiness that can directly affect crucial operations involving decision making, services rendering, optimization and overall profitability. However, due to the need for applications to understand the context of the user to be able to provide good services, there is now a rapidly increasing need for user data sharing and use for a variety of changing purposes to improve cloud-based services personalization across mobile devices and the Internet of Things (IoT). For example, Google Home or Alexa needs

[5] https://ai.googleblog.com/2017/04/federated-learning-collaborative.html

extensive personal data to understand the current user's context and correctly interpret user voice commands; otherwise, the interaction with the device quickly becomes frustrating for the user. Another example is self-driving cars, which can be called by people "Uber-style" across different providers, could benefit from seamlessly sharing data about location, payment methods, user seating preferences, driving style settings, and favourite routes.

Often, the data owner's consent is required to be asked again, which can be obtrusive to a user who does not see what there is to gain. In addition, the consent forms are long and opaque; they give no option to users to select the data they are willing to share and prevent sharing other data; it is "take it or leave it." Users do not even read them but scroll down and click "Agree" since otherwise, they cannot use the service. Thus, it becomes hard or even impossible for users to remember what consent they have given to which enterprise and to keep track of who accesses their data and for what purpose. A flexible mechanism for obtaining and renewing consent for data use and sharing is required that provides appropriate and meaningful incentives to capitalize from data sharing and ensures transparency for users to be aware of which of their dataset has been accessed, by whom, for what purpose, and under what conditions.

Besides, people often wonder how they can facilitate communication without trusting anybody and possibly replacing the centralized controlling authority (Zuboff, 2019). For example, the future of ridesharing is indisputably moving from the traditional taxicab service and centralized ridesharing (Lyft, Uber) towards decentralized transportation with the drivers taking home most of the profits and not the platform, and which is also less expensive for the riders (Duy et al., 2018). It has been demonstrated that the creativity and the advancement of technology have given birth to many computational backbones to ensure privacy and a data-sharing model that includes cloud computing services, intelligent computing, etc. However, these services are often criticized with regards to security, centrality issues, and the credibility of the services being offered. Trust residing within centralized service providers for all the storage and management of data holds the risk of data being misused or sold to third parties and even destroyed if the service providers go bankrupt.

In the past few years, distributed ledger technologies, such as blockchain, have evolved as promising means to offer immutability features for trusted records in various use cases, including healthcare, agricultural research, tourism, etc. In addition, many blockchain systems provide a technology called "smart contract" that allows building automatic verification of the conditions for access or modification of each data entity. Smart contracts can be deployed to encode allowed

purposes of data use, allowed software apps, people or businesses who can access the data, time limitations, price for access, etc. Therefore, distributed ledger technologies provide a new type of platform as a possible solution to the privacy, user control and incentives problems of the common data governance models. A usable blockchain-based system enables sharing user data of both kinds (user models and user-contributed data). It can allow users to create a proof of ownership and provenance of their data, share data without losing control and ownership of it, provide/receive incentives for sharing and give users full transparency and control over who accesses their data, when, and for what purpose. The question arises of whether blockchain technology is scalable, mature, and stable enough to enable massive sharing of user data. The most frequent criticisms to blockchain-based approaches to date relate to their performance and scalability, yet the rapid development of the technology allows, through thoughtful combinations of blockchains, to achieve acceptable performance. A harder problem emerges related to the usability and user acceptance of blockchain technology for storing data, and in general, in non-crypto-currency domains.

1.2 Research Aim

This thesis aims to provide a usable and trusted framework for sharing three categories of user data: descriptive (user profiles, user behavior model data), user-contributed/owned document (data consciously created by the user), and research data (to produce original research results). In pursuit of this goal, the thesis reviews the latest developments in user-controlled access and sharing of user data through a survey of the most relevant literature of the related areas. Considering issues of security, privacy, user transparency and control, and incentives for data sharing, this research aims to develop novel contributions to the area by proposing blockchains and smart contracts technologies to support immutable and trusted records in various fields, including e-commerce, the research community, and tourism domains. To ensure the feasibility and usability of the proposed solution, the research also aims to examine the indicators that affect the user's acceptance of the system and analyze the performance. Another objective of the thesis is to conceptualize user-controlled data sharing practices and develop guidelines and methods for embedding trust, privacy, user transparency, security, control, and incentives for data sharing. I will propose a framework in this thesis to support a data-sharing platform that users can trust to protect their data. The

framework will also allow users to have control over their data and how it is shared—for what purposes, with whom and under what conditions.

1.3 Research Questions

The research questions as per the aim of this research can be formed as below:

- How can full transparency be offered over who accesses user data, when, and for what purpose?
- How can the users be allowed to specify the purposes of data sharing, what kinds of data can be shared, and which applications or companies can access the data?
- How can the users be provided with an incentive for sharing their data (in terms of payment through blockchains for the use of the data by applications, as specified by the contracts)?
- How can such platforms be constructed and then evaluated based on the performance metrics and the user experience models?
- How can the role of the *perceived usefulness* construct in relation to *perceived usability* and *quality of the system* factors be elucidated on such blockchain-based data-sharing platforms?
- How can the dimension and contribution of the *trust* construct concerning *perceived security* and *perceived privacy* factors be elucidated on such data-sharing platforms?
- How can the multidimensional constructs affect the attitude and behavioral intention of the user to adopt blockchain-based systems?

1.4 Research Objectives

Based on the aims of this research, the following research objectives are defined:

1. To conduct a literature review and background study:

   The first research objective is to review the most relevant literature in the field of user data-sharing platforms that incentivize users for their contribution to discover gaps in the existing systems and methods towards user-controlled privacy-preserving user data sharing approaches.

2. To define and classify user data:

   The second objective of this research is to develop a standard definition of user data and classify user data to support all kinds of user-profiles data, user documents (user-generated/created data) and research data.

3. To develop a decentralized user data sharing framework:

   To enable a decentralized systematic approach towards sharing user data, this research proposes a novel framework – the DUDS framework – to design blockchains and smart contracts-based Decentralized User Data Sharing (DUDS). The DUDS framework provides one possible solution for the main problems of the existing common data governance model: protection or privacy, security, access control, ownership, and rewards.

4. To analyze the preliminary factors that affect acceptance of a prototype system developed with the DUDS framework:

   To develop the user-controlled privacy-preserving user data sharing system, this research aims to design a user model that enables examining the role and dimensions of various antecedents of the behavioral intention of users to adopt the data sharing platforms such as the DUDS platform, with the help of an extended technology acceptance model (TAM). The results of the model are used as input to develop usable real-life blockchain-based applications for sharing user data that also incentivize the users to share their data. Furthermore, it also aims to explore the current state of the art to conceptualize the constituents of digital trust in the realm of security and attitudinal privacy for blockchain-based platforms and examine how they characterize the behavioral intention of users towards the adoption of such system.

5. To implement the DUDS framework:

   Three implementations of the DUDS framework for sharing user data are performed in example scenarios in three different domains: tourism, research, and online shopping cart.

6. To evaluate the user experience model and the DUDS framework:

   The thesis sets out to measure the performance metric in the private blockchain network along with evaluating the user experience model. The data-sharing activity requires exceptionally low latency for optimal performance. The user acceptance or experience

study on real-life applications is also essential to evaluate technological solutions and observe the effects of different variables using theory-backed models. The initial adoption of such blockchain-based systems is necessary for continued use of the services, but their user acceptance study with implemented applications has not been well investigated in the literature. This research tries to evaluate the user behavioral model and DUDS framework for its usefulness and trustworthiness through its performance metrics evaluation and user experience study on a real-life blockchain-based e-commerce application. The ultimate results are used as the basis to postulate guidelines and methods for incorporating security, privacy, user transparency, control, and incentives from the start in the design of the data-sharing framework.

1.5 Thesis Structure

The final thesis is organized into seven chapters. Chapter 2 starts with a focus on the research methodology adopted in the thesis. It presents a brief background on general Research Framework and Design Science Research. It is followed by an overview of distinct categories of user data that delves into user modeling, user-generated/created data, and research data. It then presents an extensive multidisciplinary review of the literature on user data sharing frameworks and approaches. Moreover, it gives some background on the heuristic behavioral model based on the technology acceptance model (TAM), privacy model, and trust model within the cognitive schemas. Finally, it presents a background of blockchain and smart contract technologies and their role in the integrated data-sharing model. Chapter 3 gives a detailed description of the DUDS framework. In chapter 4, the heuristic cognitive-behavioral model with reference to our published article entitled "User Acceptance of Usable Blockchain-Based Research Data Sharing System: An Extended TAM-Based Study" (Shrestha & Vassileva, 2019b) is presented. This model aims to identify constructs affecting an initial adoption of a prototype based on the DUDS framework. Furthermore, it conceptualizes the role of principal antecedents of digital trust—perceived privacy and perceived security on the behavioral intention to accept the DUDS framework. Chapter 5 focuses on the proof of concept of the DUDS framework applied in three different domains: tourism, research, and online shopping cart with reference implementations from my published articles "A Blockchain Platform for User Data Sharing Ensuring User Control and Incentives"

(Shrestha et al., 2020), "Blockchain-Based Research Data Sharing Framework for Incentivizing the Data Owners" (Shrestha & Vassileva, 2018a), and "Customer Data Sharing Platform: A Blockchain Based Shopping Cart" (Shrestha et al., 2020). Chapter 6 evaluates the DUDS framework for efficacy, trustworthiness, and success in scrutinizing results of the behavioral user model within cognitive schemas. Finally, Chapter 7 concludes the thesis, summarizes the main findings and contributions of this research, provides recommendations, and discusses limitations and future work on the topic of modeling and analysis of decentralized user data sharing frameworks.

Table 1. 1. Outline of the thesis chapters

Chapter	Title	Description of Chapter Contents
Chapter 2	Research Background	Research Framework, Design Science Research, Methodology investigated and adopted in the thesis; Defines distinct categories of the user data, user modeling, participative web and research data; Extensive multidisciplinary review of the literature on user data sharing frameworks and approaches, Extended TAM, Privacy Model, Trust Model
Chapter 3	Blockchain and Smart Contracts for Data Sharing: DUDS Framework	The architecture of the DUDS framework
Chapter 4	Extended TAM and Conceptualizing Trust Article: "User Acceptance of Usable Blockchain-Based Research Data Sharing System: An Extended TAM-Based Study"	Heuristic cognitive-behavioral model to identify constructs that affect the adoption of a prototype user data sharing system based on DUDS framework; Conceptualizes the role of perceived privacy and perceived security of user trust construct on the behavioral intention towards accepting such blockchain-based DUDS platforms

Chapter 5	Proof of concept of the DUDS framework Article 1: “A Blockchain Platform for User Data Sharing Ensuring User Control and Incentives” Article 2: “Blockchain-Based Research Data Sharing Framework for Incentivizing the Data Owners” Article 3: “Customer Data Sharing Platform: A Blockchain Based Shopping Cart”	Proof of concept implementations of the DUDS framework in three different domains—tourism, research and online shopping cart; Performance evaluation.
Chapter 6	Evaluation of the DUDS framework for trustworthiness	Evaluation of the DUDS framework for efficacy and trustworthiness and success in scrutinizing results of the behavioral user model within cognitive schemas
Chapter 7	Conclusion and Future Work	Summary of the main findings and contribution made, Discussion on recommendations, limitations, and Future work on the topic of modeling and analysis of decentralized user data sharing frameworks.

2 RESEARCH BACKGROUND

Many potential methodologies and research paradigms are available in the literature for designing successful research. This chapter starts by providing a background on the research worldview and Design Science Research (DSR) approach adopted by my research. The chapter then presents the standard definition of user data and its classification based on the systematic investigation of the literature to obtain support for all kinds of user profiles and user-generated or created data. This research categorizes user data into three types: (1) User profile data, (2) User documents, and (3) Research data. This categorization is the first step toward realizing the objectives of my research to form a body of knowledge defining the terms, study scope and concerns/ factors to be taken into consideration. Subsequently, this chapter identifies the state of art in the field of sharing user data by reviewing different proposed architectures, frameworks, and methods. The focus is to review literature, investigate systematic approaches and study concepts towards the design of the decentralized user data sharing framework for incentivizing the data owners, thereby narrowing the study domain to a socio-technical approach. The research problems in the study domain are identified, and approaches to solve the problems are discussed. The background on the heuristic behavioral model is then presented based on the technology acceptance model (TAM), privacy model, and trust model within the cognitive schemas. Moreover, this chapter provides an overview of blockchain and smart contracts as promising tools to develop the DUDS (Decentralized User Data Sharing) framework. This analysis leads the chapter to delve more into details of various types of blockchains and smart contracts as the distributed ledger technologies.

2.1 Research Framework

Research may involve non-empirical or empirical studies. A non-empirical study is often regarded as theoretical research conducted without data that relies on previous theories, postulates, and logic to deduce new theoretical constructs. On the other hand, empirical studies usually carry out data collection or experiments and analyze quantitative and qualitative data. Quantitative research involves deductive approaches by measuring numerical data and examining the relationships between constructs to conceptualize an idea from a generalized principle (Creswell, 2014).

Qualitative research involves researchers performing inductive reasoning with non-numerical data such as reviews, interviews, observations, etc., to move from the specific sub-instances into a generalized conclusion (Fink, 2000). There are also mixed methods involving both quantitative and qualitative research approaches and incorporating entire exclusive designs that consist of theoretical frameworks and philosophical assumptions (Creswell, 2014).

Any research approach that tends to be quantitative, qualitative, or mixed usually falls in the intersection of the three components: philosophical worldviews or paradigm, research designs and research methods, as shown in Figure 2.1, which was adapted from (Creswell, 2014; Denzin & Lincoln, 2005; Greene et al., 1989; Mertens, 2005).

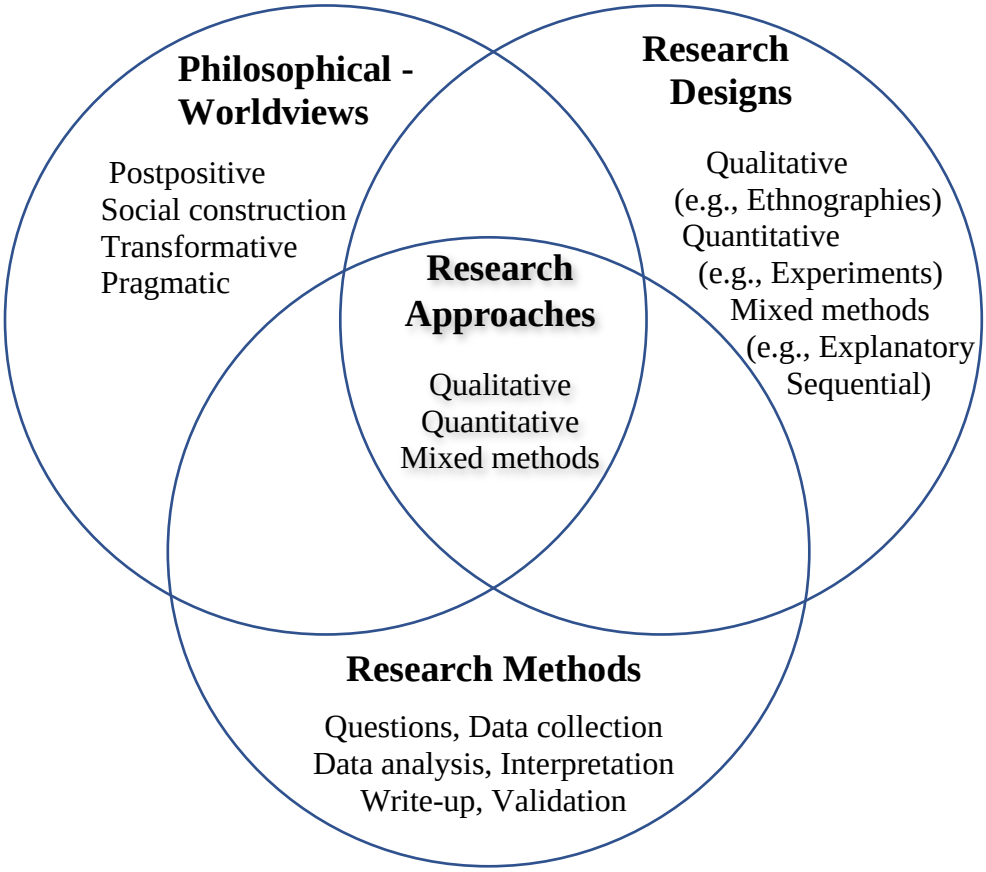

Figure 2. 1. Research framework (Creswell, 2014)

The research problem of my study requires investigating technical and user behavioral aspects of the decentralized user data sharing frameworks. Therefore, the mixed approach is suitable for my research. My research worldview follows a pragmatic philosophy as it involves breaking the problem into smaller parts and choosing an appropriate approach to deal with each part. In my

research, both quantitative and qualitative approaches are concurrently implemented in the same phase of the study, so the design element of my research follows a convergence strategy. Both the quantitative and qualitative elements have the same priority, and the research mixes the results at the interpretation phase of the method. The research method follows the mixed approach with both emerging and predetermined methods, both open and closed-ended questions, both textual and numeric data with statistical analysis.

2.2 Design Science Research

Design science research (DSR), as depicted in Figure 2.2, is the conceptual framework that combines behavioral science and design science worldviews to gain knowledge, understand the problem, execute, and evaluate the research (Hevner et al., 2004). DSR also follows the pragmatic worldviews allowing researchers to freely choose multiple methods from qualitative and quantitative assumptions. My research employs the mixed approach, so the adoption of DSR enables a systematic method towards solving the problems identified by my research. DSR has guidelines for conducting research and can be studied from three cycles – relevance cycle, design cycle and rigour cycle.

The relevance cycle ensures the novelty of my research approach, the design cycle assures the working of my proposed research approach, and the rigour cycle confirms that the research fits into the current application domain (Hevner et al., 2004). Figure 2.3 presents the activity diagram of my research incorporating the DSR framework, which was adapted from Offermann et al. (2009).

2.2.1 Relevance Cycle

DSR is intended to improve the environment by developing new artifacts (Hevner & Chatterjee, 2010). As shown in Figure 2.2, the environment defines the problem space consisting of people, organization, and technology. In the relevance cycle, my research aims to identify problems in the context of the decentralized user data sharing domain. Here, my research pre-evaluates the relevant state of the art and defines the objectives for a solution towards potential improvements in the application domain.

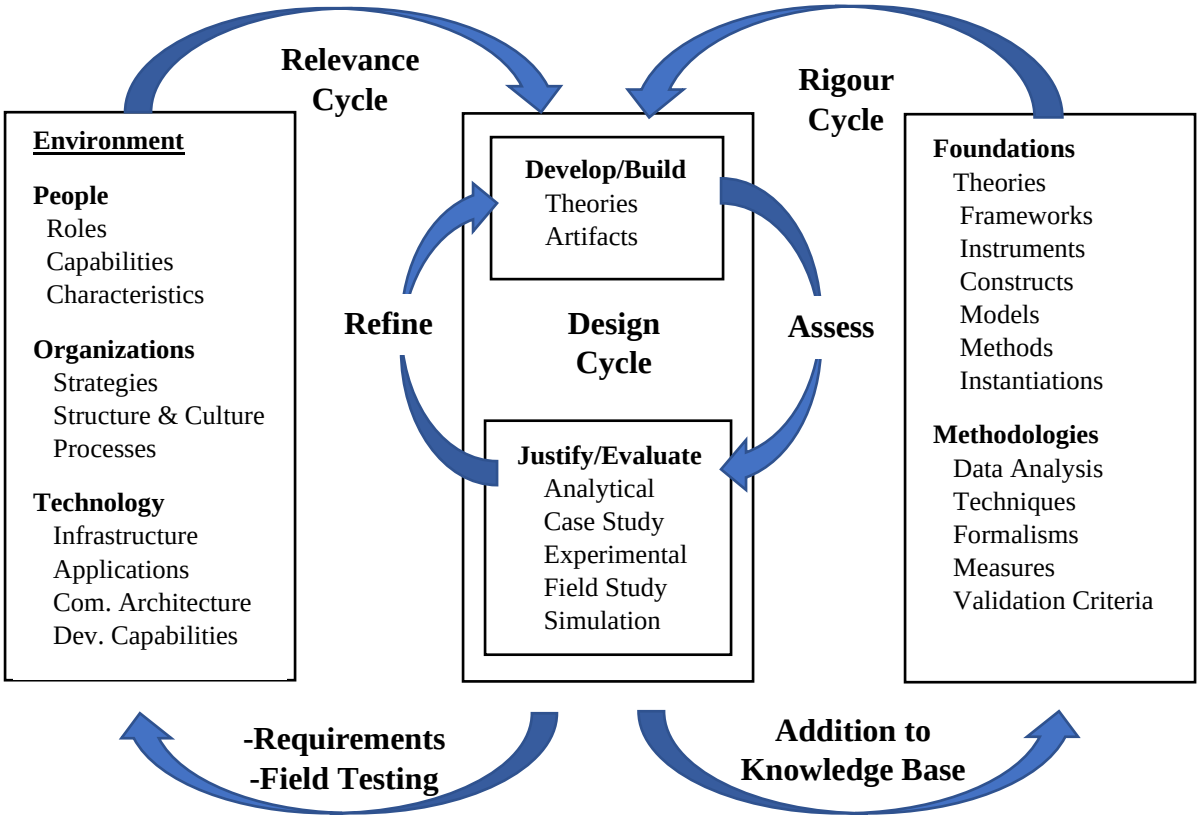

Figure 2. 2. Design science research cycle (Hevner, March, Park, & Ram, 2004)

2.2.2 Design Cycle

The design cycle of the DSR has the intention to produce a research artifact and ensure a balance between development and evaluation (Hevner & Chatterjee, 2010). This cycle deals with defining the user data and explaining the decentralized user data sharing (DUDS) framework. The DUDS framework is explained in Chapter 3.

Moreover, the design cycle performs content analysis on the findings of the empirical studies to design the artifact (*user behavior model)* that conceptualizes the constituents of digital trust for blockchains and smart contract-based platforms and examines the behavioral intention of users towards the adoption of such platforms. This leads to the collection of quantitative and qualitative data for the rigour cycle. Moreover, the artifact-*DUDS platform* will be implemented and validated through the reference implementations for sharing user data in different contexts that offer rewards to the data owners in terms of payment through blockchains for the use of the data by applications, as specified by the contracts.

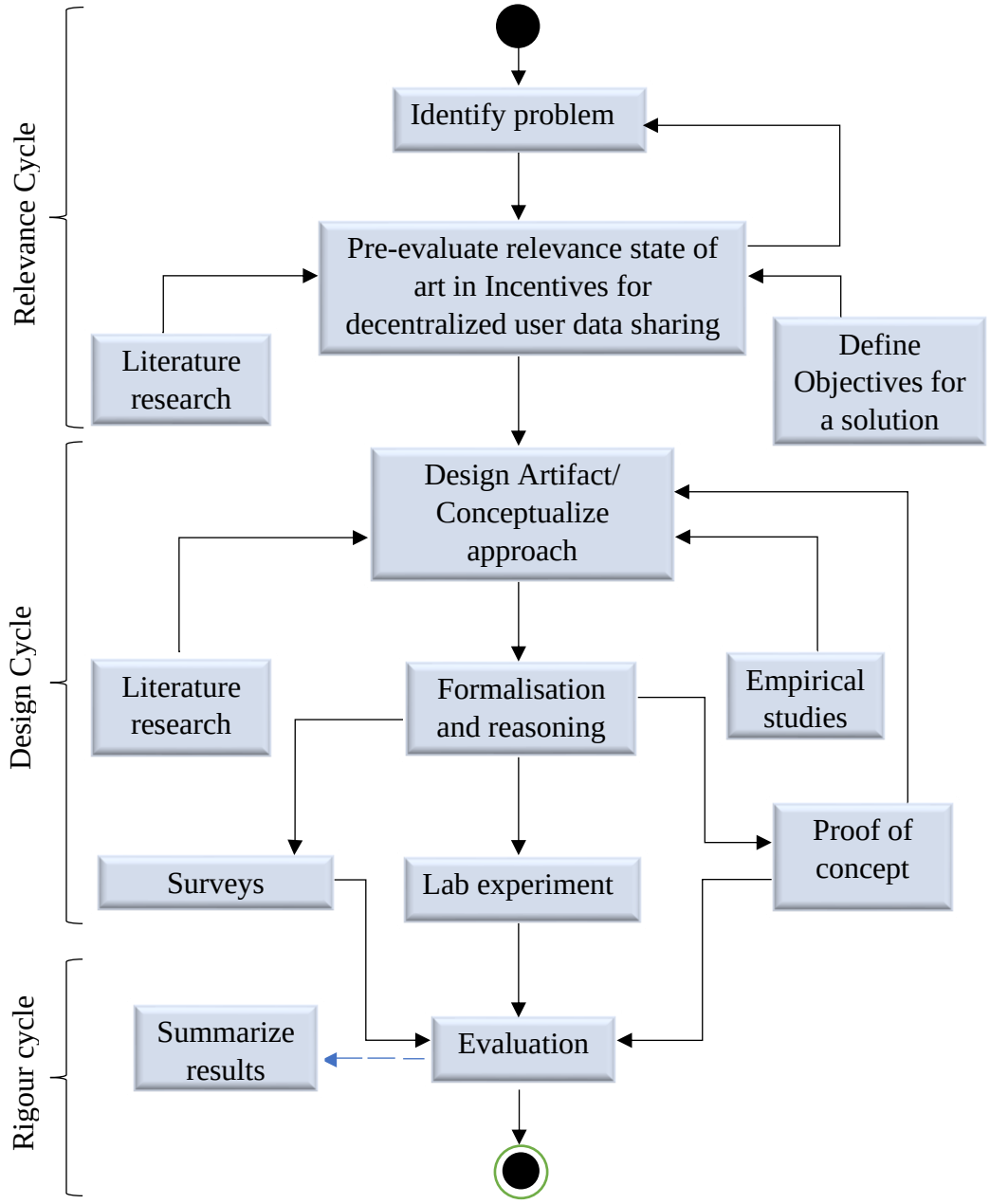

Figure 2. 3. Research methodology (Offermann et al., 2009)

2.2.3 Rigour Cycle

The rigour cycle of the DSR deals with the knowledge base of theories and engineering procedures (Hevner & Chatterjee, 2010) to evaluate the user behavioral model and DUDS framework for its usefulness, trustworthiness with the findings from the performance evaluations and user survey

studies. The ultimate results are compared with the available literature and application domain to ensure that they can be added to the knowledge base as the guidelines and methods for incorporating security, privacy, user transparency, control, and incentives from the start in the design of decentralized user data-sharing platforms.

2.3 User Data

This section presents the standard definition of user data and its classification based on the systematic investigation of the literature to obtain support for all kinds of user profiles and user-generated or created data. This step is the first in realizing the objectives of my research, as stated earlier in section 1.4. This research categorizes user data into three types: (1) User profile data, (2) User documents, and (3) Research data. Subsection 2.3.1 describes user profile data as the first category of the user data along with the discussion on the user modeling, personalization, sharing of the user model data, ownership status with regard to the data and incentives for the data sharing. The second category of user data is user documents, which are presented in subsection 2.3.2. Subsection 2.3.3 presents the last category of user data, research data. This section forms a body of knowledge constituting the meaning of the terms, study scope and concerns/factors to be taken into consideration for this research.

The data, resource, and user are three crucial elements of the data-sharing framework. According to Davoust (2015), the user–data relationship can be represented by a graph, as in Figure 2.4. The data model can be structured as in a resource description framework (RDF) or relational model, or semi-structured as in the JSON format. These technical details about the data model are important while storing the user data. Despite any data model, the data can be decomposed into nodes and edges or tuples, which are considered as a set of atomic resources connected by relationships in the user-data graph. These resources are created, read, updated, or deleted by the user. This implementation level description is not important to my work. The important enabler of the data-sharing framework is that the users are in control of the fixed resources, and its notion is strongly represented by the relationships in the graph, as in Figure 2.4.

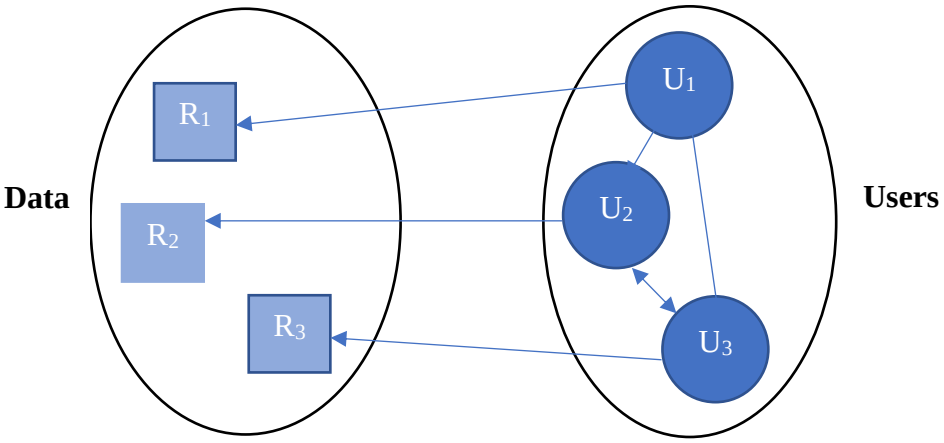

Figure 2. 4. User – Data model (Davoust, 2015)

2.3.1 User Profile Data

The characteristics of the user are defined as the user profile data, which shows who the users really are. According to Kiertz (2014), the user profile information can be collected in real-time through the application or outside the application and imported to the analytics platform. Some applications can anonymously collect a large amount of user profile data without requiring the user to sign up for the application. For example, an app like Google Weather can record home locations and other favourite places from the user without signing in. The user profile data can be divided into four categories, which are given below, along with some of their attributes (Kiertz, 2014):

1) Static data: Name, Birthday, Gender, Birthplace, Annual Income
2) Interest Categories: Hobbies, Sports Teams, Music Artists, News, Movies
3) In-app info: Registered User, Subscriber Type, Linked Social Network Account
4) Outside Channel attributes: Rewards Status, In-store Purchaser, Frequent Buyer Member

Therefore, user profile data is the explicit representation of users' identities, which usually exist on websites such as social networking sites, recommender systems, bulletin boards or any software applications. A user model is defined as the data structure used to obtain a person's characteristics and preferences, and the process involved in capturing the user model data is known as user modeling.

2.3.1.1 User Modeling

To share the user data, the proper collection and refinement of the data are two of the foremost important steps. In an unobstructed data collection process, a balance must be found between user privacy and the amount and nature of data that is required to deliver a better personalization (Kobsa, 2007). Successful personalization ensures the tailoring of improved services to fit the needs of individuals or specific groups of individuals. A breach of user privacy is a serious threat to successful user modeling and personalization.

The user modeling process is used to capture human-computer interactions to escalate the conceptual understanding of the user when interacting with an application (Biswas & Robinson, 2010).

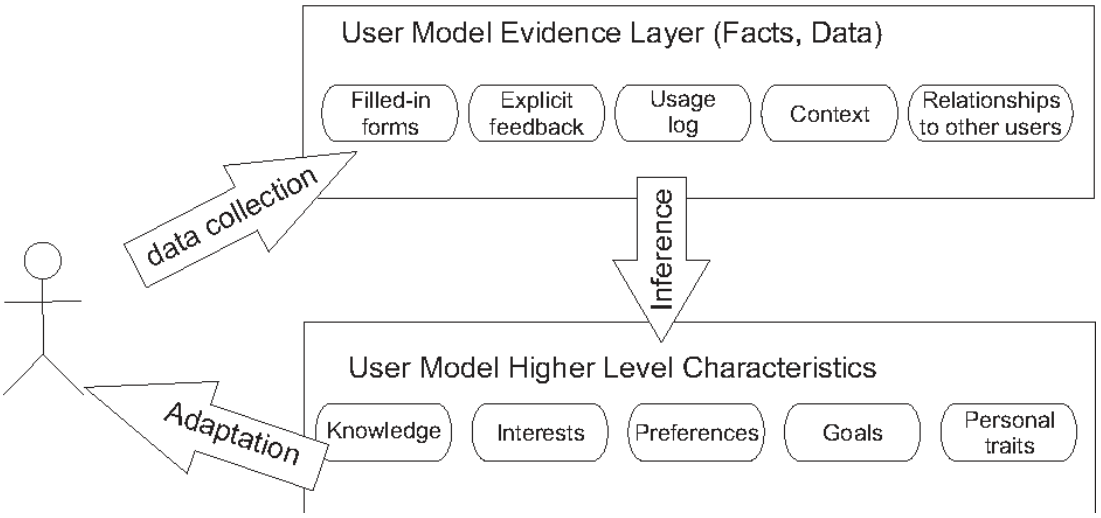

Figure 2. 5. User modeling life cycle process (Barla, 2011)

The user modeling components must be able to share user data/models. The overall process of user modeling has three prominent phases (Barla, 2011): *data collection, inference,* and *adaptation,* which are demonstrated in Figure 2.5.

The data collection phase has an adaptive system capable of capturing user profile data along with other useful data (Barla, 2011) through survey questionnaires and forms. The collected data are stored in the user model evidence layer. More specifically, the explicit data collection process is based upon the user profile data collected via questionnaires or forms (Fink & Kobsa, 2000). On the other hand, the implicit data collection process constructs the user models through observation of user activity within the system, feedback given by the user, logs of user activity or inference drawn based on historical knowledge about the user or user group (Brusilovsky, 2001).

The inference phase generates the estimates of the real user characteristics after an adaptive system processes the data from the user model evidence layer into the user model higher-level characteristics such as personal traits, preferences, etc. (Barla, 2011). Not all the trace data needs to be interpreted, or at least it can be interpreted in many ways depending on the context of use.

Often interpretation only happens at the time adaptation is undertaken. The final adaptation phase deals with the actual use of the evidence layer and higher-level characteristics layer so that it can offer personalized services to the users while using the adaptive system (Barla, 2011). To keep the model fresh and support better personalization, the adaptive system continuously builds up and modifies the user model by regularly collecting new data about the user.

Whenever a user interacts with some social web or mobile application, the digital traces about the user's activities, interests and contexts are left in the system, and on many occasions, those user data remain open and accessible. So, the exploration of such data can be very useful to deal with the cold start problem (Dim & Kuflik, 2012), which refers to an adaptive system with insufficient user profile data of a new user interacting with it. Therefore, collecting user profile data from many subscriber applications and their respective context offer an opportunity to create a comprehensive and complex adaptive system. Often, the cold start problem is resolved through the use of stereotypes as the initial user model, followed up by refinements to the model based on interpretations of trace data of actual user behavior.

Successful personalization depends upon whether a single user's behavior is used, or aggregates of many users' data are used in creating the user models (Balabanović & Shoham, 1997). Personalization using data from a single user's behavior that is abstracted into user models (i.e., apps that learn from the individual user's behavior) is usually known as a content-based recommendation. Personalization using aggregates of many users' data (i.e., correlations between likes or purchases) is typical for collaborative recommendation systems. A content-based recommender system builds user profiles containing abstract features that may be applied for adaptation in broader contexts, such as learning about the interests of a user from analyzing the queries, or the text of emails, or reviews. On the other hand, a collaborative recommendation system uses all or part of the entire user model, raw interaction data such as navigation paths, clickstreams or user ratings etc., and computes correlations with the data of other users, or computes "neighbourhoods" of similar users. Therefore, the data of the user is meant to make decisions only in the context of other users' data; it is hard to generalize it for a different purpose

of application. Therefore, most of the big personalized systems, such as Facebook and Google, which are multi-user, adopt both methods in combination, which is known as hybrid personalization (Herder & Kärger, 2008). It is clear that the more data is collected in this way about many users, the better the personalization and the higher the quality of services provided. Since both Facebook and Google provide an ecosystem of services, they all share the user data accumulated by all of them in Google and Facebook servers, which gives them a great competitive advantage over stand-alone, independent services. However, this centralization leads to unfairness, prevents competitiveness, and creates a single point of failure.

User modeling facilitates personalization by enabling an application to interact with the user and adapt to the user's needs arising in the specific context. However, automatically personalizing a user's interactions entails gathering considerable amounts of data about them (Hagen et al., 1999). Therefore, there must be proper support for the user model interoperability-related functionalities. Carmagnola et al. (2011) provided a review on the analysis of the level of support for user model interoperability in the existing systems along with five aspects, known as '*PRICE'*: (1) Privacy, (2) Representation of the exchanged data, (3) Integration of the exchanged data, (4) Communication, and (5) Exchanged data. The summary of the classification of UM systems in their different levels of support for user model interoperability is available in Appendix I.

2.3.2 User Documents

This section discusses issues related to sharing *user-generated content* (UGC) or *user-created content* (UCC), i.e., user documents, which can be represented in different media, such as text, videos, audio, photos. The data recorded by other parties, or the data obtained as a by-product of users' behaviour can be termed "user-generated content".

UGC evolved primarily in the business world with many discussions around free content for organizations and got recognition with the emergence of participative web or "*Web 2.0*". Therefore, some professionals coined a new term UCC that indicates all the original content made by users (J. Kim, 2010). The Organization for Economic Cooperation and Development (OECD) has identified user-created content as the content available on the internet that reflects some creative efforts after being created outside of the professional routines and practices (Graham & Sacha, 2007). However, both UGC and UCC refer to the same content (Bruns, 2016) which is "a generic term comprising a wide range of media and creative content types that were created or at least substantially co-

created by contributors (users) working outside of conventional professional environments". Therefore, I have used a common term "*user documents*" to indicate UGC/UCC in this paper.

There are various types of user documents and more specifically we have categorized them as shown in Table 2.1, which has been reproduced from the OECD (Graham & Sacha, 2007). The user documents can be recognized from different perspectives, such as creative content, collaborative content, or small-scale documents.

Table 2. 1. Types of user documents

Category	User Document	Description
Creative content	Text, novel, and poetry	Original writings or expanding on other texts, novels, poems.
	Photo/Images	Photos or images originally taken or modified by users and posted online.
	Music and Audio	User's own audio content originally recorded and/or edited and published, syndicated, and/or distributed online.
	Video and Film	User's own video content originally recorded and/or edited and published, syndicated, and/or distributed online.
Collaborative content	Citizen journalism	Content such as news stories, blog posts, and photos or videos of current events are posted online by ordinary users. E.g., globalvoices[6].
	Educational content	Content created for educational purposes in schools or universities. E.g., Wikibooks[7].
Small scale document	Mobile content	Content created on mobile phones such as text messaging, or photos and videos that are sent to other users via email or uploaded to the Internet. E.g., photos/videos of natural calamities.
	Virtual content	Content created within the context of an online virtual environment or integrated into it. E.g., user-created games.

[6] https://globalvoices.org/
[7] https://en.wikibooks.org/wiki/Main_Page/

The creative contents include videos, photography or images, text and audio that are distributed on social web platforms such as YouTube, Google Photos[8], Facebook, Twitter, Instagram, blogs, etc. They require an intellectual exchange platform, which ensures the proper user incentives mechanisms for sharing them.

The collaborative contents include educational materials, citizen journalism that is often undertaken and handled by small groups of users across open-source software, such as Apache Wave, Google Docs, and Wikipedia. Regardless of being free or proprietary software, they are mostly centralized services. The collaborative platforms use either synchronous services such as operational transformation (OT) technology or asynchronous services such as version control (e.g., GIT) wikis to allow a range of collaborative functionalities, including application sharing.

In addition, the greater range of small-scale tools enables users to operate and modify existing data sets that have already been built from creative content. Most of the time, mobile apps are used as the interface to operationalize small-scale tools. Therefore, user-created content has enabled new business models to rise for hosting the content and for enabling the data owner to monetize their content.

2.3.2.1 Sharing User Documents

The practice of sharing user-created content across social media platforms is common amongst many connected users. The users have been producing content at a higher rate. Both the user modeling and the participative webs can leverage this opportunity of ever-growing user-generated content.

Users are always the owners of the content they create. Most of the social web platforms maintain user profiles based upon the profile data collected from users during the sign-up process, stating that this data is to provide services such as personalization and adaptation. Those web platforms also offer privacy settings to allow users to control the privacy levels of their information, and in principle, the information that is not publicly available will not be traded with third parties. Most of the time, users create and share content without expecting to receive any monetary rewards. They, however, get recognition for their contributions. In the past, with traditional media

[8] https://photos.google.com/

publishing, it was a difficult and expensive undertaking to publish and distribute user documents, as the associated value chain depended on the number of entities, as demonstrated in Figure 2.6. It shows an offline traditional media publishing value chain for content such as audio, videos, or texts, where the publisher spent excessive time in selecting the user's work, getting consent forms signed, advertising, and finally signing the distribution channels, which were also very limited. The feedback loop generates customers' experiences and behavioral intentions that help to select future content. This also influences amateur users to contribute to creating new content or building on the existing content.

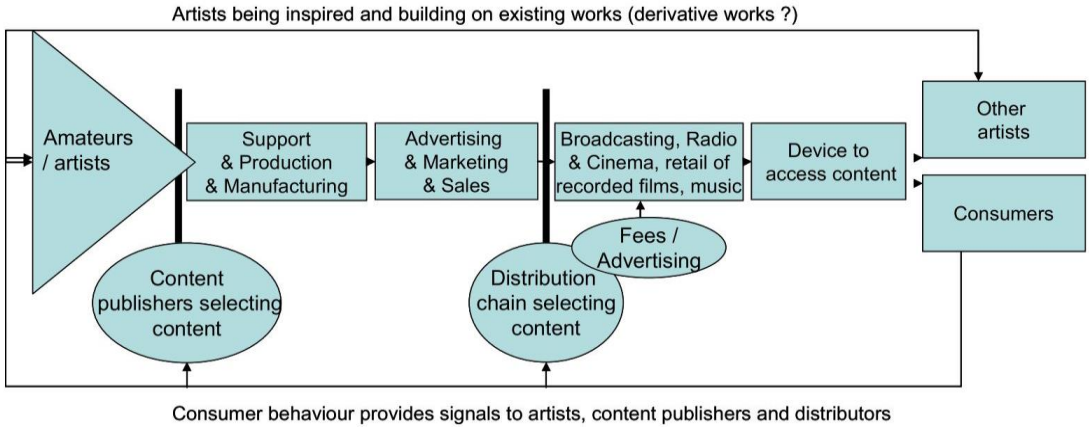

Figure 2. 6. Traditional media publishing value chain (Graham & Sacha, 2007)

In contrast to the traditional model, Figure 2.7 represents the present internet value chain for user-created content. This model enables all the users to create and publish content through access providers. The content creation and sharing processes consume comparatively very little time. Like the traditional media chain, the present internet value chain for user-created content is also influenced by the feedback loop to work on the existing content. The feedback loop comprises recommendations and ratings, which ultimately leads to giving some recognition to the content creators.

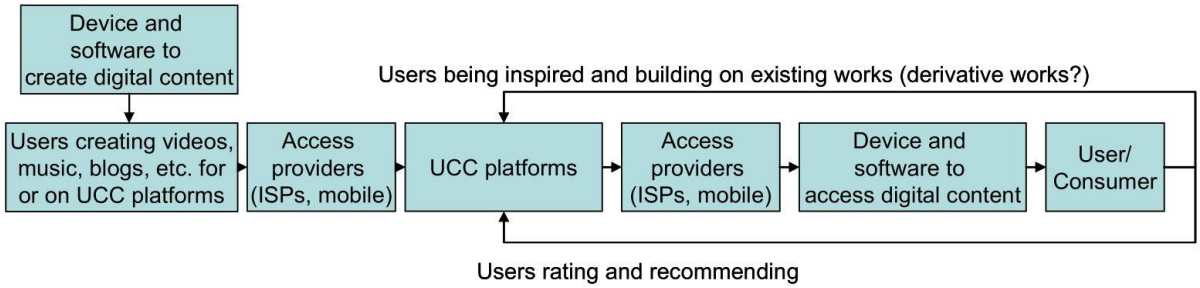

Figure 2. 7. Internet value chain for user-created content (Graham & Sacha, 2007)

However, different web platforms employ different approaches to govern the economic value chain for monetizing user-created content. There are five models for the web platforms (Graham & Sacha, 2007):

(1) Voluntary contribution-based model,

(2) Charging users for a service-based model,

(3) Advert based model,

(4) Licensing content to third parties-based model, and

(5) Online sales-based models.

The voluntary contribution-based models enable users to share their content in online social networking sites such as wikis and various blogs where they do not directly get any monetary incentives from sharing their user documents. The users can share contents publicly, and the motivating factors in these models include acquiring a certain level of reputation, fame, self-expression, and networking with peers. Sometimes, other users/audiences voluntarily offer donations for hosting the site and its maintenance or for the contents as such.

A subscription-based or pay-per-item-based model charges viewers for the services. Users can opt for a "free" subscription to receive a few basic services from a subscription model such as the *Weebly* web hosting service or pay for a "Pro" account to receive unlimited or enhanced features. Similarly, pay-per-item-based models, such as Shutterstock, restrict freeriding and charge users for services like downloading photos. These models may also remunerate content creators or make the posting of content completely free.

Advert-based models offer an advertising-supported presentation to the users (customers) and receive feedback on their preferences and overall behavior that helps to select future content. The users can receive the service free, and the owners of the service serve adverts to their audience, who may possibly click on the ad and visit the advertiser.

A large amount of user-owned original documents (e.g., audio, videos, images, etc.) are uploaded to personal blogs or other platforms and are subsequently redistributed. The original content owners are rarely attributed for their effort in creating such content. On the other hand, most of the UCC platforms specify that they retain intellectual property rights (IPRs) in their contents (e.g., images, audio, videos, text, and graphics, created by site layout) under copyright (Graham & Sacha, 2007). They often set the terms and conditions in such a way that the users must agree to allow the sites a license to collect and use the user's contents before using their services.

Some of those sites may specify that the user who shared their documents retains ultimate ownership to their content, but the sites also receive a limited perpetual license (and right to sublicense) to reproduce, modify, and distribute such contents. Sometimes the agreement enforces the user-posted content subject to the Creative Commons license and may be commercially exploited. The users do not receive any payment for their contents and contributions, but the UCC sites receive all the revenue from the monetization. UCC sites can also legally agree to allow third parties to use their services, thereby hold the right to sublicense user content to the third parties.

Moreover, the online sales models enable user monetization by allowing them to sell goods and services to the online community. Successful UCC sites often have a huge user community, and they can cooperate with third parties to monetize their audience. Some business models, as per their terms of services, may sell anonymized information about users and their preferences and behavioral intentions to market research and other commercial firms.

2.3.3 Research Data

Research data can be user profile data, user documents, or any factual records identified in a research activity and used as primary sources for analysis to produce original research results. The research data cannot be *private*; rather, it is considered as "owned" by the researcher or "public" based upon the rules within the data governance policy under which the data was collected. These rules are often deliberately vague about the outcomes of the research. The research data has always been a way to validate research outcomes and can be combined with other information to accelerate new research findings. Most of the time, the research data are collected, observed, and recorded, or created in digital format such as spreadsheets, images, videos, survey data, experimental data, sensor data, artifacts, scripts, etc., which make them easier to share. The research data can be in different states throughout the life cycle of the research, which may vary as per the research disciplines.

2.3.3.1 Stages of Research Data

We have presented four stages of the research data. Figure 2.8 represents a general category of research data. The very first data produced during the research work is usually in a raw stage that requires further processing to generate the research-ready formatted data. The formatted research

data must be completely annotated or labeled with a standard scale of readings. Then, the data will be ready for analysis. Furthermore, the selected dataset can be reformatted for publication purposes following the detailed analysis of the research-ready data. Good research practice finally produces the catalogue of the published output research dataset.

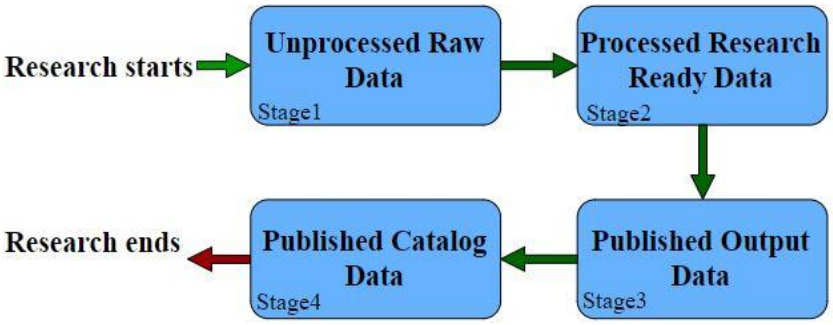

Figure 2. 8. Stages of research data (Whyte & Pryor, 2011)

Table 2. 2. Research materials at different research cycle stages (Whyte & Pryor, 2011)

Research cycle stages	Research materials
Conceptualization	Resumes, user profiles
Design	Proposal drafts, study protocols
Analysis	Metadata, workshop papers, posters
Documentation	FAQs, readme files,
Publication	Conference/ Journal articles, white paper
Translation	Web applications, general articles
Infrastructuring	Databases, web services, software tools

Moreover, other various materials are being produced besides the core research findings at different stages of the research cycle, such as research materials gathered during conceptualization and design of the research project, none of which is individually included in our general categorization. The classification, as shown in Figure 2.8, is a more optimized version, including the core research findings. A summary of those research material outputs at different stages of the research cycle, as obtained from Whyte and Pryor (2011), is given in Table 2.2. It considers seven stages of the research cycle: *Conceptualization, Design, Analysis, Documentation, Publication, Translation* and *Infrastructuring*.

2.3.3.2 Scope of Research Data

We have classified the research data scope into two groups: *Centralized* and *Decentralized*, in terms of the level of control, degree of openness, and availability of research data. These factors are helpful to study various dimensions associated with the storage and sharing of research data, the aggregation of original researchers and research data consumers, and different data sharing policies of the institutions or funding agencies. Table 2.3 summarizes the basic two categories of the research data scope.

Table 2. 3. Category of research data

Dimensions	Centralized	Decentralized
Description	Company and University repositories: depositing research data in an institutional repository Funding agency repositories: depositing research data in an institutional repository Journals: submitting research data to a journal to support a publication	Cloud services: depositing research data with a specialist data centre, data archive or data bank Web/Email services: making research data available online via a project or institutional email/website USB stick/external hard drive: making research data available offline
Control of data	Exerted by a single entity	Shared among independent entities
Data redundancy	Minimum (for non-distributed)	Maximum (for distributed)
Maintenance	Very easy	Easy
Recovery of lost data	Very hard	Easy
Updates to dataset	Immediately delivered to every end-user	Not immediately received by every end-user

The centralized notion indicates that the research data should be kept in only one database file, located at a single point at a given period on a given network and the control of data is exerted by just one entity. The single authorized entity manages the central filesystem that must be capable of maintaining all the received data from the researchers and responding to every single query coming from different agents by themselves. Mostly the centralization in research data sharing is non-distributed with the same proprietary physical storage for the data files. On the other hand, the

underlying idea of the decentralized dimension is that the research data can be stored in multiple physical locations at a given time on any network, and the control is shared among various entities. The distributed storage strategy gives resiliency and possesses a high degree of availability. The researcher can share research data files with other interested agents, who contribute to the replication of the data in a different physical location.

2.3.3.3 Sharing Research Data

"*No longer a hypothetical or occasional occurrence, the use of research data by individuals other than those who originally gathered the data, is currently encouraged or mandated by parallel efforts in the legislature through the 21st Century Cures Act, biomedical journal leadership through the draft data-sharing policy of the International Committee of Medical Journal Editors, charitable foundations such as the Wellcome Trust and the Bill and Melinda Gates Foundation, and the National Institutes of Health (NIH) in its recent request for information on data management and sharing strategies*" (Bierer et al., 2017).

Sharing research data enables researchers to start a new collaboration. However, the *ownership* of the data as an important asset to the researcher in a competitive research environment creates negative incentives for sharing. Most researchers, on an individual level, may feel reluctant to share their research data; however, they appreciate the overall benefits of data sharing, which was also concluded from the qualitative interviews-based study conducted in Whyte and Pryor (2011) and Van Den Eynden and Bishop (2014). Those studies also recognized six different means of data sharing such as private management sharing, peer exchange, community sharing, collaborative sharing, sharing for transparent government, and public sharing. The researchers who collected/generated the data must feel motivated towards sharing their research data since some data are significantly important and valuable resources beyond their purposes. We have compiled some of the important motivations or influential factors for the original researcher to share research data. They are as follows:

1. Sharing research data is an integral part of the research activity.
2. The standard guidelines, policies, data services and research disciplines that they follow.
3. Transparency, accountability, scrutiny of research outcomes.
4. Expectations of funder and publisher.
5. Increase in the research's impact and its visibility.

6. New collaboration can lead to career benefits.
7. Direct credits and attributions are received for their efforts in collecting the data.
8. The duplicate research trials are reduced.
9. Validation of research methods is encouraged.
10. Educational and training materials can receive resources.

Proper data organization is equally important not just for original researchers to validate their research results but also to facilitate other researchers for accessing the research data and working on them. Although research data at different stages have different structures, researchers must maintain the minimum documentation, such as the organized version control commitments, to keep track of the collected data and ascertain that they are usable. Besides, the documentation must also include the overall structure of the research dataset, relationships between various data files and information on data confidentiality. Moreover, the mechanism for querying data and providing answers rather than full sharing of data is very much needed. However, current services are missing these characteristics since the data are not uniform, and they would generally require complex and difficult processes (Wei et al., 2014) to provide such services.

2.3.4 Summary

As shown in the preceding sections, user data covers a wide range from user profiles to user-generated or created data and to research data. The classification developed in Section 2.3 helps to focus attention on a specific domain to identify user data and to manage and share them appropriately. The thesis provides a data-sharing framework that can be used in conjunction with a user profile or descriptive data sharing while supporting any applicable compliance requirements, such as data privacy acts.

2.4 Privacy Compliance and Data Governance

Privacy is an important concept that has been addressed by the legislation of most developed countries. In Canada, the Office of the Privacy Commissioner of Canada (OPC) oversees compliance with two federal privacy laws:

1. Privacy Act: Applies to the Canadian federal government departments and agencies while handling personal information about individuals.

2. Personal Information Protection and Electronic Documents Act (PIPEDA): For businesses' personal information-handling practices.

Most importantly, to comply with the law, businesses must receive the users' consent for collection, usage, or disclosure of the users' personal data. In addition, they must safeguard users' personal data and use them only for predefined purposes. If the data is to be used for any other purpose, the users' consent must be obtained beforehand. The users own the right to access their data anytime.

The United States does not have any formal privacy-related legislation at the federal level but ensures the protection of personal information about individuals through the following laws:

1. Privacy Act of 1974: Covers the personal information-handling practices of federal agencies.
2. Health Insurance Portability and Accountability Act of 1996 (HIPPA or Kennedy–Kassebaum Act): Provides a legal mechanism to modernize the flow of healthcare information, stipulates how Personally Identifiable Information (PII) maintained by the healthcare and healthcare insurance industries should be protected from fraud and theft, and addresses limitations on healthcare insurance coverage.
3. Safe Harbor Act: Known as the U.S.-EU Safe Harbor Framework. It establishes a code of fair information practices for companies to transfer personal data from the EU to the United States.

In Europe, the EU *General Data Protection Regulation* (GDPR), as of May 25, 2018, regulates the processing of the personal data of individuals. Service providers and businesses that collect data from European citizens must meet certain restrictions and requirements presented in the GDPR, so they now need to change their data processing practices to be compliant. As explained in the GDPR, there are several rights defined for individuals:

1. The right to be informed.
2. The right of access.
3. The right to rectification.
4. The right to erasure.
5. The right to restrict processing.
6. The right to data portability.
7. The right to object.

8. Rights in relation to automated decision-making and profiling.

There are some initiatives from working groups and organizations, such as *Kantara,* around data privacy and user authentication. Kantara has developed the GDPR-ready Consent Receipt specification (Lizar & Hughes, 2018) to give control of the data back to the users by changing the way users consent to data disclosure. The consent receipt standard can be used between a business and the consumers when user profile data is provided in order to standardize a GDPR-compliant consent format. With the GDPR in place, Facebook, Google, and other social sites changed their terms and conditions, and due to the GDPR restrictions on tracking cookies, most websites started providing "cookie consent" before allowing users to see anything on the sites.

At the University of Saskatchewan, the data governance framework has delineated roles and aligned data stewardship accountabilities to support the university's Data Management Policy with industry standards in research data identification, access, and quality dimensions. The framework identifies five designated roles within the university that have different responsibilities for decision-making regarding university data. The five roles are Data Trustees, Data Stewards, Data Custodians, Data Guardians, and Data Users. The relationship between these roles and their responsibilities is shown in Figure 2.9.

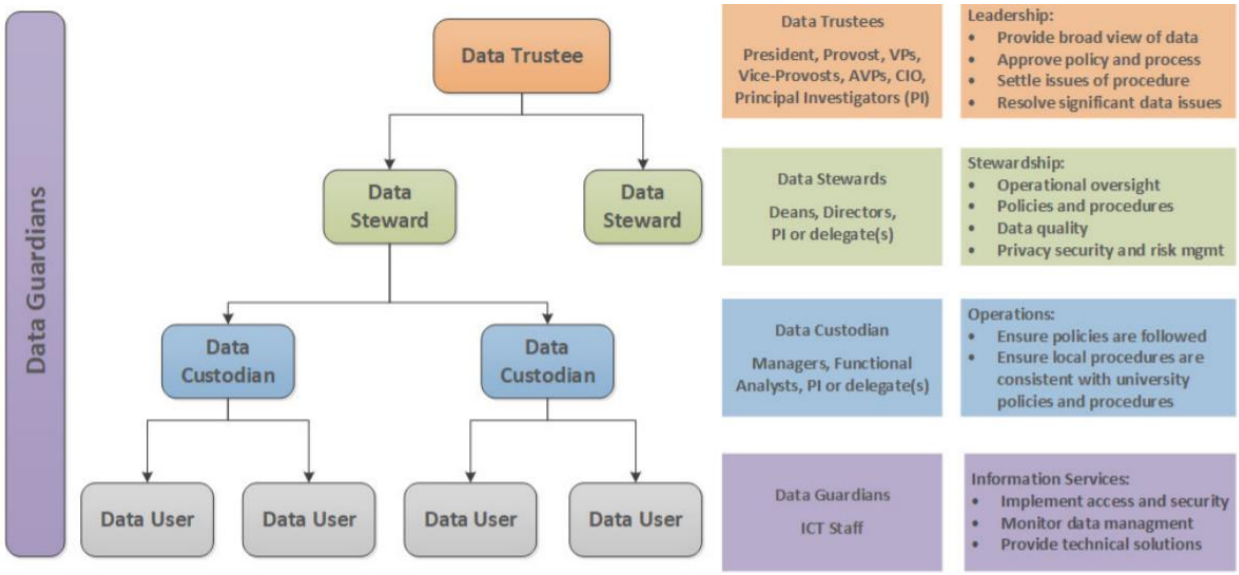

Figure 2. 9. Roles and responsibilities at the University of Saskatchewan[9] (EDUCAUSE, 2015)

[9] https://www.usask.ca/avp-ict/documents/data-governance-framework.pdf

Different teams of an enterprise or an institution in its different roles might have divergent views towards data sharing and management, leading to controversial data governance. So, it is very important to have use cases in data governance programs that can continue to expand with new technologies to tackle new threats and challenges. A common well-designed data governance model is shown in Figure 2.10. The most common challenge for data governance is to break down data silos (Seaman, 2003) and ensure data consistency, compatibility, privacy, security, access control, ownership, and rewards for sharing.

When we combine data governance with ethics and transparent data sharing architecture that aligns trust and trustworthiness among the stakeholders, then we get the notion of data trust—a nascent sub-area of Data Management (O'Hara, 2019). The necessity to incorporate data trust to support the secure and mutually beneficial data exchange portal was also proposed for the growth of UK AI industries in the 2017 report by Hall and Pesenti (2017). Data trust is a way to support data sharing by capturing user trust with transparent, accountable governance structures over digital property and rights. A data trust is similar to fiduciary trust (O'Hara, 2019) since it works within the law and requires the data controller (data trustee) to obtain permissions from a data subject or generator (data trustor) on transparent processing and sharing of the data among data collectors (data beneficiaries). There are three models of data trust—data-centric, collector-centric, and generator-centric—based on who created the data, how data stewardship is designed, where decision making lies and who receives the share of the value generated from the data (Mills, 2019). Data trust would be greatly beneficial since this is what the whole area of data management is caring about: ensuring data to feed data mining with increasing legal protections for privacy and sustain the underlying ownership of the data and digital rights.

Data trusts are a recent (over the last couple of years) and very active direction of research related to the Open Data Institute (TheODI)[10], spearheaded by Sir Tim Berners Lee. This initiative is interdisciplinary and involves legal, ethical, medical, geomatics, AI, and data mining, as well as many other researchers. It has spun many projects already in the UK and is now expanding into many other countries. The core ideas underlying my work predated TheODI. My work emphasizes a particular aspect of sharing data, transparency, control, and incentives for the user, and it investigates a specific type of technical platform, distributed ledgers with smart contracts. Another

[10] https://theodi.org/

distinguishing feature of my work is that it investigates not only the technical aspects (feasibility and performance) but also it explores and evaluates the usability, usefulness, and trustworthiness of the technology through user studies using established state-of-the-art methodologies.

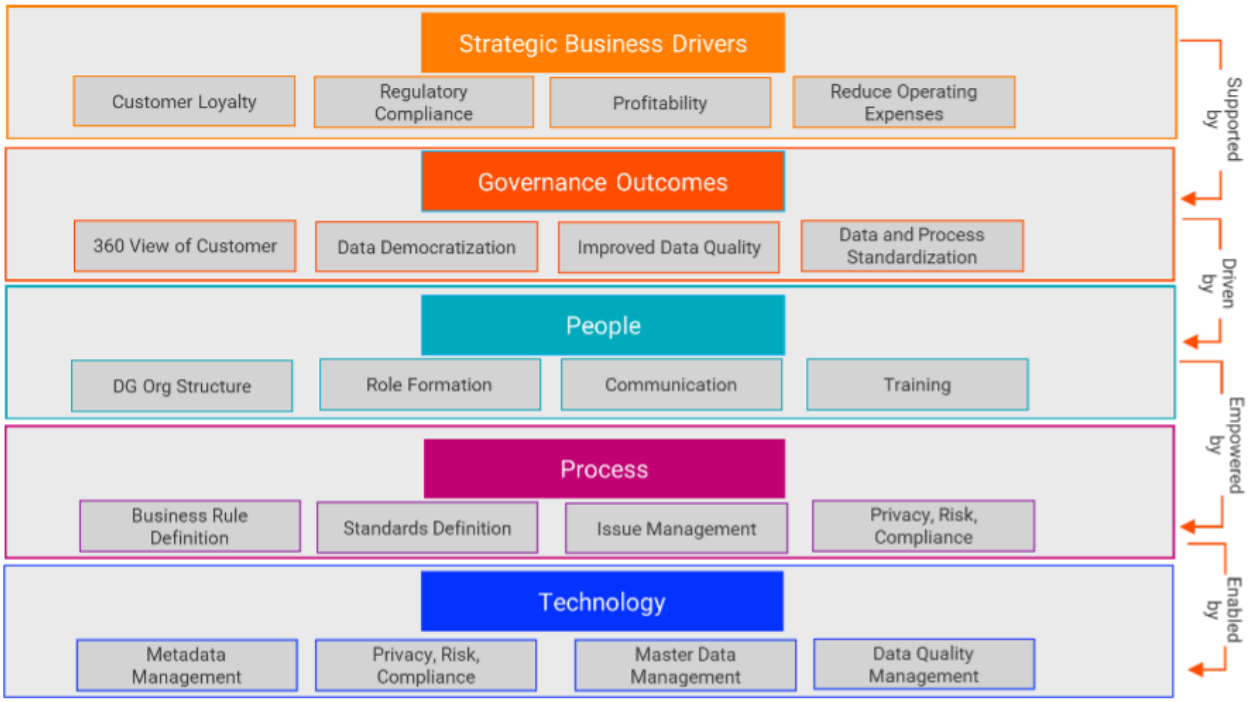

Figure 2. 10. Building blocks for common data governance model (adopted from Informatica[11])

2.5 Data Sharing Patterns

The technological advancement in the last few decades has brought many enterprises to collaborate in a better way while making intelligent decisions. The use of Information Technology tools in obtaining data of people's everyday life from various autonomous data sources allowing unrestricted access to whole data has emerged as an important practical issue and has given rise to legal implications. Data sharing models at an enterprise level are the networked information systems allowing users to create a profile, store data, and make it accessible to others as per the agreement. I have presented different patterns for sharing user data considering three different systems: The technological advancement in the last few decades has brought many enterprises to

[11] https://www.informatica.com/ca/resources/articles/data-governance-framework.html

improve collaboration while making intelligent decisions. The use of Information Technology tools in obtaining data of people's everyday lives from various autonomous data sources allowing unrestricted access to whole data has emerged as an important practical issue and has given rise to legal implications. Data sharing models at an enterprise level are the networked information systems allowing users to create a profile, store data, and make it accessible to others as per the agreement. I have presented different patterns for sharing user data considering three different systems:

1. Centralized,
2. Decentralized (agent-based, peer-to-peer, service-based), and
3. Blockchain-based (decentralized + incentives).

2.5.1 Centralized System

The most dominant companies in the information technology industry, such as Google, Facebook, etc., use their own ecosystems for both the collection and usage of the user profile data. Schmidt (2018) outlined several experiments showing the massive collection of user data by Google, which then targets the users with paid advertising. Google collects various data sets for various purposes whenever the user interacts with it by using their platforms (e.g., Chrome, Android, etc.), applications (e.g., Google Maps, YouTube etc.), publisher tools (e.g., AdSense, G analytics) and other tools (e.g., AdWords, AdMob). These single corporate entities (Google, Facebook, Twitter, and Microsoft) have contributed to the data transfer project (DTP), allowing any other service to use their existing APIs and authorization mechanisms to access data. Users can back up (download) their data, leave one service and try out new services whenever they want. The DTP offers an open-source service-to-service data portability platform that enables users to move their data between online service providers. The contributors to the project have felt that interoperability is central to innovation for data portability.

The centralized architecture, in most cases, doesn't collect and share the diverse fragments of user data coming from the autonomous and independent entities (applications, agents, devices, sensors, services) in service-oriented, mobile, and ubiquitous computing environments (Dolog & Vassileva, 2005). A centralized architecture is limiting since it forces a unified logical structure (ontology) for the user model and thus loses the contextual information that exists in the various applications closer to the user data. If all incoming data from the various applications have to adhere

to the information structure (database schema, annotation dictionary, ontology) used by the server, it has to process/interpret the data, generalize it, and thus lose some of its specifics, for example, how it was collected, for what purpose, etc., that may be valuable in a different context. Such loss of information can be a source of bias in the system.

Different centralized architectures are still predominant in business enterprises because the efficient client-server feature can be incorporated into the creation of numerous user modeling servers. In fact, the physical storage of user data at one central point does not necessarily imply centralized user modeling. There are cases in which the user data is stored in distributed storage spaces, but the components of the user modeling are structured centrally (Carmagnola et al., 2011). Thus, the storage structures are a different level of granularity from the conceptual level oriented around the end use of the data.

Often centralized user modeling architectures have a predefined point of access that leads to the central point of failure. Replication of the data via mirroring the servers could be the option, but that usually comes with high communication costs as well. Therefore, decentralized approaches for user modeling have achieved the research trend to overcome the limitations brought by the centralized user modeling architecture, as mentioned in the earlier section too.

Several alternative approaches to the centralized user modeling approach have been proposed. The Houdini framework (Hull et al., 2004) enables the sharing of context-aware and privacy-conscious user data for global computing. It comes with a method to collect data from various sources focusing on how and when to share them. It is built with an infrastructure to manage principles focusing on the preferences for data sharing conditions, such as what kinds of data to share and when to share them. The main aspect of the infrastructure is the self-provisioning of data sharing preferences by allowing the users to provide such conditions using web-based forms.

2.5.2 Decentralized (Agent-Based, Peer-to-Peer, Service-Based)

Iyilade and Vassileva (2013) presented a decentralized architecture for life-log sharing and reuse by multiple applications. All the life logs (e.g., clickstreams, events, interests, etc.) from different systems are gathered by agents, which then forward the information to a centralized broker, which is responsible for user modeling, that comprises request analysis, source selection, source connection, semantic mapping, data integration, and response transformation. An online social network called Persona (Starin et al., 2009) allows users to choose and define the rules for whom

they want to share their personal information or photographs with. It uses attribute-based encryption and public-key cryptography to hide data and provide the flexibility needed. For the decryption and authentication by groups and users, it uses group-based access policies. Persona can perform just as well as the existing general online social networks with added privacy features. It can also browse through highly sensitive data on web pages.

A few prominent examples of data sharing systems include different frameworks to achieve the interoperability of distributed model with a centralized server such as Mypes (Abel et al., 2013), online P2P file-sharing networks and data management systems, collaborative repositories such as Wikidata (Vrandeč & Krötzsch, 2014) etc. Almost all of these systems implement different architectures, and their evaluation is based on different non-functional requirements, such as efficiency, scalability, or reliability (Davoust, 2015). Accordingly, most of the research relevant to their design framework is focused on the optimization of those properties, and the technical performance of a data-sharing system alone does not guarantee the practicality of the systems.

Sweeney (1997) demonstrated a computer program called Datafly that offered a practical means of maintaining anonymity or confidentiality in medical data by automatically generalizing, substituting, and removing information as appropriate without losing many of the details found within the data. Decisions are made at the field and record level at the time of database access, so the approach can be used on the fly in role-based security within an organization and in batch mode for exporting data from an organization. Often organizations release and receive medical data with all explicit identifiers such as name, address and phone number being removed with the incorrect belief that patient confidentiality is maintained because the resulting data look anonymous. However, the remaining data can often be used to re-identify individuals by linking them to other databases or by looking at unique characteristics found in the fields and records of the database itself. When these less apparent aspects are considered, each released record is ambiguously mapped to many other possible people so that only the user can determine a level of anonymity. Datafly explicitly quantifies “trust” in the recipient, so the associated risk becomes clear. Usually, it is very hard to infer what other systems/people can infer from user data. When a company states its privacy policy, it is important not only for it to say how it will guard the data but also what purposes it will use the data for (or not). The purpose, in the end, determines the *meaning* of the data and the risk the user is at if they share the data.

Moreover, Ozzie et al. (2009) have provided an architecture that enables user-controlled access to user profile information. A user is allowed to selectively mask (expose) portions of her profile to third parties. Advertisers and content providers can offer incentives or enticements to a user to encourage the user to expose larger portions of their profile. The architecture comprises profile management utilizing a profile component for facilitating the creation and storage of an electronic profile of a user and a control component under the control of the user for controlling access to the profile. Machine learning and reasoning are also provided to make inferences and automate aspects thereof.

Hu et al. (2011) stated that the online social networks (OSNs) in the then-recent years offer not only attractive means for virtual social interactions and information sharing but also raise many security and privacy issues. Although OSNs allow a single user to govern access to her data, they currently do not provide any mechanism to enforce privacy concerns over data associated with multiple users, leaving privacy violations largely unresolved and leading to the potential disclosure of information that at least one user intended to keep private.

In the IoT domain, there are systems such as MobiTribe (Thilakarathna et al., 2014), which have a distributed model. Nevertheless, it focuses on mobile devices and uses the centralized content management system as a moderator for the exchange of information between the devices and the applications. PersonisAD (Assad et al., 2007) is another active, distributed, scrutable model that gathers information from different sensors associated with different users and combines their preferences to provide a richer experience. As mentioned in Dim and Kuflik (2012), the distributed user model was presented from the single method standalone agents, which store a single attribute of a user model with the holistic vector. In other models (Niu et al., 2004; Vassileva et al., 2003), the model is decentralized, held by different agents, and the information is gathered from different agents only temporarily, for a given purpose of adaptation.

Furthermore, it is observed that accommodating the conflicting interests among the users is not separable from the architectural design that applies optimization of the specific system properties and involves trade-offs with the participants' autonomy (Davoust, 2015). In a structured peer-to-peer design such as Stoica et al. (2003), participants are not given the privilege to connect to the peers of their choice, but rather they have to store data with arbitrary peers.

Well-developed incentive mechanisms positively motivate the users in the virtual communities to willingly engage in data sharing with others (Chia-Shen et al., 2012). For cloud storage, an

incentive mechanism was introduced into rational secret sharing schemes, and a fair data access control scheme was proposed (Liu et al., 2017). In the scheme, the decryption key reconstruction activity is to be formalized, and then its security, fairness, and correctness are defined. Afterward, the decryption key obfuscation is performed with a generation of many fake keys over the shared data. During the exchange of shares, they adjust the action order through the agreed term. As a result, the users' selfishness is restricted and fair data access control in cloud storage is achieved. For example, a user is required to first send her shares when she deviates from the prescribed scheme and thereby accesses the shared data together. Fabian et al. (2015) proposed an attribute encryption-based architecture to enable selective access authorization and secret password sharing in a paper on collaborative and secure sharing of healthcare data in multi-clouds.

To study incentives used in the clinical domain, Stovel et al. examined 251 articles in five medical education journals (Stovel et al., 2018) and found that only 8% of papers described the incentive mechanism employed in their system. They found that the incentives architectures were not well discussed in the literature. Xu and Van Der Schaar (2014) proposed a rating-based approach to incentivize users to produce and sharing their content in a user-generated content UGC network for enhanced social welfare. In their system, the service administrator divides users into separate groups and assigns the responsibility to produce content to certain groups for a certain time so that other users can enjoy freeriding during that period. Freeriding is when users passively acquire others' content without contributing on their own. The ratings are only earned as the users comply with the guidelines; otherwise, their rating declines. The lurkers or the users acting poorly carry lower ratings and are given limited access to the content. As most of the users prefer freeriding, their study suggests that introducing a fixed degree of freeriding behavior could offer better social welfare than incentivizing all the users to produce content. Finally, their analysis shows that the heterogeneity of users in their content valuation and cooperation strongly influences users' commitment and the incentives that they would find attractive. For a similar case, Zohar and Rosenschein (2009) proposed a protocol for a P2P file-sharing system to discourage agents from consuming the content without sharing a certain portion of the file they were downloading.

Golle et al. (2001) brought up the issue of freeriding with an example scenario of the then P2P file-sharing network Napster where users would not get any incentive for UGC. They highlighted the importance of incentives for sharing in P2P networks and introduced their internal currency,

named “points,” which the agents could buy with money or by making a substantial contribution to the networks.

In the medical field, the research community is increasingly recognizing the importance of sharing patients’ data from clinical trials to maximize the knowledge gain from the research effort (Lo & DeMets, 2016). European Medicines Agency (EMA), several drug companies and one other trial funder have already implemented a data-sharing framework. However, the issue with them is to address the appropriate and meaningful incentives to capitalize on the promise of data sharing and to ensure proper data privacy and security. In 2018 alone, data breaches in healthcare exceeded 15 million records, with the third parties continuing to pose risks to healthcare providers (Protenus, 2019). Therefore, an effective privacy-preserving mechanism must be considered while designing user modeling and personalization systems (e.g., while storing and sharing medical records or other user data).

2.5.3 Blockchain-Based (Decentralized + Incentives)

Blockchain is the first fully functional suite of distributed digital ledgers with an immutable record of every transaction that has ever taken place. These records are organized in ‘blocks’ that are linked together by cryptographic validation, and thus the technology is named blockchain. There is a cryptographic signature (Blanchette, 2006) to identify each block in the blockchain, and every block refers to the signature of the previous block in the chain. Cryptocurrencies such as Bitcoin are the first-ever use case of blockchain technology. I have provided a detailed explanation of blockchain in section 2.8.1.

MedRec (Azaria et al., 2016) uses blockchain technology for the first time to preserve the privacy of user data while handling electronic medical records (EMRs) in a completely distributed P2P network. The participating medical stakeholders (researchers, public health authorities, etc.) in the network act as blockchain “miners.” The system provides the participants with access to aggregate, anonymized data as mining rewards in return for sustaining and securing the network as miners. However, it only collects static data from medical examination records and will be inefficient to support metadata change while sharing data streams generated continuously from sensors and other monitoring devices. Samaniego and Deters (2016) presented an idea of using blockchain as a service for IoT that manages device configuration, stores sensor data, and enables micro-payments.

CreditCoin (Li et al., 2018) uses blockchain technology to design an effective vehicular announcement network where users are motivated with incentives to share traffic information in the vehicular ad hoc networks (VANETs). The user has a credit account at several addresses that contain reputation points called "coins" in the CreditCoin platform. The authors of CreditCoin also suggested that users who are concerned with their privacy forward any messages in VANETs if there is a risk that their privacy will be breached (Li et al., 2018). Zhang and Chen (2019) also proposed the data security sharing and storage system based on the consortium blockchain for the VANETs. However, it is inevitably necessary to allocate welfare to data owners through an incentive platform for data sharing.

Furthermore, there have been many studies in the academic literature conducted about online privacy, security, and trust (Rios et al., 2017). For blockchain-based solutions, the problem remains with the acceptance of blockchain technology, which remains mostly used as deeply hidden background technology in the banking sector rather than in real-life applications involving end-users. Numerous researchers (Herrera-Joancomartí & Pérez-Solà, 2016; Sas & Khairuddin, 2017; Henry et al., 2018) have talked about the mistrust people have in Bitcoin and blockchain-based coins and that their adoption even as currencies is limited to tech people and speculators. Bian et al. (2018) highlighted different risks associated with investing in cryptocurrencies as millions of dollars are lost every day due to fraud cases that range from Ponzi schemes to fake initial coin offering and pump and dump schemes (Baum, 2018). One such fraud scheme is OneCoin, which was considered by The Times as one of the biggest Ponzi schemes. OneCoin was promoted as a cryptocurrency by Bulgaria-based offshore companies, resulting in the theft of 5.5 billion dollars from unsuspecting investors. Furthermore, Prashanth et al. (2018) and Kshetri (2017) argued in their comprehensive survey that numerous privacy and security-related issues have arisen in the adoption of blockchain-based decentralized applications. So, the problem of making this technology usable—transparent, controllable and trusted by people—is still an open problem. It is therefore important to have a user study as a method of identifying and quantifying user trust and privacy concerns (Buchanan et al., 2007; Rios et al., 2017) and user attitudes towards using and accepting such blockchain-based solutions.

2.6 Augmented Technology Acceptance Model

The Technology Acceptance Model (TAM), developed in social psychology, was based on the Theory of Reasoned Action (TRA) (Fishbein & Ajzen, 1975), which claims that behavioral intention is a strong indicator of actual behavior. The TAM model, as shown in Figure 2.11, has been used as a conceptual framework in the information systems literature to study the potential users' behavioral intention to use a particular technology. The behavioral intention is defined as "the degree to which a person has formulated conscious plans to perform or not perform some specified future behavior" (Warshaw & Davis, 1985), which is therefore in line with the TRA.

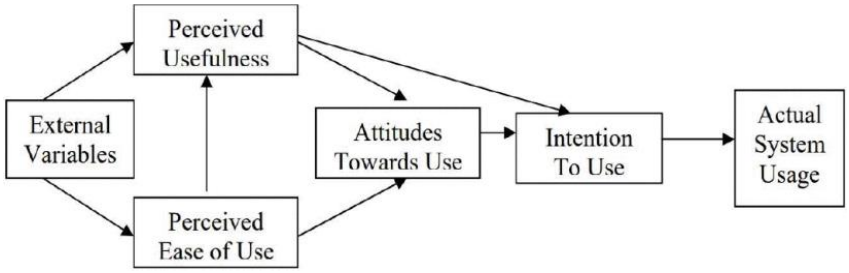

Figure 2. 11. A classical TAM model (Davis, 1986; Davis, 1989)(Davis, 1989)

The classical TAM focuses on using technology, where perceived ease of use (PEOU) and perceived usefulness (PU) are two factors or antecedents to influence user acceptance behavior. PEOU is defined as the degree to which a person believes that using a particular system would be free of effort. PU is the degree to which a person believes that using a particular system would enhance his or her job performance. TAM hypothesizes that the actual use of the system is determined by behavioral intention to use (ITU), which is the degree to which a person has behavioral intention to adopt the technology. ITU is, in turn, influenced by the user's attitude toward using the system and their perceived usefulness and perceived ease of use of the system, as represented in Figure 2.11. Attitude towards use is the degree of belief to which a person uses the system as guided by valuations (Shin, 2019; Shin, 2017).

TAM is widely used to understand how users come to accept and use information technology. However, there is no literature on TAM in the context of blockchains and smart contracts-based applications, indicating a significant knowledge gap. Our research applies the extended TAM to distributed ledger technologies to fill this gap. Due to the limitation of classical TAM, because many key factors are not included in the model (Melas et al., 2011), many researchers often extend

TAM by adding external constructs depending upon the contexts. Perceived enjoyment (PEnj) (Davis et al., 1992), quality of system (QOS) (Koh et al., 2010), trust (T) (Wu & Chen, 2005), and behavioral control (BC) (Bhattacherjee, 2000) are some of the constructs that have been added as influential variables to user acceptance of the information technology and are therefore inevitable for evaluating my proposed blockchain-based data-sharing approach as well. Furthermore, a privacy model (Buchanan et al., 2007) with perceived security (Shin, 2010) is also very useful to study the usability and trust users can have in the technology.

2.7 Privacy, Security and Trust Model

Privacy is defined as the right to be left alone (Warren & Brandeis, 1890). Furthermore, privacy has been considered as the right to prevent the disclosure of personal information to others (Westin, 1968).

Later, privacy has become recognized as multidimensional (Burgoon et al., 1989; DeCew, 1997), as it includes informational privacy along with accessibility privacy, physical privacy, and expressive privacy.

1. Informational privacy: "How, when, and to what extent information about the self will be released to another person" (Burgoon et al., 1989; DeCew, 1997). E.g., the user is asked for too much personal information while using online services.
2. Accessibility: "Acquisition or attempted acquisition of information involves gaining access to an individual" (DeCew, 1997). E.g., User information might be left in the old system.
3. Physical privacy: The "degree to which a person is physically accessible to others" (Burgoon et al., 1989). E.g., Viewing the user screen in an unauthorized way.
4. Expressive privacy: "Protects a realm for expressing one's self-identity or personhood through speech or activity" (DeCew, 1997). It restricts extrinsic social control over choices and improves intrinsic control over self-expression. E.g., User data may be inappropriately forwarded to others.

Introna and Pouloudi (1999) developed a framework of principles for the first time to study privacy concerns while exploring the interrelations of interests and values for various stakeholders. The study found that different users have distinct levels of concern about their own privacy. Smith et al. (1996) developed a scale for the concern for privacy that measured unidimensional aspects of privacy, such as collection, errors, secondary use, and unauthorized access to information

factors. Malhotra et al. (2004) also presented a model to consider multiple aspects of privacy, such as identifying attitudes towards the collection of personally identifiable information, control over personal information, and awareness of privacy practices of companies gathering personal information. However, all these studies just focused on informational privacy, so the scales to measure privacy were also based on a unidimensional approach and were not validated. Furthermore, the issue regarding the benefit of giving up privacy, such as offering personalization, enhanced security, etc., was not addressed by those studies.

Hence, to address the multidimensionality of privacy, it is particularly important to consider privacy-related behaviors while studying privacy concerns and user attitudes towards privacy in user data-sharing platforms, as shown in Figure 2.12.

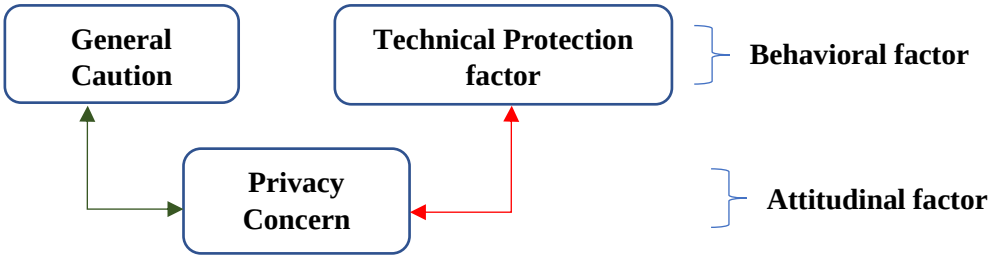

Figure 2. 12. Privacy model (Buchanan et al., 2007)

The constructs presented by Buchanan et al. (2007) are validated and considered both privacy concerns and user behavior models. Behavioral items include General Caution and Technical Protection of privacy. The attitudinal item includes general concern about privacy. They found that the Privacy Concern correlates significantly with General Caution (represented by green double arrow line) but not most significantly with the Technical Protection factor (represented by red double arrow line).

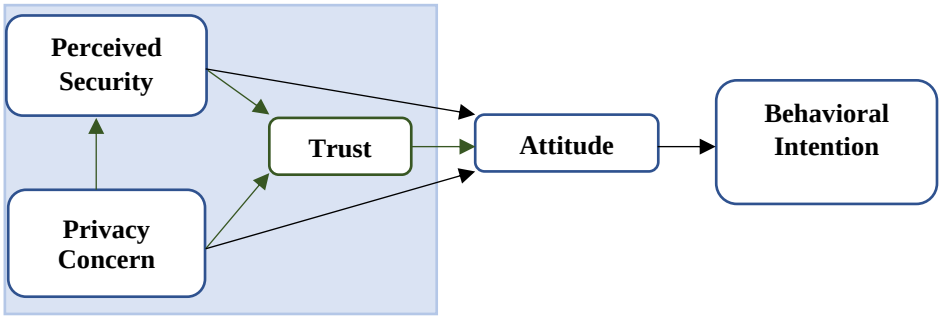

Figure 2. 13. Trust model

The user acceptance behavioral model as presented in Figure 2.13 (Shin, 2010; Rios et al., 2017) for the theoretical social network services is also useful for conceptualizing the role of perceived security and perceived privacy (Privacy Concern) on the digital trust of the blockchain-based data sharing frameworks. Perceived security is the degree to which a user believes that the online service has no predisposition to risk (Yenisey et al., 2005) or it will be risk-free to use the system (Shin, 2010). Perceived security here not only means technical security but the user's subjective feeling of being secure in the network (Roca et al., 2009).

Similarly, trust is an important contributing factor for users to do a certain task that can make them vulnerable and yet hope the service provider on the other end can fully comply with the set of protocols to complete a transaction (Dwyer et al., 2007) and eventually develop a new relationship (Coppola et al., 2004; Jarvenpaa & Leidner, 1999; Piccoli & Ives, 2003). In a virtual environment, as the users do not have any control over the outcome of their actions, trust becomes one of the prime factors for them to ground some firm belief in the reliability to engage with the other party (Hoffman et al., 1999). In e-commerce, when information is disclosed, users tend to trust the service provider more (Metzger, 2004), resulting in users being free of doubt and more likely to engage with the other party (Hoffman et al., 1999).

2.8 Distributed Ledger Technologies

Distributed ledger technology (DLT) is a data structure used to create a public or private distributed digital transaction ledger that, instead of resting with a single provider, is shared among a distributed network of computers. Blockchain is one sub-class of DLTs, which is used to store, distribute, exchange and track anything of value between users, for example, storing critical assets in the supply chain to tracking their ownership and changes in state. The basic software pattern of blockchain was introduced in the original source code for the digital cash system, Bitcoin (Nakamoto, 2008) and implemented in 2009 by mining it for the very first time. However, Nakamoto's paper did not use the term '*blockchain'* to describe the technology component underlying Bitcoin. Several variations of DLTs such as Hashgraph[12], IOTA's Tangle network[13],

[12] https://www.hedera.com/cryptocurrency/
[13] https://www.iota.org/research/meet-the-tangle/

NANO[14] (formerly known as RaiBlocks), and Peaq[15] have been created to support the continual growth and change of the crypto world. However, blockchain technology holds promise to transform data management and business models in many domains. Initially, blockchain was also considered for only powering virtual currency, but the applications of blockchain technology have since quickly evolved to numerous use cases.

Figure 2.14 presents the percentage of start-ups (dataset of 1140 applications) in different industry sectors using blockchain technology. This study by Friedlmaier et al. (2016) shows that the finance and ICT sectors mostly dominated blockchain uses and that blockchain technology is not just limited to cryptocurrencies.

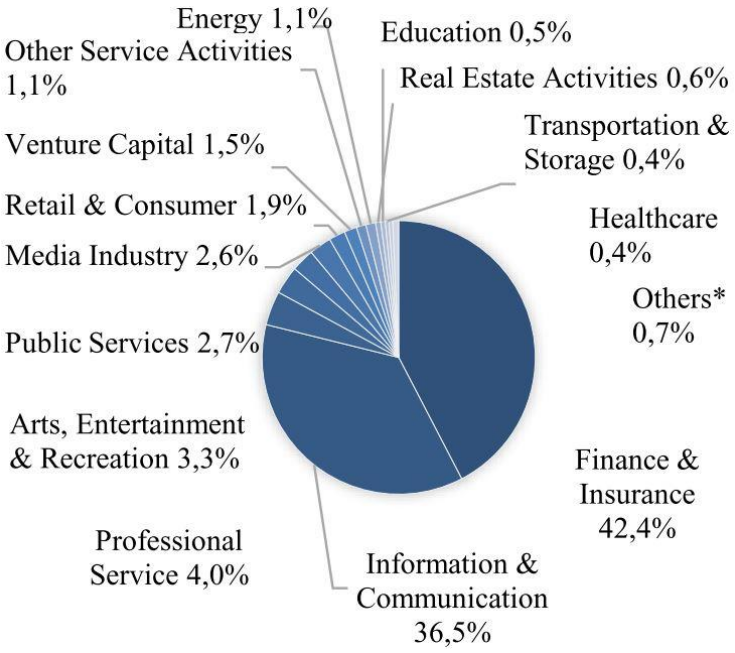

Figure 2. 14. Percentage of start-ups operating in various industry sectors with blockchain (Friedlmaier et al., 2016)

Presently, many industries, including fintech and banks, commercial supply chains, healthcare, agriculture, tourism, building industries, etc. are working on incorporating blockchain (distributed ledgers) technology as a core of their data-management systems (Bullock & Bannigan, 2016; Feng Tian, 2016; McGhin et al., 2019; Shrestha & Vassileva, 2016, 2018a). A Google trends chart, as

[14] https://nano.org/en
[15] https://peaq.io/

shown in Figure 2.15, also highlights the relative popularity of the *blockchain* technology between July 2017 to July 2021 in comparison with the *Internet of Things (IoT), Cloud Computing*, and *Big Data*. From the given chart, the Google search trend for "Blockchain" went very high between December 2017 and January 2018. Even in other periods, the average search trend for blockchain technology stayed substantially above the other three most hyped technologies.

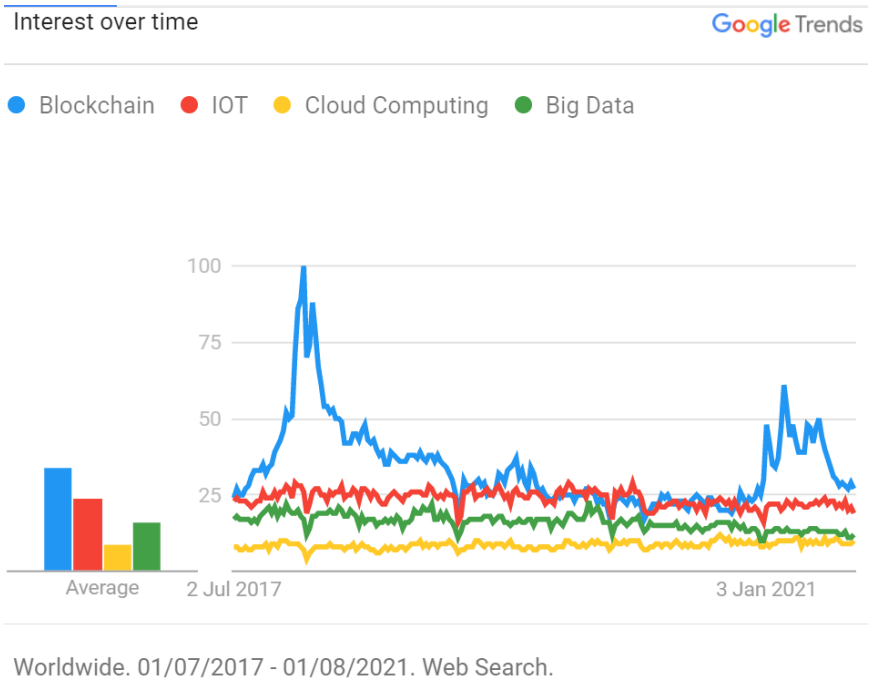

Figure 2. 15. Google Trends chart for emerging technologies

By 2025, the IoT will exceed 100 billion connected devices, each with a dozen or more sensors collecting and sending data (Bryzek, 2015). IoT devices, such as Google Assistant, Siri, Alexa, Bixby, Cortana, etc., interact with us, but before long, they will be talking to each other. The Economist in 2017 penned the meme, as "Data is the new oil." If data, like oil, is the new medium value exchange, then there should be a protocol to move that value between devices. Distributed Ledger Technology (DLT) could be the new medium. DLT is about enabling global and open platforms for human economic coordination, trust minimization, censorship resistance, and decentralization.

2.8.1 Blockchain

Blockchain technology represents a digital ledger—a database with an immutable record of every transaction that has ever taken place (Crosby et al., 2015). These records are organized in 'blocks' that are linked together by cryptographic validation. Each block aggregates a timestamped batch of transactions to be appended in the chain. There is also a cryptographic signature to identify each block. Each block refers to the signature of the previous block in the chain, and that chain can be traced back to the very first (Genesis) block created in the chain.

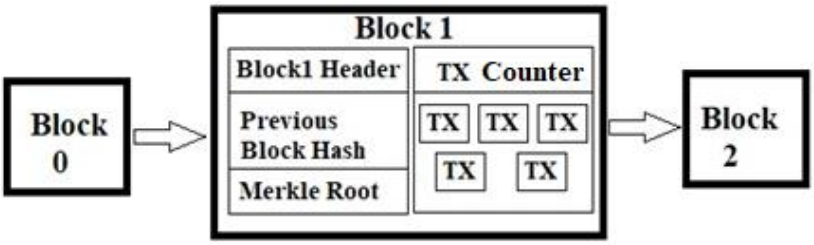

Figure 2. 16. An example of blockchain

Figure 2.16 represents a blockchain consisting of a sequence of blocks. The block in the blockchain is composed of several transactions, which depends on the size of the block. The block contains vital information in its block headers such as the previous block's hash, which points to the previous block, Merkle tree root hash, which is the aggregate hash value of all the transactions' hashes in the block, and timestamp, which records the time in the UNIX Epoch time.

2.8.1.1 Construction of Blockchain

Assuming that there are n blocks on the chain, the height of the blockchain is n-1. The construction of the blockchain is as below (Nakamoto, 2008):

Blockchain:=Block0 ‖ Block1 ‖ … ‖ $Block_{(n-1)}$

Arranging the transactions of S (when a consensus round is ended, an agreement is reached for all transactions recorded in candidate set S) in an ascending sort, constituting a Merkle Hash Tree and calculating a hash value of the block, the new block's structure is described as below:

$Block_{new}$:= (rt, num, info, $hash_{pre}$, $hash_{new}$, $hash_{next}$)

In $Block_{new}$, $hash_{pre}$ is a hash value of a recent block. rt is a Merkle tree root of the present block.

$hash_{new}$ is a hash value of the present block where $hash_{new}$=Hash(rt, num, info, $hash_{pre}$).

$hash_{next}$ is a hash value of the next block, num is the number of transactions in set S, info is a series of transactions.

After reaching the consensus, the new block will be added to the end of the blockchain, as follows:

Blockchain:=Block1 ‖ Block2 ‖ … ‖ Blockn-1 ‖ $Block_{new}$

The key idea is that this ledger is neither stored in a centralized location nor managed by any single entity; rather, multiple distributed parties come to a consensus, which is committed into the ledger and thereafter can be accessed by anyone. Computationally, it is impracticable for any corrupted node (unless the number of such nodes is a higher majority consensus) to go back and alter the history because the blockchain represents a chronological chain of events recorded in a distributed network. There is no single point of failure in a blockchain because the redundancy of the system ensures many backups, and the lack of a central storage place ensures there is no one target for hackers. Because the data is stored in a distributed and redundant fashion and each node verifies each transaction, it is extremely hard for malicious parties to attack and manipulate the data to their advantage. Thus, blockchain does not require a trusted central server or entity and is called "trustless." Many blockchain systems provide technology called "smart contracts," which add the terms of agreements or contract rules in the self-executing computer codes.

To ensure the success of the blockchain platform, incentives play a particularly important role in encouraging participation. Usually, the nodes are incentivized with mining rewards for taking part in the transaction validation activities, such as Bitcoin/ Ethereum miners. Additionally, there are some proposals for using blockchain incorporating access control measures to ensure the privacy of data (Zyskind et al., 2015). Not only can blockchain technology manage access control, but also store and share off-chain files (Shrestha & Vassileva, 2018a). User modeling can benefit from a platform using distributed ledgers and smart contracts to ensure user-controlled privacy and data-sharing policies encoded in smart contracts. In contrast to the centralized system, blockchain technology can be transparent to the users and very promising to incentivize users for data sharing. It also naturally supports building up incentives for users to share their data in terms of rewards (micro-payments or credits) encoded in the smart contracts. In this way, users become owners of their data and can decide how their data is collected, used, and shared and benefit not only in terms of improved personalized experience with the service but also directly, for example, by

participating in the share of the advertising revenue generated by the service provider. However, different business use cases can only be adopted after considering the cost-benefit analysis and by choosing the right blockchain technology.

2.8.1.2 Blockchain Types

There are three categories of blockchain, each with a slightly distinct set of protocols and consensus mechanisms. The consensus is to achieve agreement across validators (or miners) in a network on every new ledger of transactions. The blockchain is usually equipped with consensus protocols to tolerate unreliable involved parties or malicious nodes.

The first category of blockchain is *public blockchain*, in which anyone can participate in the chain and contribute to the consensus process. The read permission or the right to see the public blockchain is always open to anyone with internet access. The second category is *consortium blockchain*, in which a pre-selected set of nodes control the consensus process. The right to see the consortium blockchain remains either public or restricted to the participants. The last category is *private blockchain*, in which the transactions are contained within a closed community and are of interest to only the members of the community present in the chain, e.g., MultiChain, Hyperledger, or Sawtooth. The *private blockchain* adopts the core idea of blockchain as a distributed ledger technology (DLT) but assigns the private validator, which is a member of a consortium or separate legal entities of the same organization. The right to see the private blockchain remains restricted to the participants.

A blockchain can also be referred to as *permissioned* or *permissionless*, each with slightly different properties. A permissioned blockchain is a faster, trusted network that offers managed upkeep and private membership such that members can contribute to the consensus process only after meeting some criteria. On the other hand, a permissionless blockchain is a slower, trust-free, open, transparent, and public membership network such that any members can contribute to the consensus process without any restriction (Wood, 2016). Therefore, depending upon the consensus mechanism, different blockchains are suitable for various types of business use cases. A summary of the criteria to be considered as features for choosing a proper blockchain is available in Appendix I. Some emerging protocols or algorithms for achieving consensus in a blockchain are: (i) Proof of Work (PoW), (ii) Proof of Stake (PoS), (iii) Delegated Proof of Stake (DPoS), (iv) Practical Byzantine Fault Tolerance (PBFT), (v) Istanbul Byzantine Fault Tolerance (IBFT), (vi) Federated

Byzantine Agreement (FBA), (vii) Prof of Elapsed Time (PoET), (viii) Raft-based Consensus. Their short descriptions are also available in Appendix I.

2.8.1.3 Blockchain for Data Management

According to Hileman and Rauchs' study on global blockchain benchmarking, the finance sectors have dominated the blockchain use cases (Hileman & Rauchs, 2017). Currently, we have seen some significant research efforts using blockchain for data aggregation and collaboration in many other fields, including engineering, science and government. Section 2.4.4 has already talked about the literature on blockchain-based data sharing. The research efforts on using blockchain for data storage and management in decentralized P2P networks are mostly aimed to keep track of the flow of data and preserve data ownership.

In the healthcare sector, for example, there has been a lot of interest in researching the potential use of blockchain-based applications for data management and access control services (Azaria et al., 2016; Rouhani et al., 2019; Yang et al., 2019). Different blockchains adopt different protocols for storing data, metadata or limited lengths of arbitrary data associated with the transactions.

In the case of the Bitcoin blockchain, transactions can carry a limited length of arbitrary data onto the chain when they are indicated as "unspent." *Namecoin* (decentralized domain name system) is the first fork of the Bitcoin with the completely new Genesis block claimed to store metadata onto the blockchain. The on-chain data storage in the public blockchain leads to low throughput, as the blockchain process requires the redundancy of each data object over the entire network and thus gets limited by the huge data size.

In the context of permissioned blockchain, MultiChain can optionally store any published data off-chain, which saves storage space and bandwidth (Greenspan, 2013). It can hash up to 1 GB per item (Off-chain data) into the blockchain, with the data itself delivered rapidly over the P2P network. The same idea of storing data in the private and permissioned blockchain has also been proposed in works such as Shrestha et al. (2017) and Yang et al. (2019).

Even in the public blockchain, *virtualchain* allows some third-party storage to be connected onto the chain for storing payload along with the data owners' signatures. This process enables the blockchain to store only the minimal metadata, such as digital fingerprints of the files and state transitions. This concept has been applied in some research works, such as Shafagh et al. (2017) and Mcconaghy et al. (2016). These studies are for large-scale IoT. The main underlying idea is

that the metadata is stored in the off-chain decentralized storage systems such as Distributed Hash Tables (DHTs), Inter-planetary file system (IPFS[16]), and the pointer pointing to the address of the file on the off-chain storage systems is committed on the blockchain.

To motivate users for data storage and retrieval services, IPFS incentivizes the participating nodes with *Filecoin*[17] (cryptocurrency) for hard drive space instead of computing power. The proof of replication consensus model requires miners to prove to a verifier that they have created different copies of the files on the network. *Sia*[18] and *Storj*[19] blockchains also support distributed data storage by shredding the user uploaded file, encrypting each segment and spreading the file ciphertext to the participating nodes across the network. Here, the nodes are also incentivized with Sia coins and Storj coins, respectively.

Similarly, IOC offers a decentralized I/O name server (DIONS[20]) that enables document and identity storage on the blockchain with AES 256 encrypted messaging, along with a complete alias system. It initially supports PoW to generate enough coins, then incorporates PoS for the data storage job. The system fees are redistributed to all active nodes who put their coins at stake in the network.

Besides, Zyskind et al. (2015) provide an off-chain storage solution along with a decentralized personal data management model using blockchain that allows the data owner to share secret keys to the data requester via some secure channel such that only the users holding secret keys can access the data.

Table 2.4 compares different types of blockchain platforms according to various dimensions matching the needs for sharing user data. All the listed blockchain platforms have their own technical aspects, design, and consensus algorithms. However, they are similar in the way they try to provide a blockchain-based solution for the on-chain or off-chain user data storage through an incentivizing model in the competitive marketplace.

[16] https://ipfs.io/
[17] https://filecoin.io/
[18] https://sia.tech/
[19] https://storj.io/
[20] https://github.com/IOCoin/DIONS

Table 2. 4. Different blockchain for data storage

Blockchain Types	**Consensus Process**	**Consensus Decide**	**Smart Contracts**	**Efficiency**	**Data Storage**
Bitcoin	Public-Permissionless	Miners (Proof of Work)	No	Very Low	Tx Data
Hyperledger (Sawtooth)	Private-Permissioned	Selected Nodes (Proof of Elapsed Time)	Yes	Very High	Tx Data/ Metadata
MultiChain	Private-Permissioned	Selected Nodes (Practical Byzantine Fault Tolerance and round-robin)	No	Very High	Local storage
Ethereum	Public-Permissionless	Miners (Proof of work)	Yes	Low	Tx Data/ Metadata
Storj	Public-Permissionless	ERC20 Token (Proof of work)	Yes	High	Decentralized storage
Swarm	Private-Permissionless	Miners (Proof of work)	-	High	Decentralized storage
Sia	Private-Permissioned	Miners (Proof of Work + Storage)	Yes	High	Decentralized storage
Quorum	Consortium-Permissioned	Quorumchain	Yes	High	Local store
IPFS+ Filecoin	Public-Permissioned	Proof of replication	Yes	-	Decentralized storage
Blockstack[21]	Public-permissioned	Miners (Proof of Work)	No	Low	Decentralized storage (Gaia[22])
IOC[23]	Public-permissionless & permissioned	Proof of Work & Proof of Stake	Yes	low	1MB storage capabilities

[21] https://blockstack.org/
[22] https://github.com/blockstack/gaia
[23] https://iodigital.io/

Since different blockchains offer distinctive features, blockchain interoperability is also very important to construct a proper system for the storage and sharing of user data. Some of the desirable characteristics of an ideal system to enable users to store and share user data are:

(i) Rewards for honesty: An incentive mechanism that provides rewards to each of the participating entities (data producer, data provider, and data consumer) for doing their part honestly or penalties for their malicious behaviors, in such a way that rational parties are persuaded to be truthful.

(ii) Authorization: The system should define access policies and specify access rights and privileges to each party involved in the system

(iii) Integrity: None of the participants is authorized to alter the audits and the agreed terms and conditions.

(iv) Auditability: The system must provide complete audit features to trace back data's state and route. It should be possible to track every action performed by the participants.

Currently, the big social media giants such as Facebook, Google, and Amazon hoard user data whenever users do anything online and use it to sell targeted advertisements (their main source of revenue). These social media networks manage the entire internet and do not give users a real choice or awareness of what data about them are kept, nor how it will be used. They provide very few control options and no rewards for users in exchange for their data except for free access to their limited services. Moreover, all websites and enterprise applications normally run on the current version of the centralized internet. Hence, many technology experts have suggested that the blockchain will make the internet decentralized and become a crucial component of the next generation of the internet (Domingue, 2017).

In the context of decentralizing the internet, Ethereum blockchain provides the Ethereum Virtual Machine (EVM) to enable enterprise applications called *dapps*[24] that can run on the decentralized network. Ethereum is one of the world's biggest blockchain platforms supporting smart contracts with which users can code, deploy, and execute their contracts to deal with assets such as user data, commodities, or goods in the supply chain marketplace. Ethereum has *'ether'* as its own virtual currency, which can be used to pay a transaction fee and to provide a primary liquidity layer for

[24] https://www.investopedia.com/terms/w/wei.asp

exchanging digital assets. The next sections briefly explain Ethereum, smart contracts, and MultiChain.

2.8.2 Ethereum

Ethereum is based on the same underlying Bitcoin blockchain standards and proof of work consensus protocol. However, it is usually considered as an open-source platform to create dapps (decentralized applications leverage on the blockchain) where users interact with the online services in a distributed peer-to-peer manner that takes place on a censorship-proof foundation. Developers can create interfaces and business logic with any of the known programming languages and tools.

There are "messages" in Ethereum that can be created either by an external entity or internally by a contract, unlike the Bitcoin transaction, which can only be created externally (Buterin, 2015). There is also an explicit option for Ethereum messages to contain data and the recipient of Ethereum messages to return a response. Ethereum also has "transactions" as the signed data package that stores a message to be sent from an externally owned account. Transactions contain the recipient of the message, a signature identifying the sender, the amount of ether, and the data to be sent, as well as two values called STARTGAS and GASPRICE. STARTGAS is the maximum number of computational steps the transaction execution can take, and GASPRICE is the fee per computational step which the sender pays in "*Wei*," which is the smallest denomination of ether in the Ethereum network (1 Wei = ether/1e18)[25].

The state in Ethereum is made of accounts, each consisting of a 20-byte address and state transitions. An account contains four fields (Buterin, 2015), which are:

i. The nonce: a counter used to make sure each transaction can only be processed once.
ii. The account's current ether balance
iii. The account's contract code if present
iv. The account's storage (empty by default).

Since Ethereum supports smart contracts, it can be used as a semi-financial application such as on-blockchain escrow, which allows users to enter into contracts and manage them using their ether

[25] https://www.investopedia.com/terms/w/wei.asp

to deal with non-monetary assets such as the user profile data. The issues with the public Ethereum blockchain have always been scalability, a very long block validation time and GASPRICE. Currently, Ethereum imposes a PoW consensus algorithm, which uses a high computational power (i.e., electricity). Some of the experts in this domain are now admitting that there are problems with the underlying architecture that need to be addressed before they can deliver production-ready solutions. It is expected that the Ethereum community will adopt the PoS consensus protocol by 2021, eliminating the problem of high computational power.

According to Vitalik, co-founder of Ethereum, public blockchain is a far less efficient worldwide computer and ledger than technologies that have existed for over four decades. This is true in the context of validating all the blockchain transactions with the current protocol and infrastructure. We can investigate a mathematical derivation for the efficiency of the public blockchain, assuming the current Ethereum network size to be around 15000 nodes. The average block interval for Ethereum is around 12s as obtained from *eth gas-station webpage*, and it takes tentatively 200ms for a computer (CPU) to process. So, by assuming a block of 200ms of CPU equivalence time, we will have Safety factor as:

Safety factor = block interval / block verification time = 12s/200ms= 60.

The safety factor of 60 is large. Therefore, a node in the Ethereum network spends about (1/60) of its time to do the computation, which is needed to keep the *uncle* rate down. *Uncles* occur when a valid block propagates slowly and cannot make it into the long-term consensus. It is caused by high network latency, DDOS attacks or some other network interference. This dramatic increase in the uncle rate indicates that the block gas limit is too high, thus making the blocks too large to propagate efficiently. Then, the net overhead will be:

Net overhead = network size * safety factor = $0.9*10^6$

It confirms that everything is processed tentatively with a 900k factor difference for efficiency. Therefore, the public blockchain does not give raw execution efficiency.

However, in return for this inefficiency, we can have the following goals that can be achieved with the public blockchain: (i) Censorship resistance, (ii) Fraud resistance, (iii) Transparency, (iv) Robustness, (v) Integrity, (vi) Interoperability.

(i) Censorship resistance: Resistance to the transaction being interfered with by a third party, such as banks, governments, or internet service providers (ISPs).

(ii) Fraud resistance: Resistance from interference by the first party with whom we interact, such as exit scams.

(iii) Transparency: It is possible to see all the interactions and trace back the actions performed by the system to the actual initiating entity.

(iv) Robustness: The whole service remains online even when a few computers fail for some reason.

(v) Integrity: No one can modify the audits and agreed-upon terms and conditions.

(vi) Interoperability: We can design systems that interoperate with other applications regardless of whether those other applications actively cooperate. The Ethereum contract Application Binary Interface (ABI) is one of the most successful open APIs available, as anyone can launch a contract, and once they launch, anyone else can programmatically call that contract and interact with it.

Besides, we should not be using blockchain for everything. Blockchain is needed only to run the core business logic. Most of the computation of applications does not have to be done by every single node in the network inside of consensus. Instead, it can be done by individual users on the client-side or through some sub-system that is connected to a blockchain.

2.8.3 Smart Contracts

Research in smart contract technology has evolved from conceptual-based architectures to application-oriented scenarios. According to the systematic mapping study conducted by Alharby et al. (2018), 64% of a total of 188 relevant papers on smart contracts in 2018 were from the applications category: a significant increase compared to 2017 (Alharby & Moorsel, 2017) when they found only 24 papers in total. Their 2017 study shows that about 66% of the papers focused on the conceptual level finding and tackling smart contract issues. We have now seen many academic researchers taking up smart contract technologies in actual building applications on top of the blockchain.

Smart contracts are now recognized as instances of contracts deployed on the Ethereum blockchain (Buterin, 2015), although it was originally coined in a paper by Szabo (1997) to design electronic commerce protocols between strangers on the internet. A smart contract stores the rules that

(a) Negotiate the terms of the contract,

(b) Automatically verify the contract, and

(c) Execute the agreed terms.

A smart contract consists of different functions that might be called from outside of a blockchain or by other smart contracts. Blockchain coupled with smart contract technology removes the reliance on a central system between the transaction parties (Shrestha et al., 2017). Since the smart contracts are stored on the blockchain, all the connected parties in the network will have a copy of them (Shrestha & Vassileva, 2018b).

A smart contract can execute an agreed stored process when triggered by an authorized or agreed event. All contract transactions are stored in chronological order for future access, along with the complete audit trail of events. If any party tries to change a contract or transaction on the blockchain, all other parties can detect and prevent it. If any party fails, the system continues to function with no loss of data or integrity. It, therefore, creates a single large secure computer system logically, without the risks, costs, and trust issues of a centralized model, although it does have its own issues with security and deployment cost.

The Solidity[26] programming language is used to write smart contracts because it only allows performing basic operations on its basic data types, resulting in lightweight code. The EVM (Ethereum Virtual Machine)[27] code is used in the contracts, which consist of bytes, each representing an operation. The code can access the amount of *Wei*[28] sent in the transaction and data of the incoming message, block header data, and return a byte array of data as an output. With the implementation of EWASM (Ethereum Web Assembly)[29] in the near future, smart contract development can be done in any other programming language besides Solidity, which will speed up the function call between Web Assembly and JavaScript (JS).

2.8.4 MultiChain

On-chain data storage can be successfully achieved with a limited number of peers in a private blockchain. MultiChain, for example, being a private blockchain, has the potential to replace the

[26] https://solidity.readthedocs.io/en/v0.5.3/
[27] http://ethdocs.org/en/latest/introduction/what-is-ethereum.html#ethereum-virtual-machine
[28] https://www.investopedia.com/terms/w/wei.asp
[29] https://github.com/ewasm/design

traditional centralized databases used to store user data in a decentralized manner, offering more cryptographic auditing features (Greenspan, 2013). It allows users to optionally store any published data in an off-chain, saving storage space and bandwidth. It can hash up to 1 GB per item (off-chain data) into the blockchain, with the data itself delivered rapidly over the P2P network. MultiChain provides the privacy and control required in an easy-to-configure and deploy package (Greenspan, 2013). It supports UNIX and Windows servers and comes up with a rich JSON-RPC API for easy integration with existing systems. The private blockchain also tries to solve privacy and openness issues through integrated management of user permissions with some objectives (Greenspan, 2013), such as:

(1) To enable selected participants to see the blockchain's activity,

(2) To ensure that only selected transactions are permitted,

(3) To securely conduct mining without proof of work and its associated costs.

2.9 Conclusion

This chapter discussed different philosophical frameworks for research, design, and research methods and provided the reasoning behind the mixed research approach adopted by my research. The different cycles of the design science research were summarized and presented with an activity diagram in the context of my research. Each of the steps after the relevance cycle presented as the research methodologies from the design and rigour cycles is provided in detail in the next chapters.

Also, this chapter presented an overview of the characteristics of three distinct categories of user data: (1) User profile data, (2) User documents, and (3) Research data. First, it discussed the user modeling, personalization, sharing of the user model data, ownership status with regard to the data and incentives for data sharing. Secondly, it provided user documents and their evolution in the context of the participative web. Lastly, it identified various stages and scopes of research data along with the motivations for data sharing.

A review of the literature on sharing user data of each of the three types was presented. The heuristic cognitive-behavioral models were presented based on the extended technology acceptance model (TAM), privacy model, and trust model. Different data sharing approaches based on a single system, centralized servers, decentralized (agent-based, peer-to-peer, service-based) systems, and blockchain networks (decentralized + incentives) have been reviewed. No deployed

system was found that enabled micro-payments or reputation aggregation at the scale needed for blockchain-based file-sharing networks. Adopting a systematic way to empirically design and develop the blockchains and smart contracts-based user data sharing framework for incentivizing the data owners is a promising direction of future work to fill the gap.

Finally, the chapter described blockchain and smart contract technologies along with the presentation of the literature survey on the different existing blockchain-based systems for data storage and sharing. The next chapter introduces the DUDS framework for decentralized user data sharing with the aid of the permissioned Multichain blockchain and smart contracts supported Ethereum blockchain.

3 BLOCKCHAIN AND SMART CONTRACTS FOR DATA SHARING

This chapter identifies blockchain and smart contracts as promising tools to develop the platform that utilizes my proposed DUDS (Decentralized User Data Sharing) framework. Subsequently, the chapter provides details of the design principles of the DUDS framework by incorporating public blockchain, permissioned blockchain, and smart contracts as the distributed ledger technologies. The DUDS framework enables supporting user incentives for sharing user data in a decentralized P2P fashion.

3.1. DUDS – Decentralized User Data Sharing Framework

This section describes the DUDS framework, which ensures the preservation of user control over their data while supporting data sharing among different enterprises in a decentralized manner. Some of the issues of data sharing are with whom the data is going to be shared and by what means, and how can the data owners be incentivized. With the smart contracts stored in the blockchain, users can retain the ownership of data and are incentivized as per the agreed terms for sharing their data. In addition to privacy, user control and incentives for sharing, this framework ensures security and scalability.

The privacy-related legislation, General Data Protection Regulation (GDPR[30] in EU) as of May 25, 2018, regulates the processing of the personal data of individuals and demands personal data erasure. However, data on the blockchain are always immutable. This contradiction could be a challenge to adopting blockchain as part of the solution. Therefore, I present a general architecture of the user data sharing system based on blockchain and off-chain data storage, as shown in Figure 3.1. Then I present my proposed DUDS framework, as shown in Figure 3.2. The DUDS framework comprises a permissioned blockchain as a solution to both on-chain and off-chain data storage, encryption, hashing, and tracking of data, together with public blockchain and smart contracts for access control and enabling transactions with digital tokens or virtual credits.

[30] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679

The general architecture ensures that the actual user data is never exposed to the public blockchain. The user data is first hashed and encrypted before uploading into the off-chain storage. The data owners from their client applications can directly store them on the off-chain storage. The terms and conditions regarding the access to user data are encoded in smart contracts along with the metadata and hash of the data and published on a blockchain platform (Ethereum). The hashes of the data on the blockchain prevent the middleware from tampering with the data. The content-based addressing makes hashes of data serve as their identifier for retrieval. When the data consumer invokes the smart contracts for accessing the user data, only the successful invocation of the contracts results in the release of the key for decrypting the user data. The trusted program (Ning et al., 2018) then extracts the hash from the blockchain, uses this hash to retrieve the data from the off-chain storage, decrypts it, and releases the data to the data consumer while settling the incentives for the data owner.

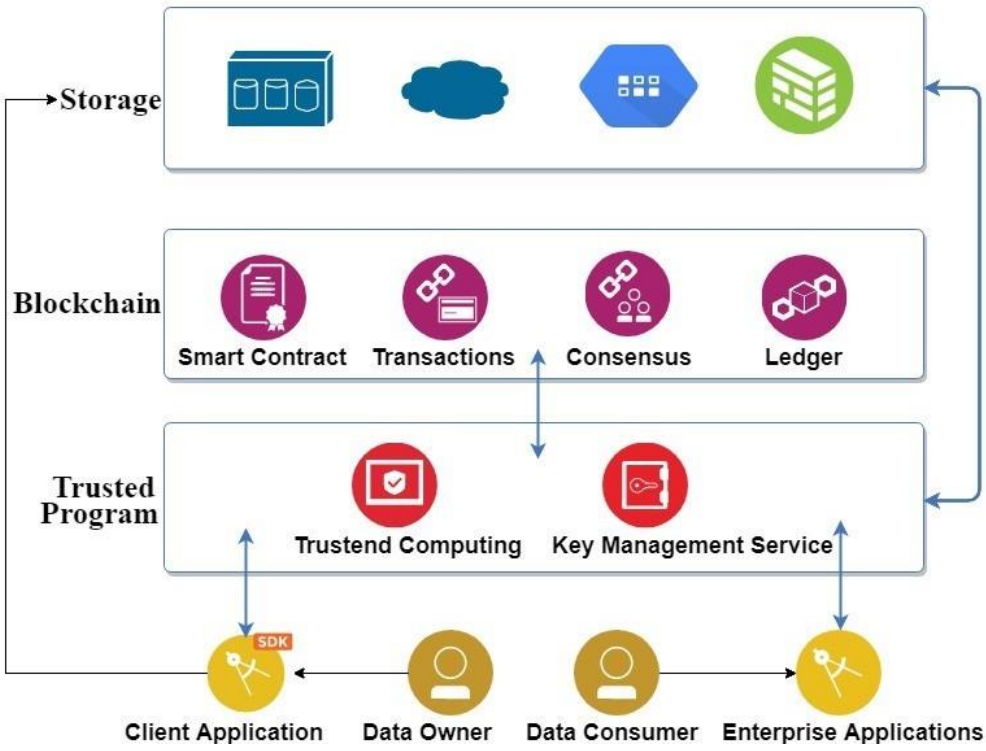

Figure 3. 1. User-controlled privacy-preserving data sharing architecture

Blockchain and smart contracts support users by giving the users full transparency over who accesses their data, when and for what purpose, allowing the users to specify a range of purposes of data sharing, kinds of data that can be shared, and classes of applications/companies that can

access the data, and providing an incentive to the users for sharing their data (in terms of payment for the use of the data by applications, as specified by the contracts). The general architecture presents the underlying off-chain user data storage mechanism, which could be a centralized data store hosted by a trusted party. When trust resides within a centralized service provider for all the storage and management of data, it is hard to mitigate the various risks, for example, of data being misused, hacked, or sold to any other bodies without user consent and even destroyed when the company defaults.

Therefore, I present a new platform with a separate private permissioned blockchain, MultiChain as a solution to both on-chain and off-chain data storage, encryption, hashing and tracking of data, together with Ethereum (for access control). Off-chain blockchain implementation with user data storage can be successfully achieved with the limited number of peers in the MultiChain (Greenspan, 2013). Users can optionally store any published data off-chain, which saves storage space and bandwidth. A similar idea of storing off-chain data and accessing them via MultiChain has also been proposed in other research works (Yang et al., 2019; Ferrer-Sapena & Sánchez-Pérez, 2019). MultiChain nodes handle key operations, such as hashing and encrypting the user data, storing the encrypted file locally (outside of blockchain), committing the hash of the file on the blockchain, searching the required data, verifying the data and delivering the data.

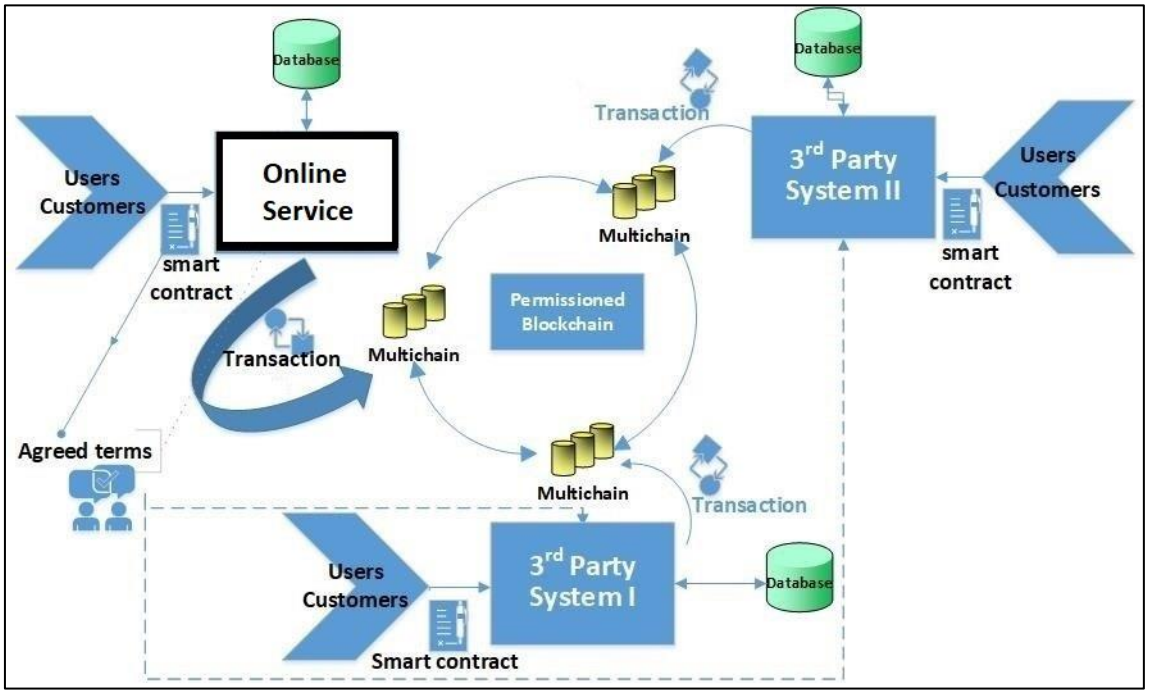

Figure 3. 2. DUDS framework

The DUDS framework, as shown in Figure 3.2, uses two blockchains: MultiChain to share user profile information among enterprises in their private network, and Ethereum to store securely the access-control policies and user data sharing preferences as smart contracts. The main design principles for the DUDS framework are ascertained by the following design choices:

- Only selected nodes can connect to the permissioned blockchain network.
- Each node stores data locally.
- Data sharing is completely decentralized.
- Users get full transparency over who accesses their data, when, and for what purpose.
- Users specify the purposes of data sharing, which kinds of data can be shared, and which nodes can access the data.
- Users receive an incentive for sharing their data (in terms of payment for the use of the data by applications, as specified by the contracts).

A user registers in the system by providing her basic profile information and activates the smart contracts on an Ethereum network node by simply selecting data sharing preference on the web form. The Ethereum node automates the functionality to support the user-controlled privacy regarding (a) with whom the user's data will be shared and (b) how the user will be incentivized once her data are being shared with other third parties. The communication with the Ethereum network node is made directly with individual hosted enterprise nodes. Since the smart contracts are stored on Ethereum, all the users and hosted nodes have their own ether addresses, which they use to safely store ethers and transfer them into their own cold wallet (offline wallet) and pay gas fees for the transactions. Once the users' data are being used by any other participating organization, the corresponding users (data provider) will be incentivized with ether as per the smart contracts. The user does not have to know all the technical details happening at the DUDS level.

3.1.1 Data Sharing Solution

For sharing user data among enterprises, the MultiChain blockchain is installed on each of the participating enterprise nodes or nodes that are involved in the ultimate data sharing process. Public key encryption is an underlying technology of MultiChain, so all the connected nodes generate their own pairs of public addresses and private keys. The nodes publish public keys associated with their own nodes, encrypted user data as items and secret keys encoded with the recipient's public

key into respective streams (append-only ledger for any data) in accordance with the smart contracts.

Since the MultiChain is a permissioned blockchain in nature, the node that starts the chain with the genesis block remains the admin of the consortium network by default. The admin node grants other nodes admin privileges, too, as per their mutual agreement. The standard permissions for other nodes in the permissioned blockchain: 'connect,' 'send,' 'receive,' 'issue,' 'create,' 'mine,' 'admin,' and 'activate' are set by the admin nodes. The consortium network limits the blockchain access to a group of permitted users by expanding the "handshaking" process when two blockchain nodes connect governed by the following Algorithm A.

ALGORITHM A: Peer-to-peer (P2P) connection (Greenspan, 2013)

1. A node is identified as a public address.
2. A node verifies that the other's address is on its own version of the permitted list.
3. A node sends a challenge message to the other party.
4. A node sends back a signature of the challenge message to prove their ownership of the private key linked to the public address they presented. If either node is not satisfied with the results, it aborts the P2P connection.

The basic chain parameters set in the Multichain are as below:

#Basic chain parameters

1. chain-protocol = multichain # Chain protocol
2. chain-description = MultiChain model # Chain Description
3. root-stream-name = root # Root stream name
4. root-stream-open = true # Allow anyone to publish in root stream
5. chain-is-testnet = false # Content of the 'testnet' field of API responses, for compatibility.
6. target-block-time = 15 # Target time between blocks (transaction confirmation delay), seconds. (5 - 86400)
7. maximum-block-size = 8388608 # Maximum block size in bytes.

The basic chain parameters are set to limit permissions to any newly added nodes. Similarly, the global permissions in the Multichain are as shown below:

#Global permissions

1. anyone-can-connect = false # private blockchain.
2. anyone-can-send = false # transaction signing is not restricted by address.

3. anyone-can-receive = false # transaction outputs are restricted by address.
4. anyone-can-receive-empty = true #without permission grants, asset transfers and zero na$
5. anyone-can-create = false # selected can create new streams.
6. anyone-can-issue = false # selected can issue new native assets.
7. anyone-can-mine = false # selected can mine blocks (confirm transactions).
8. anyone-can-activate = false # selected can grant or revoke connect, send and receive permissions.
9. anyone-can-admin = false # selected can grant or revoke all permissions.
10. support-miner-precheck = true # Require special metadata output with cached scriptPubKey for input, to support advanced mine$

The multichain daemon is created using the following command with the chain name model:

multichain-util create model

multichaind model –daemon

This creates the MultiChain Core Daemon of the existing version so that other nodes can connect to the starter node using its IP and port number:

multichaind model@[ip]:[port], (E.g. multichaind model@192.168.204.132:8353)

The creation of the nodes offered the individual addresses for those new nodes that are acknowledged by the admin node to grant a “connection” permission to them into the permissioned blockchain.

multichain-cli model grant [address] connect, send, ...

This is the first step in creating the blockchain. While granting the connection permission, other standard permissions could also be set for other nodes.

Once the node is in the P2P network, it can take part in the data-sharing process. Through the smart contract, only the selected eligible nodes can access or consume the shared data by subscribing to the corresponding published streams. The data providers and the nodes offering user data are incentivized as per the negotiation made between the stakeholders. The functions (methods) are confined in the smart contract as per the roles of the participating entities to successively execute different tasks between the data provider and the consumer, as explained in the next section.

The stream in MultiChain is used for general data storage and retrieval. It offers data integrity via immutable timestamped append-only ledger and authenticity via digital signatures as every

transaction is signed; however, confidentiality is to be ascertained through some form of encryption techniques (Greenspan, 2013). To ensure data confidentiality, I employed a combination of symmetric and asymmetric cryptography for encrypting the data. This method utilizes three blockchain streams:

1. Public-keys stream: This append-only archive or directory contains the participants' public keys under the RSA public-key cryptography.
2. Data-item stream: This append-only archive offers space to the participants to publish the encrypted data using a symmetric AES cryptography scheme.
3. Access-item stream: The data provider creates stream-entry with a secret key to the encrypted data, encoded with the data consumer's public key to provide data access.

Therefore, only a subset of blockchain participants with the associated private key can decode the encoded secret and then fully access the encrypted data by decrypting it with the secret key.

The application accesses all the MultiChain Community commands using the JSON-RPC API, as they are available under an open-source license[31]. All the participants in the system with their Ethereum account addresses are in the MultiChain network. The data are stored in the local storage of the nodes. Once the data sharing is completed, the smart contracts get triggered, and with their successful execution, the tokens are transferred from the data consumer's Ethereum address to the data providers' account while delivering the requested data (with verified hashes) to the local storage of the consumer node using the same path.

3.1.2. User Incentives for Sharing

The smart contracts are committed into the Ethereum, which guarantees the user receives the incentive when the user data is consumed by the consumers or participating enterprise nodes. All the users (data providers) and participating enterprise nodes (data consumers) have Ethereum addresses that interact with the smart contracts. The users can decide which consumers (applications or companies) can access their data. Here, the application sets the terms and conditions that the users agree to allow the enterprise a license to collect and use the contents before using their services and specifies that the users who share their data retain ultimate ownership to

[31] https://github.com/MultiChain/multichain/blob/master/COPYING

their content, but the enterprise node also receives a limited perpetual license (and right to sublicense) to distribute such contents (see Appendix II for smart contracts). Smart contracts give users the ability to set data-sharing preferences. So, in the beginning, only the selected enterprises access the user data by subscribing to the published streams in the MultiChain. The incentive is given either in the form of virtual credit or a digital token by transferring ethers to users' ether addresses or both. Figure 3.3 illustrates a flowchart for the workflow logic with the process cycle, including the creation of contracts for handling successful payments for data sharing.

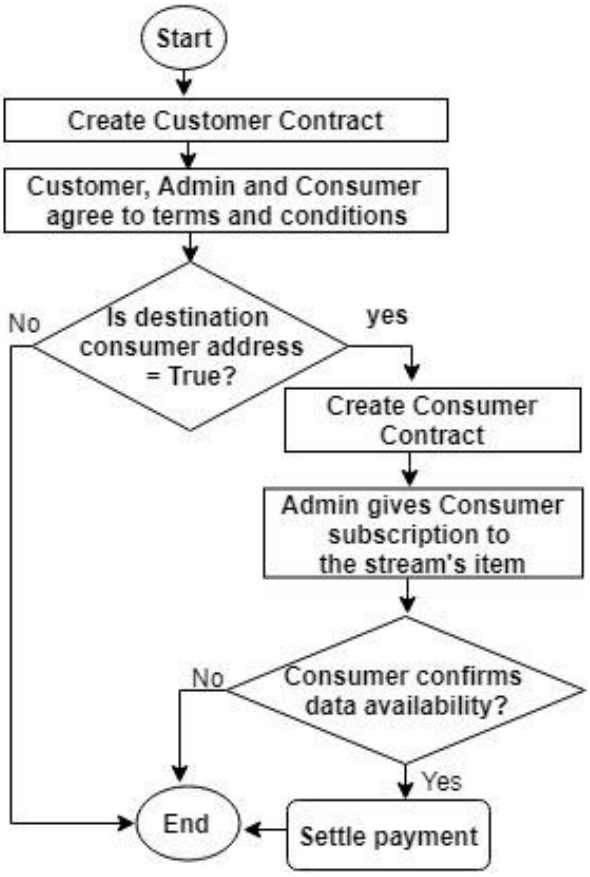

Figure 3. 3. Flowchart for workflow logic of smart contracts

The functions (methods) are confined in the smart contract as per the roles of the participating entities to successively execute different tasks between the data provider and the consumer. The steps to generate the general smart contract codes are described in *Algorithm B* and *Algorithm C*.

ALGORITHM B: signedTermsAndConditions

Input: A_{cu}, A_{co}, $A_{cc,}$ deposit, dataPrice, contractState

1. A_{cu}, A_{co}, A_{cc} are the set of all ether addresses of customers (data provider), data consumers and contracts respectively.
2. Grant access to only $a_{cu} \in A_{cu}$, $a_{co} \in A_{co}$ who got registered into the system.
3. Change the contract state to Created.
4. a_{cu} deposits e_d.
5. Set data price to ep such that $e_p = ½\ e_d$.
6. Contract balance of a_{cc} is $r_b = e_d$, where $a_{cc} \in A_{cc}$.
7. Allow a_{co} to choose the customer data of its interest.

ALGORITHM C: confirmedDataConsume

Input: A_{cu}, A_{co}, $A_{cc,}$ deposit, dataPrice, contractState

1. A_{cu}, A_{co}, A_{cc} are the set of all ETH addresses of customers (data provider), consumers and contracts respectively.
2. a_{co} decides to consume the customer data acu, pays $2e_p$ such that consumer's deposit = e_p.
3. Contract balance of acc is $r_b = e_d + 2 * e_p$.
4. Change the contract state to Locked.
5. Grant a_{co} to access the customer data.
6. a_{co} confirms data availability
7. Change the contract state to Inactive.
8. Transfer deposit e_p from a_{cc} back to a_{co}.
9. Remaining Contract balance of a_{cc} becomes $r_b = e_d + e_p$.
10. Transfer r_b to a_{cu}.

The variables hold the Ethereum addresses, incentives, and contract state. A setter function is created to make its parent or child change the state of the contract if needed, and the compiler automatically generates getter functions for all public variables. Modifiers are added in the contracts to support access control and inheritance by restricting the functions with some conditions. Events are added to keep arguments in the transaction logs that notify clients about what is happening with the contracts. This model does not include the attestation of the smart contracts, which might be required to ensure the reliability of the contracts. However, the attestation authority, as needed, can be added accordingly into the system.

The contracts (see Appendix III) following the algorithms B and C were deployed on the Remix. Thus, as seen from the algorithms, after accessing the user data, the corresponding user is incentivized by the data consumer. This DUDS framework delivers a usable blockchain-based model for a collection of user data, providing accountability of access, maintaining the complete and updated information with a verifiable record of the provenance, including all accesses/sharing/usages of the data.

3.2. Conclusion

This chapter introduced the DUDS framework for decentralized user data sharing with the aid of the permissioned MultiChain blockchain and smart contracts supported Ethereum blockchain. Chapter 5 covers the proof of concept of the DUDS framework with the example scenarios for sharing user data in three different areas: tourism, research, and an online shopping cart that offer rewards to the data owners in terms of payment through blockchains for the use of the data by applications, as specified by the contracts. But it is especially important to understand the user behavioral model before developing the minimum viable products. Chapter 4 delves into the formulation of the user behavioral model to determine the factors that affect the intention to adopt such platforms based on the DUDS framework.

4 EXTENDED TAM AND CONCEPTUALIZING TRUST

A critical aspect of developing a DUDS framework-based platform is user acceptance. I carried out a study that helps to identify the factors affecting the users' behavioral intention to adopt such a system. Section 4.1 of this chapter presents the findings of a study of user experience with a prototype system based on the DUDS platform. The extended technology acceptance model (TAM) was used to evaluate the user experience model. The content of this section is based on my published article (Shrestha & Vassileva, 2019b). Furthermore, section 4.2 provides the findings from a literature survey to conceptualize trust and its constituents in the realm of security and attitudinal privacy for blockchain-based platforms and examine the behavioral intention of users towards the adoption of such systems.

4.1 User Acceptance of Usable Blockchain-Based Research Data Sharing System: An Extended TAM-Based Study

In the light of the findings from Chapters 2 and 3, blockchain technology has evolved as a promising means to transform data management models in many domains, including healthcare, agricultural research, tourism domains, etc. A usable blockchain-based system allows users to:

- create a proof of ownership and provenance of the conducted activities,
- share data without losing control and ownership of it, and
- provide incentives for sharing with full transparency and control over who accesses their data, when, and for what purpose.

To study the user acceptance of the DUDS framework, I used the case of sharing research data. For data sharing in the scientific research domain, I provided a usable blockchain-based prototype model that incentivizes the dataset owners with digital tokens, proper acknowledgment, or both, while giving access to the detailed information of all data to them in an immutable and incorruptible database. This would be the first in the area of research data sharing systems, as there was no deployed system enabling micro-payments or reputation aggregation for research data-sharing

networks based on blockchain until 2018. In 2018, I worked in collaboration[32] with ARTiFACTS, a US-based company that was working with blockchain-based collaboration and attribution platforms for scholarly research. Our collaboration was intended to study the value of applying a blockchains-based approach and its usability and usefulness for sharing scholarly work. ARTiFACTS initially used the Ethereum Ropsten Test network[33] and now they are experimenting with the Hyperledger network and Bloxberg blockchain[34]. I set up the Hyperledger sawtooth permissioned node on the ARTiFACTS network at the University of Saskatchewan for both organizations to study blockchain's efficacy in sharing academic papers[35]. I conducted two empirical studies on usability and usefulness, one on our blockchain-based prototype approach (Shrestha & Vassileva, 2019b) and another on the ARTiFACTS system. The initial adoption of such blockchain-based systems is necessary for continued use of the services, but their user acceptance behavioral model was not well investigated in the literature. So, I used the Technology Acceptance Model (TAM) as a foundation and extended the external constructs to uncover how the perceived ease of use, perceived usability, quality of the system, and perceived enjoyment constructs influence the user's behavioral intention to use the blockchain-based system.

I used a TAM validated questionnaire as a tool to evaluate the user acceptance and the usage of a prototype of a DUDS platform. The results show that all the individual constructs of the behavior model significantly influence the intention to use the system, while their collective effect is found to be insignificant. The quality of the system and the perceived enjoyment have a stronger influence on the perceived usefulness than the perceived ease of use. The following section 4.1.1 is based on my published paper (Shrestha & Vassileva, 2019b).

4.1.1 Introduction

Chapter 3 already highlighted that the blockchain and smart contracts could provide a new type of platform that would be useful for sharing research data by providing solutions to the problems of privacy of user data and compliance to ethics standards and user consent agreements, as well as

[32] https://medium.com/@ARTiFACTS/artifacts-and-the-university-of-saskatchewan-collaborate-on-use-of-blockchain-to-improve-scholarly-7a3cd390eeee
[33] https://ropsten.etherscan.io/
[34] https://bloxberg.org/
[35] https://github.com/artifactsofresearch/sawtooth

researcher control and incentives for sharing. The most important criticisms to blockchain-based approaches to date relate to their performance and scalability; yet the rapid development of the technology allows it through thoughtful combinations of blockchains to achieve acceptable performance. A harder problem emerges related to the user acceptance of blockchain technology in non-currency-related application domains.

For example, due to the lack of familiarity with blockchain, it is not clear if researchers would be receptive to using blockchain technology in regulating access and sharing of research data. It is therefore important to study the user acceptance of blockchain-based applications for example if users understand the blockchain and smart contracts technologies and if they can competently share research data. Many studies have evaluated the performance of blockchain-based systems; however, to our best knowledge, no studies have focused on user acceptance of the blockchain-based system. To bridge this gap and advance research in blockchains- and smart contracts-based systems, we adopted the extended Technology Acceptance Model (TAM), which has been one of the most influential models to examine the indicators that affect the user's acceptance of the system (Venkatesh et al., 2003). We based this study on user evaluation of the DUDS framework on a prototype platform with the TAM validated tool deployed as a research instrument. We chose to investigate the influence of the perceived ease of use, perceived usefulness, perceived enjoyment, and quality of the system, on the participants' intention to use the system. We also analyzed the influence of the perceived ease of use, perceived enjoyment, and quality of the system on perceived usefulness.

We used the structural equation modeling technique and observed the values of path coefficient (β), p and R^2. The results of our investigation showed a stronger influence of quality of the system and perceived enjoyment on the intention to use construct for the blockchain-based research data sharing system while perceived usefulness and perceived ease of use had moderate and weaker effects respectively. Moreover, our results showed that the combined effect of all four antecedents, perceived ease of use, perceived usefulness, quality of the system, and perceived enjoyment on intention to use was found to be insignificant. Specifically, our results showed that the effect of perceived ease of use on perceived usefulness was insignificant. However, the quality of the system and perceived enjoyment had a stronger influence on perceived usefulness.

4.1.2 Background and Related Work

Most researchers, on an individual level, may feel reluctant to share their research data; however, they appreciate the overall benefits of data sharing, which was also concluded from the qualitative interviews-based study conducted by Whyte and Pryor (2011), Van Den Eynden and Bishop (2014). Those studies recognized six different ways of data sharing: private management sharing, peer exchange, community sharing, collaborative sharing, sharing for transparent government, and public sharing.

There are proposals in the literature (most prominently by Zyskind et al. (2015)) to use blockchain as an access control platform to ensure the privacy of data. Molina-Jimenez et al. (2019) mentioned some of the technical challenges present in the proposals adopting blockchain as part of their solutions. In my solution, I proposed a usable blockchain- and smart contracts-based platform for research data sharing. This platform is the basis for the user-acceptance study using extended TAM that I present in the next section.

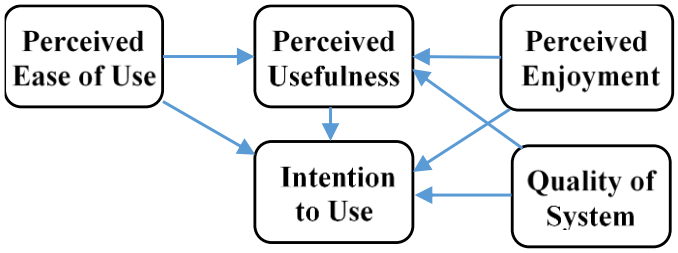

Figure 4. 1. An extended TAM model for our study

The classical TAM as mentioned in section 2.6 hypothesizes that the actual use of the system is determined by behavioral intention to use, which is, in turn, influenced by the user's attitude toward using the system and perceived usefulness and perceived ease of use of the system. The classical TAM is often criticized for excluding important context-based latent variables and external factors in the model (Melas et al., 2011). Since there is no literature on TAM in the context of blockchains and smart contracts-based applications, our research applied the extended TAM to distributed ledger technologies. We extended TAM, in our case, by adding external constructs: perceived enjoyment (Davis et al., 1992) and quality of the system (Koh et al., 2010), which are often added as influential variables to user acceptance of the information technology. Table 4.1 presents the definition of different study constructs for the extended TAM that we used in our research. The

corresponding extended TAM is shown in Figure 4.1. Based on the literature (Davis, 1989; Davis et al., 1992; Koh et al., 2010), we conducted a similar study, in our case, to investigate the user perception on the acceptance of blockchain-based applications.

Table 4. 1. Constructs and definition (Davis, 1989), (Davis et al., 1992), (Koh et al., 2010)

Construct	Definition
Perceived Ease of Use (PEOU)	The degree to which a person believes that using a particular system would be free of effort.
Perceived Usefulness (PU)	The degree to which a person believes that using a particular system would enhance his or her job performance.
Quality of System (QOS)	The degree to which a person is pleased, hence reducing users' psychological objection to the system or the loss of volition.
Perceived Enjoyment (PEnj)	The degree to which the use of technology is seen to be enjoyable.
Intention to use (ITU)	The degree to which a person has a behavioral intention to adopt the technology.

4.1.3 Methodology

In this section, I present our research hypotheses, research questions, measurement instruments and the demographics of participants.

Research Hypotheses

We based our study on user evaluation of prototypes of the blockchain-based research data-sharing framework with the questionnaire deployed as a research instrument to collect data. We set several hypotheses, based on the literature review, to investigate the constructs as given in Table 4.1, which are as follows:

1. H1: The *perceived ease of use* will significantly influence the *perceived usefulness* of the blockchain-based research data-sharing framework.
2. H2: The *perceived enjoyment* will significantly influence the *perceived usefulness* of the blockchain-based research data-sharing framework.

3. H3: The *quality of system* will significantly influence the *perceived usefulness* of the blockchain-based research data-sharing framework.
4. H4: The *perceived ease of use* will significantly influence the *intention to use* the blockchain-based research data-sharing framework.
5. H5: The *perceived usefulness* will significantly influence the *intention to use* the blockchain-based research data-sharing framework.
6. H6: The *quality of system* will significantly influence the *intention to use* the blockchain-based research data-sharing framework.
7. H7: The *perceived enjoyment* will significantly influence the *intention to use* the blockchain-based research data-sharing framework.
8. H8: The combined effect of *perceived ease of use, perceived usefulness, quality of the system and* perceived *enjoyment* will significantly influence the *intention to use* the blockchain-based research data-sharing framework.

Research Design

The study was approved by the Behavioral Research Ethics Board of the University of Saskatchewan. To contextualize the extended TAM tool, we provided participants at the beginning with a description of our blockchain-based DUDS framework for sharing research data (prototype model). Thereafter, we presented the participants with an online survey through SurveyMonkey. The survey instrument adapted to the context of our study was based on constructs validated by (Davis, 1989), Davis et al. (1992), and Koh et al. (2010). The instrument consists of six items for perceived ease of use, six items for perceived usefulness, four items for quality of system, three items for perceived enjoyment and four items for intention to use. The respective items (questions) in the constructs are shown in Table 4.2. See Appendix IV for the participants' consent form. We measured the responses to the items on a 7-scale Likert scale from 1 = extremely unlikely to 7 = extremely likely. A total of 22 participants took part in the study, but upon data cleaning, 20 participants' responses were left for analysis. We recruited participants from academia, who had some research experience. Specifically, around 47% of participants were somewhat familiar, and

53% were highly familiar with other research content sharing social networks such as ResearchGate[36], Mendeley[37] or Orchid[38]. Table 4.3 highlights the demographics of the participants.

Table 4. 2. Constructs and items (Davis, 1989), (Davis et al., 1992), (Koh et al., 2010)

Construct	Items
Perceived Ease of Use (PEOU)	peou1 - Learning to operate this system would be easy for me.
	peou2 - I would find it easy to get this system to do what I want it to do.
	peou3 - My interaction with this system would be clear and understandable.
	peou4 - I would find this system to be flexible to interact with.
	peou5 - It would be easy for me to become skillful at using this system.
	peou6 - I would find this system easy to use.
	How confident are you in the ratings made on this page?
Perceived Usefulness (PU)	pu1 - Using this system would enable me to accomplish data sharing tasks more quickly.
	pu2 - Using this system would improve my performance with regard to sharing research data.
	pu3 - Using this system would increase my productivity.
	pu4 - Using this system would increase my effectiveness.
	pu5 - Using this system would make it easier to share the data.
	How confident are you in the ratings made on this page?
Quality of System (QOS)	qos1 - I would be satisfied with the research paper sharing methodology of this system.
	qos2 - I would be satisfied with the feature of creating proof of the existence of the research work (ownership).
	qos3 - I would be satisfied with the feature of allowing users to set permissions for the way to share their data.
	How confident are you in the ratings made on this page?
Perceived Enjoyment (PEnj)	penj1 - I would be satisfied to use this system to share research data
	penj2 - I would like to use this system to share research data.
	How confident are you in the ratings made on this page?

[36] https://www.researchgate.net/
[37] https://www.mendeley.com/
[38] https://orcid.org/

Intention to use (ITU)	itu1 - I believe it is worthwhile to use this system to share research data.
	itu2 - I will use this system to share research data.
	itu3 - I intend to use this system to share research data in the future.
	How confident are you in the ratings made on this page?

Table 4. 3. Participants' demographics

Respondents' characteristics [(Female, male) = (45%, 55%)]	**Criteria**	**Percentage**
Age	18 to 24	9.09%
	25 to 34	72.73%
	35 to 44	18.18%
Highest education level	Grad-High school	4.55%
	Bachelors	22.73%
	Masters	54.55%
	PhD	18.18%
Current occupation	Student	59.09%
	Researcher	31.82%
	Faculty	4.55%
	Other	4.55%
Ever served as a reviewer	Yes	42.86%
	No	57.14%
Familiar with blockchain technologies and smart contracts	Extremely familiar	13.64%
	Very familiar	27.27%
	Somewhat familiar	27.27%
	Not so familiar	31.82%
Familiar with research/ social networks (e.g., ResearchGate, Mendeley, ORCID)	Extremely familiar	9.52%
	Very familiar	42.86%
	Somewhat familiar	47.62%

4.1.4 Result

In this section, I first present and briefly analyze the collected data with descriptive statistics. Then, I present our results of the structural equation model (SEM), which includes the measurement models (internal consistency, composite reliability, average variance extracted, KMO, and Bartlett's test of sphericity), structural models (exploratory factor analysis, regression analysis) and

brief analysis of the results. For the second part, I started by fitting the measurement models to the data, and later I tested the underlying structural models. The calculations of descriptive statistics in this study were carried out using MS Excel and IBM SPSS Statistics 25.

Descriptive Statistic

Since I measured the responses to the items on a 7-level Likert scale, I categorized the scale in seven score ranges to analyze the score for each item and overall impression of the construct. Table 4.4 provides the category of different score ranges of the 7-scale Likert scale. Table 4.5 to Table 4.9 summarize data collected for all the items in perceived ease of use, perceived usefulness, quality of system, perceived enjoyment, and intention to use constructs of our model respectively.

Table 4. 4. Categorization for score range

Score range	Category
$6 < x \leq 7$	Extremely High
$5 < x \leq 6$	Quite High
$4 < x \leq 5$	Slightly High
$3 < x \leq 4$	Neither
$2 < x \leq 3$	Slightly Low
$1 < x \leq 2$	Quite Low
$0 < x \leq 1$	Extremely Low

The obtained scores for different selected constructs indicate that user perceptions on the benefits of using the proposed system should be maintained or enhanced by making improvements in order to achieve a higher level of score category. The preliminary descriptive statistic of the obtained data shows that all the constructs provide a significant impression in the context of user acceptance of the usable blockchain-based research data sharing prototype. Figure 4.2 shows the average results of the constructs, which are all in the range 5 to 6; therefore, they qualify for the *quite high* category.

Table 4. 5. Analysis of Perceived Ease of Use (PEOU)

Indicators	Score	Std. Deviation
Ease of Learning	6	0.726
Controllable	5.65	1.04
Understandable	5.55	0.945
Flexible	5.7	1.129
Effort to Skillful	5.75	0.911
Easy to Use	5.8	1.057
Total Average	5.742	
Category	Quite High	

Table 4. 6. Analysis of Perceived Usefulness (PU)

Indicators	Score	Std. Deviation
Work More Quickly	**5.65**	0.934
Job Performance	**5.3**	1.261
Increase Productivity	**5.1**	1.411
Effectiveness	4.95	1.539
Makes Job Easier	5.95	0.999
Useful	6.05	1.191
Total Average	5.5	
Category	Quite High	

Table 4. 7. Analysis of Quality of System (QOS)

Indicators	Score	Std. Deviation
Satisfy with research file sharing method	5.8	0.834
Satisfy with retaining ownership	5.9	0.789
Satisfy with setting permission for data sharing	6.05	0.888
Satisfy with receiving incentives for data sharing	5.9	0.912
Total Average	5.913	
Category	Quite High	

Table 4. 8. Analysis of Enjoyment (ENJ)

Indicators	Score	Std. Deviation
Satisfy to use the system	5.8	0.895
Use the system	5.85	0.989
Enjoy using the system whenever needed	5.7	0.865
Total Average	5.784	
Category	Quite High	

Table 4. 9. Analysis of Intention to Use (ITU)

Indicators	Score	Std. Deviation
Worthwhile to use	6.15	0.813
Use for sharing research data	5.85	0.876
Intend to use for sharing research data in future	5.75	0.911
Necessary to use to share research data	5.15	1.226
Total Average	5.725	
Category	Quite High	

Figure 4. 2. Analysis of all the constructs

Measurement Models

I checked the measurement model with the exploratory factor analysis by testing the internal data consistency, reliability, and validity of the constructs.

Exploratory factor analysis: Based on the recommendation of Hair et al. (Hair et al., 2014), factor loadings greater than 0.50 can be considered as significant. I checked the factor loadings in the measurement model to see whether the items in each variable loaded highly on their own construct over the other respective constructs. Table 4.10 presents the factor loadings and their corresponding Squared Multiple Correlation (SMC) for our study. All the indicators in the measurement models had a factor loading greater than 0.50.

Convergent Validity: I observed the convergent validity for each construct measure by calculating Average Variance Extracted (AVE) and Composite Reliability (CR) (Hair et al., 2014) from the factor loadings (see Table 4.11). AVE for each construct exceeded the recommended level of 0.50, so over 50% of the variances observed in the items were accounted for by the hypothesized constructs. Similarly, CR should also be above 0.75 to publish the result. In our study, CR for each construct was above 0.80.

Reliability of the Measures*:* I checked the internal consistency for estimating the reliability of a measure by evaluating the within-scale consistency of the responses to the items of the measure. Since our study has multiple-item measurement instruments, I used Cronbach (Coefficient) Alpha (Cortina, 1993) for estimating the internal consistency. "*Coefficient Alpha assumes: (i) unidimensionality, and that (ii) item are equally related to the construct; therefore, interchangeable*" (Cortina, 1993). In practice, CR does not assume factor loadings to be the same for all items but takes into consideration the varying factor loadings of the items, whereas Alpha assumes factor loadings to be the same for all items. As can be seen in Table 4.11, the Alpha coefficient for each of the four antecedent construct measures is greater than 0.8 (good) while Alpha for intention to use is greater than 0.75 (acceptable) based on the recommendation of (Cronbach, 1971). CR and Alphas are related to each other based on factor loadings, as more factor loadings fluctuate among items, the higher the discrepancy between the values of CR and Alpha will be.

Table 4. 10. Exploratory factor analysis

Item	PEOU	PU	QOS	PEnj	ITU	SMC
peou1	0.852	-	-	-	-	0.722
peou2	0.850	-	-	-	-	0.722
peou3	0.815	-	-	-	-	0.664
peou4	0.758	-	-	-	-	0.574
peou5	0.734	-	-	-	-	0.538
peou6	0.707	-	-	-	-	0.499
pu1	-	0.694	-	-	-	0.481
pu2	-	0.892	-	-	-	0.795
pu3	-	0.786	-	-	-	0.617
pu4	-	0.836	-	-	-	0.698
pu5	-	0.752	-	-	-	0.565
pu6	-	0.876	-	-	-	0.767
qos1	-	-	0.891	-	-	0.793
qos2	-	-	0.882	-	-	0.777
qos3	-	-	0.858	-	-	0.736
qos4	-	-	0.812	-	-	0.659
penj1	-	-	-	0.941	-	0.885
penj2	-	-	-	0.917	-	0.840
penj3	-	-	-	0.913	-	0.833
itu1	-	-	-	-	0.879	0.772
itu2	-	-	-	-	0.868	0.753
itu3	-	-	-	-	0.799	0.638
itu4	-	-	-	-	0.663	0.439

Table 4. 11. Reliability analysis

	PEOU	**PU**	**QOS**	**PEnj**	**ITU**
Cronbach's α	0.87	0.89	0.882	0.913	0.792
AVE	0.621	0.654	0.742	0.853	0.651
CR	0.907	0.919	0.92	0.946	0.881

KMO and Bartlett's Test: I performed the KMO test for suitability of data for factor analysis based on (M.-Y. Chen et al., 2003) and found the KMO measure > 0.5 (acceptable), as can be seen in Table 4.12. Similarly, based on (M.-Y. Chen et al., 2003), I then performed Bartlett's Test of sphericity to check the homogeneity of variance for our structural models: ANOVA and regression models. Our result showed that the significance level was smaller than 0.05 as recommended (see Table 4.12), which suggested the factor analysis would be useful with our data.

Table 4. 12. Data suitability analysis

		PEOU	**PU**	**QOS**	**PEnj**	**ITU**
KMO Measure		0.63	0.82	0.778	0.747	0.661
Bartlett's Test	**χ2**	59.58	69.30	39.87	36.64	29.49
	df	15	15	6	3	6
	Sig.	0	0	0	0	0

Structural Models

I built a structural model for the general population to begin our Structural Equation Modeling (SEM) analysis (Gefen et al., 2000), as shown in Figure 4.3. The model is characterized by coefficients of determination ($R^{2'}$s) and path coefficients (β's). R^2 determines the variance of a given construct explained by antecedents, while β captures the strength of the relationship between the selected constructs. The structural model shows different paths linking perceived ease of use, perceived usefulness, quality of system, perceived enjoyment, and intention to use constructs in the context of a blockchain-based research data sharing prototype.

Table 4.13 shows the standardized path coefficient (β), t-statistics, p-value, and R^2 across selected constructs. According to Chin's guideline [20], a path coefficient should be equal to or greater than 0.2 to be considered relevant. Based on [21], we normally refer a model to be statistically somewhat significant (*p) when p-value < 0.05, statistically quite significant (**p) when p-value < 0.01 and statistically highly significant (***p) when p-value < 0.001. In our study, we found that the combined effect of perceived ease of use, perceived usefulness, quality of system, and perceived enjoyment on intention to the blockchain-based research data sharing system were insignificant at p > 0.05. The path coefficients ranged from -0.044 to 0.480. However, the

individual influence of quality of system ($\beta = 0.83$, $p < 0.001$) and perceived enjoyment ($\beta = 0.75$, $p < 0.001$) on intention to use was highly significant while there was a moderate and weaker influence of perceived usefulness ($\beta = 0.5$, $p < 0.01$) and perceived ease of use ($\beta = 0.56$, $p < 0.05$) respectively on intention to use. Hence, hypotheses (H4 - H7) were supported, whereas H8 was not supported.

Moreover, we found that perceived ease of use did not relate positively to perceived usefulness ($\beta = 0.453$, $p > 0.05$). However, the quality of system ($\beta = 0.99$, $p < 0.001$) and perceived enjoyment ($\beta = 0.75$, $p < 0.01$) had significant positive effect on perceived usefulness. So, our hypotheses H2 and H3 were also supported, whereas H1 was not supported. Table 4.14 summarizes the validation of our study's hypotheses.

Table 4. 13. SEM analysis

Structural Path		**β**	**T Statistics**	**P-Value**	**R^2**
PU ← PEOU		0.453	1.563	0.135	0.11
PU ← PE		0.755	3.550	0.002	**0.41**
PU ← QOS		0.993	4.578	0.000	**0.54**
ITU ←	PEOU	-0.044	-0.236	0.816	0.86
	PU	0.053	0.341	0.737	
	QOS	0.364	1.266	0.224	
	PE	0.480	2.295	0.036	
ITU ← PEOU		0.562	2.876	0.01004	0.31
ITU ← PU		0.5	3.674	0.00173	0.42
ITU ← QOS		0.834	5.803	0.000	0.65
ITU ← PE		0.751	6.46	0.000	0.7

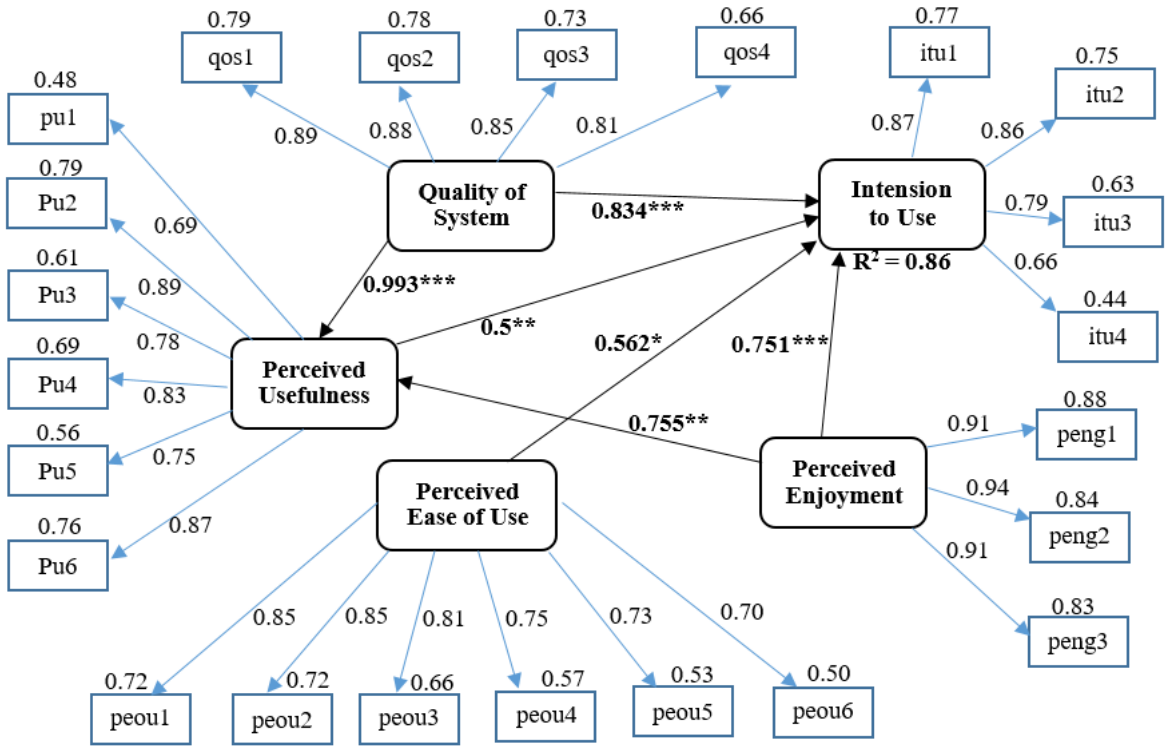

Figure 4. 3. Structural model showing test results.

*p < 0.05; **p < 0.01; ***p < 0.001

Table 4. 14. Validation of study's hypotheses

H	Hypothesis	Result
1	The perceived ease of use will significantly influence the perceived usefulness.	×
2	The perceived enjoyment will significantly influence the perceived usefulness.	√
3	The quality of system will significantly influence the perceived usefulness.	√
4	The perceived ease of use will significantly influence the intention to use.	√
5	The perceived usefulness will significantly influence the intention to use.	√
6	The quality of system will significantly influence the intention to use.	√
7	The perceived enjoyment will significantly influence the intention to use.	√
8	The combined effect of perceived ease of use, perceived usefulness, quality of system and perceived enjoyment will significantly influence the intention to use.	×

√ = True; × = False

Participants' Comments

Table 4.15 shows the comments from participants related to the initial adoption of the DUDS prototype. Most of the participants did not provide any comments, but those who provided the comments are focused mostly on the policy level and privacy. Participants wanted to see the real-life implementation of the system to uncover the comparative usefulness of the proposed DUDS approach with other systems.

Table 4.15. Participants' comments related to adoption

Sample comments from participants on the DUDS prototype
1. "Love to learn more about this system."
2. "Privacy can be explained more as it is yet not clear from the explanation on the first page. This is my number 1 concern."
3. "It would be good to see comparisons with other systems."
4. "If the system can accomplish the user friendly state of the art as Mendeley, then it has great potential"
5. "I am concerned about the ethical requirements of some universities that might prevent some researchers from being able to share their data. How would this be dealt with? And who receives the reward (e.g., monetary) for sharing your data: the student? the supervisor? the university? etc."

4.1.5 Discussion

We achieved our research goal to introduce external constructs perceived enjoyment and quality of system on the classical TAM in the context of blockchain-based research data sharing system and explored whether users would be willing to adopt the system. Our study validated most of the proposed hypotheses.

Quality of system is the most significant determinant that influences perceived usefulness and intention to use. When users receive greater satisfaction with the quality of the blockchain-based system that helps researchers to share their data while maintaining ownership over the data, set permissions for data sharing, and receive incentives for sharing the data, the system's perceived usefulness will be higher as well as the user's intention to use it. Furthermore, when users enjoy

and get satisfied with the quality of system during their interaction with the prototype system with known benefits for sharing research data, they are quite likely to find the system more useful and extremely likely to adopt the system.

Previous research by (Rovai, 2004), (I.-F. Liu et al., 2010) shows that the UI design is the most significant external construct that affects perceived ease of use, and since our study used a prototype rather than an actual working blockchain based-system, most subjects may have experienced difficulty in relating the actual user-interface. Thus, the effect of easy-to-use may not be reflected on the users' belief in finding it to be more useful, which explains our result on not supporting H1 (which is predicted by the classical TAM).

Thus, this chapter provided the importance of user study in designing the platforms based on the DUDS framework. The empirical study was carried out to build a cognitive-behavioral user model with the aid of an extended TAM. As a result, it identified and reviewed distinct constructs affecting the end users' intention to accept the platform.

The main limitation of this study was that our findings were based on a small sample size with a prototype system, and the participants were not representative of the general user population. Further studies are needed to confirm that the findings will generalize for the larger population of users in a real-life implemented system for sharing user data using blockchain and smart contracts-based DUDS framework. Yet, the methodology for doing a larger study in the context of a real system would be the same. The main challenges regarding the acceptance of distributed ledger-based systems such as DUDS platforms are skills gap, insufficient organizational awareness, and lack of trust on the security of the underlying technology itself.

4.2 Trust for DUDS Framework

It is important for software designers to recognize that there are privacy, security and trust issues related to user data sharing systems, which could affect user behavior and attitude towards the use of the system. As the next step towards achieving the objectives of the thesis research, this section identifies two important antecedents of Trust—Privacy and Security—which ultimately affect the user behavioral intention to use the platform based on the DUDS framework. From its early inception, this research determined to perform the user study on a real blockchain-based system, not just on the prototype, to understand the role in those systems of the user's trust. Chapter 2 has already discussed the multi-faceted notion of privacy—attitudinal privacy and privacy concern.

This section describes user study on the platform based on the DUDS framework to uncover the relationship between the *privacy concern*, *general caution*, and *technical protection* factors on the user acceptance of such platforms. Furthermore, the study aims to reveal the responsibility of trust in incorporating *security* and *privacy concern* factors in the DUDS framework. Recently, Shin et, al. (2019) presented validated constructs to measure *Trust* in blockchain technology using previously validated constructs for measuring trust in the Information Technology domain. Shin et al. did not use a real blockchain platform, neither did they consider the service conditions and service quality. However, their validated constructs are useful to test on the platform based on my DUDS framework and the findings would enhance the final artifact and contribute to the knowledge base for the application domain.

Table 4. 16. Constructs and items for privacy (Buchanan et al., 2007)

CONSTRUCT	**ITEMS**
General Caution (Behavioral)	I am confident that I only register for online services that have a privacy policy.
	I am confident that I read the privacy policy of the online system before I register my information.
	I am confident that I look for a privacy certification of the online system before I register my information.
	I am confident that I read license agreements fully before I agree to them.
Technical Protection (Behavioral)	I am aware of removing browser cookies.
	I am aware of using a pop-up window blocker.
	I am confident that I check the computer for spyware.
	I am confident that I clear browser history regularly.
Privacy Concern (Attitudinal)	I am confident that I know all the online organizations as they claim they are while collecting information I provide during the use of the system.
	I am aware of the exact nature of the information that will be collected during the use of the system.
	I am confident that the information I submitted on the blockchains could not be misused.
	I am not concerned that the Ether account that I used could be intercepted by someone else.

Table 4. 17. Constructs and items (Shin, 2017)

CONSTRUCT	ITEMS
Perceived Trust	The blockchain-based system is a trustworthy service.
	I can count on the blockchain-based system to protect my privacy.
	The blockchain-based system can be relied on to keep its promises.
Perceived Security	I believe the information I provide with blockchains will be handled by appropriate processes.
	I am confident that the information I provide will be secured.
	I believe only legitimate organizations may view the information I provide with the blockchain-based system.
Attitude	I would have positive feelings towards blockchain-based systems in general.
	The thought of using blockchain-based systems is appealing to me.
	It would be a good idea to use blockchain-based systems.

Based on the literature (Buchanan et al., 2007), I identified the constructs with validated questions (those with the highest factor loading), as shown in Table 4.16., suitable for constructing the privacy model for the real DUDS-based system. Moreover, Table 4.17 presents the validated constructs suitable to study Trust, Security and Attitude towards the adoption of the platform. Chapter 6 presents the final user study with these identified latent variables to examine their significance to influence the user's attitudes and intention towards the initial adoption of the implemented prototype DUDS-based platforms.

4.3 Conclusion

User studies are much needed to evaluate technological solutions and observe the effects of different variables using theory-backed models. In this chapter, I presented an extended TAM-based model to measure the relationship between perceived usefulness, perceived ease of use, quality of system, perceived enjoyment, and intention to use constructs for a prototype research data sharing system based on blockchain-based DUDS framework. Although these constructs have been much investigated previously as antecedents to user acceptance of different technologies in

various domains, this work was the first to investigate the use of TAM for analyzing the factors influencing user acceptance of blockchain-based applications for sharing data, in this case, research data among researchers. Using the methodology of theory-based model building and evaluation through a user study and statistical analysis, it was possible to discover the factors that influence the intention to use, and the adoption of a platform based on the DUDS framework. This opened new directions to study distributed ledger technologies and decentralized applications from the user behavioral modeling perspective. I implemented the descriptive statistic, measurement models, and structural models to present the results and used SEM analysis to observe the users' acceptance of the proposed blockchain-based system. This helped to build actual DUDS platforms that support decentralization with user control and incentives. However, in this study, the investigation of the roles and the dimensions of perceived privacy, perceived security, and user trust on the blockchain-based system were missing. Therefore, this study advocated further investigation towards conceptualizing trust with privacy and security elements for the DUDS framework with a larger and broader participants' pool. This would help to analyze the framework for identifying any issues in relation to users' behavioral intention to use the blockchain and smart contracts-based data-sharing platform. To overcome the limitations of the study described in this section, further investigations were needed in an actual deployed system, rather than in a prototype with sketchy UI. Therefore, the next section presents the proof of concepts of the DUDS framework with implemented prototype systems.

5 PROOF OF CONCEPT IMPLEMENTATION OF THE DUDS FRAMEWORK

This chapter presents the content from my three published articles including a demo paper of the user data sharing system to provide the proof of concept demonstrating the working of the DUDS framework (Shrestha, Vassileva, et al., 2020), (Shrestha & Vassileva, 2018a), (Shrestha, Joshi, et al., 2020). Three different implementation prototypes and the minimum viable products (MVPs) (Münch et al., 2013) are included in this chapter to test the optimum features of the framework while enabling sharing of different types of user data. The three implementations include three domains—tourism, research and online shopping cart—that offer rewards to the data owners in terms of payment through blockchains for the use of the data by applications, as specified by the contracts. Moreover, this chapter aims at measuring and investigating the performance of the MVPs based on the DUDS framework and preparing the framework for the final user study to evaluate its usability. Also, this chapter presents the complete design using conceptual DUDS architecture that will be helpful for the stakeholders and blockchain engineers to design the system for their own platforms.

5.1 A Blockchain Platform for User Model Data Sharing

This section is based on my published article (Shrestha et al., 2020). I designed a '*new platform'* for user modeling with the DUDS framework that allows users to share data without losing control and ownership of it and applied the framework to the domain of travel booking (Tourism). This new platform provides the solution to three important problems: ensuring privacy and user control and incentives for sharing. It tracks who shared what, with whom, when, by what means, and for what purposes in a verifiable fashion. This section of the chapter presents a case study of applying the framework for a hotel reservation system as one of the enterprise nodes of Multichain that collects users' profile data and allows users to receive rewards while sharing their data with other travel service providers according to their privacy preferences expressed in smart contracts. The user data from the repository is converted into an open data format and shared via streams with the blockchain so that other nodes can efficiently access, process, and use the data. The smart contract

verifies and executes the agreed terms of use of the data and transfers digital tokens as a reward to the user. The smart contract imposes double deposit collateral to ensure that all participants act honestly. Furthermore, I also conducted a performance evaluation of this new platform by analyzing latency and memory consumption with selected three test scenarios for different numbers of nodes and measuring the transaction cost for smart contracts deployment. The results show that the node responded quickly in all our cases with a befitting transaction cost. The smart contract execution takes a reasonable amount of time given the current Ethereum blockchain consensus architecture, but it is acceptable for our purpose of deploying data sharing policies over smart contracts and receiving incentives for sharing. So, at this stage, I did not measure usability but just observed various performance metrics of the new platform based on the DUDS framework for user profile data sharing.

Throughout this section, I have used the term '*new platform'* to represent the DUDS-based MVP for the tourism domain that uses multiple distributed ledger technologies. Following the DUDS framework, I used MultiChain (Greenspan, 2013) and Ethereum (Buterin, 2015) blockchains to provide an uneditable private record of all transactions made with user data. We can optionally store any published item off-chain that saves storage place and bandwidth. MultiChain, however, due to its current architecture, cannot support an access control mechanism, which the proposed platform needs, to provide users means to control how their data is shared and the rewards they get. Therefore, I used Ethereum, which supports smart contracts and commits the contracts' transactions. Smart contracts govern the accountability of access and provide incentives to the users for sharing user data. Combining the security and immutability of data stored in the blockchain, with the specific strengths of two popular blockchains—MultiChain and Ethereum—combines their advantages: proper data storage and data sharing, and smart contracts for access control mechanisms. In this way, this new platform addresses the shortcomings of traditional centralized user models used by internet businesses, which have security vulnerabilities, lack accountability, and take away the ownership, control, and incentives for users to share their data.

5.1.1 DUDS Platform in Tourism Domain

To enable effective user data sharing among enterprises, this section used a travel industry that covers travel and tour agencies, hotels and resorts, airlines, restaurants, etc. The travel industries within the hospitality domain usually want to compete successfully, and they must do so by using

technologies to drive value to all the parties associated with them (Cassidy & Chae, 2006). By sharing real-time data about users, which is being updated simultaneously by the different participating entities of the consortium, including travel agencies, hotels, resorts, airlines, restaurants, shopping malls etc., each of which can offer personalized services after analyzing their customers' preferences. To share the user data among enterprises, the MultiChain blockchain is installed on each participating enterprise end, which can publish the user data as items into the stream and share them in the network according to the smart contracts set by the customers.

I deployed a general hotel reservation web application on one UNIX machine. The web application was developed in PHP with an Apache server and MySQL as a backend. The machine running MultiChain serves as one of the nodes that collect customers' data with proper validation. The public Ethereum blockchain stores and executes smart contracts. All the registered customers and data consumers have Ethereum addresses, which are used to transfer the ether while sharing the data as per the smart contracts. Users create their profile in the hotel reservation system as the first enterprise node (Node1, for instance, Grandee Hotel) in the consortium blockchain by registering with their information and simultaneously choosing which of their data can be shared with third parties. The data stored in the repository is converted into an open data JSON format, which is published via stream in the MultiChain. The three streams in MultiChain: public-key stream, data-item stream and access-key stream are used for general data storage and retrieval. Other imaginary third parties of the tourism industry, for instance, Saskatoon Travel and Tours, and Saskatoon Shopping Mall have their own nodes, which also contain Multichain in their systems and generate their own pairs of public addresses and private keys using RSA public-key cryptography. All these nodes are given permission to be in the closed network of the consortium blockchain.

As shown in Figure 5.1, I first initiated a Multichain in the Hotel reservation system for the first node Grandee Hotel (Node1) in the blockchain. This node got an administrator role to grant associated access privileges to other nodes. In our case, the permissions for other nodes were set by the first admin node, and they could be made true for all the nodes while setting chain parameters.

After that, I created the multichain daemon using the chain name *model*. It led to the creation of the MultiChain Core Daemon, with a specific version as build 1.0 alpha 27 protocol 10007 in this case, such that other industry nodes could connect to this starter node. I then created two other

nodes—Node2 and Node3—representing Saskatoon Travel and Tour and Saskatoon Shopping Mall, respectively, as imaginary independent firms in the travel domain. The creation of the nodes offered the individual addresses for those new nodes, which were acknowledged by the first node to grant a "connection" permission to them into the MultiChain since it is the consortium blockchain network. Back on the first admin node server, I added connection permissions for other node addresses. This was the first step in creating the blockchain. While granting the connection permission, further other permissions could also be set for other nodes.

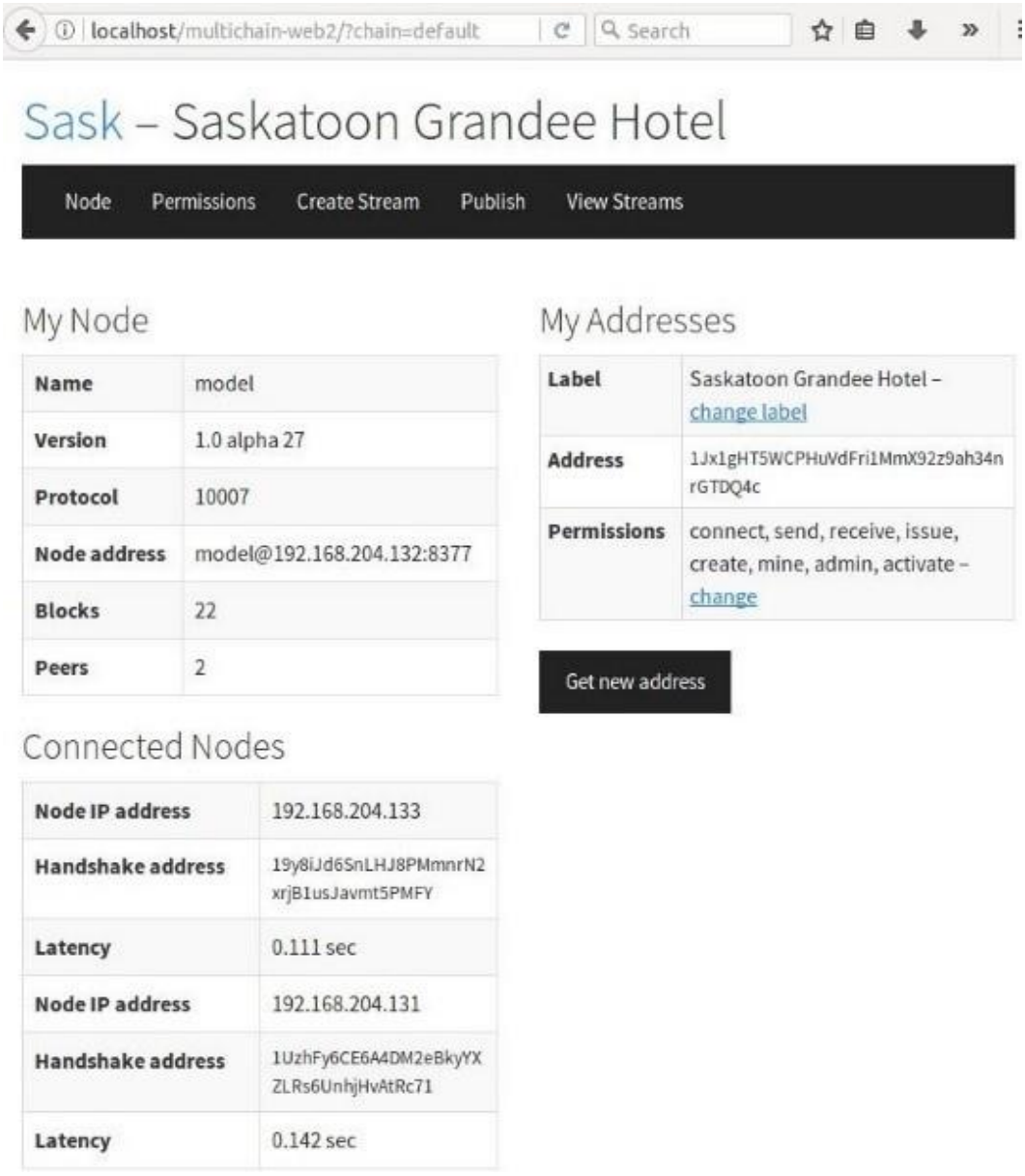

Figure 5. 1. All connected nodes as seen from Node1-Grandee hotel

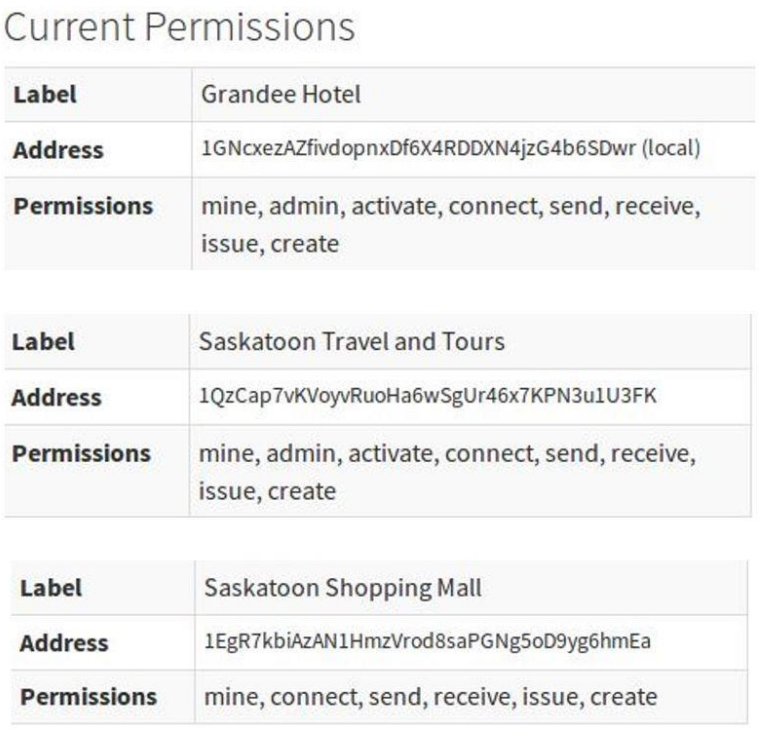

Figure 5. 2. Permissions set for connected nodes as seen from hotel reservation system (Node1)

As shown in Figure 5.2, Node2 (Saskatoon Travel and tour) was granted 'connect,' 'send,' 'receive,' 'issue,' 'create,' 'mine,' 'activate,' and 'admin' permissions, and Node3 (Saskatoon shopping Mall) was granted all permissions except 'admin' and 'activate.' This means Node2 in the blockchain could also act as admin, but Node3 could not. Figure 5.3 shows how to publish the stream by publishing the items containing user data into the data-item stream and share them with other consortium enterprise nodes. The collected customer data at Node1 from the off-chain database was converted into JSON format before being published as items into the stream. In fact, every node in the MultiChain could access any stored raw data. To ensure data confidentiality, streams from the Multichain were used with a combination of symmetric and asymmetric cryptography for encrypting the data before being published into the streams. The encryption process starts with the generation of the RSA private-public key pair on the nodes using the OpenSSL[39] and then publication of the public key into the public-key stream. Before publishing user data, they were first encrypted using the OpenSSL with the symmetric AES cryptography

[39] https://www.php.net/manual/en/ref.openssl.php

scheme. Then, the secret to decrypt the data was encoded with the recipient's public key in accordance with the data sharing preferences updated by the customer, and finally, both the encoded secrets and encrypted data were published in the respective streams.

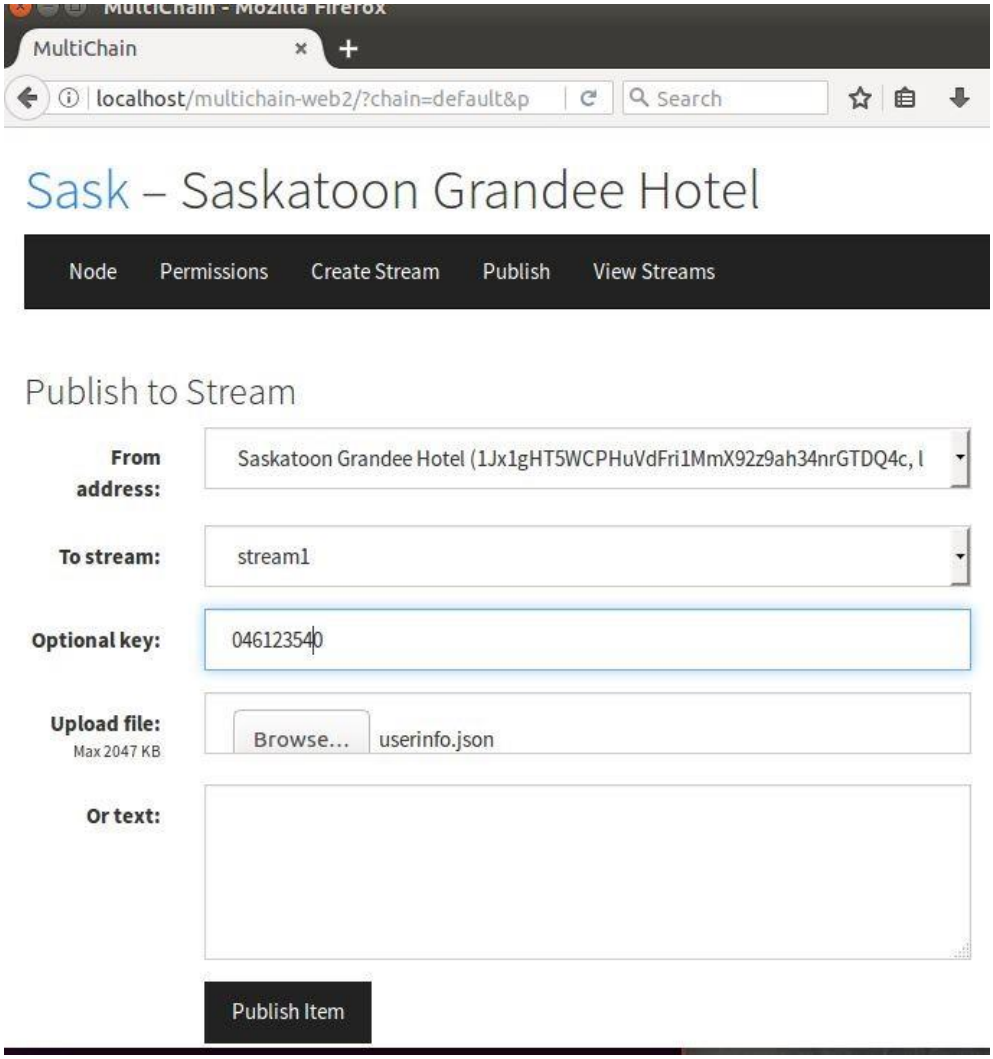

Figure 5. 3. Publishing stream of items

Figure 5.4 shows the list of the streams created by the Node1-Grandee Hotel. The first node was the Hotel Reservation System, which collected the users' data during the hotel booking process. The system at Node1 collected the basic user information such as name, the purpose of visit, nationality, and duration of stay along with the data sharing preferences through the web form. The customer information is particularly useful to other tourism enterprises including a shopping mall, travel and tour operators, etc. since they can provide attractive offers to the customers during their stay.

Figure 5. 4. List of the streams created by Node 1

The collected useful data about the customers were shared among the enterprise consortium as per the predefined agreed terms in the smart contracts. There could be any number of streams and the data published in every stream were stored in full or referenced by a hash inside the transactions. Only the nodes with permission were able to view the contents of the streams. The eligible recipient nodes in the network subscribed to the streams, used the OpenSSL with their private key to decode the secret and finally used the OpenSSL with the decoded secret to decrypt the encrypted data, which could be then converted and stored into their own local repositories. Once the data access was completed, the smart contracts got triggered and with their successful execution, the tokens were transferred from the data consumer's Ethereum address to the customer's account while delivering the requested data (with verified hashes) to the local storage of the consumer node. The application accessed all the MultiChain Community commands using the JSON-RPC API, as they are available under an open-source license[40].

[40] https://github.com/MultiChain/multichain/blob/master/COPYING

5.1.2 Incentivizing Customers for Data Sharing

I used the Ropsten test network for the Ethereum blockchain to implement the smart contracts through an online Remix IDE[41] because the Ropsten is an easy-to-use test network with the same proof of work (PoW) consensus mechanism for the block validation as in the Mainnet of Ethereum, and Remix is a free IDE to deploy any untrusted codes before going live. In addition to that, anyone can use Etherscan[42] to explore the Ropsten blockchain to search for any transactions taking place on the blockchain. Also, I used metamask[43] to deploy the Injected web3 environment to connect the contracts with the Ethereum account addresses. The contracts (see Appendix III) were deployed on the Remix. After accessing the user data, the corresponding user would be incentivized by the data consumer. My approach, therefore, delivered a usable blockchain and smart contracts-based DUDS platform, which enables users to maintain control over the conditions of access to their data, thus providing a possibility for including flexible data-sharing policies with incentives for sharing data.

5.1.3 Performance Metrics

I conducted the system performance evaluation by analyzing the performance metrics that mostly affected user experience (UX). I did successive experiments on the freshly created node to evaluate the performance of the system by setting four goals to find out:

1. How long it takes an enterprise Multichain node to get connected to the network.
2. How long it takes the enterprise Multichain node to respond to the actions (like starting stream, viewing a stream item, loading, or publishing the items into the stream).
3. How much memory the node consumes when the blockchain daemon gets started, and
4. How much gas the transactions use (validation cost) to complete the execution.

Since there were two blockchain platforms in my system, I considered measuring latency and memory consumption parameters among the private blockchain nodes because the data sharing is

[41] https://remix.ethereum.org
[42] https://ropsten.etherscan.io/tx/0xa278a0b05cea83b45bc1879246113265e34e72d37f66dfb5bf2ed2660d6d6902
[43] https://metamask.io/

performed in the private network, which requires very low latency for optimal performance, and the storage delay can also play a role in the increased latency and poor performance. For the smart contracts' execution, we always try to make transaction costs low, so I measured the efficiency of our smart contracts in terms of the transaction cost. To evaluate the implementation prototype, I set an evaluation plan to simulate real-world interactions. I categorized the first three goals as implementation under the data-sharing model and the last goal as the implementation under the user incentive model.

5.1.3.1 Experiments

To achieve the first three goals, the evaluation involved three scenarios to simulate different levels of concurrency, while monitoring latencies in the Windows and the UNIX machines. The three scenarios are shown in Table 5.1. The scenarios, experimental models, and reasoning are available in Appendix V. Additionally, I carried out another experiment to observe the memory consumption for the nodes when the corresponding multichain core daemon started on that particular node. A total of five observations were carried out, one of which is available in Appendix V.

Table 5. 1. Test scenario description

Scenario	Descriptions
S1	Two enterprise nodes connected
S2	Three enterprise nodes connected
S3	Eight enterprise nodes connected

To achieve the last goal, the evaluation involved deploying the codes with the Remix IDE on the Ethereum Ropsten testnet. Currently, the smart contracts require a gas fee to deploy their code and commit the transactions into the Ethereum blockchain, and the actual cost is paid in ether. ETH Gas Station[44] provides three categories of gas prices. They are SafeLow (less than 30 minutes), Standard (less than 5 minutes), and Fast (less than 2 minutes). The gas limit is helpful to optimize the gas used to provide a safety mechanism, as sometimes code with bugs might keep on consuming unnecessary gas for the execution. I used a gas price of 25 Gwei (2.5e-8 ether) which was the then-

[44] https://ethgasstation.info/

price for achieving a faster transaction. In fact, the cost of a transaction always increases when the gas price goes higher. I have provided one of the instances of the contract creation with the transaction hashes on the Ropsten Ethereum explore, the details of which are available in Appendix V.

Thus, I evaluated the performance of the new platform with implementation under the data-sharing model and implementation under the user incentive model. The result shows that the data sharing DUDS model has very low latency for an enterprise MultiChain node to get connected to the consortium network and to respond to actions like starting stream, viewing a stream item, loading or publishing the items into the stream. It also consumes less memory when the blockchain daemon gets started. In addition to that, the user incentive model has an acceptable transaction cost for executing smart contracts (see Appendix V).

5.1.4 Conclusion

I provided a quantitative study of the new platform for sharing user profile data based on the DUDS framework that allows users to have control over their data and earn rewards. This platform uses blockchain technology for user-controlled privacy with data-sharing policies encoded in smart contracts. It naturally supports building up incentives for users to share their data, in terms of rewards (micro-payments). In this way, users become owners of their data and can decide how their data is collected and used as well as shared. To share users' profile data in a distributed fashion, the concept of streams from the MultiChain was successfully interpreted in the travel booking domain. I presented a hotel reservation system as one of the enterprise nodes of MultiChain that collects users' profile data and allows users to receive rewards while sharing their profile data with other travel industries according to their privacy preferences expressed in smart contracts. The user data from the repository was converted into an open data format and shared via stream in the blockchain so that other nodes efficiently processed and used the data. The smart contract verified and executed the agreed terms of use of the data and transferred digital tokens as an incentive to the data provider. The smart contract imposed double deposit collateral to ensure that all participants act honestly. This section has provided a basic use of smart contracts on privacy-preserving data sharing and management models. It combined blockchains and off-blockchain repositories to create a data sharing and management model focused on security and privacy. This blockchain-coupled user data sharing model is not just limited to the travel domain but is also

applicable to other similar domains such as eCommerce, education, health, etc. The section also evaluated the performance of the new DUDS-based platform, and it met the expectations in terms of the latency, memory consumption, and transaction cost for smart contracts deployment. The node responded quickly in all our cases with a befitting transaction cost. This section concludes with the suitability of the DUDS-based platform for the next study of evaluating the usability and usefulness of the approach, and the trust users could have in the system, which is presented in Chapter 6.

5.2 Blockchain-Based Research Data Sharing Framework

This section briefly presents another usable data sharing platform based on the DUDS framework for the collection of researchers' data, providing proof of existence and ownership of data, and maintaining a verifiable record of all accesses/ sharing/ usages of the data. This section has the content from my paper published at the 2018 International Conference on Blockchain (Shrestha & Vassileva, 2018a), which was the first in the area of sharing research data using blockchain technology, and it generated a lot of interest. Data sharing practices are much needed to maximize knowledge gain by researchers. However, when and what data should be shared with whom, and how credit should be awarded to the data owner needs to be clearly addressed to create an individual incentive for data owners to share their data. A platform that allows owners to control and get rewards from sharing their data would be an important enabler of research data-sharing since presently, such incentives for researchers to share their data are largely missing.

Data owners not only enjoy increased transparency and protection of data from falling into the wrong hands, but they are also incentivized with digital tokens, acknowledgment, or both, to share their data with the interested data seekers, thus becoming active participants that stand to benefit from the research data economy. The DUDS-based platform for sharing research data allows researchers a proper way of creating proof of the existence of their research work and tracking the sharing of their extensive research data and samples while receiving incentives in real-time in the form of digital tokens (with monetary value) or attribution for their research work (credit, citation, or a collaboration offer) or both. They can protect their intellectual properties and provide provenance of their ownership indisputably. This would also solve the problem faced by other

individuals (farmers, plant breeders, customers, etc.) who could benefit by sharing their data with researchers.

5.2.1 Background and Related Works

Over the last decade, there has been a huge technological innovation bringing many research consortiums to use data-driven approaches and to collaborate in making intelligent decisions to improve their scientific research activities. Data Analytics methods can significantly improve the quality of services, but they depend on collecting, sharing, and mining research data.

According to Bierer et al. (Bierer et al., 2017), research data sharing is the "use of research data by persons other than those who originally gathered the data, for no longer a hypothetical or occasional occurrence." Most of the research on data sharing is relevant to the design framework that focuses on the optimization of those properties. However, the technical performance of a data-sharing system alone does not guarantee the practicality of the system. Decentralized approaches for data sharing aim to overcome the limitations brought by the centralized architecture, which has a predefined point of access that leads to the central point of failure.

Naz et al. (Naz et al., 2019) adopted an architecture similar to the DUDS framework for trading the research data. In their research, they incorporated IPFS for the storage of a file where payment is made in *ether* for the access of the data. Similarly, Shen et al. (Shen et al., 2020) provided a design of the multiple clouds- and blockchain-based data sharing system using the concept of Shapley value that dynamically incentivizes the data owners with revenue in *ether* for sharing their data (patient data). In their design, since the data is not encrypted before placing into the cloud, it would be an issue for some patients to hide their sensitive medical information.

Dong et al. (Dong et al., 2019) also proposed a data-sharing enabler similar to the DUDS framework. In their research, they adopted the IPFS with the map-reduce model to store data, conceptualized the access control policy with a smart contract, presented a price compensation and reputation management mechanism in order to achieve self-control and price balance in sharing data. They highlighted the challenges in relation to compensating data owners in a low privacy consumption rate to allow easy and valid access to their data.

5.2.2 Solution Framework and Discussion

Figure 5.5 presents our general solution architecture that introduces a new way of incentivizing the users for sharing their research data. I introduced blockchains to share the data among registered parties/enterprises in their private network by incorporating automatic contracts so that access-control policies would be stored securely on the blockchain. A user can register into the system by providing her basic profile information and public wallet-address to activate the smart contract, which automates the functionality to support the user-controlled privacy. The system is thus able to (a) give the user full transparency over who accesses their data, when and for what purpose, (b) allow the user to specify a range of purposes of data sharing, kinds of data that can be shared, and classes of applications/ companies that can access the data through the smart contract, and (c) provide an incentive to the user for sharing their research data (in terms of payment for the use of the data by applications, as specified by a smart contract). This user-incentive model with the blockchain is run by the public (Ethereum) blockchain network nodes.

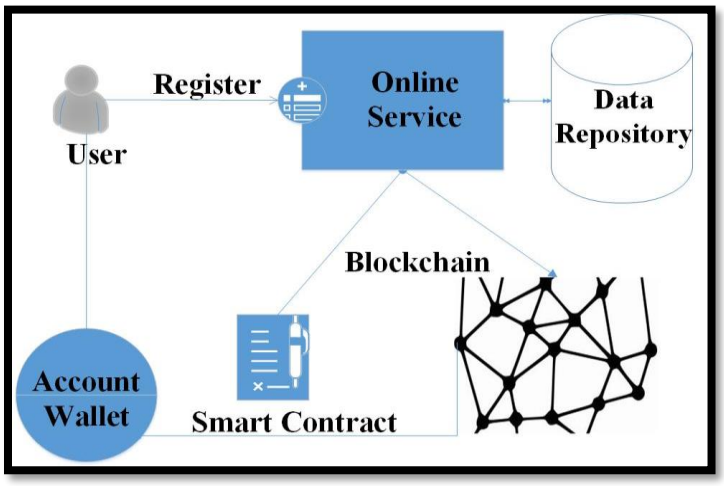

Figure 5. 5. General user-controlled privacy-preserving data sharing architecture.

Since the smart contract is stored on the public blockchain, the users should have their own digital token addresses safely stored in their personal wallets. Once the users' data are being used by any other participating parties, then the corresponding users will be incentivized with the digital tokens (ETH). And similarly, for sharing the data among enterprises, consortium (MultiChain) blockchains are installed on each participating registered node, which publishes the items (research dataset) into the stream to be shared among other nodes in the network.

Regarding this proposed general model based on the DUDS framework, the actual implementation is portrayed with the solution framework in Figure 5.6. One of the elements of data sharing would be to whom the data are available for sharing and by what means, and how can the researchers/ data owners be incentivized either with digital tokens or with acknowledgment of their efforts in collecting the data. Our system clearly guides registered users about what smart contracts do with their data. With the smart contract in the public Ethereum blockchain, researchers can retain the ownership of data with themselves and are incentivized as per the agreed terms. Any academic or industrial unit as a data seeker with valid credentials and approval from a local institutional review board (IRB) is eligible to access the data. The local IRB must also be enlisted in the system by providing the certification that it is bound by regulations to look at scientific methods proposed by a node (data seeker) for accessing the research data. Through the smart contract, only the selected eligible nodes can access the items (dataset) by subscribing to the corresponding published streams. The data owner is incentivized as per the negotiation made on the options between the two parties. An acknowledgment can be given to the data owner during the publication of the research article and/ or a predefined incentive is offered in the form of the digital token by transferring ETH to the data owners' Ether addresses. An escrow service can be optionally added into the system to bind the users with legal obligations. The access-control policies are stored securely on the blockchain while retaining the same user interface.

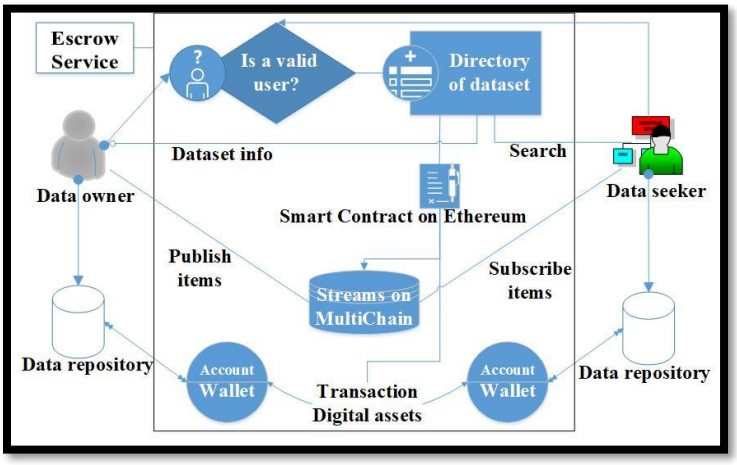

Figure 5. 6. User-controlled privacy-preserving research data sharing model.

The smart contract is deployed just once for each node on the Ethereum blockchain that stores _billingAddress. The smart contract developed with Solidity contains the following functions:

```
contract ShareResearch is tested {
  function ShareResearch(address _billingAddress)
  function getStatus(uint externalIncentiveID) constant
returns (string)
  function getPrice(uint externalIncentiveID) constant
returns(uint)
  function startNewIncentive(uint externalIncentiveID,
uint price) onlyOwner
  function pay(uint externalIncentiveID) payable
  function finish(uint externalIncentiveID) onlyOwner }
```

To provide ETH to the data owner (say node1) for accessing the data, a participating eligible data seeker at some node (say node2) queries the system to use the specific filename. Public key cryptography is implemented to ensure the authenticity of the eligible users requesting the file. This results in the execution of the startNewIncentive function of the smart contract with the incentiveID and total incentive to be paid to the data owner. The incentiveID is generated for the data owner during registration. Node2 invokes the pay function of the smart contract with the incentiveID of Node1 and the ETH to be sent as an incentive to the data owner. The contract verifies the two parameters and then it receives the ETH and updates the status accordingly. It then calls the getStatus function to get the status and with the confirmation of ETH being provided by Node2, data is made available to the data seeker and finally calls the finish function to transfer ETH to the _billingAddress. The ETH is made available in node1's account since the incentiveID is paired with the ETH address of the data owner. Thus, the data seeker is entitled to the data while incentivizing the corresponding data owner. As demonstrated under section 5.1 about the solution to the scalability issue, any number of MultiChain nodes can join the consortium network, because every new node does not need to connect to all the existing nodes on the chain to create a fully connected P2P network. However, any new participating node should replay complete blockchain transactions starting with the genesis block, so it could delay the immediate action to be performed by the new node on the chain as it must wait before it is up to date.

5.2.3 Conclusion

In summary, this section briefly presented a decentralized framework for incentivizing researchers for sharing their research data based on the DUDS framework, which provides a way to specify/control the parameters of sharing and providing full accountability of access to such data. The security, scalability, and privacy of those systems are gracefully realized through the implementation of the smart contract and blockchains, which offer a secure, immutable, tamper-proof, distributed research data-sharing network that guarantees the proof of existence and ownership of the researcher data.

5.3 A Blockchain-Based Shopping Cart

This section presents the blockchain-based online shopping cart as the third reference implementation of the flexible DUDS framework. This section has the content from my demo paper published at the 2020 IEEE International Conference on Blockchain (Shrestha et al., 2020).

This new free-eCommerce platform with blockchains allows customers to connect to a seller directly, share personal data without losing control and ownership of it. This platform provides a solution to four important problems: private payment, ensuring privacy and user control, and incentives for sharing. It allows the trade to be open and transparent with immutable transactions that can be used for settling any disputes.

This section presents a case study of applying the framework for a shopping cart as one of the enterprise nodes of MultiChain that provides trading in *ethers* controlled by smart contracts and collects users' profiles and transaction data and allows them to receive rewards for sharing their data with other business enterprises.

5.3.1 Customer Data Sharing Platform

The DUDS framework-based shopping cart is a novel platform to ensure privacy and user control, and incentives for sharing while addressing the issue of private payment. The DUDS framework-based platform offers the following features.

1. To create a decentralized e-commerce experience for customers.
2. To enable companies to increase trust in their products and supply chains.

3. To offer direct payment with native Ethereum tokens, thereby enabling privacy and confidentiality.
4. To create proof of the existence of every transaction.
5. To enable companies to share customers' data among others in the consortium network.
6. To provide transparency over every access to the user data.
7. To provide incentives to customers in real-time for sharing their data.

This novel blockchains based platform has a 3-tier shopping cart application employing *Spring Boot* and *ReactJS* as the main building technologies, that allows users to shop online using *ether* with all the transactions stored in the blockchain eliminating the trust and to get incentivized upon permitting to share their own data as stated in the smart contracts.

5.3.2 Background and Related Works

There are a few open-source decentralized marketplace projects such as *OpenBazaar* (Taaki & Hoffman, 2016) that support peer-to-peer transactions with cryptocurrencies. However, they do not have any provision for offering incentives to users for sharing their data in the digital marketplace.

Therefore, this DUDS platform is different in the context that it has used an online marketplace domain and presented the work for a shopping cart using both the Ethereum Smart contracts and MultiChain blockchain to offer a novel platform for customers to make payment with digital tokens and receive incentives for sharing their personal data among enterprises. The platform is proposed to combine a payment mechanism through the *ether* and a mutual agreement between customers and sellers via smart contracts. It automatically registers the immutable timestamped metadata about transactions, which can be useful to settle any possible disputes among stakeholders in the future. Furthermore, the enterprises share their customer data among their consortium network through a secure permissioned blockchain network that keeps track of who shared what, to whom, when, for what purpose and under what condition.

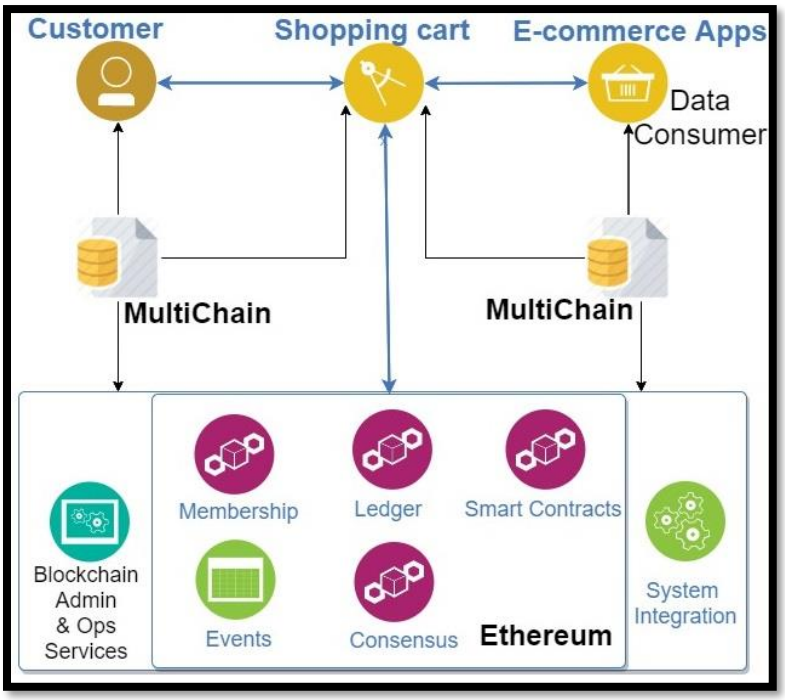

Figure 5.7. Blockchains based shopping cart with a data-sharing platform

5.3.3 System Development

The system has a 3-tier restful maven shopping cart application that uses Spring Boot as the backend technology with certain dependencies such as Lombok and starter-data-jpa, allowing opinionated configurations yet connecting to a MYSQL database and ReactJS. The application is built with a node.js runtime environment covering the frontend part written in the ES6 version of JavaScript providing dynamic refreshment of pages upon changes with a flexible, responsive, and intuitive user experience. Furthermore, the system has the permissioned MultiChain as a solution to both on-chain and off-chain data storage, encryption, hashing and tracking of data, together with Ethereum for access control and enabling transactions with *ethers*. Figure 5.7 is a blockchain diagram representing the interaction between the customer (data provider) and other e-commerce apps (data consumers) of the DUDS framework-based user-controlled data sharing scheme.

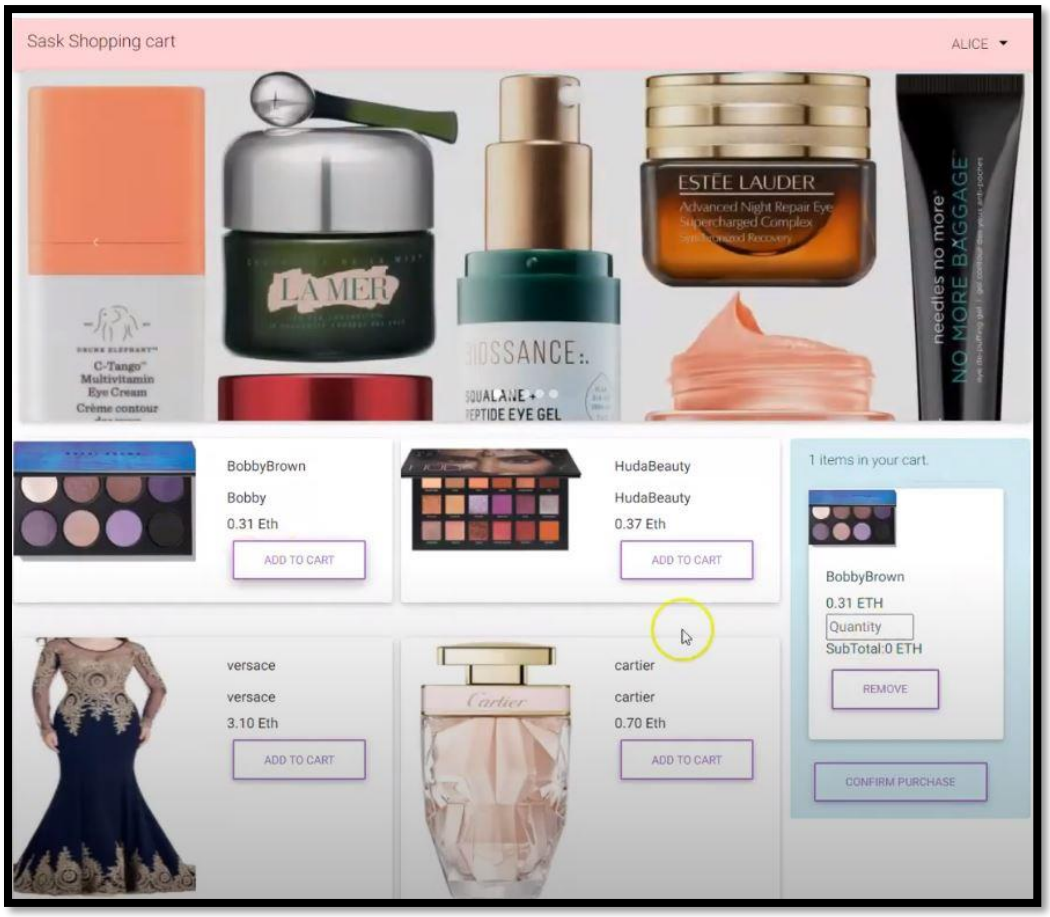

Figure 5. 8. Implementation testing – user interface

The interface of the web application for the customers is shown in Figure 5.8. This Blockchain-based 3-tier online shopping website is created via *Spring initializr*, which allows us to include the necessary dependencies along with an option to choose either a maven or a *gradle* project. The application built is of maven type with spring boot on the backend, which is an extension of an application development open-source framework called spring framework. Spring boot helps to drastically reduce the development time by eliminating the boilerplate configurations required for setting a spring application, which increases productivity. It decides which default dependencies and packages to use for the configuration. It has an embedded server tomcat to reduce complexity in deployment and test applications. However, the configurations can be changed by listing the needed dependencies in the pom.xml file.

Spring-boot-starter-parent is a dependency used to mainly provide plugin management for Spring Boot-based applications. It contains the default versions of Java to use, the default versions of dependencies that Spring Boot uses, and the default configuration of the Maven plugins.

Subsequently, Spring-boot-starter-data-jpa is an abstraction of JPA that helps to access data between relational databases and Java objects/classes. It allows an API to process queries and transactions in terms of objects with the database not represented in tables and columns. To allow the web application to use restful web services and Tomcat as the default embedded server, dependency Spring-boot-starter-web is included with all other dependencies related to web development.

Spring Boot provides default configurations to the H2 database. In order to connect to MySQL, Mysql-connector-java is listed in the pom file and those configurations are overridden and database properties are defined in the application.properties file.

Properties include 'jdbc:mysql://localhost:3306/projectname' as the spring.datasource.url and 'com.mysql.cj.jdbc.Driver' as the spring.datasource.driver-class-name along with username and password to MYSQL.

Lombok is another dependency used to generate getters and setters for data/model objects automatically by using annotations. It also reduces boilerplate codes and keeps them clean.

Since it is a maven project, Spring-boot-maven-plugin uses the public static void main() method as a runnable class. The application is created with REST Architecture, which governs a set of rules that a developer follows on the server so that clients can communicate to it. The request from the client contains an endpoint, header, method, and data. The method defines the type of request: Get, Post, Put, Patch, and Delete to carry out operations like create, read, update, and delete. Basic authentication is carried out using a username and Bcrypt encoded password and JWT for transmitting information from client to server as a JSON object. Authorization and authentication, routing between services are all carried out securely based on the permission given to the token. JWT is created with an encoded header, payload, and signature signed.

For the front-end part, *ReactJS* is the main technology used. To set up the project, Nodejs, which is a runtime environment based on the V8 JavaScript engine, is adopted to build a skeleton of *ReactJS* application using the create-react-app command in the command line. 'Npm start' loads the project on the browser. With Nodejs, Create-react-app gets installed to set up a development environment to use java features, tools to start a project and optimize for production. It employs Babel and Webpack under the hood. Babel converts the ES6 version of JavaScript into ES5 because most browsers do not support ES6 yet. On the other hand, Webpack is an asset bundler. It collects all the assets (codes and files) and creates a big bundle that can be sent from the server to a client's

browser. ‘Npm create-react-app appName’ is the command line that creates a new *ReactJS* application. Like in the backend, spring boot dependencies can be viewed and changed in the pom.xml file, and dependencies in Reactjs can be viewed in the package.json file.

Axios is a promise-based HTTP client for browsers and nodes that grants asynchronous HTTP requests to rest endpoints to generate CRUD operations. It can be used either in plain JavaScript or using libraries like React and Vue. Furthermore, React-dom is a library for web apps that lets a developer manipulate a DOM. It provides methods like render() which manipulates the content passed into the React node. React-router-dom operates react-router at the core that helps to render components depending on the route being used in the URL, and those navigations take place without refreshing the page. Since enhancing the presentation of the components can enhance the user experience significantly, Reactstrap, which is a component library, provides prebuilt bootstrap components to allow flexibility and prebuilt validation.

This platform follows the DUDS framework that allows the customer to purchase items with the virtual currency *ether*, which gets transferred from the customer’s account to the seller’s account. All the participants in the system must have their Ethereum account addresses, which they access via Metamask wallet as shown in Figure 5.9 to make payment.

Moreover, it offers the web form as shown in Figure 5.10 to the customers allowing them to select their data sharing preferences and execute the associated smart contracts. This removes the need for coding the smart contracts by the customers.

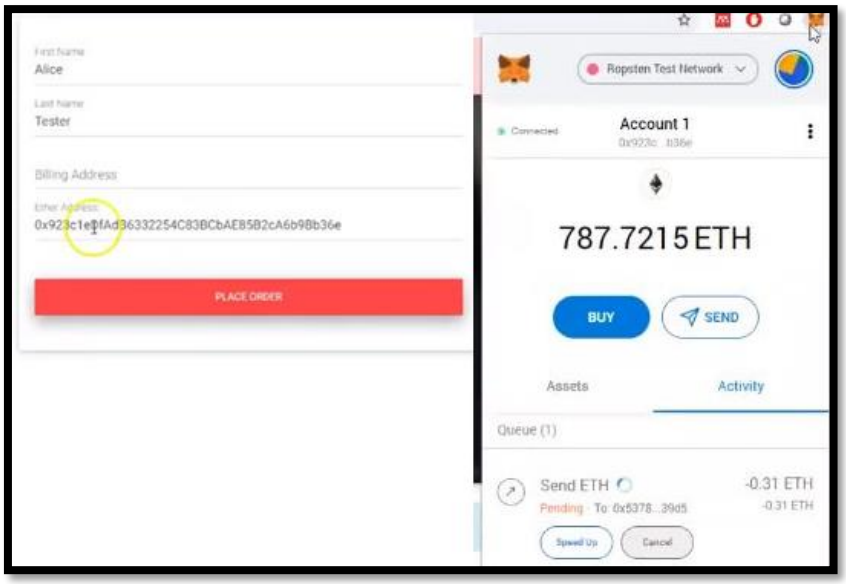

Figure 5. 9. Implementation testing – purchase item

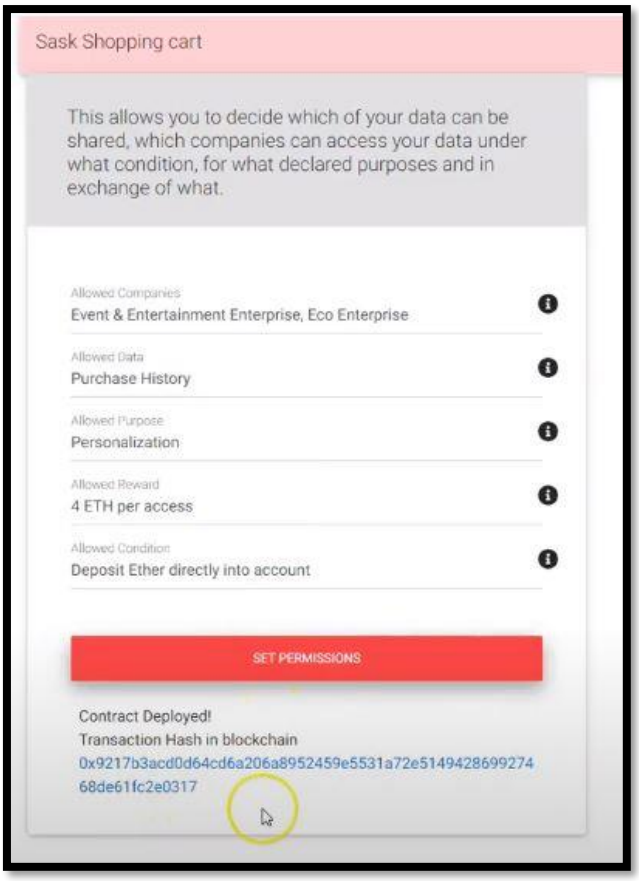

Figure 5. 10. Implementation testing – deploy contract for a new consent

The system also has the private permissioned MultiChain data-sharing consortium network that allows the off-chain storage of customer data (profile data and transaction history) at the shopping cart node. Only the participating companies such as Shopping Cart enterprises are in the consortium network. The MultiChain node (company) enables encrypting the data, creating a transaction with their hashes and committing it into the blockchain. It also allows searching and delivery of the data over the consortium network. Once the data access is successful on the recipient enterprise node, the smart contracts get triggered and with their successful execution, the tokens are transferred from the data consumer's Ethereum account to the customer's account. As shown in Figure 5.11, the consumer node is the Event and Entertainment (E&E) Enterprise that executes final smart contracts while receiving the requested data (with verified hashes) to its local storage and providing incentives to the customers for their data. The customer data sharing mechanism between the companies or participating enterprises is similar to that explained in section 5.1.

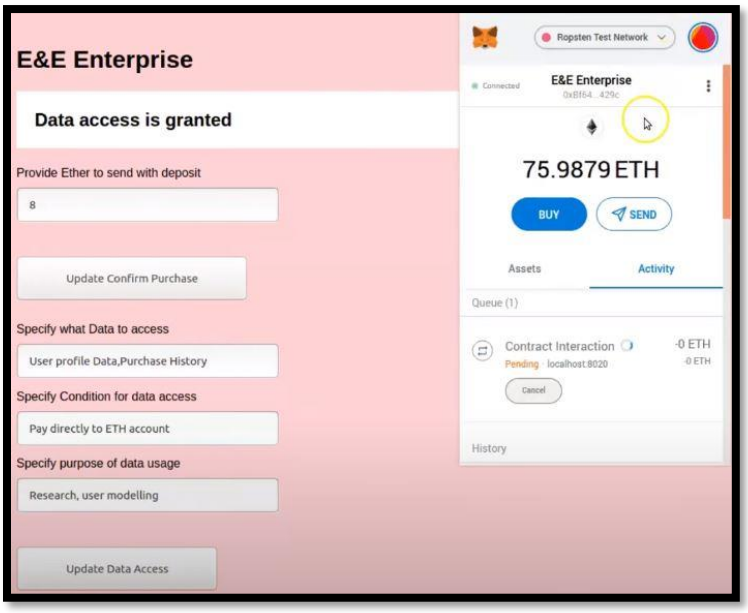

Figure 5. 11. Implementation testing – grant data access

5.3.4 Smart Contracts Deployment

Direct connection to different testnet[45] (test blockchain network) or mainnet[46] (main blockchain network) without downloading the blockchain onto the system can save time and storage space. *Geth*[47] can be used for the Ethereum blockchain node, but *Infura*[48] has made the development of Dapps even easier and simpler with its APIs that help to connect the application to the blockchain with just a key. The smart contracts can easily be deployed on the testnet via Infura. In our case, the Infura Project ID was created and then the associated project secret was obtained.

For the Ropsten endpoint, the associated URL was https://ropsten.infura.io/v3/<secret>. We began by initializing the Truffle project on one directory and installed the truffle-HDwallet package using npm. Truffle requires a running Ethereum client that supports the standard JSON RPC API. The seed words from the Metamask wallet and the *Project ID* and *Secret* from Infura are needed to configure the truffle project. I added the following lines in the truffle-config file.

const HDWalletProvider = require('truffle-hdwallet-provider');

[45] https://ropsten.etherscan.io/
[46] https://etherscan.io/
[47] https://geth.ethereum.org/
[48] https://infura.io/

const mnemonic = Metamask seed words;

provider: () => new HDWalletProvider(mnemonic, associated URL)

The initial templates of the smart contracts were constructed with Solidity language (see Appendix VI). Then, we created migration files for the smart contracts and deployed them into the Ropsten Testnet blockchain with the migrating command: *truffle.cmd migrate –network Ropsten.*

The successful deployment, followed by compilation and migration at some addresses in the Ropsten network, resulted in JSON files in the build directory that consisted of network information including contract address, transaction hash, and event. This information was used to connect the application with the smart contracts that were deployed on the blockchain.

5.3.5 Smart Contract Execution

Figure 5.12 presents the sequence diagram of the smart contract execution. The customer is the data owner who is responsible for constructing and deploying the smart contracts. The data owner can set the permission for data access and must deploy the smart contract with a deposit. Initially, the state of the contract is set to 0. If any issues are encountered with the contract, the data owner can abort the contract by changing its state to 3, which will refund the deposit. When the data consumer wants to access the data, then the request for data access is initiated by sending the deposit plus the data price that was set by the data owner, which will change the contract state to 1. Once the data is accessed, the data consumer makes another request to get the refund and changes the state of the contract to 2. The smart contract sends the deposit back to the data consumer. Finally, the data owner requests the incentives and initial deposit, which will be refunded by the smart contract. The smart contract executes everything autonomously as per its state and predefined conditions.

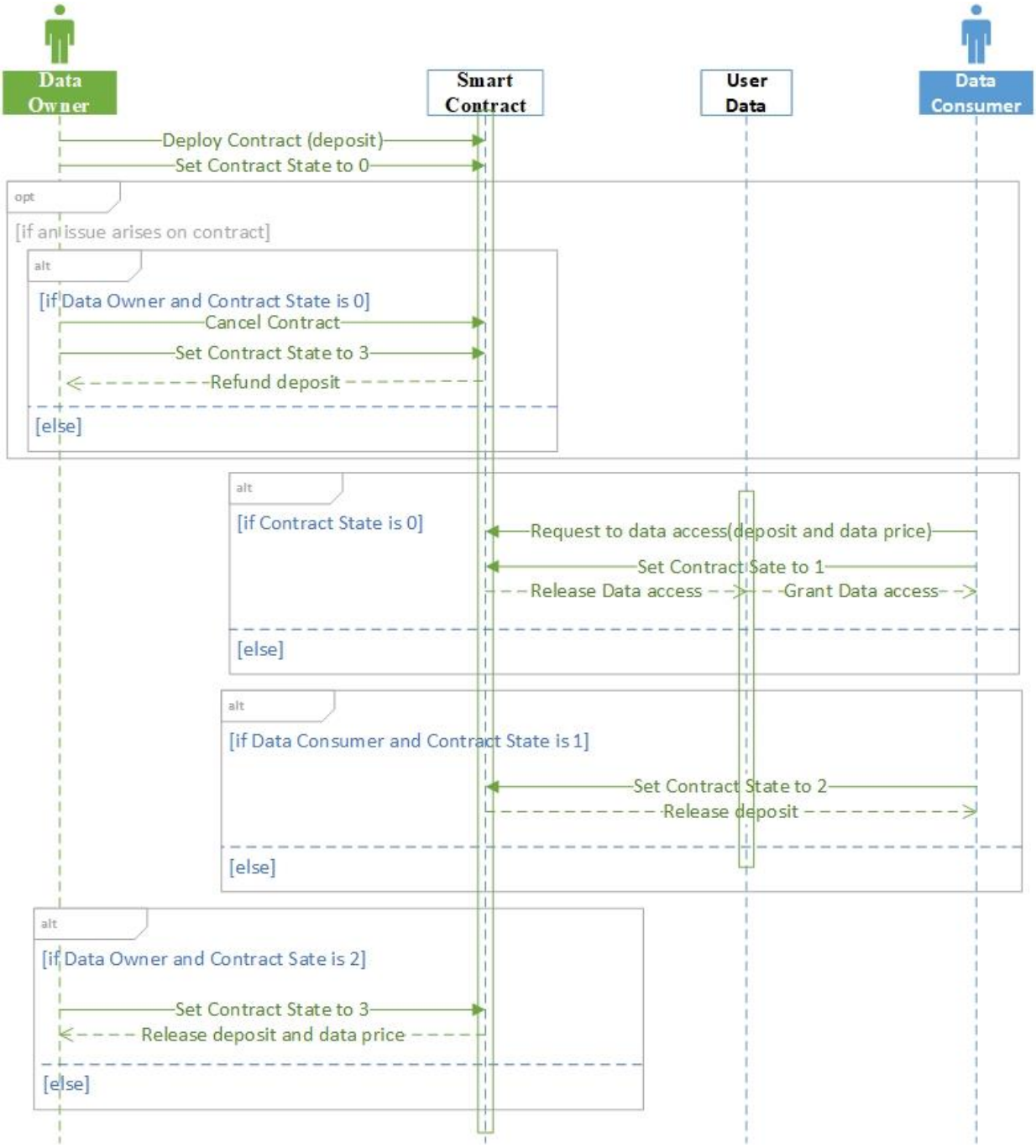

Figure 5. 12. Smart contract sequence diagram

5.3.6 Summary

This section presented a novel platform for the online shopping cart based on the DUDS framework that enables user-controlled privacy and data-sharing policies to be encoded in smart contracts. The use of Ethereum blockchains enables the purchase of the items using virtual cryptocurrency and

ultimately builds up a verifiable record of the provenance, accountability of access, and incentives for customers to share their data in terms of rewards. Data sharing is done in the private blockchain network and there is no storage of user profile data on the public blockchain, so the data deletion rule in a compliant manner can delete each category of data. Moreover, there is transparency over which enterprise has access to the relevant customer data, when, and for what purpose. I then ran the usability experiments with this version of the DUDS platform. The usability statistics generated from people using the version of my system are presented in Chapter 6.

5.4 Conclusion

This chapter provided the design and implementation of the DUDS platform with example scenarios for sharing different user data types in three different domains—tourism, research and online shopping cart—that incentivize data owners in terms of payment through blockchains for the use of the data by applications. The illustration of the conceptual DUDS architecture in this chapter offers guidance to the blockchain designers to design a similar system for the intended platforms. Furthermore, the performance of the DUDS platform was also analyzed through a set of comparative experiments to calculate the latency and memory consumption for the consortium network, and gas consumption and transaction cost for the smart contracts' deployment and execution. The test results indicated that the nodes responded quickly in all cases with befitting transaction costs. This chapter also advocated the suitability of the user study for evaluating the DUDS framework and building more abstract designs of the user-controlled privacy-preserving decentralized user data-sharing model. In the next chapter, I present the final study to evaluate the DUDS framework using the augmented TAM with trust model on the real-life blockchain-based system.

6 EVALUATION OF THE DUDS FRAMEWORK

This chapter presents the final evaluation of the user acceptance of an implemented prototype system based on the DUDS framework by observing various latent variables affecting the development of users' attitudes and intention to use the system. It also aims to uncover the dimensions and role of trust, security and privacy alongside the primary Technology Acceptance Model (TAM)-based predictors and their causal relationship with the users' behavior to adopt such DUDS platforms. The content of this chapter is based on my published article (Shrestha et al., 2021).

6.1 Augmenting The Technology Acceptance Model With Trust Model for The Initial Adoption of A Blockchain-Based System

I tested the augmented TAM with Trust Model on a blockchain-based system (BBS), which is the online shopping cart version from the previous Chapter 5.3 that offers rewards to the data owners in terms of payment through blockchains for the use of the data by applications, as specified by the contracts. This BBS comprises two subsystems: A Shopping Cart System (SCS) - a system-oriented towards end-users and a Data Sharing System (DSS) - a system-oriented towards system administrators. I set research questions and hypotheses and conducted online surveys by requesting each participant to respond to the questionnaire after using the respective system. The main study comprises two separate sub-studies: the first study was performed on SCS and the second on DSS. Furthermore, each study data set comprises initial pre-test and post-test data scores. I analyzed the research model with partial least square structural equation modeling. This empirical study validates our research model and supports most of the research hypotheses. Based on the findings, I deduced that TAM-based predictors and trust constructs cannot be applied uniformly to BBS. Depending on the specifics of the BBS, the relationships between perceived trust antecedents and attitudes towards the system might change. For DSS, trust is the strongest determinant of attitudes towards the system, while SCS has perceived privacy as the strongest determinant of attitudes towards the system. The quality of the system shows the strongest total effect on the intention to use SCS, while perceived usefulness has the strongest total effect on intention to use DSS. Trust has a positive significant effect on users' attitudes towards both BBS, while security does not have any significant effect on users' attitudes toward BBS. In SCS, privacy positively affects trust, but

security has no significant effect on trust, whereas, in DSS, both privacy and security have significant effects on trust. In both BBS, trust has a moderating effect on privacy that correlates with attitudes towards BBS, whereas security does not have any mediating role between privacy and attitudes towards BBS. Hence, the research findings recommend that while developing BBS, particular attention should be paid to increasing user trust and perceived privacy.

6.1.1 Background

As suggested by (Cunningham, 1967), the evaluation process is crucial in studying the user perception of the adoption of new information technology services. I already presented that the Technology Acceptance Model (TAM) proposed by (Davis, 1989) has been used widely in the literature to examine whether users understand the underlying technology and can competently use the services (Granić & Marangunić, 2019). In addition to that, many studies extend TAM by adding external constructs depending upon the contexts to explain the critical relationship between customers and their adoption of the new technology (Melas et al., 2011). With the rapid development of the use cases of blockchain in recent years, a few studies have already been conducted considering the user acceptance of an abstract blockchain-based system. Although numerous extensive systematic studies have been conducted on evaluating the performance of blockchain-based systems, to the best of our knowledge, no study has been conducted in the context of users' acceptance of real-life blockchain-based applications except for bitcoin as financial technology (Folkinshteyn & Lennon, 2016). Previous works, including mine, have evaluated user acceptance of the prototype blockchain-based system using an extended Technology Acceptance Model (TAM) in (Kern, 2018; Shrestha & Vassileva, 2019b) and the trust model in Shin (2019). The previous studies suggest that the blockchain-based system will be accepted if it is perceived as trustworthy, convenient, and useful.

The major contribution of this final study of my research is that it expands the previous work by conducting a new user study on a real-life blockchain-based system (BBS), described in section 5.3. This final study includes the augmented TAM by incorporating additional constructs—User Trust, Perceived Security and Perceived Privacy—in the technology adoption study and presents the total effect and mediation analyses. We have used the validated constructs for our study. Before conducting the main study, we did a pilot study with 14 researchers to evaluate the suitability of

adapting the already validated questionnaire. The findings from this final study are informative and potentially useful for designing new blockchain-based systems.

The BBS of this study is the MVP (Minimum Viable Product) as presented in Chapter 5.3, which is the implemented general-purpose blockchain-based system providing a solution to four important problems: private payment, ensuring privacy and user control, and incentives for sharing. The BBS was constructed for the online shopping cart, which also allows customers to connect to the seller directly and share personal data without losing control and ownership of it. The BBS can be viewed as the combination of two subsystems, a customer-specific shopping cart system (SCS) and a company-specific data sharing system (DSS). SCS allows customers to set their data sharing preferences and deploy them via smart contracts. Similarly, DSS allows companies to check data integrity, get tamper-proof records and proof of the existence of every transaction while sharing data in the consortium blockchain network. Therefore, the BBS used in the study was a novel decentralized application that covered the aspects of both the customer and company. So, its in-depth analysis to examine the factors of the Trust model and the TAM indicators that mostly affect the user acceptance of the BBS was crucial to provide an opportunity for a broad debate and perspective on potential uses of blockchain- and smart contract-based DUDS framework for the eCommerce domain along with other different important industries such as healthcare, agriculture, tourism, and research fields.

Therefore, this final usability study is based on the user evaluation of the DUDS framework-based SCS and DSS, before and after using those sub-systems by the selected participants, using the validated constructs of the TAM and the Trust model. This new augmented model incorporates both TAM with perceived ease of use, perceived usefulness and quality of system, and Trust model with security and privacy variables, and it can be applied to evaluate the acceptance of the general blockchain- and smart contracts-based systems. Here, using the PLS-SEM on augmented TAM, I hypothesized and validated various causal relationships between the constructs of interest and intention to use the BBS.

6.1.2 Blockchain-Based System (BBS)

The term BBS for a general blockchain-based system was initially used in the literature (Jun, 2018) without any detailed explanation. BBS in our study represents the blockchain-based service that we have developed with an engineering-oriented approach to address trust-aware business

processes in an e-commerce domain in the context of an online shopping cart system (Shrestha et al., 2020). The requirements for the BBS are:

- To enable companies to increase trust in their products and supply chains.
- To offer direct payment with native Ethereum tokens, thereby enabling privacy and confidentiality.
- To create proof of the existence of every transaction.
- To give the users full transparency over who accesses their data, when, and for what purpose.
- To enable companies to share customers' data among others in the consortium network.
- To provide incentives to customers in real-time for sharing their data.

This BBS has a 3-tier architecture (Fernandez et al., 2008) employing *Spring Boo*t and *React* as the main building technologies. The system uses permissioned MultiChain as a solution to both on-chain and off-chain data storage, encryption, hashing and tracking of data, together with Ethereum. Ethereum is used for access control and enabling transactions with *ethers* that allow users to shop online with all the transactions stored in the blockchain and get incentives for permitting them to share their data as they specify in the smart contracts. Figure 5.7 presents the interaction among the customer (data provider) and other e-commerce companies/apps (data consumers) of the BBS. The system comprises two subsystems: Shopping Cart System (SCS) and Data Sharing System (DSS). SCS is used in the online shopping cart enterprise. It has a payment mechanism supporting cryptocurrency, *ether,* and manages the mutual agreement between customers and enterprise through smart contracts. SCS automatically registers the immutable timestamped information about the transactions that acts as proof of existence and can be useful to settle any disputes between the stakeholders in the future. Moreover, SCS deploys smart contracts that allow customers to provide their data sharing preferences on a template form without needing them to write the code for the smart contracts. The smart contracts support users in the following ways (Shrestha & Vassileva, 2019a):

- Give users full transparency over who accesses their data, when and for what purpose.
- Allow users to specify the purposes of data sharing, which kinds of data can be shared, and which applications or companies can access the data.
- Provide an incentive to users for sharing their data (in terms of payment for the use of the data by applications, as specified by the contracts).

DSS is used for sharing user data among the companies that provide the shopping cart system to the customers. DSS allows enterprises to form a consortium blockchain network in the MultiChain environment so that user data are only shared with the particular node that has been given the data access permission as defined in the smart contracts when deployed by customers on SCS. DSS offers tamper-proof encrypted data storage, publication, and provenance mechanisms with a transparency of the event log mechanism in collaborative processes where different enterprises use published/shared data.

6.1.3 Augmented Technology Acceptance Model

This study applies the augmented TAM with trust model to the BBS that we implemented, with participants who used the system before answering the survey questionnaires. Our study also uncovers the individual mediating effects of trust, security, and perceived usefulness.

In classical TAM, the main design constructs, such as perceived ease of use and perceived usefulness, have shown significant influence on the behavioral intention of the user to adopt the information systems, and the latest study (Shin, 2019) shows the necessity of considering the Trust-Security-Privacy factors in the decision model of the blockchain-based-solution adoption. So, we adopted the partial least square structural equation modeling (PLS-SEM) analyses on the augmented TAM, as it is a useful technique to estimate complex cause-effect relationship models with latent variables, and we aimed to model the latent constructs under conditions of non-normality and small sample sizes (Wong, 2013).

Although in the software engineering domain, security and privacy are regarded as part of QoS, in this study, we have presented perceived security and perceived privacy as separate constructs.

QOS refers to the technical details of the system interface and system's quality that produces output response such that the technology attributes singularly or jointly influence user satisfaction. Hence, it is assumed that the QOS affects user satisfaction and that directly or indirectly through PU, affects users' intention to use the system (DeLone & McLean, 1992; Shrestha & Vassileva, 2019b).

Moreover, perceived privacy and perceived security have critical roles in the acceptance of the technologies, as the prior research suggests they have a significant effect on users' attitudes that positively influence their intention to use the technologies.

Perceived privacy, which is the attitudinal privacy or privacy concern, undoubtedly plays a critical role in user accepting technologies (Hoffman et al., 1999; Poon, 2008). It sheds light on the possibility of unauthorized use and access to the personal and financial information of the users by the companies whose services they are intending to use (Dwyer et al., 2007).

Regarding perceived security, the protected financial and personal information may get compromised by theft and fraudulent activities leading to vulnerability on the internet. Because of this, a sense of security becomes a major concern for the customers when asked to hand out their details(Gefen, 2000; Shrestha, 2014; Wang et al., 1998). Authors (Linck et al., 2006) have argued that a lack of subjective security in the user's mind will create hesitation to use systems.

Research has shown that user trust has a positive significant impact on attitude and intentions to use systems (Papadopoulou, 2007). With greater trust, users question the authenticity of online services less. Trust gives users the feeling that the service is credible, enabling their greater engagement and further recommendation. Current state of art examines trust from the perspective of different variables such as privacy and security that determine a user's behavioral intention both before and during the user interaction with an online service. The user acceptance behavioral model for theoretical social network services is very useful for conceptualizing the role of perceived security and perceived privacy (privacy concern from attitudinal privacy) on user trust. Their findings revealed that perceived security has a moderating effect on perceived privacy that correlates significantly with trust the user can have in the system.

6.1.4 Related Work

Numerous studies have been conducted to examine the factors that determine the acceptance of information technology in the context of an extended TAM and Trust model. We cover a cross-section of those studies that are related to our work.

To the best of our knowledge, (Folkinshteyn & Lennon, 2016) conducted a very first user study with TAM in the context of the adoption of bitcoin as financial technology. Their findings revealed both positive and negative factors associated with the acceptance of bitcoin, the first real-life application of blockchain technology. They have also argued that the cryptocurrency offers borderless and efficient transactions with significant positive factors in Perceived Ease of Use (PEOU) and Perceived Usefulness (PU), giving users full control over their currency. However, it is also extremely volatile without being lenient with security breaches or errors (Folkinshteyn &

Lennon, 2016). So, it has both risks and benefits that affect the overall adoption of the cryptocurrency. Their findings also suggested exploring other aspects beyond TAM variables to consider the underlying risk and trustworthiness constructs associated with blockchain-based applications. Previous research on an abstract blockchain-based application model suggested that the blockchain-based system can be accepted if it can sustain enough trust in the user and is perceived as convenient and useful in a highly competitive market. Almost all of the existing research so far is limited to blockchain-based prototype systems, using an extended TAM (Kern, 2018; Shrestha & Vassileva, 2019b) and Trust model. Our current study extends the research contribution of the prior study (Shrestha & Vassileva, 2019b) by conducting a new user study on the real-life blockchain-based system, BBS.

Gefen et al. have previously explored a mixed model with a TAM and Trust model to study the adoption of the online shopping setting (Gefen et al., 2003). Their model presented the use of the online system into both system attributes, such as perceived usefulness and perceived ease of use and trust in e-vendors. Their model resulted in the integrative indication of the TAM and Trust constructs as good predictors for the output response, which was the behavioral intention to use the online shopping system. Therefore, the current study adopts a similar model and presents it as augmented TAM, which comprises an extended TAM and Trust model.

As online activities such as online shopping generate a plethora of real-time transactions of all kinds of assets and information, they are prone to security and privacy-related risks (Roca et al., 2009). A privacy issue mostly occurs when there is unwarranted access to the users' personal data but that does not necessarily involve security breaches. It can happen due to poor access control mechanisms in the system allowing malicious actors to control the system. Privacy breaches are critical issues and they often exist on the online services where users typically feel hesitant to provide private information over the internet (Hoffman et al., 1999). Shin (2010) previously explored the impact of security and privacy in the acceptance of social networking sites. Later, Shin (2019) presented the role and dimension of user trust in the emerging blockchain context. At the same time, (Siegel & Sarma, 2019) argued that it had not been investigated how privacy/security factors affect user's behavioral cognitive process of accepting the blockchain-based systems.

My study, in addition to previous TAM-validated constructs, explores the users' perception towards the security and privacy aspect of the BBS and their influence on intention to use the BBS

by using the moderating effects of trust on attitudes towards the system. Besides, the current research aims to answer the following research questions when exploring the relationship between different indicators of the augmented TAM with the trust model:

- RQ1: Which of the design attributes is/are the strongest antecedents of the attitudes towards BBS?
- RQ2: Which of the design attributes is/are the strongest antecedents of the intention to use BBS?
- RQ3: Is the influence of privacy on attitudes towards BBS mediated by security and/or trust?
- RQ4: Is the influence of security on attitudes towards BBS mediated by trust?
- RQ5: Is the influence of ease of use/quality of system on intention to use BBS mediated by perceived usefulness?

6.1.5 Research Model and Hypotheses

Figure 6.1 presents the structural model with the main constructs and their associated structural paths. Fourteen research hypotheses are thus constructed for our research model based on the findings of the literature review presented briefly in the previous section.

Perceived Usefulness and Perceived Ease of Use (Davis, 1989; Davis et al., 1992)
H1: Perceived ease of use significantly influences the perceived usefulness of BBS.
H2: Perceived ease of use significantly influences the intention to use BBS.
H3: Perceived usefulness significantly influences the intention to use BBS.

Quality of System (Koh et al., 2010)
H4: Quality of system significantly influences the perceived usefulness of BBS.
H5: Quality of system significantly influences the intention to use BBS.

Attitude Towards BBS (Shin, 2017)
H6. Attitude towards BBS significantly influences the intention to use BBS.

Trust (Dennis et al., 2012); (Jian et al., 2000)
H7. Trust positively affects users' attitudes toward BBS.

Perceived Privacy (Buchanan et al., 2007)

H8. Perceived privacy has a positive effect on the users' trust in BBS.

H9. Perceived privacy has a positive effect on the users' attitudes toward BBS.

H10. Perceived privacy positively or negatively affects users' perceived security.

H11: Privacy concern positively affects users' behavior on general caution.

H12: Privacy concern positively affects users' behavior on technical protection.

Perceived Security (Shin, 2010)

H13. Perceived security positively affects users' trust in BBS.

H14. Perceived security positively affects users' attitudes toward BBS.

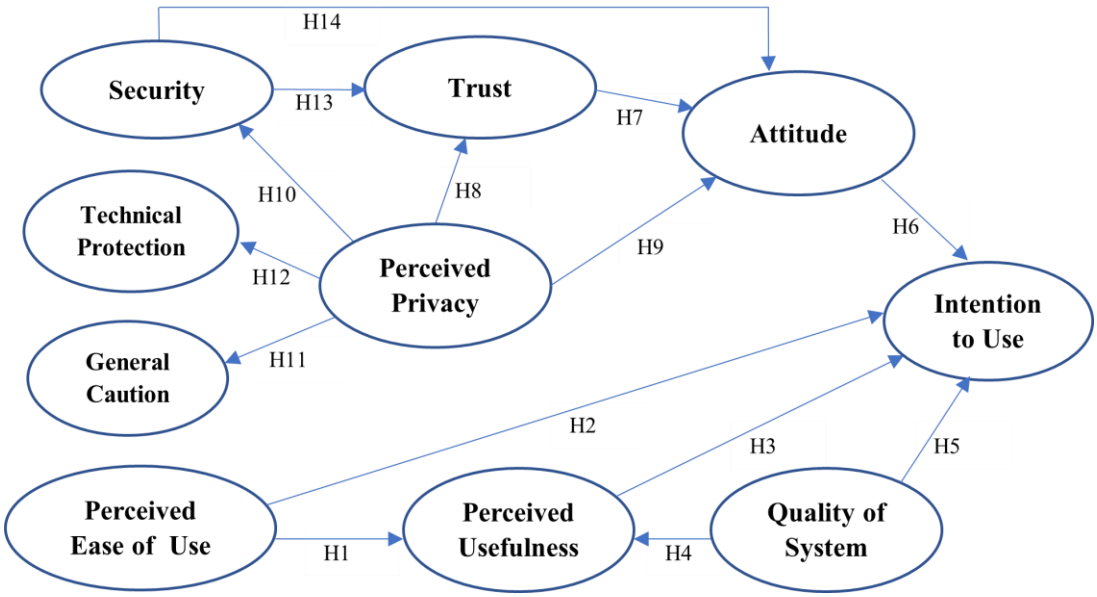

Figure 6. 1. An Augmented TAM with trust model

The main study comprises two separate sub-studies: the first study was performed on the SCS and the second on the DSS. Furthermore, each study data comprises pre-test and post-test data scores. The pre-test defines the data collected from participants before they use the system, whereas post-test data is collected after participants use the system.

The pretest study can be considered as the study associated with the prototype model. Since, the present study follows the previous research work from (Shrestha & Vassileva, 2019b), the pretests for the current study do not include the constructs from classical TAM, as they were already evaluated in the previous study. So, the pretests of the current study do not present data for hypotheses H1 – H6. The post-tests for both SCS and DSS do not have *behavioral privacy-general*

caution and *-technical protection* constructs, as they are only evaluated once, during the pre-test. So, the post-test data do not test hypotheses H11 – H12.

6.1.6 Materials & Methods

The present study was approved with delegated review by the University of Saskatchewan Behavioural Research Ethics Board (Beh-REB). The approval with reference number Beh #ID2106 was given for behavioural application/amendment form, consent form, and survey questionnaire. We conducted an initial pilot study with 14 quantitative research experts at the University of Saskatchewan to evaluate the feasibility and duration and improve upon the study design of our research approach. The participants in the pilot study provided feedback with their opinion of the survey in general. Based on the pilot test outcomes and the review of quantitative research experts, the final survey questionnaires were modified and restructured, and then the research model was empirically tested by collecting survey data. The design of the research instrument, sample organizations and sample demographics are described below.

6.1.6.1 Research Instrument Design

We conducted online surveys through SurveyMonkey by requesting each participant to respond to the questionnaire on different constructs. The survey instrument adapted the constructs that were validated in prior studies by (Buchanan et al., 2007; Davis, 1989; Davis et al., 1992; Dennis et al., 2012; Jian et al., 2000; Koh et al., 2010; Shin, 2010; Shin, 2017). The instrument consists of six items for perceived ease of use, six items for perceived usefulness, four items for quality of system, three items for perceived enjoyment, four items for intention to use, three items for perceived security, nine items for trust, four items for attitudinal privacy (perceived privacy), four items for behavioral privacy-general caution, four items for behavioral privacy-technical protection, and three items for attitude towards BBS. For our later analysis, we did not consider data related to perceived enjoyment. All the respective items (questions) of the constructs are provided in the appendices, which are mentioned in the next section. We measured the responses to the items on a 7-level Likert scale from 1 = strongly disagree to 7 = strongly agree.

6.1.6.2 Sample Organizations

We recruited participants through a website announcement on the University of Saskatchewan's PAWS homepage and on the social networking site, LinkedIn. Participation was entirely voluntary. The participants had to read and accept the consent form (see Appendix VII) to participate in the study. No real identities and email addresses were collected during the data-gathering phase in the surveys. The consent for participation was obtained via an implied consent form. By completing and submitting the questionnaire, participants' free and informed consent was implied and indicated that they understood the conditions of participation in the study spelled out in the consent form.

To contextualize the surveys for SCS, we provided participants at the beginning of the pre-test survey questionnaire (Appendix VIII) with a video[49] with a brief description of blockchain technology and BBS. The inclusion criteria for the SCS survey were simple: any individual with knowledge about the internet could participate. After participants completed the pre-test survey, we presented them with another video[50] about using the SCS and hosted a remote session allowing them to use the SCS for fifteen minutes. They were given a similar task about creating an account, exploring the functionalities, and conducting the purchase of the items using DLT over the SCS as in the same video that they watched. We did not record the participants during their tasks due to the nature of our study approval by the Beh-REB, but we noted their comments and confusion during their interaction with the system. Thereafter, we presented them with a post-test survey questionnaire (Appendix IX) to measure different constructs of our Augmented TAM with Trust model. Similarly, we conducted the pre-test and post-test surveys for the DSS part as well. The post-test survey questionnaire for DSS is presented in Appendix X. Each participant in the DSS survey was also asked to use the DSS remotely for fifteen minutes after allowing them to watch a video[51] about using the DSS. The inclusion criterion for the DSS survey was that the participants should be from a technical (computer science or engineering) background because the DSS includes technical aspects that only a software developer or system administrator can understand well. Most of the participants completing DSS surveys also took part in the SCS surveys.

[49] https://youtu.be/HlqadcFDhU4
[50] https://youtu.be/6qdfMU3aKZA
[51] https://youtu.be/J7xik-6wGeQ

6.1.6.3 Participant Demographics

A total of 66 participants took part in the SCS study and 53 participated in the DSS study. However, upon cleaning, 63 valid responses for SCS and 50 for DSS were left for the analysis. We used a partial least square nonparametric bootstrapping procedure to test the statistical significance with 5000 subsamples (Hair et al., 2013) so that the resampling process would create subsamples with observations randomly drawn from the original set of data.

For the study, we based our survey on collecting data from the participants who understood at least something about the blockchain and smart contract technologies after watching the video that we prepared on blockchain technology and BBS. We recruited participants from Academia (students) using the University Bulletin Board. We also recruited a few participants from industry for the evaluation of the DSS (system administrators and developers). The mean score suggests that for SCS, 79% of participants have basic knowledge and 19% have advanced knowledge of blockchain technology; whereas for DSS, 68% of participants have basic knowledge and 28% have advanced knowledge of blockchain technology. The demographics of the participants are available in appendix XI.

6.1.7 Results

We used SPSS version 26 to process the collected data with descriptive statistics (see Appendix XII for survey data). We analyzed the research model with structural equation modeling using smartPLS (Partial Least Squares). PLS is a well-established technique for estimating path coefficients in structural models and has been widely used in research studies to model latent constructs under conditions of non-normality and small to medium sample sizes (Wong, 2013). The structural equation model (SEM) as suggested by (Hair et al., 2013) includes the testing of the measurement models (exploratory factor analysis, internal consistency, convergent validity, divergent validity, Dillon-Goldstein's rho) and the structural models (regression analysis). We started by fitting the measurement models to the data and later we tested the underlying structural models.

We applied the path weighting structural model scheme in smartPLS (Wong, 2013), which provides the highest R^2 value for endogenous or dependent latent variables. The purpose of PLS regression is to combine features from principal component analysis (PCA) and multiple regression

(Roca et al., 2009). PLS-SEM is applicable for all kinds of PLS path model specifications and estimations. We first used 300 maximum iterations for calculating the PLS results with 7 stop criterion values (i.e., 0.0000001 is the default amount) as recommended by (Hair et al., 2013) in order to seek convergence on a solution. This enabled the PLS algorithm to stop when the change in the outer path weights between two consecutive iterations was smaller than the 7-stop criterion value for fine tolerance (i.e., converges at very low levels of iterative changes in the latent variable scores) (Hair et al., 2013). We then used a nonparametric bootstrapping procedure to test the statistical significance of various PLS-SEM results that include path coefficients and R^2 values. Bootstrapping is a resampling technique with replacement from the sample data to generate empirical sampling distribution. In our case, we used 5000 subsamples and a two-tailed test type with a 0.1 significance level (Hair et al., 2013).

Descriptive Statistic

We had a 7-level Likert scale for the responses to the items, so we categorized the scale in seven score ranges (from Table 4.4 in Chapter 4) to observe the overall impression of the constructs. We collected scores for all the items in perceived ease of use, perceived usefulness, quality of system, trust, security, privacy, attitudes, and intention to use constructs of our model.

The scores obtained for selected constructs suggest that user perceptions on the benefits of using BBS should be maintained by making improvements to achieve a higher level of score category. The preliminary descriptive statistic of the obtained data is shown in Figure 6.2, which indicates that the average results of the constructs range between 5 and 6, so they qualified for the quite high category (Shrestha & Vassileva, 2019b). The comparatively lower pre-test scores indicate that participants developed confidence and trust towards the overall usefulness, usability, attitudes, and intention to use the BBS after they used the SCS and DSS. Furthermore, higher scores for PEOS, PU, QOS for SCS over DSS signify that the participants find SCS easier to use compared to the participants who participated in the DSS part of the study. However, all the selected constructs in our study provided a significant impression in the context of both BBS.

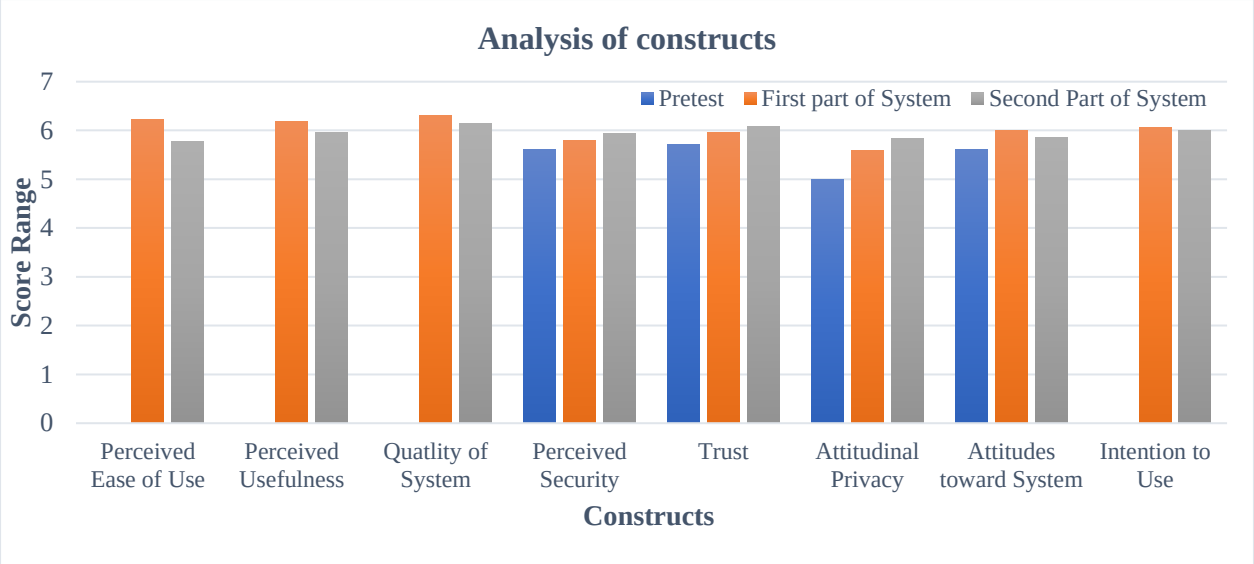

Figure 6. 2. Analysis of constructs

Measurement Validation

We checked the measurement model with the exploratory factor analysis by testing the convergent validity, reliability of measures, and discriminant validity.

For Exploratory Factor Analysis, we first checked the factor loadings of individual items, which are available in Appendix XI, to see whether the items in each variable loaded highly on their own construct over the other respective constructs. According to (Chin et al., 2008), factor loadings exceeding 0.60 can be considered significant. In our study, all the indicators in the measurement models had a factor loading of value greater than 0.60 except for Item 4 in the construct Behavioral Privacy-Technical Protection (BP-TP4). Since the square of factor loading is directly translated as the item's reliability, the item BP-TP4, "I regularly clear my browser's history" with a very low loading value of 0.39 indicated that its communality value would be only 0.15, and thus should be avoided in the model. Although we used the validated constructs, our exploratory analysis detected that the item BP-TP4 had a weak influence on the Behavioral Privacy construct.

For the Convergent Validity of each construct measure, we calculated the Average Variance Extracted (AVE) and Composite Reliability (CR) from the factor loading. AVE for each construct should exceed the recommended level of 0.50 so that over 50% of the variances observed in the items were accounted for by the hypothesized constructs, and CR should also be above 0.75 to publish results (Hair et al., 2014). In our study, the AVE reported in *Table 6.1* exceeds 0.50 for all

the constructs except for Behavioral Privacy-Technical Protection (BP-TP). However, CR for each construct was above 0.75 (acceptable), confirming that it measures the construct validity of the model. Since the BP-TP had the item BP-TP4 of very low factor loading along with an AVE value of 0.469, it suggests that the factor BP-TP did not bring significant variance for the variables (items/questions) to converge into a single construct which means BP-TP items are a less-than-effective measure of the latent construct. We also justify this with the exceptionally low rho_A value for the construct BP-TP.

Table 6.1 shows the calculated rho_A value (Dillon-Goldstein's rho) for checking the internal consistency to justify the reliability of each measure. The rho_A evaluates the within-scale consistency of the responses to the items of the measures of constructs and is a better reliability measure than Cronbach's alpha in SEM (Demo et al., 2012).

Table 6. 1. Constructs reliability and validity

	Pretest			**SCS**			**DSS**		
Construct	**rho_A**	**CR**	**AVE**	**rho_A**	**CR**	**AVE**	**rho_A**	**CR**	**AVE**
Attitudes Towards System	0.967	0.976	0.932	0.94	0.962	0.893	0.944	0.964	0.899
Intention to Use	X	X	X	0.873	0.917	0.787	0.875	0.92	0.794
Perceived Ease of Use	X	X	X	0.923	0.93	0.69	0.928	0.939	0.721
Perceived Usefulness	X	X	X	0.834	0.876	0.541	0.917	0.933	0.7
Atd Privacy or Privacy	0.875	0.909	0.714	0.843	0.894	0.678	0.838	0.884	0.657
Quality of System	X	X	X	0.872	0.9	0.695	0.928	0.942	0.801
Security	0.894	0.934	0.825	0.84	0.903	0.756	0.895	0.935	0.826
Trust	0.942	0.944	0.652	0.919	0.931	0.599	0.948	0.953	0.693
Beh Privacy-General Caution	0.93	0.924	0.753	X	X	X	X	X	X
Beh Privacy-Technical Protection	*0.285*	0.766	*0.469*	X	X	X	X	X	X

In our study, as recommended, rho_A for each construct was greater than 0.70 except for BP-TP which had a 0.28 rho value. Therefore, this also supports our decision to remove the behavioral privacy constructs from the post-tests for both SCS and DSS. We assumed that using the BBS simply does not influence the user's behavioral perception of privacy. So, we were interested to see if there is any significant effect on the attitudinal aspect of privacy.

Table 6. 2. Discriminant validity

Pretest								
Construct	**AP**	**ATS**	**BP-GC**	**BP-TP**	**S**	**T**		
Atd Privacy	**0.845**							
Attitudes Toward BSS	0.58	**0.965**						
Beh Privacy-General Caution	0.465	0.285	**0.868**					
Beh Privacy-Technical Protection	0.068	0.187	0.375	**0.684**				
Security	0.637	0.535	0.303	-0.065	**0.908**			
Trust	0.65	0.762	0.275	0.162	0.728	**0.808**		
Shopping Cart System (SCS)								
	ATS	**ITU**	**PEOU**	**PU**	**P**	**QOS**	**S**	**T**
Attitudes Toward SCS	**0.945**							
Intention to Use	0.61	**0.887**						
Perceived Ease of Use	0.557	0.543	**0.831**					
Perceived Usefulness	0.553	0.691	0.615	**0.736**				
Privacy	0.617	0.587	0.482	0.489	**0.823**			
Quality of System	0.509	0.69	0.508	0.691	0.299	**0.834**		
Security	0.374	0.398	0.233	0.344	0.654	0.282	**0.87**	
Trust	0.677	0.653	0.599	0.56	0.748	0.395	0.61	**0.774**
Data Sharing System (DCS)								
	ATS	**ITU**	**PEOU**	**PU**	**P**	**QOS**	**S**	**T**
Attitude Towards DSS	**0.948**							
Intention to use	0.766	**0.891**						
Perceived Ease of Use	0.661	0.573	**0.849**					
Perceived Usefulness	0.672	0.725	0.782	**0.837**				
Privacy	0.596	0.496	0.666	0.656	**0.81**			
Quality of System	0.688	0.631	0.685	0.762	0.708	**0.895**		
Security	0.563	0.447	0.592	0.68	0.824	0.7	**0.909**	
Trust	0.705	0.557	0.718	0.713	0.8	0.775	0.777	**0.832**

To assess the Discriminant Validity of measures, we calculated the square root of the AVE (along the diagonals) of each construct as shown in *Table 6.2*. To lean towards discriminant validity, (Fornell & Larcker, 1981) recommended having low correlations between the measure of interest and the measures of other constructs. In our model, we observed those diagonal values for each construct exceeded other corresponding values, which are the intercorrelations of the given

construct with the other remaining constructs. This pointed out that the measures of each construct, which were theoretically supposed to not be overlapping with measures of other variables, are, in fact, unrelated in our model.

Partial Least Square Path Modeling

To begin our Structural Equation Modeling (SEM) analysis, we built the models for the general population in the context of the pre-test (prototype model) and two subsystems SCS and DSC. We characterized the models by looking into coefficients of determination (R^2's), path coefficients (β's) and corresponding P-value. R^2 determines the variance of a given construct explained by antecedents, β captures the strength of the relationship between the selected constructs and P-value determines the statistical significance of the models (Shrestha & Vassileva, 2019b). According to Chin's guideline (Chin et al., 2003), a path coefficient should be equal to or greater than 0.2 to be considered relevant. A model is statistically somewhat significant (*p) when p-value < 0.1, statistically quite significant (**p) when p-value < 0.01 and statistically highly significant (***p) when p-value < 0.001.

6.1.8 Validation of Hypotheses

For the pre-test in the context of the prototype model, the model presented in Figure 6.3 shows a causal relationship between perceived attitudinal privacy, behavioral privacy-technical protection, behavioral privacy-general caution, perceived security, trust and attitude towards BBS. *Table 6.3* shows the standardized path coefficient (β), t-statistics, p-value and R^2 across selected constructs for the pre-test. The indirect and total effects of one construct over another construct in the presence of mediating constructs were also computed. Considering the direct effects, attitudinal privacy (privacy concern) had very high significant effects on security ($\beta = 0.64$; $P < 0.001$) and trust ($\beta = 0.313$; $P < 0.001$), but an insignificant effect on attitudes towards the system ($\beta = 0.176$; $P > 0.05$). In addition, attitudinal privacy also positively affected behavioral privacy-general caution ($\beta = 0.465$; $P < 0.001$) but had an insignificant effect on behavioral privacy-technical protection ($\beta = 0.068$; $P > 0.1$). The effect of security on trust was also highly significant ($\beta = 0.529$; $P < 0.001$), but insignificant on attitudes towards BBS ($\beta = -0.104$; $P > 0.1$). Finally, trust had a high significant positive effect on attitudes towards BBS ($\beta = 0.724$; $P < 0.001$). Thus, hypotheses H7, H8, H10,

H11 and H13 were supported, but H9, H12, and H14 were rejected in the context of the pre-test. Moreover, trust, privacy and security explain 59.8% of variance in attitudes towards BBS (R^2 = 0.598), security and privacy explain 58.8% of variance in trust (R^2 = 0.588), privacy explains 40.6% of variance in security (R^2 = 0.406), whereas attitudinal privacy explains very low, 21.6%, of variance on behavioral privacy-general caution (R^2 = 0.216) and 0.5% on behavioral privacy-technical protection. R^2 value higher than 0.26 indicates a substantial model (Muller & Cohen, 1989).

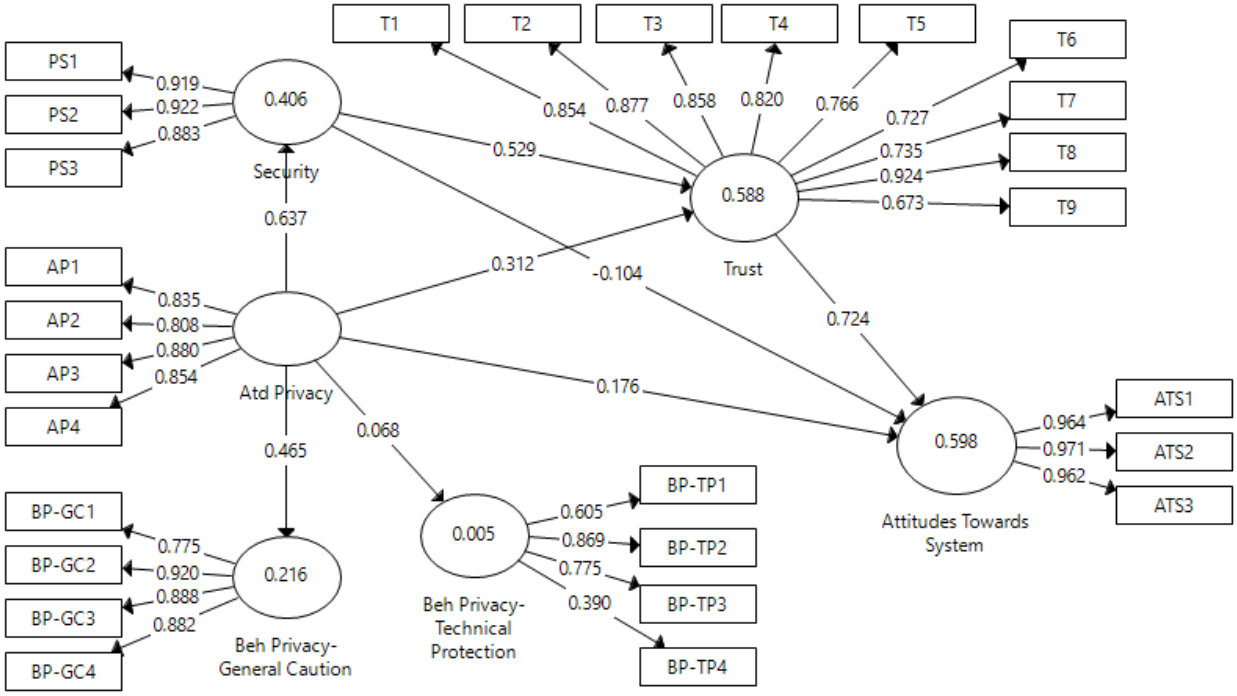

Figure 6. 3. Pretest direct effect

Table 6. 3. Structural estimates (hypotheses testing) for Pre-test.

Direct effect column is when all latent variables are present in the model without any exclusion. R^2 (Attitude=0.598; BP-GC=0.216; BP-TP=0.005; Security=0.406; Trust=0.588)

Structural Path	**Direct effect**			**Total effect**			**Indirect effect**			
	Std β	**T**	**P**	**Std β**	**T**	**P**	**Std β**	**T**	**P**	**VAF**
Atd Privacy → Attitudes Towards System	0.176	1.4	*0.162*	0.586	6.403	0	0.412	4.466	0	0.703
Atd Privacy → Beh Privacy-General Caution	0.465	5.123	0	0.474	5.123	0				
Atd Privacy → Beh Privacy-Tech Protection	0.068	0.366	*0.715*	0.036	0.366	*0.715*				
Atd Privacy → Security	0.64	9.94	0	0.64	9.94	0				
Atd Privacy → Trust	0.313	3.8	0	0.654	10.921	0	0.341	5.301	0	0.521
Security → Attitudes Towards System	-0.104	0.865	*0.387*	0.283	2.396	0.017	0.391	4.348	0	1.382
Security → Trust	0.529	6.446	0	0.529	6.446	0				
Trust → Attitudes Towards System	0.724	6.33	0	0.732	6.33	0				

For the post-test study in the context of SCS, the model presented in Figure 6.4 shows causal relationship between perceived ease of use, perceived usefulness, quality of system, security, privacy, trust, attitude towards SCS and intention to use SCS constructs. *Table 6.4* shows the standardized path coefficient (β), t-statistics, p-value and R^2 across selected constructs for SCS. The indirect and total effects of one construct over another construct in the presence of mediating constructs were also computed. Considering the direct effect, perceived ease of use had quite significant effect on perceived usefulness ($\beta = 0.356$; $P < 0.01$) but insignificant effect on intention to use ($\beta = 0.058$; $P > 0.1$); therefore, H1 was supported and H2 was rejected. Perceived usefulness had relevant but somewhat significant effect on intention to use ($\beta = 0.284$; $P < 0.1$); thus, H3 was also supported. Quality of system had positive significant effect on perceived usefulness ($\beta = 0.509$; $P < 0.001$) and somewhat significant effect on intention to use SCS ($\beta = 0.338$; $P < 0.1$); therefore, H4 and H5 were supported. Attitude towards SCS had relevant but somewhat significant effect on intention to use ($\beta = 0.25$; $P < 0.1$); therefore, H6 was supported. The effect of trust was highly significant on attitude towards SCS ($\beta = 0.534$; $P < 0.001$); therefore, H7 was supported. Perceived privacy had positive significant effects on trust ($\beta = 0.609$; $P < 0.001$), attitudes towards SCS ($\beta =$

0.325; P < 0.01) and perceived security (β = 0.654; P < 0.001); therefore, H8, H9 and H10 were supported. Perceived security had insignificant effect on trust (β = 0.212; P > 0.1) and attitudes towards SCS (β = -0.165; P > 0.1); therefore, H13 and H14 were rejected. In the following, the explained variances include perceived usefulness (R^2 = 0.571), security (R^2 = 0.428), trust (R^2 = 0.585), attitude towards SCS (R^2 = 0.5) and intention to use (R^2 = 0.612). Therefore, R^2 value higher than 0.26 indicated a substantial model for SCS (Muller & Cohen, 1989).

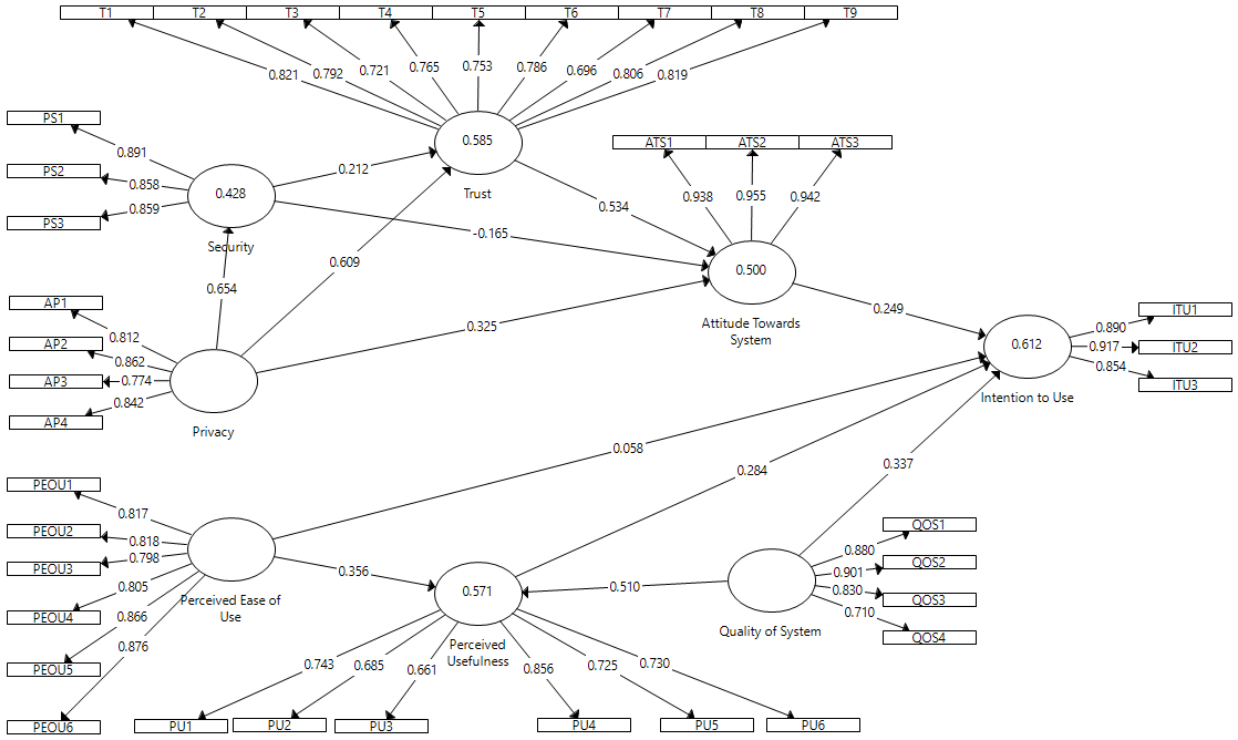

Figure 6. 4. Shopping Cart System (SCS) direct effect

Table 6. 4. Structural estimates (hypotheses testing) for SCS

Direct effect column is when all latent variables are present in the model without any exclusion. R^2 (Attitude= 0.5; Intention=0.612; Perceived Usefulness=0.571; Security=0.428; Trust=0.585)

Structural Path	**Direct effect**			**Total effect**			**Indirect effect**			
	Std β	**T**	**P**	**Std β**	**T**	**P**	**Std β**	**T**	**P**	**VAF**
Attitude Towards System → Intention to Use	0.25	1.713	0.087	0.25	1.713	0.087				
Perceived Ease of Use → Intention to Use	0.058	0.449	*0.653*	0.153	1.201	*0.23*	0.108	1.363	*0.173*	*0.706*
Perceived Ease of Use → Perceived Usefulness	0.356	2.902	0.004	0.362	2.902	0.004				
Perceived Usefulness → Intention to Use	0.284	1.743	0.081	0.297	1.743	0.081				
Privacy → Attitude Towards SCS	0.325	2.828	0.005	0.62	7.626	0	0.293	2.23	0.026	0.473
Privacy → Security	0.654	8.333	0	0.664	8.333	0				
Privacy → Trust	0.597	4.58	0	0.752	11.855	0	0.155	1.225	*0.221*	*0.206*
Quality of System →Intention to Use	0.338	2.116	0.034	0.49	4.116	0	0.152	1.557	*0.12*	*0.31*
Quality of System → Perceived Usefulness	0.509	4.731	0	0.509	4.731	0				
Security → Attitude Towards SCS	-0.165	1.536	*0.125*	-0.049	0.349	*0.727*	0.122	1.191	*0.234*	*-2.489*
Security → Trust	0.212	1.339	*0.181*	0.229	1.339	*0.181*				
Trust → Attitude Towards SCS	0.534	3.349	0.001	0.538	3.349	0.001				

Similarly, for the post-test study in the context of DSS, the model presented in Figure 6.5 shows causal relationship between perceived ease of use, perceived usefulness, quality of system, security, privacy, trust, attitude towards DSS and intention to use DSS constructs. *Table 6.5* shows the standardized path coefficient (β), t-statistics, p-value and R^2 across selected constructs for DSS. The indirect and total effects of one construct over another construct in the presence of mediating constructs were also computed. Considering the direct effect, perceived ease of use had significant effect on perceived usefulness ($\beta = 0.488$; $P < 0.001$) but insignificant effect on intention to use ($\beta = -0.173$; $P > 0.1$); therefore, H1 was supported and H2 was rejected. Perceived usefulness had

relevant but somewhat significant effect on intention to use (β = 0.495; P < 0.1); thus, H3 was also supported. Quality of system had positive significant effect on perceived usefulness (β = 0.427; P < 0.01), but insignificant effect on intention to use DSS (β = -0.009; P > 0.1); therefore, H4 was supported and H5 was rejected. Attitude towards DSS had relevant and positive significant effect on intention to use (β = 0.554; P < 0.001); therefore, H6 was supported. The effect of trust was highly significant on attitude towards DSS (β = 0.637; P < 0.001); therefore, H7 was supported. Perceived privacy had positive significant effects on trust (β = 0.495; P < 0.001) and perceived security (β = 0.82; P < 0.001), but insignificant effect on attitudes towards DSS (β = 0.097; P > 0.1); therefore, H8 and H10 were supported but H9 was rejected. Perceived security had significant effect on trust (β = 0.369; P < 0.01) but insignificant effect on attitudes towards DSS (β = -0.012; P > 0.1); therefore, H13 was supported but H14 was rejected. In the following, the explained variances include perceived usefulness (R^2 = 0.708), security (R^2 = 0.679), trust (R^2 = 0.683), attitude towards SCS (R^2 = 0.5) and intention to use (R^2 = 0.678). Therefore, R^2 value higher than 0.26 indicated a substantial model for DSS (Muller & Cohen, 1989). *Table 6.6* summarizes the validation of our study's hypotheses.

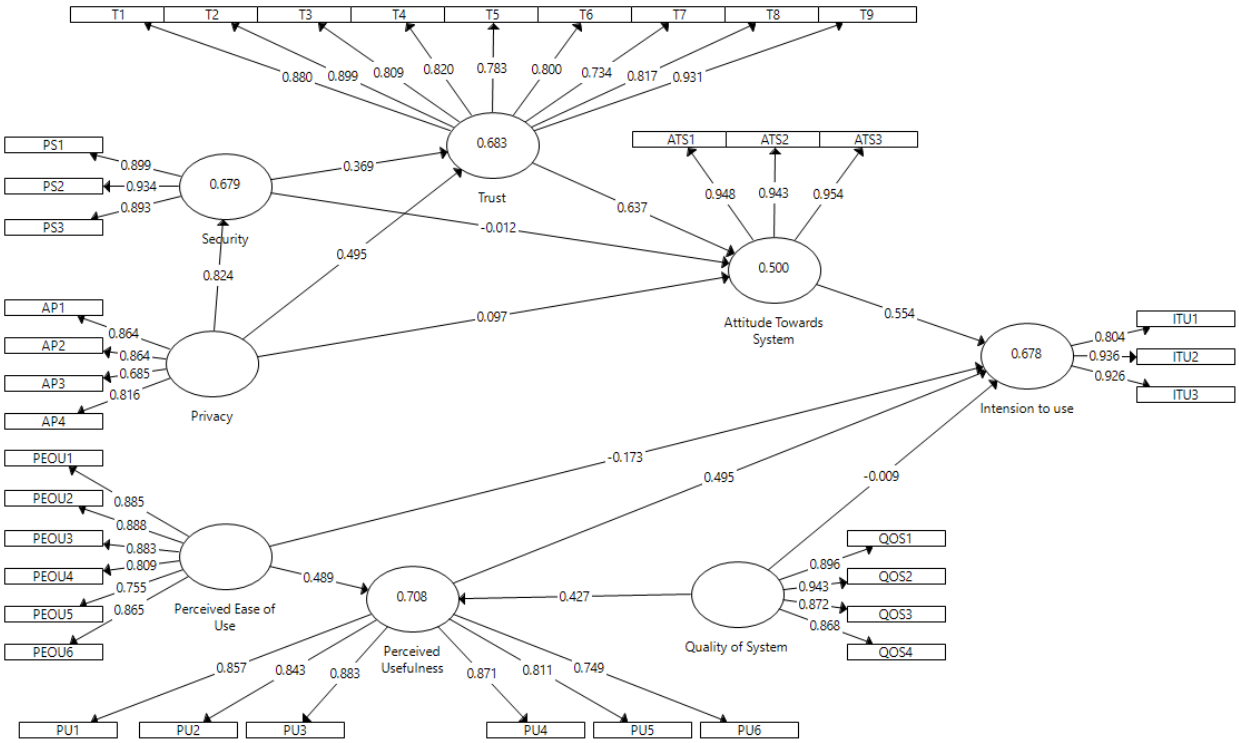

Figure 6. 5. Data Sharing System (DSS) direct effect

Table 6. 5. Structural estimates (hypotheses testing) for DSS

Direct effect column is when all latent variables are present in the model without any exclusion R^2 (Attitude=0.5; Intention to use=0.678; Perceived Usefulness=0.708; Security=0.679; Trust=0.683)

Structural Path	**Direct effect**			**Total effect**			**Indirect effect**			
	Std β	**T**	**P**	**Std β**	**T**	**P**	**Std β**	**T**	**P**	**VAF**
Attitude Towards System → Intention to Use	0.554	3.967	0	0.554	3.967	0				
Perceived Ease of Use → Intention to Use	-0.173	0.979	*0.327*	0.069	0.5	*0.617*	0.242	1.946	0.052	3.507
Perceived Ease of Use → Perceived Usefulness	0.488	3.456	0.001	0.489	3.456	0.001				
Perceived Usefulness → Intention to Use	0.495	2.354	0.019	0.495	2.354	0.019				
Privacy → Attitude Towards System	0.097	0.383	*0.701*	0.596	5.531	0	0.5	2.319	0.02	0.839
Privacy → Security	0.82	15.161	0	0.824	15.161	0				
Privacy → Trust	0.495	3.54	0	0.8	14.958	0	0.304	2.59	0.01	0.38
Quality of System →Intention to Use	-0.009	0.046	*0.964*	0.202	1.063	*0.288*	0.212	1.846	0.065	1.05
Quality of System → Perceived Usefulness	0.427	2.667	0.008	0.427	2.667	0.008				
Security → Attitude Towards System	-0.012	0.053	*0.958*	0.223	0.998	*0.318*	0.235	2.113	*0.035*	*1.05*
Security → Trust	0.369	2.566	0.01	0.369	2.566	0.01				
Trust → Attitude Towards System	0.637	4.316	0	0.637	4.316	0				

Table 6. 6. Validation of the study's hypotheses

√ = True; × = False; √→ = Mediated by latent variable; ×→Not mediated by latent variable

	Hypothesis	**Pre-test**	**SCS**	**DSS**
H1	Perceived ease of use significantly influences perceived usefulness of BBS.		√	√
H2	Perceived ease of use significantly influences intention to use BBS.		× (×→Usefulness)	× (√→Usefulness)
H3	Perceived usefulness significantly influences intention to use BBS.		√	√
H4	Quality of system significantly influences perceived usefulness of BBS.		√	√
H5	Quality of system significantly influences intention to use BBS.		√ (×→Usefulness)	× (√→Usefulness)
H6	Attitude towards BBS significantly influences intention to use BBS.		√	√
H7	Trust positively affects users' attitudes toward BBS.	√	√	√
H8	Perceived privacy has a positive effect on the users' trust in BBS.	√	√	√
H9	Perceived privacy has a positive effect on the users' attitudes toward BBS.	× (√→Trust) (×→Security)	√ (√→Trust) (×→Security)	× (√→Trust) (×→Security)
H10	Perceived privacy positively or negatively affects users' perceived security.	√	√	√
H11	Privacy concern positively affects users' behavior on general caution.	√		
H12	Privacy concern positively affects users' behavior on technical protection.	×		
H13	Perceived security positively affects users' trust in BBS.	√	×	√
H14	Perceived security positively affects users' attitudes toward BBS.	× (√→Trust)	× (×→Trust)	× (×→Trust)

6.1.9 Total Effect Analysis

To address the first research question, RQ1 (Which of the design attributes is/are the strongest antecedents of the attitudes towards BBS?), we presented the total effect of antecedents from the trust model on attitudes towards BBS as shown in Figure 6.6. In the pre-test model, trust had the strongest total effect on attitudes towards BBS ($\beta = 0.732$; $P < 0.001$), followed by privacy on attitudes towards BBS ($\beta = 0.586$; $P < 0.001$) and security on attitudes towards BBS ($\beta = 0.283$; $P < 0.1$), which was marginally significant. In the SCS model, privacy had the strongest influence on attitudes towards SCS ($\beta = 0.62$; $P < 0.001$), followed by trust on attitudes towards SCS ($\beta = 0.538$; $P < 0.001$), while security had no significant total effect on attitudes towards SCS ($\beta = -0.049$; $P > 0.1$). Finally, the total effect statistic for the DSS model was similar to that of the pre-test model, with respect to first two strongest design constructs, which were trust ($\beta = 0.637$; $P < 0.001$), followed by privacy ($\beta = 0.596$; $P < 0.001$). Security turned out to have no significant effect on attitude towards DSS ($\beta = 0.223$; $P > 0.1$).

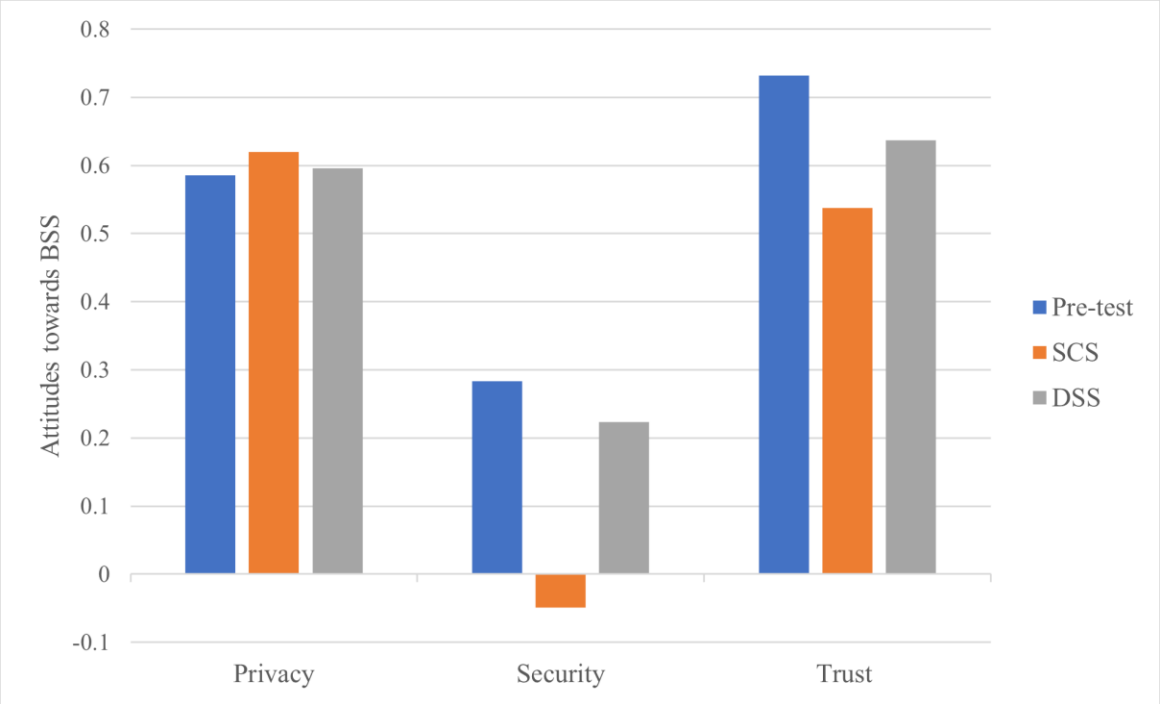

Figure 6. 6. Total effect of the trust design constructs on attitudes towards BBS

To address the second research question, RQ2 (Which of the design attributes is/are the strongest antecedents of the intention to use BBS?), we present the total effect of the perceived design

constructs on intention to use from SCS and DSS model as shown in Figure 6.7. In SCS model, quality of system had the strongest total effect on intention to use SCS ($\beta = 0.49$; $P < 0.001$). Perceived usefulness had a weak total effect on intention to use SCS ($\beta = 0.297$; $P < 0.1$), while privacy had no significant total effect on intention to use SCS ($\beta = 0.156$; $P > 0.1$), followed by perceived ease of use on intention to use SCS ($\beta = 0.153$; $P > 0.1$) and security on intention to use SCS ($\beta = -0.01$; $P > 0.1$). In the context of DSS, Perceived usefulness had the strongest total effect on intention to use DSS ($\beta = 0.495$; $P < 0.01$), followed by trust on intention to use DSS ($\beta = 0.353$; $P < 0.001$) and privacy on intention to use DSS ($\beta = 0.33$; $P < 0.001$). Quality of system had no significant total effect on intention to use DSS ($\beta = 0.202$; $P > 0.1$), followed by security on intention to use DSS ($\beta = 0.124$; $P > 0.1$) and perceived ease of use on intention to use DSS ($\beta = 0.069$; $P > 0.1$).

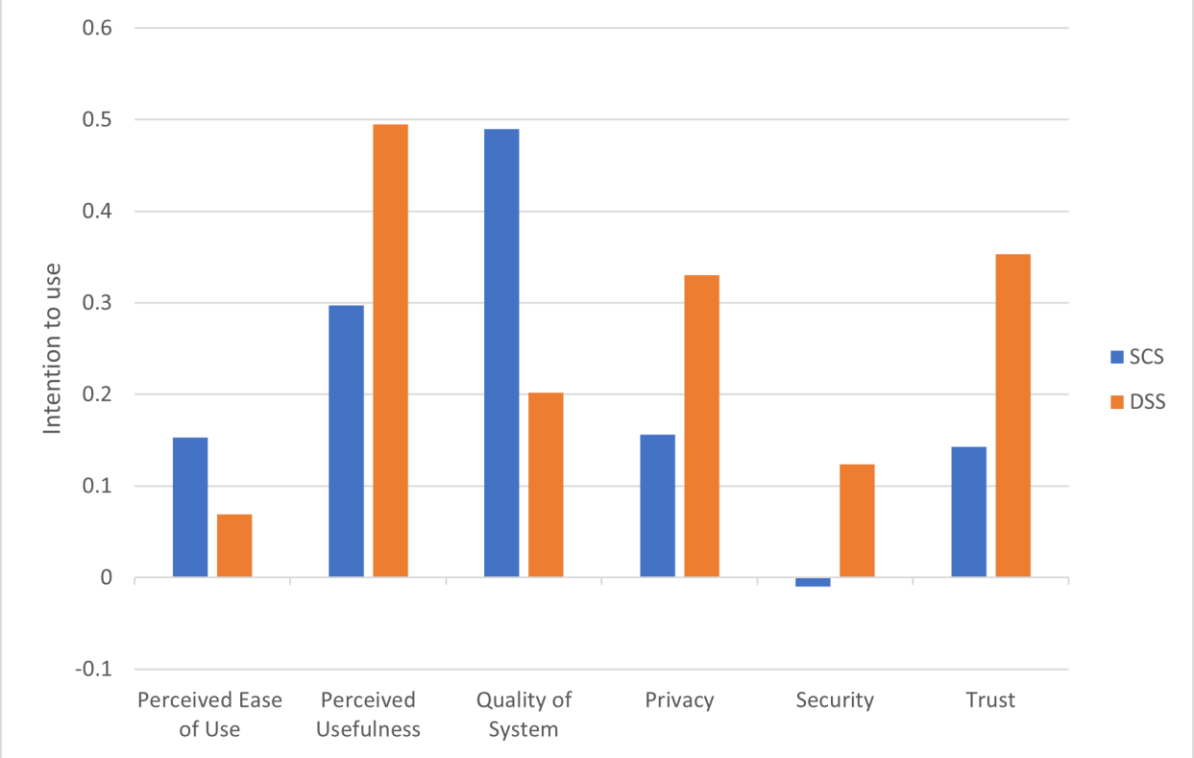

Figure 6. 7. Total effect of predictors on intention to use

6.1.10 Mediation Analysis

To address our third, fourth and fifth research questions (RQ3: Is the influence of privacy on attitudes towards BBS mediated by both security and/or trust?, RQ4: Is the influence of security

on attitudes towards BBS mediated by trust?, RQ5: Is the influence of ease of use/quality of system on intention to use BBS mediated by perceived usefulness?), we carried out the indirect effect analysis. We first investigated the mediating effect of security and/or trust over the relationship between privacy and attitudes towards BBS, then investigated the mediating effect of trust over the relationship between security and attitudes towards BBS, and finally investigated a similar mediating effect of perceived usefulness over the relationship between ease of use/quality of system on intention to use BBS. According to (Baron & Kenny, 1986; Hair et al., 2014), there is no need to check for the indirect effect if the direct effect is insignificant in the model.

So, for the pre-test model as presented in *Table 6.3*, we found the observed indirect effects for the selected predictors in the presence of mediating variables in the pre-test model. In the presence of mediating effect of both trust and security, the effect of privacy on attitude towards BBS slightly decreased from ($\beta = 0.584$; $T = 6.868$; $P < 0.001$ while excluding both trust and security) to ($\beta = 0.412$; $T = 4.466$; $P < 0.001$) with the *variance accounted for* (VAF) value of 0.703. The VAF is calculated as the ratio of the indirect path coefficient to the total path coefficient. With 70.3% VAF, trust and security had a partial mediation effect between privacy and attitude towards BBS. While analyzing the individual mediating effects between privacy and attitudes towards BBS, trust alone had positive significant effect ($\beta = 0.226$; $T = 3.151$; $P < 0.01$), but security alone had no significant effect ($\beta = -0.068$; $T = 0.852$; $P > 0.1$). So, our finding suggested that only trust played a crucial mediating role while security had no significant effect between privacy and attitudes toward BBS. Similarly, in the presence of mediating effect of trust, the effect of security on attitude towards BBS slightly decreased from ($\beta = 0.538$; $T = 6.14$; $P < 0.001$ while excluding Trust) to ($\beta = 0.391$; $T = 4.348$; $P < 0.001$) with the *variance accounted for* (VAF) value of 1.382. With 138% VAF, trust had a perfect mediation effect between security and attitude towards BBS.

For the SCS model, as presented in *Table 6.4*, in the presence of mediating effect of trust and security, the effect of privacy on attitude towards SCS slightly decreased from ($\beta = 0.619$; $T = 8.504$; $P < 0.001$ while excluding both trust and security) to ($\beta = 0.293$; $T = 2.23$; $P < 0.1$) with the *variance accounted for* (VAF) value of 0.473. With 47.3% VAF, trust and security had a partial mediation effect between privacy and attitude towards SCS. While analyzing the individual mediating effects between privacy and attitudes towards SCS, trust alone had positive significant effect ($\beta = 0.326$; $T = 2.683$; $P < 0.01$), but security alone had no significant effect ($\beta = -0.108$; $T = 1.466$; $P > 0.1$). So, our finding suggested that only trust played a crucial mediating role while

security had no significant effect between privacy and attitudes toward SCS. In the same SCS model, with the presence of mediating effect of trust, the effect of security on attitude towards SCS became insignificant from ($\beta = 0.384$; $T = 3.304$; $P < 0.001$ while excluding trust) to ($\beta = 0.122$; $T = 1.191$; $P > 0.1$) with the variance accounted for (VAF) value of -2.489. Therefore, trust had no mediating effect between security and attitude towards SCS since after adding trust predictor as a mediator, the indirect effect on attitude towards SCS became non-significant while the direct effect was also insignificant. Furthermore, in the same SCS model, with the presence of mediating effect of perceived usefulness, the effect of quality of system on intention to use SCS became insignificant from ($\beta = 0.705$; $T = 11.127$; $P < 0.001$ while excluding usefulness) to ($\beta = 0.152$; $T = 1.557$; $P > 0.1$) with the variance accounted for (VAF) value of 0.31. Therefore, perceived usefulness had no significant mediating effect between quality of system and intention to use SCS since after adding perceived usefulness predictor as a mediator, the indirect effect on intention to use SCS became non-significant while the direct effect was still significant.

In addition, with the presence of mediating effect of perceived usefulness, the ease of use on intention to use SCS became insignificant from ($\beta = 0.559$; $T = 7.035$; $P < 0.1$ while excluding usefulness) to ($\beta = 0.108$; $T = 1.363$; $P > 0.1$) with the variance accounted for (VAF) value of 0.706. Therefore, perceived usefulness had no significant mediating effect between ease of use and intention to use SCS since after adding perceived usefulness predictor as a mediator, the indirect effect on intention to use SCS became non-significant while the direct effect was also non-significant.

Finally, in the DSS model, as presented in *Table 6.5*, in the presence of mediating effect of trust and security, the effect of privacy on attitude towards DSS slightly decreased from ($\beta = 0.6$; $T = 6.134$; $P < 0.001$ while excluding both trust and security) to ($\beta = 0.5$; $T = 2.319$; $P < 0.1$) with the variance accounted for (VAF) value of 0.839. With 83.9% VAF, trust and security had a partial mediation effect between privacy and attitude towards DSS. While analyzing the individual mediating effects between privacy and attitudes towards DSS, trust alone had positive significant effect ($\beta = 0.316$; $T = 2.726$; $P < 0.01$), but security alone had no significant effect ($\beta = -0.01$; $T = 0.053$; $P > 0.1$). So, our finding suggested that only trust played a crucial mediating role while security had no significant effect between privacy and attitudes toward DSS. In the same DSS model, no mediation effect was observed for trust between security and attitudes towards DSS. Furthermore, in the same DSS model, with the presence of mediating effect of perceived

usefulness, the effect of quality of system on intention to use DSS reduced from (β = 0.66; T = 7.82; P < 0.001 while excluding usefulness) to (β = 0.212; T = 1.846; P < 0.1) with the variance accounted for (VAF) value of 1.05. Therefore, perceived usefulness had a complete significant mediating effect between quality of system and intention to use DSD since after adding perceived usefulness predictor as a mediator, the indirect effect on intention to use DSD became significant while the direct effect was insignificant. This suggested that the indirect significant path between quality of system and intention to use DSD was contributed by perceived usefulness predictor construct. In addition, with the presence of mediating effect of perceived usefulness, the effect of ease of use on intention to use reduced from (β = 0.588; T = 6.385; P < 0.001 while excluding usefulness) to (β = 0.242; T = 1.946; P < 0.1) with the variance accounted for (VAF) value of 3.507. Therefore, perceived usefulness had a complete mediating effect between ease of use and intention to use DSD since after adding perceived usefulness predictor as a mediator, the indirect effect on intention to use DSD became significant while the direct effect was non-significant. This suggested that the indirect significant path between perceived ease of use and intention to use DSD was contributed by perceived usefulness predictor construct.

6.1.11 Discussion

This section reviews and evaluates the augmented TAM with trust model with respect to the initial adoption of the blockchain-based system (BBS), employing the DUDS framework. In this study, there was a customer-specific system: SCS, and a company-specific system: DSS.

6.1.11.1 Effectiveness of TAM with Trust Model

The purpose of this study was to evaluate the user acceptance of a working blockchain-based system (BBS) by observing the attributes affecting the development of users' attitudes and intention to use the system. We achieved the goal of our research by testing the augmented TAM with a Trust model on our application (BBS) that is built using blockchain technology. The empirical study validates our research model and supports most of the research hypotheses that were set considering the aim of this study. We also identified different issues influencing users' attitudes and intentions to adopt BBS by considering observed facts from the causal relationships and their implications. According to (Gefen et al., 2003), extending TAM with trust model is well

justified for its effectiveness in improving the predictive power of the explored issues associated with the acceptance of online services. BBS can be considered as a set of online services, so applying the TAM augmented with trust, as we did in our study, is justified. The major contribution of our study to the existing literature of blockchain and distributed ledger technologies is to uncover the dimensions and role of trust alongside primary TAM-based design predictors and their causal relationship with users' attitudes and behavioral intention to accept such technologies.

Moreover, the empirical results of our study also confirm a significant positive effect of the users' attitudes on their intention to use the BBS and suggest that the most important antecedent of attitudes towards using BBS is trust, which is also supported by the previous studies, confirming that the trust predictor significantly influences the user's decision to adopt the online services. Therefore, familiarity with the significance of the underlying blockchain technology and the honesty of the companies to keep their promises of protecting privacy, securing information, and incentivizing customers for sharing their data bring a higher level of trust and stimulate positive attitudes of customers towards using the SCS. Similarly, trusting the blockchain technology for its integrity and dependability significantly improves the company's attitudes towards adopting the DSS.

6.1.11.2 Multidimensionality of Privacy

A previous study by Buchanan et al. (2007) suggested that attitudinal privacy, in the privacy model, correlates significantly with behavioral privacy-general caution but not significantly with the technical protection factor. The findings of our current research indicate comparable results. Users who are concerned with their data privacy tend to be more cautious and careful about protecting it; however, if the users are technically competent, they have already used tools to protect their privacy such as clearing the browser's cache and history, using spyware etc., so they become less concerned about their privacy infringement.

6.1.11.3 Effect of Usability on Intention to Use

Based on our research findings, usability (or perceived ease of use) does not impact behavioral intention to use the actual BBS unlike in our previous study on the blockchain-based prototype model, where ease of use was significant in the initial stage (Shrestha & Vassileva, 2019b). This

result is because users perceive BBS, a user-friendly web application, as easier to learn and operate. Based on representative literature such as (I.-F. Liu et al., 2010), UI design is the most significant item that affects perceived ease of use. Users, instead of being more concerned about learning to use the system, are concerned about the usefulness and overall performance of the BBS. Previous studies by Venkatesh et al. (2003), Chan and Lu (2004), Pikkarainen et al. (2004) and Roca et al. (2009) confirm that usability remains non-significant to develop an intention to use the system.

6.1.11.4 Usability and Quality of System to Predict Usefulness

According to the results of our study, we deduce perceived ease of use and quality of BBS as significant predictors of the usefulness construct. When users find BBS easier to use and believe they can be skillful in using it, they will consider the system as more useful to improve their performance and productivity. This is also confirmed by previous studies (Gefen et al., 2003; I.-F. Liu et al., 2010). In our system, SCS allows customers to set their data sharing preferences and receive incentives for sharing their data as per the smart contracts, while DSS guarantees companies that the customer data they access have integrity and confirm provenance. So, the users of each system, who feel more satisfied with these features, develop a higher understanding of its perceived usefulness. Eventually, with positive feelings about the usefulness of the BBS, users develop a stronger behavioral intention to accept the system. Since the quality of the system has an insignificant direct effect on the intention to use the system for DSS, its effect through perceived usefulness is found out to be a significant positive effect in our study, which is in agreement with the suggestions made by (DeLone & McLean, 1992).

6.1.11.5 Direct Effect of Perceived Security and Perceived Privacy

According to Shin (2010), trust has a moderating effect on perceived security and perceived privacy when it comes to adopting social networking sites. Perceived security has a mediating effect on perceived privacy that correlates to trust (Rios et al., 2017). The findings from our study suggest that perceived security has a direct effect on trust in the context of the prototype model and DSS. Outside of this, there is no significant relationship between security and other constructs. Perceived privacy has a direct effect on user trust and perceived security, which reinforces the findings by Rios et al. (2017) that claims perceived security and perceived privacy are related. Based on our findings,

the direct effect of perceived privacy on users' attitudes towards BBS is only significant for SCS and is moderated by trust in all pre-tests, SCS and DSS models.

6.1.11.6 Context Dependent Effects of Augmented TAM-Based Predictors

Every participant who completed the post-survey for DSS also completed the post-survey for SCS with 66 participants completing the post-survey for SCS and 53 participants completing the post-survey for DSS. There was no major difference between the users of each system that could lead to the difference between the responses of the two surveys. With the SCS, the user engaged with the point of view of a customer, whereas the DSS had the user engage the system as an enterprise's system administrator. According to the results from our research, TAM-based predictors or trust constructs cannot be applied uniformly to BBS. Depending on the specifics of the BBS, the relationships between perceived trust, perceived security, perceived privacy, and attitudes towards the system might change.

Our findings suggest that the influence of perceived privacy and perceived security depends strongly on which blockchain-based system users interact with. When answering the initial pre-test survey, participants have no system to base their ideas on. So, security, privacy, trust, and BBS become abstract concepts. As abstract concepts, participants believe privacy affects security, security and privacy affect trust, and trust affects their intention to use the system. However, they are not aware of any direct effect of privacy and security on their choice to use the system.

In our study, we see that after using the SCS, there is a significant effect of perceived privacy on the user's attitude towards BBS. Yet, the pre-test and DSS survey results show that participants feel perceived privacy does not positively affect their attitudes towards BBS. Perceived privacy's effect on user's attitudes towards BBS is only significant with a customer-specific BBS like SCS but not significant with a company-specific BBS like DSS. However, trust has either a partial or complete mediating role in all kinds of BBS, which is consistent with prior research.

Based on the initial pre-test survey results, we deduce that participants feel security protection mechanisms are an important indicator to trust the system. This initial result might be affected by external factors such as the media, since ransomware and other ICT security breaches have been a big deal in the media recently. However, after using the SCS, we learn that perceived security tends to be an insignificant predictor of trust. For the DSS, the effect of perceived security on trust is once again significant. It may be because after experiencing the real-life blockchain-based system,

respondents using the SCS become aware of the underlying security infrastructure of blockchain and smart contracts, but once they learn that the business process models deployed via smart contracts are committed on a public blockchain, they may focus more on privacy and think less about underlying security. As they are not concerned about security, they want to have control over their data instead, so the relative significance of perceived privacy to trust SCS for these users is higher. On the other hand, respondents experiencing DSS to access customer data may not care much about privacy since they are already putting their information through transparent processes for customers and other enterprises. Instead, they may care more about secure transactions, mitigating anomalies and malicious behavior in their consortium network and cyber-resilient smart contracts. Therefore, perceived security may significantly affect trust in an abstract context, but with a specific context, it may be insignificant for a model like SCS but may remain significant for a model like DSS.

6.1.11.7 Effects of Trust and Its Predictors on Prototype System and Concrete System

Prior research on the effect of perceived security and perceived privacy on user trust are mixed. Shin (2019) found a significant moderating effect of security on trust, but participants had no real interaction with an actual system. Studies on non-blockchain online services had comparable results. McCole et al. (2010) found that perceived privacy and perceived security moderates the effect of trust. Eastlick et al. (2006) empirically showed that the relationship between privacy concerns and trust was the third strongest of all relationships studied. Chellappa and Pavlou (2002) argued that perceived security is a stronger predictor of trust. All four of these studies were abstract and did not have participants engage with a real system before answering their survey. These results support our initial pre-test results. Without interacting with any system, participants often consider privacy, security, and trust to be strongly related.

In previous studies, where participants engaged with online services such as online shopping, perceived security had a stronger effect. Both Belanger et al. (2002) and D. J. Kim et al. (2008) found that perceived security had a stronger effect than perceived privacy on consumer behavior. Roca et al. (2009) found that perceived privacy did not influence trust, but they did not consider the influence of security factors moderating privacy concerns in their model based on extended TAM. These findings do not align with ours from when participants used the SCS. Our study found

security has no significant relationship to trust in real SCS, while privacy significantly affects trust and attitudes towards both BBS.

The discrepancy between results from abstract studies and studies with concrete systems shows how important it is to focus on the latter. Although the abstract studies show a strong relationship between trust, privacy, and security, studies with actual eCommerce systems have mixed and inconclusive results. Furthermore, studies on eCommerce systems focus on the customer. Few relevant studies focus on the company's trust and its intention to use the technology. Therefore, we cannot find other results to compare to the current study's finding that for DSS, perceived security positively affects trust in BBS, and trust completely mediates the influence of privacy on attitudes towards adopting the BBS. Also, based on our pre-test and post-test results, there is no mediating effect of security over the perceived privacy on the users' attitudes towards BBS. Further study is needed with specific types of BBS to see if there are more BBS types other than customer-specific and company-specific and to better understand which trust construct is significant for each type of system.

6.1.11.8 Methodological Contributions

Our study also brings a methodological contribution to the literature with the use of partial least square structural equation modeling (PLS-SEM) to analyze the user acceptance of the concrete blockchain-based application. PLS is component-based and can model the latent constructs under conditions for smaller sample sizes by maximizing the explained variance of dependent indicators and using multiple regressions to observe the effect of predictors on the response variables (Chin et al., 2008; Hair et al., 2013). Furthermore, this study contributes to the methodology by adopting Dillon-Goldstein's rho for estimating internal consistency reliability, which is suggested as an always better choice than conservative Cronbach's alpha in the presence of skew items and smaller samples (Demo et al., 2012).

6.1.12 Limitations

The main limitation of our study is that our findings are based on a relatively small, targeted population size and only on two specific types of BBS. Therefore, the results may not generalize to the broader population and to any type of BBS. Further study may consider using a larger sample

with specific types of BBS, to explore BBS types other than customer-specific and company-specific, and to better understand which trust antecedent is significant for each type of system. Another limitation is not tracking the users for a longer time while they perform authentic activities on BBS, making important decisions about things they care about, with genuine worries about the privacy of their information. Moreover, an obvious limitation comes from using the same participants for both systems. Most respondents who participated in the DSS study also completed the SCS study, since they also satisfied the inclusion criteria of DSS while doing the SCS study. The DSS study had participants only with a technology background. By taking on separate roles, these participants may have experienced different motivations that skewed their survey results, so a further study is needed to draw any conclusions about the role users take and what factors influence their desire to use the specific BBS. Another limitation is the demographics of the participants: almost 79% of the participants for SCS had a basic knowledge of blockchain technology, while only 19% had advanced knowledge, and some of the participants belonged to academia. To address this, we need to consider an underlying effect of participants' background on their behavioral intention to use BBS. Therefore, this study offers an opportunity for future exploration of BBS from conducting longitudinal studies to considering multigroup analysis based on participants' demography and background knowledge when analyzing the endogenous and exogenous variables, which will further explain the user acceptance of the BBS.

6.1.13 Conclusion

In this chapter, I presented the augmented TAM with trust model on the real-life blockchain-based system (BBS), which comprises two subsystems: Shopping Cart System (SCS) and Data Sharing System (DSS). The main contribution of this study to the body of knowledge is that, to the best of our knowledge, this study is the first to examine the augmented TAM with trust model using real-life concrete blockchain-based applications. The empirical study validated the research model and supported most of the research hypotheses that I set based on my research. The findings suggested that TAM-based predictors and trust constructs cannot be applied uniformly to BBS. Depending on the specifics of the BBS, the relationships between perceived trust, perceived security, perceived privacy, and attitudes towards the system might change. In SCS, privacy was the strongest determinant of attitudes towards the system, but in DSS, trust was the strongest determinant of attitudes towards the system. Quality of system had the strongest total effect on intention to use

SCS, while perceived usefulness had the strongest total effect on intention to use DSS. Trust significantly influenced the users' attitudes towards both types of BBS, while security did not have any effect on users' attitudes toward BBS. In SCS, privacy positively affected trust, but security had no significant effect on trust, whereas, in DSS, both privacy and security significantly influenced trust. In both BBS, trust had a moderating effect on privacy that correlated directly with attitudes towards BBS, whereas security had no mediating effect between privacy and attitudes towards BBS.

Hence, we recommend that while implementing and upgrading blockchain-based solutions, the decision-makers should carefully consider the trust patterns and address the associated privacy challenges of the users. Designers and decision-makers for the industries should know that the effect of trust antecedents is context-dependent whether it is customer or company-oriented. For the development of customer-oriented BBS, the effect of a privacy-aware system to influence users' attitudes toward BBS is relevant. For the development of a company-oriented BBS, additional security measures must also be carefully addressed to significantly influence users' trust in BBS, which in turn positively leads to a higher intention to adopt the system. In future work, I plan to investigate multigroup analysis based on participants' background knowledge when analyzing the latent variables and performing the qualitative analysis based upon the respondents' feedback, which will further explain the user acceptance of the BBS.

7 CONCLUSION AND FUTURE WORK

The currently dominant ownership model over user data usually encoded in the service license agreements presumes that data ownership is transferred from the user to the enterprise collecting it and, if shared, to the entire network of businesses. There are privacy and security problems associated with storing user data. Even the most prominent online services have experienced security breaches and data theft. When trust resides within a centralized service provider for all the data storage, it could be affected by centrality issues such as intentionally deleting the user data or not delivering the user data due to a technical failure. This centralization impedes the initiatives of the data privacy and protection legislation.

Sharing user data across applications and enterprises helps to improve the personalization of functionality, interface and options and thus creates a better user experience. In scientific research, sharing research data helps to strengthen research activities, reduce duplicative trials, and checks the research's validity. However, there are problems associated with the security, privacy and user control of sharing user data. Security of data sharing has been addressed by standard security techniques as well as experimental approaches, for example, carrying out all the communication without trusting anybody and possibly replacing the centralized controlling authority. Various advanced technologies have been deployed as computational backbones to collect and share user data, such as cloud computing services, as well as various security technologies to protect the collected user data from hackers. One can also use Google Federated learning to mine data scattered in distributed locations.

In light of addressing these challenges, I adopted the pragmatic philosophy worldviews and convergence strategy by implementing both quantitative and qualitative approaches. I followed Design Science Framework to perform mixed research approaches under relevance, design, and rigour cycles. As a result, this thesis presented a solution—the DUDS framework—to address all the security, privacy, user transparency, and control issues and provide incentives for data sharing, which are the most common challenges of general data governance. The thesis offers user-controlled privacy-preserving incentives-enabled data-sharing policies encoded in smart contracts. The smart contracts are constructed with data-sharing choices provided on the web form, allowing users to select their preferences and execute the associated smart contracts. This process considers

DLT as a service, reducing the need for organizations to invest significantly in technical talent and for users to learn codes, narrowing the skills gap.

This thesis also presents the implementation of the DUDS framework to share user data in a decentralized fashion under different scenarios, including different businesses providing services in the travel-booking domain, e-commerce, and research data sharing domain. I also conducted usability studies of real-life applications based on the DUDS framework. I uncovered different constructs to consider while designing and developing a new platform based on the DUDS framework. The DUDS framework supports creating incentives for users to share their data in terms of rewards (micro-payments or credits). Thus, users become owners of their data and can decide how their data is collected and used and shared. Users benefit not only in terms of improved personalized experience with the service but also directly, for example, by participating in the share of the advertising revenue generated by the service provider.

This chapter concludes the thesis with a summary of the contributions to the knowledge along with research challenges and possible future works.

7.1 Research Contributions

The significant contributions of my research include the following:

1. I performed a review of the literature in sharing user data. I found the need for a platform that could enable incentives (including reputation aggregation) or micro-payments at the scale needed for realizing decentralized data sharing networks (Chapter 2).
2. I developed a standard definition of user data and classified user data to support all kinds of user profiles, user-generated/created data and research data (Chapter 2).
3. I developed the blockchains- and smart contracts-based user data sharing framework for incentivizing the data owners as the solution to the identified problems and named it *DUDS* for the Decentralized User Data Sharing framework (Chapter 3).
4. I collected quantitative and qualitative data to construct the heuristic cognitive-behavioural model based on the extended TAM. As a result, distinct constructs affecting the end users' intention to accept the prototype based on the DUDS framework were identified and reviewed. The analysis further advocated a need for an investigation towards conceptualizing

digital trust with privacy and security elements for blockchain-based systems, including the DUDS framework (Chapter 4).

5. I provided the designs for the user data sharing platforms based on the DUDS framework employing the smart contracts, permissioned MultiChain and Ethereum blockchains (Chapter 5). Three different implementation prototypes with their MVPs were designed for tourism (hotel booking), research data sharing, and e-commerce (online shopping cart) domains that offer user-controlled privacy and rewards to the data owners in terms of payment through blockchains for the use of the data by applications, as specified by the smart contracts. The performance of the DUDS platform was also analyzed through a set of comparative experiments to calculate the latency and memory consumption for the consortium network and gas consumption and transaction cost for the smart contracts' deployment and execution. The test results indicated that the nodes responded quickly in all test cases with befitting transaction costs for given well-defined requirements. This outcome led to the final evaluation of the framework by conducting a final user experience study with the augmented TAM on such real-life MVPs based on the DUDS framework.
6. Finally, I evaluated the user behavioral model and DUDS framework for their usefulness and trustworthiness through the final user study. The ultimate results were used as the basis to postulate guidelines and methods for incorporating security, privacy, user transparency, control and incentives from the start in the design of the data-sharing framework (Chapter 6). The main contribution of this study to the body of knowledge is that, to the best of our knowledge, this study is the first to examine the augmented TAM with trust model using real-life concrete blockchain-based applications.

7.2 Discussion

This section reviews and evaluates the DUDS framework with respect to its technical solutions and usability.

7.2.1 Unlocking The Properties of DLTs

After a decade-long development phase, blockchain as a part of DLTs has demonstrated a high potential for data management applications in many industries, including healthcare, agriculture,

tourism, and research fields. Blockchain technology is likely to disrupt many traditional centralized business models because it is a decentralized, immutable, tamper-proof, and transparent process supporting autonomous, distributed and unalterable smart contracts. This thesis presents an engineering approach and a framework to create a distributed ledger-based solution for sharing user data among stakeholders using blockchain technology and smart contracts. The proposed framework provides users with flexibility in expressing their data sharing preferences and requiring rewards/incentives for sharing, thus enabling users to maintain control over the conditions of access to their data. The framework also provides users with proof of existence and data ownership, ensures their control, and operationalizes incentives over data sharing. Although users of the DUDS-based applications do not have to be technology experts, they still should be familiar with the basics of blockchain and smart contracts execution.

7.2.2 Transparency, Confidentiality and Rewards For Sharing

The DUDS framework proposed in this thesis can be viewed as a tool that provides a transparent mechanism to control and support the collection of required data for the identified purpose, explain where, how, and when the user data was collected, used and shared, offer meaningful fresh consent for every new purpose, and reward the owners of data. Smart contracts manage the distribution of the digital tokens or acknowledgement for data sharing to the participating stakeholders. Although all the transactions are digitally signed in the blockchain network, the system may still suffer a loss of data confidentiality as data is not encrypted by default. DUDS framework uses public-key cryptography to protect the confidentiality of the data while maintaining its integrity.

7.2.3 DUDS Framework in Data Governance Framework

The DUDS framework fits nicely into the existing joint data governance for every organization by addressing the challenges of breaking down data hoarders (Congosto et al., 2017), ensuring data integrity, confidentiality, and availability. Data governance policies depend upon organizational structures and business goals. They are usually defined by the organization itself, by the government (e.g., compliance regulations for privacy, laws for Intellectual Property), by another organization (e.g., creative commons licensing, rules for sharing research data by government funding agencies), or by the users. The DUDS framework works well with the models that allow

users to choose their data sharing preferences. For travel booking systems and e-commerce applications, the DUDS framework can effectively address the challenges regarding data sharing from the existing data protection legislation (e.g., PIPEDA). The university's data-sharing regulations about Intellectual property can be appropriately integrated within the DUDS framework for sharing research data.

7.2.4 Data Access

Since data trust is non-deterministic, one of the possible scenarios for data sharing is the case with the data controller, trustees, and beneficiaries, all to be among the trusted nodes (such as data cooperative, data collaborative)[52] that agree to share the pooled data with all the valid nodes in the network. This scenario is realized with the arrangement of the permissioned blockchain network from the DUDS framework that enables specific trusted nodes to perform the verification process and offers a data-sharing mechanism via modifiers of the smart contracts. The modifies are used to restrict certain functionalities including data access to certain nodes. Another scenario to structure the data trust would allow users to decide how their data should be shared with informed consent. This scenario is represented with the DUDS framework that offers data sharing preferences encoded in smart contracts, thereby addressing the issue of consent regimes. Those smart contracts, once deployed, automatically represent the users' consent over their data to be searched and accessed by the selective data consumers.

7.2.5 User Acceptance Studies for DUDS-Based System

Furthermore, concerning the trustworthiness of the experimentation carried out with students and IT professionals, it was also imperative to maintain a low level of variance in the knowledge of the participants who took part in the studies. I conducted the consistency and reliability tests to measure such factors that limit the inconsistency and unreliability in the results of the user studies. I applied the inclusion criteria to filter out undesirable participants from the studies. One of such inclusion criteria was that the participants should be from a technical (computer science or engineering)

[52] https://blog.equinix.com/blog/2019/09/26/how-to-speak-like-a-data-center-geek-data-trusts/

background for the user study on the company-specific DUDS platform because the system included technical aspects that only a software developer or system administrator could understand well.

7.2.6 Augmented TAM in User Studies

Although I provided the arguments in the thesis based on the two specific types of DUDS platforms and goal-oriented requirements, this general approach could also be applied on a larger scale for sharing personal data in social networks, where much of the data are contributed voluntarily by the user; others are obtained by the system from observation of user activities or inferred through advanced analysis of volunteered or observed data. In addition to that, the user behavioral experience model and the augmented TAM with Trust model presented in this thesis can be used in other areas of software engineering and computational science. These models are not restricted to the modeling, designing and analyses of goal-oriented requirements engineering.

7.2.7 Trust and Privacy Challenges

Considering experimentation with customer-specific and company-specific applications based on DUDS platforms, my study results show that TAM-based predictors or trust constructs cannot be applied uniformly to all the DUDS platforms. The relationships between perceived trust, perceived security, perceived privacy and attitudes towards the initial adoption of the system might change depending on the specifics of the platforms. The decision-makers, especially IT infrastructure architects, should carefully consider the trust patterns and address the associated privacy challenges of the users to obtain a balanced design that will most likely address the solution requirements to an acceptable extent despite not being necessarily able to satisfy all functional requirements.

7.3 Limitations

In the thesis, I have identified several issues and provided some solutions to them, with some of them remaining as limitations to be addressed in future works.

The current version of the DUDS framework does not solve all the challenges of data governance. The use of the blockchain can create redundancy in the system. Currently, the

framework has no documented controls to guarantee proper and ethical data usage. Furthermore, like any other technology, DLTs also have some limitations and are not the silver bullet that can be incorporated to demonstrate any business values of data governance. The existing literature has presented numerous privacy and security-related issues when decentralized applications and smart contracts are adopted. As per the current state of the art, although blockchain supports authenticity, integrity, immutability, availability, and peer-to-peer security, decentralized applications and smart contracts themselves are vulnerable to security breaches and privacy infringements. It is vital to identify and conduct careful analysis and evaluation of various constructs affecting the collaborative business model built on top of blockchain and smart contracts technologies.

There are also limitations related to the usability study presented in Chapter 6. One of the limitations is the smaller-sized participants' pool and experimentation with only two specific types of DUDS platforms—customer-specific and company-specific. Also, the results from a single study may not generalize to the broader population and another type of platform.

Furthermore, the user studies during my research included both students and industry people from various affiliations. While this pool offered the diversity of participants' worldviews in generalizing the results to a broad domain, it is noteworthy that each participant may have varying answers for the same question under different circumstances, which may lead to unreliable and inconsistent results. Also, there might be an issue over how much we could trust the results from students, who usually do not have industrial experience. Further, the questions weren't related to the business model. So, the students with sufficient system technical knowledge were expected to answer those questions, but some studies in the literature demonstrate that a student sample may or may not introduce bias in the final result. I haven't conducted two separate studies based on a student sample and a non-student sample. It would be desirable to conduct follow-up studies based upon different demographic groups (such as gender, age, computer experience, nationality, etc.) to strengthen the results of the studies. Also, longitudinal studies should be conduced to evaluate how the user trust evolves in time.

7.4 Future Work

Following up on my current research, I intend to improve the current user experience model for the DUDS framework through the findings from the study on users' attitudes to data sharing and the

incentives to which they would be receptive. I further intend to conduct my research with deep learning and the theory of reasoned action framework, which is similar to the augmented TAM, by further exploring and incorporating other constructs that explain the relationship between attitudes and behaviours within human action towards adopting the blockchain-based computational systems. The aim is to investigate the potential users' behavioral intention to adopt such systems. My future research will develop the heuristic cognitive-behavioural model that enables the identification of other predictors that affect user trust and the acceptance of such computational systems. This goal is due to the increasingly prominent issue of trust in technology. The study results will be useful tools that incorporate additional factors affecting trust, security, privacy, transparency, control, etc. from the start in the design and development of such blockchain-based computational platforms that support existing common data governing procedures.

Also, my initial research with user acceptance modeling was carried out for a short period without follow-ups. I would like to carry out a longer-term experiment to increase the reliability of the findings and explore the effect of other variables. I intend to conduct several longitudinal studies for a longer period and among several target populations in future research. Ideally, these studies will involve users in their real-world contexts as they make authentic decisions that have real consequences for them.

Another important future work would be to manage security, privacy, transparency, and trust issues in the blockchain-based usable computational models and identify and exclude any faulty or Byzantine members as defined in (Abraham et al., 2016). I am specifically aiming to study the performance of the blockchain-based solutions in a collective reinforcement learning and deep-learning scenario (such as offering more meaningful news and notifications to users, autonomous driving, etc.) while offering proper incentives to the participating parties and sharing obtained local gradients in the collaborative training process. The results of such studies, both in the presence and absence of Byzantine members, can be helpful to the system designers to present a model that can address most of the stakeholders' requirements. Furthermore, complete audit trails of the training processes with audit features to trace back the data's state and route (representing audited learned knowledge) help preserve fairness in deep learning model training.

Another future work direction would be to improve the DUDS Framework for developing the integrated trust model for sharing patient data between the stakeholders in the medical domain. I have done some preliminary studies on how to incorporate blockchain-based solutions. In the

medical domain, there are open-source common data models like OMOP[53], PCORNet[54], and standards such as SNOMED-CT[55] (Systematized Nomenclature of Medicine - Clinical Terms), used by health care providers to exchange clinical health information. The challenge in this research would be to properly implement ETL (extract, transform, load process) into the data warehouse (Vassiliadis et al., 2002) and build the integrated trust model for exchanging clinical health information that supports the medical data governance and regulatory compliances.

Some application domains have various characteristics in common, and others are different. These characteristics may be more important than the actual domain itself in determining the importance of security, privacy and trust issues. Medicine and government systems are deeply and legally bound to ensure privacy. But medicine and government systems may differ on the kind of data they accumulate with respect to the data modality, granularity, size, etc. A categorization of these characteristics and how they affect privacy and trust concerns might be useful to allow a domain with a particular set of characteristics to be categorized in terms of its security, privacy and trust issues based on the domain's characteristics. Such categorization would also address the challenge of several siloed systems to handle voluminous data and show how they are connected, searched and shared.

Moreover, it is equally important to address the challenge of continued growth in the shared value for organizations. It can be obstructive to users to deploy the smart contracts for every new consent. Technically, this can be addressed via data trust by modifying the smart contracts templates that allow users to specify acceptance of user-delegated authorities (custodian or trustee). So, another important future work would be to incorporate delegated authorities in the DUDS framework and designate their responsibilities. This would enable users to legally delegate the data sharing decision to the authority acting with their best interest in mind, who will administer it solely for the data sharing purposes specified in the contracts. This will offer the added benefit of increasing the performance of real-time consent management.

Besides, blockchain interoperability and the security of smart contracts are also important areas for future research. The standard, pre-tested, community-reviewed smart contracts packages from

[53] https://www.ohdsi.org/data-standardization/the-common-data-model/
[54] https://pcornet.org/
[55] https://www.snomed.org/

OpenZeppelin[56] could be used to construct smart contracts. This approach can benefit from utilizing the ERC (Ethereum Request for Comments) standards[57] while enabling the smart contacts to have minimal identified risks. However, there is always a risk factor intertwined with the smart contracts through bugs present in the codes, such as unexpected ether flow, unsafe inputs and re-entrancy method calls. Most of the current solutions, such as Oyente[58] and Manticore,[59] to address security auditing (bug finding) of smart contracts are mainly checkers on fuzzing execution. Therefore, a different approach to handling future issues could be a reverse engineering blend with pattern scrutinizing that decompiles the EVM (Ethereum virtual machine) bytecodes[60] and then checks for the different compliance and violation patterns.

Finally, the continuous development and advancement of distributed ledger technologies, including blockchains and smart contracts, must be considered over time according to the modified integrated user data-sharing framework to prevent losses and uncover different factors affecting its continued adoption and use while addressing the changing needs of the users.

[56] https://openzeppelin.com/contracts/
[57] https://ethereum.org/en/developers/docs/standards/tokens/
[58] https://github.com/enzymefinance/oyente
[59] https://github.com/trailofbits/manticore
[60] https://ethereum.org/en/developers/docs/evm/

REFERENCES

Abel, F., Herder, E., Houben, G.-J., Henze, N., & Krause, D. (2013). *Cross-system user modeling and personalization on the Social Web*. *23*, 169–209. https://doi.org/10.1007/s11257-012-9131-2

Abraham, I., Malkhi, D., Nayak, K., Ren, L., & Spiegelman, A. (2016). Solida: A Blockchain Protocol Based on Reconfigurable Byzantine Consensus. *Leibniz International Proceedings in Informatics, LIPIcs*, *95*. https://arxiv.org/abs/1612.02916v2

Alharby, M., Aldweesh, A., & Moorsel, A. van. (2018). Blockchain-based Smart Contracts: A Systematic Mapping Study of Academic Research (2018). *2018 International Conference on Cloud Computing, Big Data and Blockchain (ICCBB)*, 1–6. https://doi.org/10.1109/ICCBB.2018.8756390

Alharby, M., & Moorsel, A. van. (2017). Blockchain Based Smart Contracts : A Systematic Mapping Study. *Fourth International Conference on Computer Science and Information Technology (CSIT-2017)*, 125–140. https://doi.org/10.5121/csit.2017.71011

Assad, M., Carmichael, D. J., Kay, J., & Kummerfeld, B. (2007). PersonisAD: Distributed, Active, Scrutable Model Framework for Context-Aware Services. In *Pervasive Computing* (pp. 55–72). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-72037-9_4

Azaria, A., Ekblaw, A., Vieira, T., & Lippman, A. (2016). MedRec: Using Blockchain for Medical Data Access and Permission Management. *2016 2nd International Conference on Open and Big Data (OBD)*, 25–30. https://doi.org/10.1109/OBD.2016.11

Balabanović, M., & Shoham, Y. (1997). Fab: content-based, collaborative recommendation. *Communications of the ACM*, *40*(3), 66–72. https://doi.org/10.1145/245108.245124

Barla, M. (2011). Towards social-based user modeling and personalization. *Information Sciences and Technologies Bulletin of the ACM Slovakia*, *3*(1), 52–60. http://acmbulletin.fiit.stuba.sk/theses/barla-thesis.pdf

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173–1182. https://doi.org/10.1037/0022-3514.51.6.1173

Baum, S. (2018). Cryptocurrency Fraud: A Look Into The Frontier of Fraud. *Honors College Theses*. https://digitalcommons.georgiasouthern.edu/honors-theses/375

Belanger, F., Hiller, J. S., & Smith, W. J. (2002). Trustworthiness in electronic commerce: the role of privacy, security, and site attributes. *The Journal of Strategic Information Systems*, *11*(3–4), 245–270. https://doi.org/10.1016/S0963-8687(02)00018-5

Bhattacherjee, A. (2000). Acceptance of e-commerce services: the case of electronic brokerages. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, *30*(4), 411–420. https://doi.org/10.1109/3468.852435

Bian, S., Deng, Z., Li, F., Monroe, W., Shi, P., Sun, Z., Wu, W., Wang, S., Wang, W. Y., Yuan, A., Zhang, T., & Li, J. (2018). IcoRating: A Deep-Learning System for Scam ICO

Identification. *ArXiv*. http://arxiv.org/abs/1803.03670

Bierer, B. E., Crosas, M., & Pierce, H. H. (2017). Data Authorship as an Incentive to Data Sharing. *New England Journal of Medicine*, *376*(17), 1684–1687. https://doi.org/10.1056/NEJMsb1616595

Biswas, P., & Robinson, P. (2010). A brief survey on user modelling in HCI. *Proc. of the International Conference on Intelligent Human Computer Interaction (IHCI) 2010*. https://api.semanticscholar.org/CorpusID:9469040

Blanchette, J.-F. (2006). The digital signature dilemma. *Annales Des Télécommunications*, *61*(7), 908–923. https://doi.org/10.1007/BF03219871

Bobrow, D. G. (1991). Dimensions of interaction: a shift of perspective in artificial intelligence. *AI Magazine*, *12*(3), 64–80. https://dl.acm.org/doi/10.5555/123768.123779

Bruns, A. (2016). User-Generated Content. In *The International Encyclopedia of Communication Theory and Philosophy* (pp. 1–5). John Wiley & Sons, Inc. https://doi.org/10.1002/9781118766804.wbiect085

Brusilovsky, P. (2001). Adaptive hypermedia. *User Modeling and User-Adapted Interaction*. https://doi.org/10.1023/A:1011143116306

Bryzek, J. (2015). *Creating a TSensors-based Future*. TSensors Summit. http://www.meptec.org/Resources/7 - Bryzek.pdf

Buchanan, T., Paine, C., Joinson, A. N., & Reips, U. D. (2007). Development of measures of online privacy concern and protection for use on the Internet. *Journal of the American Society for Information Science and Technology*, *58*(2), 157–165. https://doi.org/10.1002/asi.20459

Bullock, A., & Bannigan, K. (2016). Effectiveness of activity-based group work in community mental health: a systematic review. *The American Journal of Occupational Therapy : Official Publication of the American Occupational Therapy Association*, *65*(3), 257–266. https://doi.org/10.5014/ajot.2011.001305

Burgoon, J. K., Parrott, R., Le Poire, B. A., Kelley, D. L., Walther, J. B., & Perry, D. (1989). Maintaining and restoring privacy through communication in different types of relationships. *Journal of Social and Personal Relationships*. https://doi.org/10.1177/026540758900600201

Buterin, V. (2015). *A next generation smart contract & decentralized application platform*. http://www.the-blockchain.com/docs/Ethereum_white_paper-a_next_generation_smart_contract_and_decentralized_application_platform-vitalik-buterin.pdf

Carmagnola, F., Cena, F., & Gena, C. (2011). *User model interoperability: a survey*. *21*, 285–331. https://doi.org/10.1007/s11257-011-9097-5

Cassidy, C. M., & Chae, B. (2006). Consumer Information Use and Misuse in Electronic Business: An Alternative to Privacy Regulation. *Information Systems Management*, *22*(3), 75–87.

http://www.tandfonline.com/doi/pdf/10.1201/1078.10580530/46108.23.3.20060601/93709.8

Chan, S. C., & Lu, M. Te. (2004). Understanding Internet Banking adoption and use behavior: A Hong Kong perspective. In *Journal of Global Information Management*. https://doi.org/10.4018/jgim.2004070102

Chellappa, R. K., & Pavlou, P. A. (2002). Perceived information security, financial liability and consumer trust in electronic commerce transactions. *Logistics Information Management*, *15*(5/6), 358–368. https://doi.org/10.1108/09576050210447046

Chen, C.-S., Chang, S.-F., & Liu, C.-H. (2012). Understanding Knowledge-Sharing Motivation, Incentive Mechanisms, and Satisfaction in Virtual Communities. *Social Behavior and Personality: An International Journal*, *40*(4), 639–647. https://doi.org/10.2224/sbp.2012.40.4.639

Chen, M.-Y., Wang, E. K., Yang, R.-J., & Liou, Y.-M. (2003). Adolescent Health Promotion Scale: Development and Psychometric Testing. *Public Health Nursing*, *20*(2), 104–110. https://doi.org/10.1046/j.1525-1446.2003.20204.x

Chin, W. W., Marcelin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research*, *14*(2). https://doi.org/10.1287/isre.14.2.189.16018

Chin, W. W., Peterson, R. A., & Brown, S. P. (2008). Structural Equation Modeling in Marketing: Some Practical Reminders. *Journal of Marketing Theory and Practice*, *16*(4), 287–298. https://doi.org/10.2753/MTP1069-6679160402

Congosto, M., Basanta-Val, P., & Sanchez-Fernandez, L. (2017). T-Hoarder: A framework to process Twitter data streams. *Journal of Network and Computer Applications*, *83*, 28–39. https://doi.org/10.1016/J.JNCA.2017.01.029

Coppola, N. W., Hiltz, S. R., & Rotter, N. G. (2004). Building Trust in Virtual Teams. *IEEE Transactions on Professional Communication*, *47*(2), 95–104. https://doi.org/10.1109/TPC.2004.828203

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications

Cronbach, L. J. (1971). Test validation. In R. L. Thomdike (Ed.), *Educational Measurement* (pp. 443–507). American Council on Education. https://ci.nii.ac.jp/naid/10017369036/

Crosby, M., Nachiappan, Pattanayak, P., Verma, S., & Kalyanaraman, V. (2015). Blockchain Technology Explained - Beyond Bitcoin. *Sutardja Center for Entrepreneurship & Technology Technical Report*, 1–27. http://www.blockchaintechnologies.com/blockchain-definition

Cunningham, S. M. (1967). The major dimensions of perceived risk. *Risk Taking and Information Handling in Consumer Behavior*, 82–111

Davis, F. D. (1986). *A Technology Acceptance Model for Empirically Testing New End-User Information Systems*. https://dspace.mit.edu/bitstream/handle/1721.1/15192/14927137-MIT.pdf

Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, *13*, 983–1003. https://www.jstor.org/stable/pdf/249008.pdf

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace1. *Journal of Applied Social Psychology*, *22*(14), 1111–1132. https://doi.org/10.1111/j.1559-1816.1992.tb00945.x

Davoust, A. (2015). *Decentralized Social Data Sharing* [Carleton University]. https://curve.carleton.ca/295070d6-5ad8-4c55-953e-d10109745f19

DeCew, J. W. (1997). Information Technology. In *In Pursuit of Privacy* (pp. 145–164). Cornell University Press. http://www.jstor.org/stable/10.7591/j.ctv75d3zc.13

DeLone, W. H., & McLean, E. R. (1992). Information systems success: The quest for the dependent variable. *Information Systems Research*. https://doi.org/10.1287/isre.3.1.60

Demo, G., Neiva, E. R., Nunes, I., & Rozzett, K. (2012). Human resources management policies and practices scale (HRMPPS): exploratory and confirmatory factor analysis. *BAR - Brazilian Administration Review*, *9*(4), 395–420. https://doi.org/10.1590/s1807-76922012005000006

Dennis, A. R., Robert, L. P., Curtis, A. M., Kowalczyk, S. T., & Hasty, B. K. (2012). Research Note: Trust Is in the Eye of the Beholder: A Vignette Study of Postevent Behavioral Controls' Effects on Individual Trust in Virtual Teams. *Research*, *23*(2), 546–558. https://doi.org/10.1287/isre.lll0.0364

Denzin, N. ., & Lincoln, Y. S. (2005). The SAGE Handbook of Qualitative Research. Third Edition. In *The SAGE Handbook of Qualitative Research*. https://doi.org/10.4324/9780203409527

Dim, E., & Kuflik, T. (2012). User Models Sharing and Reusability: A Component-based Approach. *AUM 2012: Augmented User Modeling*, *1*. http://ceur-ws.org/Vol-872/aum2012_paper_2.pdf

Dolog, P., & Vassileva, J. (2005). Decentralized, Agent Based, and Social Approaches to User Modeling (DASUM). *Workshop DASUM-05, at the 9th International Conference on User Modeling (UM'05)*. http://people.cs.aau.dk/~dolog/pub/DASUM-proceedings.pdf

Domingue, J. (2017). *Blockchains as a Component of the Next Generation Internet*. http://dilbert.com/

Dong, X., Guo, B., Shen, Y., Duan, X., Shen, Y., & Zhang, H. (2019). A self-controllable and balanced data sharing model. *IEEE Access*, *7*, 103275–103290. https://doi.org/10.1109/ACCESS.2019.2931982

Duy, P. T., Hien, D. T. T., Hien, D. H., & Pham, V.-H. (2018). A survey on opportunities and challenges of Blockchain technology adoption for revolutionary innovation. *Proceedings of*

the Ninth International Symposium on Information and Communication Technology - SoICT 2018, 200–207. https://doi.org/10.1145/3287921.3287978

Dwyer, C., Hiltz, S. R., & Passerini, K. (2007). Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace. *Association for Information Systems - 13th Americas Conference on Information Systems, AMCIS 2007: Reaching New Heights*, *3*, 1725–1735. https://aisel.aisnet.org/amcis2007/339

Eastlick, M. A., Lotz, S. L., & Warrington, P. (2006). Understanding online B-to-C relationships: An integrated model of privacy concerns, trust, and commitment. *Journal of Business Research*, *59*(8), 877–886. https://doi.org/10.1016/j.jbusres.2006.02.006

EDUCAUSE. (2015). *Establishing Data Stewardship Models. ECAR Working Group Paper*. https://library.educause.edu/~/media/ files/library/2015/12/ewg1514-pdf.pdf

Fabian, B., Ermakova, T., & Junghanns, P. (2015). Collaborative and secure sharing of healthcare data in multi-clouds. *Information Systems*, *48*, 132–150. https://doi.org/10.1016/j.is.2014.05.004

Fernandez, E. B., Fonoage, M., VanHilst, M., & Marta, M. (2008). The secure three-tier architecture pattern. *Proceedings - CISIS 2008: 2nd International Conference on Complex, Intelligent and Software Intensive Systems*, 555–560. https://doi.org/10.1109/CISIS.2008.51

Ferrer-Sapena, A., & Sánchez-Pérez, E.-A. (2019). Aplicaciones de la tecnología blockchain en la documentación científica: situación actual y perspectivas. *El Profesional de La Información*, *28*(2). https://doi.org/10.3145/epi.2019.mar.10

Finin, T. W. (1989). GUMS — A General User Modeling Shell. In *User Models in Dialog Systems* (pp. 411–430). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-83230-7_15

Fink, A. S. (2000). The Role of the Researcher in the Qualitative Research Process . A Potential Barrier to Archiving Qualitative Data. *Forum: Qualitative Social Research*, *1*(3), 1–11

Fink, J., & Kobsa, A. (2000). Review and analysis of commercial user modeling servers for personalization on the World Wide Web. *User Modelling and User-Adapted Interaction*. https://doi.org/10.1023/A:1026597308943

Fischer, G. (2001). User Modeling in Human–Computer Interaction. *User Modeling and User-Adapted Interaction*, *11*(1/2), 65–86. https://doi.org/10.1023/A:1011145532042

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behavior: An introduction to theory and research. In *Addison-Wesley* . http://people.umass.edu/aizen/f&a1975.html

Folkinshteyn, D., & Lennon, M. (2016). Braving Bitcoin: A technology acceptance model (TAM) analysis. *Journal of Information Technology Case and Application Research*, *18*(4), 220–249. https://doi.org/10.1080/15228053.2016.1275242

Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, *18*(1), 39. https://doi.org/10.2307/3151312

Gefen, D. (2000). E-commerce: The role of familiarity and trust. *Omega*.

https://doi.org/10.1016/S0305-0483(00)00021-9

Gefen, D., Karahanna, E., & Straub, D. W. (2003). Trust and Tam in Online Shopping: An Integrated Model. *MIS Quarterly: Management Information Systems*, *27*(1), 51–90. https://doi.org/10.2307/30036519

Gefen, D., Straub, D., & Boudreau, M.-C. (2000). Structural Equation Modeling and Regression: Guidelines for Research Practice. *Communications of the Association for Information Systems*, *4*(1). https://doi.org/10.17705/1CAIS.00407

Golle, P., Leyton-Brown, K., Mironov, I., & Lillibridge, M. (2001). Incentives for sharing in peer-to-peer networks. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *2232*, 75–87. https://doi.org/10.1145/501158.501193

Graham, V., & Sacha, W.-V. (2007). *Participative Web And User-Created Content: Web 2.0 Wikis and Social Networking*. Organization for Economic Cooperation and Development (OECD) Paris, France, France ©2007. https://dl.acm.org/doi/book/10.5555/1554640

Granić, A., & Marangunić, N. (2019). Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology*, *50*(5), 2572–2593. https://doi.org/10.1111/bjet.12864

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. In *Source: Educational Evaluation and Policy Analysis* (Vol. 11, Issue 3)

Greenspan, G. (2013). *MultiChain Private Blockchain - White Paper*. 1–17. http://www.multichain.com/download/MultiChain-White-Paper.pdf

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate Data Analysis* (7th ed.). Pearson Education Limited. www.pearsoned.co.uk

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2013). A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM) (2nd ed.). Sage Publications

Hall, D. W., & Pesenti, J. (2017). *Growing the artificial intelligence industry in the UK*. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/652097/Growing_the_artificial_intelligence_industry_in_the_UK.pdf

Henry, R., Herzberg, A., & Kate, A. (2018). Blockchain access privacy: Challenges and directions. *IEEE Security and Privacy*, *16*(4), 38–45. https://doi.org/10.1109/MSP.2018.3111245

Herder, E., & Kärger, P. (2008). Hybrid Personalization For Recommendations. *16th Workshop on Adaptivity and User Modeling in Interactive System*, 20–25. https://core.ac.uk/download/pdf/55534580.pdf

Herrera-Joancomartí, J., & Pérez-Solà, C. (2016). Privacy in bitcoin transactions: New challenges from blockchain scalability solutions. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 9880 LNAI*, 26–44. https://doi.org/10.1007/978-3-319-45656-0_3

Hevner, A. R., & Chatterjee, S. (2010). Design Research in Information Systems. In *Design Research in Information Systems* (Vol. 22). Springer US. https://doi.org/10.1007/978-1-4419-5653-8

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). DESIGN SCIENCE IN INFORMATION SYSTEMS RESEARCH. In *Design Science in IS Research MIS Quarterly* (Vol. 28, Issue 1)

Hileman, G., & Rauchs, M. (2017). 2017 Global Blockchain Benchmarking Study. *Ssrn*. https://doi.org/10.2139/ssrn.3040224

Hoffman, D. L., Novak, T. P., & Peralta, M. (1999). Building consumer trust online. *Communications of the ACM*, *42*(4), 80–85. https://doi.org/10.1145/299157.299175

Horvitz, E., Breese, J., Heckerman, D., Hovel, D., & Rommelse, K. (1998). The Lumiere Project: Bayesian User Modeling for Inferring the Goals and Needs of Software Users. *UAI'98 Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence*, 256–265. https://dl.acm.org/doi/10.5555/2074094.2074124

Hu, H., Ahn, G.-J., & Jorgensen, J. (2011). Detecting and resolving privacy conflicts for collaborative data sharing in online social networks. *Proceedings of the 27th Annual Computer Security Applications Conference on - ACSAC '11*, 103–112. https://doi.org/10.1145/2076732.2076747

Hull, R., Kumar, B., Lieuwen, D., Patel-Schneider, P. F., Sahuguet, A., Varadarajan, S., & Vyas, A. (2004). Enabling context-aware and privacy-conscious user data sharing. *IEEE International Conference on Mobile Data Management (MDM'04)*, 187–198. https://doi.org/10.1109/MDM.2004.1263065

Introna, L. D., & Pouloudi, A. (1999). Privacy in the information age: Stakeholders, interests and values. *Journal of Business Ethics*, *22*(1), 27–38. https://doi.org/10.1023/A:1006151900807

Iyilade, J., & Vassileva, J. (2013). A decentralized architecture for sharing and reusing life-logs. *LLUM 2013: LifeLong User Modelling*, *997*, 4–10. https://pdfs.semanticscholar.org/7398/8175fc23e72af94cea85a3cfdaa3d6c5da55.pdf

Jarvenpaa, S. L., & Leidner, D. E. (1999). Communication and Trust in Global Virtual Teams. *Organization Science*. https://doi.org/10.1287/orsc.10.6.791

Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000). Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics*, *4*(1), 53–71. https://doi.org/10.1207/s15327566ijce0401_04

Jun, M. (2018). Blockchain government - a next form of infrastructure for the twenty-first century. *Journal of Open Innovation: Technology, Market, and Complexity*, *4*(1), 7. https://doi.org/10.1186/s40852-018-0086-3

Kay, J. (2001). Learner Control. *User Modeling and User-Adapted Interaction*, *11*(1/2), 111–127. https://doi.org/10.1023/A:1011194803800

Kern, A. G. (2018). *Blockchain technology : a technology acceptance model (TAM) analysis*. http://hdl.handle.net/10400.14/25478

Kiertz, C. (2014). Know Your Users: The Difference Between Profile Data and Behavioral Data. *Localytics*. http://info.localytics.com/blog/know-your-users-what-is-the-difference-between-profile-data-and-behavioral-data

Kim, D. J., Steinfield, C., & Lai, Y. J. (2008). Revisiting the role of web assurance seals in business-to-consumer electronic commerce. *Decision Support Systems*. https://doi.org/10.1016/j.dss.2007.11.007

Kim, J. (2010). *User-generated content (UGC) revolution?* [University of Iowa]. https://doi.org/10.17077/etd.bgqt6x30

Kobsa, A. (2001). Generic User Modeling Systems. *User Modeling and User-Adapted Interaction*, *11*(1/2), 49–63. https://doi.org/10.1023/A:1011187500863

Kobsa, A. (2007). Privacy-Enhanced Web Personalization. *LNCS*, *4321*, 628–670. http://www.ics.uci.edu/~kobsa

Kobsa, A., & Wahlster, W. (1989). *User Models in Dialog Systems* (A. Kobsa & W. Wahlster (eds.); 1st ed.). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-83230-7

Koh, C. E., Prybutok, V. R., & Ryan, S. D. (2010). A Model for Mandatory Use of Software Technologies: An Integrative Approach by Applying Multiple Levels of Abstraction of Informing Science. *Informing Science: The International Journal of an Emerging Transdiscipline*, *13*, 177–203. https://pdfs.semanticscholar.org/724c/351bd184d910f24404b53341859e6bcc75d6.pdf

Kravchenko, P. (2016). *Ok, I need a blockchain, but which one?* https://medium.com/@pavelkravchenko/ok-i-need-a-blockchain-but-which-one-ca75c1e2100

Kshetri, N. (2017). Blockchain's roles in strengthening cybersecurity and protecting privacy. *Telecommunications Policy*, *41*(10), 1027–1038. https://doi.org/10.1016/j.telpol.2017.09.003

Li, L., Liu, J., Cheng, L., Qiu, S., Wang, W., Zhang, X., & Zhang, Z. (2018). CreditCoin: A Privacy-Preserving Blockchain-Based Incentive Announcement Network for Communications of Smart Vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 1–17. https://doi.org/10.1109/TITS.2017.2777990

Linck, K., Pousttchi, K., & Wiedemann, D. G. (2006). Security issues in mobile payment from the customer viewpoint. *Proceedings of the 14th European Conference on Information Systems, ECIS 2006*

Liu, H., Li, X., Xu, M., Mo, R., & Ma, J. (2017). A fair data access control towards rational users in cloud storage. *Information Sciences*, *418–419*, 258–271. https://doi.org/10.1016/j.ins.2017.07.023

Liu, I.-F., Chen, M. C., Sun, Y. S., Wible, D., & Kuo, C.-H. (2010). Extending the TAM model to explore the factors that affect Intention to Use an Online Learning Community. *Computers & Education*, *54*(2), 600–610. https://doi.org/10.1016/J.COMPEDU.2009.09.009

Lizar, M., & Hughes, A. (2018). Consent Receipt Specification. *Technical Specification*

Recommendation. https://kantarainitiative.org/confluence/display/infosharing/Consent+Receipt+Specification

Lo, B., & DeMets, D. L. (2016). Incentives for Clinical Trialists to Share Data. *New England Journal of Medicine*, *375*(12), 1112–1115. https://doi.org/10.1056/NEJMp1608351

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet Users ' The Information the Scale , and a Causal (IUIPC): *Informs*, *15*(4), 336–355. https://doi.org/10.1287/isre.l040.0032

McCole, P., Ramsey, E., & Williams, J. (2010). Trust considerations on attitudes towards online purchasing: The moderating effect of privacy and security concerns. *Journal of Business Research*, *63*(9–10), 1018–1024. https://doi.org/10.1016/j.jbusres.2009.02.025

Mcconaghy, T., Marques, R., Müller, A., De Jonghe, D., Mcconaghy, T. T., Mcmullen, G., Henderson, R., Bellemare, S., & Granzotto, A. (2016). *BigchainDB: A Scalable Blockchain Database*. https://www.bigchaindb.com/whitepaper/bigchaindb-whitepaper.pdf

McGhin, T., Choo, K. K. R., Liu, C. Z., & He, D. (2019). Blockchain in healthcare applications: Research challenges and opportunities. In *Journal of Network and Computer Applications* (Vol. 135, pp. 62–75). Academic Press. https://doi.org/10.1016/j.jnca.2019.02.027

Melas, C. D., Zampetakis, L. A., Dimopoulou, A., & Moustakis, V. (2011). Modeling the acceptance of clinical information systems among hospital medical staff: An extended TAM model. *Journal of Biomedical Informatics*, *44*(4), 553–564. https://doi.org/10.1016/J.JBI.2011.01.009

Mertens, D. M. (2005). *Research and evaluation in education and psychology : integrating diversity with quantitative, qualitative, and mixed methods* (2nd ed.). Sage Publications.

Metzger, M. J. (2004). Privacy, trust, and disclosure: Exploring barriers to electronic commerce. *Journal of Computer-Mediated Communication*. https://doi.org/10.1111/j.1083-6101.2004.tb00292.x

Mills, S. (2019). Who Owns the Future? Data Trusts, Data Commons, and the Future of Data Ownership. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3437936

Molina-Jimenez, C., Solaiman, E., Sfyrakis, I., Ng, I., & Crowcroft, J. (2019). *On and Off-Blockchain Enforcement of Smart Contracts* (pp. 342–354). Springer, Cham. https://doi.org/10.1007/978-3-030-10549-5_27

Muller, K., & Cohen, J. (1989). Statistical Power Analysis for the Behavioral Sciences. *Technometrics*. https://doi.org/10.2307/1270020

Münch, J., Fagerholm, F., Johnson, P., Pirttilahti, J., Torkkel, J., & Jäarvinen, J. (2013). Creating Minimum Viable Products in Industry-Academia Collaborations. In *Lecture Notes in Business Information Processing* (Vol. 167, pp. 137–151). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-44930-7_9

Nakamoto, S. (2008). *Bitcoin: A Peer-to-Peer Electronic Cash System*. https://bitcoin.org/bitcoin.pdf

Naz, M., Al-zahrani, F. A., Khalid, R., Javaid, N., Qamar, A. M., Afzal, M. K., & Shafiq, M. (2019). A Secure Data Sharing Platform Using Blockchain and Interplanetary File System.

Sustainability, *11*(24), 7054. https://doi.org/10.3390/su11247054

Ning, Z., Liao, J., Zhang, F., & Shi, W. (2018). Preliminary Study of Trusted Execution Environments on Heterogeneous Edge Platforms. *2018 IEEE/ACM Symposium on Edge Computing (SEC)*, 421–426. https://doi.org/10.1109/SEC.2018.00057

Niu, X., McCalla, G., & Vassileva, J. (2004). Purpose-based Expert Finding in a Portfolio Management System. *Computational Intelligence*, *20*(4), 548–561. http://julita.usask.ca/texte/Niu-et-al-final.pdf

O'Hara, K. (2019). *Data Trusts Ethics, Architecture and Governance for Trustworthy Data Stewardship*. https://www.southampton.ac.uk/wsi/index.page

O'Reilly, T. (2005). *What Is Web 2.0 - O'Reilly Media*. https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html

Offermann, P., Levina, O., Schönherr, M., & Bub, U. (2009). Outline of a design science research process. *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology - DESRIST '09*. https://doi.org/10.1145/1555619.1555629

Papadopoulou, P. (2007). Applying virtual reality for trust-building e-commerce environments. *Virtual Reality*. https://doi.org/10.1007/s10055-006-0059-x

Piccoli, G., & Ives, B. (2003). Trust and the unintended effects of behavior control in virtual teams. *MIS Quarterly: Management Information Systems*. https://doi.org/10.2307/30036538

Pikkarainen, T., Pikkarainen, K., Karjaluoto, H., & Pahnila, S. (2004). Consumer acceptance of online banking: An extension of the technology acceptance model. In *Internet Research*. https://doi.org/10.1108/10662240410542652

Poon, W. C. (2008). Users' adoption of e-banking services: The Malaysian perspective. *Journal of Business and Industrial Marketing*. https://doi.org/10.1108/08858620810841498

Poslad, S. (2009). *Ubiquitous Computing*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470779446

Prashanth Joshi, A., Han, M., & Wang, Y. (2018). A survey on security and privacy issues of blockchain technology. *Mathematical Foundations of Computing*, *1*(2), 121–147. https://doi.org/10.3934/mfc.2018007

Proctor, R. W., & Vu, K.-P. L. (2002). Human Information Processing: An Overview for Human–Computer Interaction. In A. Sears & J. A. Jacko (Eds.), *The Human-Computer Interaction Handbook Fundamentals, Evolving Technologies and Emerging Applications* (3rd ed., pp. 67–83). CRC Press. https://doi.org/10.1201/9781410606723-17

Protenus. (2019). *2019 Annual Breach Barometer Report*. https://www.protenus.com/2019-breach-barometer

Rios, R., Fernandez-Gago, C., & Lopez, J. (2017). Modelling privacy-aware trust negotiations. *Computers & Security*, *77*, 773–789. https://doi.org/10.1016/j.cose.2017.09.015

Roca, J. C., García, J. J., & de la Vega, J. J. (2009). The importance of perceived trust, security

and privacy in online trading systems. *Information Management and Computer Security*, *17*(2), 96–113. https://doi.org/10.1108/09685220910963983

Rouhani, S., & Deters, R. (2019). MediChain TM : A Secure Decentralized Medical Data Asset Management System. *ArXiv Preprint ArXiv:1901.10645*, *Section II*, 1533–1538. https://doi.org/10.1109/Cybermatics

Rovai, A. P. (2004). A constructivist approach to online college learning. *The Internet and Higher Education*, *7*(2), 79–93. https://doi.org/10.1016/J.IHEDUC.2003.10.002

Samaniego, M., & Deters, R. (2016). Blockchain as a Service for IoT. *2016 IEEE International Conference on Internet of Things (IThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, 433–436. https://doi.org/10.1109/iThings-GreenCom-CPSCom-SmartData.2016.102

Sas, C., & Khairuddin, I. E. (2017). Design for trust: An exploration of the challenges and opportunities of bitcoin users. *Conference on Human Factors in Computing Systems - Proceedings*. https://doi.org/10.1145/3025453.3025886

Seaman, D. (2003). From isolation to integration: re-shaping the serials data silos. *Serials*, *16*(2), 131–135. https://doi.org/10.1629/16131

Shafagh, H., Burkhalter, L., Hithnawi, A., & Duquennoy, S. (2017). Towards Blockchain-based Auditable Storage and Sharing of IoT Data. *Cloud Comput. Secur. Workshop (CCSW), Dallas, TX, USA*, 45–50. http://arxiv.org/abs/1705.08230

Shen, M., Duan, J., Zhu, L., Zhang, J., Du, X., & Guizani, M. (2020). Blockchain-based incentives for secure and collaborative data sharing in multiple clouds. *IEEE Journal on Selected Areas in Communications*, *38*(6), 1229–1241. https://doi.org/10.1109/JSAC.2020.2986619

Shin, D. H. (2010). The effects of trust, security and privacy in social networking: A security-based approach to understand the pattern of adoption. *Interacting with Computers*, *22*(5), 428–438. https://doi.org/10.1016/j.intcom.2010.05.001

Shin, D. H. (2017). Conceptualizing and measuring quality of experience of the internet of things: Exploring how quality is perceived by users. *Information and Management*, *54*(8), 998–1011. https://doi.org/10.1016/j.im.2017.02.006

Shin, D. H. (2019). Blockchain: The emerging technology of digital trust. *Telematics and Informatics*, *45*, 101278. https://doi.org/10.1016/j.tele.2019.101278

Shrestha, A. K. (2014). Security of SIP-based infrastructure against malicious message attacks. *The 8th International Conference on Software, Knowledge, Information Management and Applications (SKIMA 2014)*, 1–8. https://doi.org/10.1109/SKIMA.2014.7083519

Shrestha, A. K., Deters, R., & Vassileva, J. (2017). User-Controlled Privacy-Preserving User Profile Data Sharing based on Blockchain. *Future Technologies Conference (FTC)*, 31–40. https://saiconference.com/Downloads/FTC2017/Proceedings/3_Paper_127-User-Controlled_Privacy-Preserving_User_Profile_Data_Sharing.pdf

Shrestha, A. K., Joshi, S., & Vassileva, J. (2020). Customer Data Sharing Platform: A Blockchain-Based Shopping Cart. *IEEE International Conference on Blockchain and Cryptocurrency, ICBC 2020*. https://doi.org/10.1109/ICBC48266.2020.9169421

Shrestha, A. K., & Vassileva, J. (2016). Towards decentralized data storage in general cloud platform for meta-products. *Proceedings of the International Conference on Big Data and Advanced Wireless Technologies - BDAW '16*, 1–7. https://doi.org/10.1145/3010089.3016029

Shrestha, A. K., & Vassileva, J. (2018a). Blockchain-Based Research Data Sharing Framework for Incentivizing the Data Owners. In S. Chen, H. Wang, & L.-J. Zhang (Eds.), *Blockchain -- ICBC 2018* (pp. 259–266). Springer International Publishing. https://doi.org/10.1007/978-3-319-94478-4_19

Shrestha, A. K., & Vassileva, J. (2019a). User Data Sharing Frameworks : A Blockchain-Based Incentive Solution. *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON 2019*, 360–366. https://doi.org/10.1109/IEMCON.2019.8936137

Shrestha, A. K., & Vassileva, J. (2018b). Bitcoin Blockchain Transactions Visualization. *2018 International Conference on Cloud Computing, Big Data and Blockchain (ICCBB)*, 1–6. https://doi.org/10.1109/ICCBB.2018.8756455

Shrestha, A. K., & Vassileva, J. (2019b). User acceptance of usable blockchain-based research data sharing system: An extended TAM-based study. *Proceedings - 1st IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications, TPS-ISA 2019*, 203–208. https://doi.org/10.1109/TPS-ISA48467.2019.00033

Shrestha, A. K., Vassileva, J., & Deters, R. (2020). A Blockchain Platform for User Data Sharing Ensuring User Control and Incentives. *Frontiers in Blockchain*, *3*, 48. https://doi.org/10.3389/fbloc.2020.497985

Shrestha, A. K., Vassileva, J., Joshi, S., & Just, J. (2021). Augmenting the technology acceptance model with trust model for the initial adoption of a blockchain-based system. *PeerJ Computer Science*, *7*, e502. https://doi.org/10.7717/peerj-cs.502

Siegel, J., & Sarma, S. (2019). A Cognitive Protection System for the Internet of Things. *IEEE Security and Privacy*, *17*(3), 40–48. https://doi.org/10.1109/MSEC.2018.2884860

Smith, H. J., Milberg, S. J., & Burke, S. J. (1996). Information privacy: Measuring individuals' concerns about organizational practices. *MIS Quarterly: Management Information Systems*, *20*(2), 167–195. https://doi.org/10.2307/249477

Starin, D., Baden, R., Bender, A., Spring, N., & Bhattacharjee, B. (2009). Persona : An Online Social Network with User-Defined Privacy. *SIGCOMM '09 Proceedings of the ACM SIGCOMM 2009 Conference on Data Communication*, 135–146. https://doi.org/10.1145/1592568.1592585

Stoica, I., Morris, R., Karger, D., Kaashoek, M. F., & Balakrishnan, H. (2003). Chord: A Scalable Peer-to-peer Lookup Service for Internet Applications. *Networking, IEEE/ACM Transactions*, *11*(1), 17–32. http://pdos.lcs.mit.edu/chord/

Stovel, R. G., Ginsburg, S., Stroud, L., Cavalcanti, R. B., & Devine, L. A. (2018). Incentives for recruiting trainee participants in medical education research. *Medical Teacher*, *40*(2), 181–187. https://doi.org/10.1080/0142159X.2017.1395402

Suchman, L. A. (1987). *Plans and situated actions : the problem of human-machine communication*. Cambridge University Press. https://dl.acm.org/doi/10.5555/38407

Sweeney, L. (1997). Guaranteeing anonymity when sharing medical data, the Datafly System. *Proceedings : A Conference of the American Medical Informatics Association. AMIA Fall Symposium*, 51–55. http://www.ncbi.nlm.nih.gov/pubmed/9357587

Szabo, N. (1997). Formalizing and Securing Relationships on Public Networks. *First Monday*, *2*(9). https://doi.org/10.5210/fm.v2i9.548

Taaki, A., & Hoffman, B. (2016). *OpenBazaar · GitHub*. https://github.com/OpenBazaar

Thilakarathna, K., Petander, H., Mestre, J., & Seneviratne, A. (2014). MobiTribe: Cost Efficient Distributed User Generated Content Sharing on Smartphones. *IEEE Transactions on Mobile Computing*, *13*(9), 2058–2070. https://doi.org/10.1109/TMC.2013.89

Tian, F. (2016). An agri-food supply chain traceability system for China based on RFID & blockchain technology. *2016 13th International Conference on Service Systems and Service Management (ICSSSM)*, 1–6. https://doi.org/10.1109/ICSSSM.2016.7538424

Van Den Eynden, V., & Bishop, L. (2014). *Incentives for sharing research data: evidence from five European case studies*. http://www.data-archive.ac.uk/about/projects/incentive

Vassileva, J., McCalla, G., & Greer, J. (2003). Multi-agent multi-user modeling in I-Help. *User Modelling and User-Adapted Interaction*, *13*(1), 179–210. https://doi.org/10.1023/A:1024072706526

Vassiliadis, P., Simitsis, A., & Skiadopoulos, S. (2002). Conceptual modeling for ETL processes. *Proceedings of the 5th ACM International Workshop on Data Warehousing and OLAP - DOLAP '02*, 14–21. https://doi.org/10.1145/583890.583893

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly: Management Information Systems*. https://doi.org/10.2307/30036540

Vrandeč, D., & Krötzsch, M. (2014). Wikidata: A Free Collaborative Knowledgebase. *COMMUNICATIONS OF THE ACM*, *57*(10). https://doi.org/10.1145/2629489

Wang, H., Lee, M. K. O., & Wang, C. (1998). Consumer Privacy Concerns about Internet Marketing. *Communications of the ACM*. https://doi.org/10.1145/272287.272299

Warren, S. D., & Brandeis, L. D. (1890). The Right to Privacy. In *Harvard Law Review* (Vol. 4, Issue 5). https://www.jstor.org/stable/1321160

Warshaw, P. R., & Davis, F. D. (1985). Disentangling behavioral intention and behavioral expectation. *Journal of Experimental Social Psychology*, *21*(3), 213–228. https://doi.org/10.1016/0022-1031(85)90017-4

Wei, L.-Y., Hsu, Y.-T., Peng, W.-C., & Lee, W.-C. (2014). Indexing spatial data in cloud data

managements. *Pervasive and Mobile Computing*, *15*, 48–61. https://doi.org/10.1016/J.PMCJ.2013.07.001

Westin, A. F. (1968). Privacy And Freedom. *Washington and Lee Law Review*, *25*(1). https://scholarlycommons.law.wlu.edu/wlulr/vol25/iss1/20

Whyte, A., & Pryor, G. (2011). Open Science in Practice: Researcher Perspectives and Participation. *International Journal of Digital Curation*, *6*(1), 199–213. https://doi.org/10.2218/ijdc.v6i1.182

Wong, K. K.-K. (2013). Partial Least Squares Structural Equation Modeling (PLS-SEM) Techniques Using SmartPLS. *Marketing Bulletin*, *24*(1), 1–32. http://marketing-bulletin.massey.ac.nz/v24/mb_v24_t1_wong.pdf

Wood, G. (2016). *Blockchain what and why-*. https://www.slideshare.net/gavofyork/blockchain-what-and-why

Wu, I.-L., & Chen, J.-L. (2005). An extension of Trust and TAM model with TPB in the initial adoption of on-line tax: An empirical study. *International Journal of Human-Computer Studies*, *62*(6), 784–808. https://doi.org/10.1016/j.ijhcs.2005.03.003

Xu, J., & Van Der Schaar, M. (2014). Incentive design for heterogeneous user-generated content networks. *Performance Evaluation Review*, *41*(4), 34–37. https://doi.org/10.1145/2627534.2627545

Yang, J., Onik, M., Lee, N.-Y., Ahmed, M., & Kim, C.-S. (2019). Proof-of-Familiarity: A Privacy-Preserved Blockchain Scheme for Collaborative Medical Decision-Making. *Applied Sciences*, *9*(7), 1370. https://doi.org/10.3390/app9071370

Yenisey, M. M., Ozok, A. A., & Salvendy, G. (2005). Perceived security determinants in e-commerce among Turkish university students. *Behaviour and Information Technology*. https://doi.org/10.1080/0144929042000320992

Zhang, X., & Chen, X. (2019). Data Security Sharing and Storage Based on a Consortium Blockchain in a Vehicular Ad-hoc Network. *IEEE Access*, *7*, 58241–58254. https://doi.org/10.1109/ACCESS.2018.2890736

Zohar, A., & Rosenschein, J. S. (2009). Adding incentives to file-sharing systems. *Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS*, *1*, 624–631. https://dl.acm.org/doi/10.5555/1558109.1558131

Zuboff, S. (2019). The Age of Surveillance Capitalism. *Profile Books*

Zyskind, G., Nathan, O., & Pentland, A. S. (2015). Decentralizing privacy: Using blockchain to protect personal data. *Proceedings - 2015 IEEE Security and Privacy Workshops, SPW 2015*, 180–184. https://doi.org/10.1109/SPW.2015.27

APPENDIX I

Support for User Model Interoperability Functionalities

Table I.1. Summary of the level of support for user model interoperability functionalities in existing systems (Carmagnola et al., 2011)

Dimension	Description	Approach	Remarks
Privacy	How to manage privacy in the context of UM interoperability?	Different access rights, encryption, perturbation, scrutable user model, third-party guarantors and standards, pseudonymous personalization.	Most systems have not adequately addressed the privacy issue.
Representation of the exchanged data	How an exchanged user model is from a syntactic and semantic dimension?	Usage of common ontology like GUMO, UUCM or standard vocabulary like PAPI, LIP, etc. Some also use translators.	Most systems are based on a common standardized UM approach. Both standard and translator systems have good and bad sides.
Integration of the exchanged data	Is the collected user model data integrated into an existing knowledge structure?	Integrated data are merged or not merged	Most systems do not consider data integration.
Communication	What Languages and protocols are to employ for communicating with other systems to exchange user model data?	Protocols such as web service (WSDL, REST), CORBA, Java RMI. Data representation such as UserML, CSCP, JSON, KQML.	Solutions that exploit standard approaches for communication are those with more opportunities to be adopted (e.g., JSON, REST etc.)
Exchanged data	What kind of data are exchanged by the systems?	User data, usage data, environment data, domain data, inferred data, reasoning data, social data.	Less attention is given to the existing system to the exchange of domain, reasoning and social data.

Types of Blockchain

Table I.2. Blockchain Types adapted from (Kravchenko, 2016)

Trust to a Validator / **Anonymity of Validators**	**Permissionless**	**Permissioned**
Public	Consensus: (PoW). Anyone can download the protocol and validate transactions. E.g., Ethereum, Zcash[61], Bitcoin.	Consensus: PoS, DPoS. Anyone meeting certain pre-defined criteria can download the protocol and validate transactions. E.g. Peercoin[62], Ethereum 2.0 via Casper[63], BitShares[64].
Consortium	Consensus: PBFT. Selected nodes can download the protocol and validate transactions. E.g., Ripple[65], Tendermint[66].	Consensus: IBFT. Selected nodes meeting certain pre-defined criteria can download the protocol and validate transactions. E.g., Quorum[67].

[61] https://github.com/zcash/zcash

[62] https://peercoin.net

[63] https://github.com/ethereum/wiki/wiki/Casper-Proof-of-Stake-compendium

[64] https://en.bitcoinwiki.org/wiki/BitShares

[65] https://ripple.com/

[66] https://tendermint.com/

[67] https://github.com/jpmorganchase/quorum

Private	Consensus: FBA. Each validator decides which other validators they trust to form quorum slices, which eventually overlap to form a network-wide consensus. E.g., Swarm (on Stellar[68]).	Consensus: PBFT, Raft-based consensus. Only recommended validators or members of the consortium can validate transactions (Central authority). E.g., MultiChain, Hyperledger.

Consensus Mechanisms

i. *PoW*[69]: Proof of Work is for a public permissionless blockchain. Each node (miner) on the network competes with huge computing power to solve the cryptographic puzzle and reach a consensus. As a node solves the puzzle, it broadcasts the block so that other nodes can validate the correctness of the hash value. The miner solving the puzzle at first gets incentives in the form of cryptocurrency. Ethereum, Zcash, Bitcoin use PoW.

ii. *PoS*[70]: Proof of state is for a public permissioned blockchain. The block validators (forgers) are chosen based on the number of virtual coins they possess or put at stake. It supports the fact that more coins at stake give better chance to the node to be selected as one to validate the block of transactions. However, the malicious nodes lose all the coins they put at stake if they try to include faulty transactions (called slashing). It consumes very little computing power when compared to PoW. Peercoin and Ethereum 2.0 via Casper use PoS.

iii. *DPoS*[71]: Delegated Proof of Stake is also for the public permissioned blockchain. It has a certain number of validator nodes called delegates elected from the voting by token holders. Delegates broadcast blocks, resolve consensus issues, distribute block rewards proportionally to their voters and secure the network. The delegates can also be voted out, and so the reputation and loss of income persuade users to act honestly. *BitShares* uses DPoS.

iv. *PBFT*[72]: Practical Byzantine Fault Tolerance is for a consortium permissionless blockchain. However, the private permissioned blockchain such as Hyperledger Fabric also uses this consensus algorithm. The algorithm optimizes aspects of Byzantine Fault Tolerance (BFT) by assuming that some nodes are corrupt. It is to solve the Byzantine General's Problem, which emphasizes all the participants to agree on a single tactic to avoid a catastrophic system failure. It has a list of recommended validators formed by a central authority based on the company's protocol. The validators communicate with one another by sending messages back and forth and use a voting process to confirm a new block of transactions. The block is validated if more than 66% of the validators agree on it. Ripple, Tendermint use PBFT.

v. IBFT[73]: Istanbul Byzantine Fault Tolerance is for a consortium permissioned blockchain with transaction finality. It is inspired by the PBFT consensus algorithm. The system can tolerate most of F faulty nodes in a 3F+1 validator nodes network. All the validators are seen as clients, who pick one of them as the proposer, by default, in a round-robin fashion. It generates a verifiable new block instead of a group of read-write operations to the file system and any valid block remains somewhere in the main chain. Quorum uses IBFT.

vi. *FBA*[74]: Federated Byzantine Agreement is for a private permissionless blockchain. It has an open membership without a list of recommended validators. Each validator chooses other validators they trust to form quorum slices, which eventually overlap to form a network-wide consensus. Swarm (on Stellar) uses FBA.

vii. *PoET*[75]: Proof of Elapsed Time is for a private permissioned blockchain. It also solves the Byzantine General's Problems by using the Intel Software Guard Extensions (SGX) as a trusted execution environment. The validator requests a wait time using secure SGX

[68] https://www.stellar.org/
[69] https://en.bitcoin.it/wiki/Proof_of_work
[70] https://en.bitcoin.it/wiki/Proof_of_Stake
[71] https://en.bitcoinwiki.org/wiki/DPoS
[72] https://en.bitcoinwiki.org/wiki/PBFT
[73] https://github.com/ethereum/EIPs/issues/650
[74] https://en.bitcoinwiki.org/wiki/Stellar_Consensus_Protocol
[75] https://sawtooth.hyperledger.org/docs/core/releases/1.0/architecture/poet.html

instructions and is elected as a leader once the wait time appears to be the shortest. Since it is just about running normal software, there is no reward for the leader as in PoW. It does not consume huge computing power as in PoW. Hyperledger Sawtooth uses PoET.

viii. *Raft*[76]: Raft-based consensus is also used for a private permissioned blockchain such as Hyperledger Sawtooth. It is a Crash Free Tolerance (CFT) rather than Byzantine fault tolerance, so it is suitable for smaller non-Byzantine networks. It sets out the election among candidates to elect a leader, who is recognized by all the followers for a term of arbitrary time. All the nodes exchange messages to reach consensus, and the leader must wait for a majority of nodes to agree on a new block. Only the leader node publishes the blocks, and the leader is replaced if it gets out of time. Therefore, this algorithm is faster, and no forks arise on the Raft-based network.

APPENDIX II

External Library and Solidity Contracts

External Library for Solidity Contracts: Classes.sol

```
pragma solidity >=0.4.22 <0.6.0;
library Classes {
    struct Class {
        mapping(address => bool) enterprisemembers;
    }
    function addEnterprisemember(Class storage self, address addrs) public returns (bool) {
        if (self.enterprisemembers[addrs]){
            return false;
        }
        self.enterprisemembers[addrs] = true;
        return true;
    }
     function removeEnterprisemember(Class storage self, address addrs) public returns (bool) {
        if (!self.enterprisemembers[addrs]){
                return false;
         }
         self.enterprisemembers[addrs] = false;
          return true;
    }
}
```

Smart Contract: EnterpriseClasses.sol

```
pragma solidity >=0.4.22 <0.6.0;
import 'Classes.sol';
contract EnterpriseClasses {
    string public reason;
    Classes.Class admins;
    event Success(address a, string s);
    modifier onlyAdmins(){
        require(admins.enterprisemembers[msg.sender]);
        _;
    }
    constructor () public{
        Classes.addEnterprisemember(admins, msg.sender);
```

[76] https://sawtooth.hyperledger.org/docs/raft/nightly/master/introduction.html

```
    }
    function add(address addrs, string memory purpose) public onlyAdmins {
        if (Classes.addEnterprisemember(admins, addrs)) {
            reason = purpose;
            emit Success(addrs, reason);
        }
    }
    function del(address addrs, string memory purpose) public onlyAdmins {
        if (Classes.removeEnterprisemember(admins, addrs)) {
            reason = purpose;
            emit Success(addrs, reason);
        }
    }
}
```

APPENDIX III

Smart Contract: Incentives.sol

```
pragma solidity >=0.4.22 <0.6.0;
contract DataTrade {
    uint public dataprice;
    address payable public dataseller;
    address payable public consumer;
    enum contractState { S0, S1, S2, S3 }
    contractState public state;
    constructor() public payable {
        dataseller = 0x5378fa11529725cCC491bB6708f9E2F06a1639d5;
        dataprice = msg.value / 2;
        require((2 * dataprice) == msg.value, "Err: Please provide even value for deposit");
        consumer = 0x923c1eDfAdB6332254C83BCbAE85B2cA6b9Bb36e;
    }
    modifier condition(bool _condition) {
        require(_condition);
        _;
    }
    modifier onlyConsumer() {
        require(msg.sender == consumer, "Only consumer can execute it!.");
        _;
    }
    modifier onlyDataseller() {
        require(msg.sender == dataseller, "Only dataseller can perform this operation!");
        _;
    }
    modifier inState(contractState _state) {
        require(state == _state, "Err: Invalid state!");
        _;
    }
    event Cancelled(string info, address entityAddress);
    event BuydataConfirmed(string info, address entityAddress);
    event DataReceived(string info1, address consumer);
    event SettlePayment(string info2, address costumer);
    function confirmPurchase() public  inState(contractState.S0)
        condition(msg.value == (2 * dataprice))  payable {
                emit BuydataConfirmed("Access to data got successful", msg.sender);
                consumer = msg.sender;
                state = contractState.S1;
    }
    function confirmReceived() public onlyConsumer inState(contractState.S1) {
        emit DataReceived("Data availabilty confirmed by:", msg.sender);
        emit SettlePayment("Price for data given to:", dataseller);
        state = contractState.S2;
```

```
        consumer.transfer(dataprice);
        dataseller.transfer(address(this).balance);
    }
    function abort() public onlyDataseller inState(contractState.S0) {
        emit Cancelled("Contract cancelled", msg.sender);
        state = contractState.S3;
        dataseller.transfer(address(this).balance);
    }
}
```

APPENDIX IV

Consent Form I

CONSENT FORM

UNIVERSITY OF SASKATCHEWAN

DEPARTMENT OF COMPUTER SCIENCE
UNIVERSITY OF SASKATCHEWAN
INFORMED CONSENT FORM

You are invited to participate in this study entitled **"Evaluating user acceptance of the blockchain based research data/ artifacts sharing system"**. We want to evaluate a specific user interface for the Artifacts system using a standard usability and usefulness evaluation tool and find out if some responder characteristics (e.g., age, academic experience) influence their perception of usability and usefulness. Findings from this study will be used to improve the system interface design, and as to develop guidelines for design of user interfaces for blockchain-based storage and sharing systems. Please read this form carefully, and feel free to ask the researchers any questions you might have.

Title of Study: Evaluating user acceptance of the blockchain based research data/ artifacts sharing system.

Ethics Application Number: Beh # ID670

Researchers:

- Julita Vassileva, Department of Computer Science, University of Saskatchewan, (306-966-2073), jiv@cs.usask.ca
- Ajay Kumar Shrestha, Department of Computer Science, University of Saskatchewan, (7787262224), ajay.shrestha@usask.ca

Procedure: In this study, your participation will be online with duration between 30 and 45 minutes and is divided into three sessions: 1) Answering a questionnaire about your general experience and understanding of research data sharing systems, your familiarity of blockchain technology, and your attitude to research data sharing systems that use blockchain technology. 2) Exploration of the blockchain-based research data sharing system: register in the system (link to the website) by giving your name, last name and email adderess. You mayconnect your profile from ResearchGate (https://www.researchgate.net/)/ LinkedIn (https://www.linkedin.com/), or ORCID (optional). We will ask you to explore the functionalities of the system for approximately 20 minutes. 3) Finally, you will answer another short questionnaire about perceived ease of use, usefulness and quality of the system, and perceived enjoyment and intention to use the system

Potential Benefits: You may find interesting papers related to your research, besides, you will have a chance to contribute to the advance of the research on research data sharing system.

Potential Risks: There are no known risks in this study.

Confidentiality: The data collected from the two questionnaires will be kept confidential, available only to the researchers (password protected) and used only for the purposes of the study. The data that you shared on the Artifacts system will be publically available, similar to other systems for sharing research data, such as ResearchGate, Mendeley, Orchid or Google Scholar. Both the Artifacts server and the two surveys (hosted by SurveyMonkey, a UofS recommended survey tool) are located in the USA and are subjected to US laws. The privacy of the information you provide is subject also to the laws of those other jurisdictions. By participating in this survey you acknowledge and agree that your information will be stored outside of Canada and may or may not receive the same level of privacy protection.

Dissemination of Results: Results from this study may appear in aggregated form a PhD thesis and articles published in peer reviewed conferences and scientific journals.

Right to Withdraw: Your participation of this survey is voluntary, and you can decide not to participate at any time, or choose not to answer any questions you don't feel comfortable with. Survey responses will remain anonymous. Since the survey is anonymous, once it is submitted it cannot be removed.

Questions: If you have any question regarding the study, please feel free to ask the researchers at any point, including at a later time. This research project has been approved on ethical grounds by the University of Saskatchewan Research Ethics Board. You could call (306) 966-2975 or email Research Ethics Office at ethics.office@usask.ca regarding any questions on your rights as a participant.

Follow-Up: If you would like to know the results of this study, you can contact the researchers.

Consent to Participate: I have read and understood the description provided above; I have been provided with an opportunity to ask questions and my questions have been answered satisfactorily. I consent to participate in the study described above and I understand that I may withdraw this consent at any time. A copy of this consent form has been given to me for my records.

APPENDIX V

Performance Metrics and Analysis

Experimental Setup 1: Latency

I stopped all extra processes except the basic OS processes to run in the background alongside the Multichain daemon to ensure that no other process would affect the experiments. The experiments were carried out on the newly created Multichain nodes. Since MultiChain uses cryptography, it restricts block index and chainstate access to the list of permitted users; so, I created blockchain nodes as fresh ones. The block index maintains information for every block, and where it is stored on disk. The chain state maintains information about the resulting state of validation because of the currently best-known chain. The node parameters were set up as stated before to store the key-value pairs of all the block and state hashes.

The theoretical peak bandwidth of a network connection is fixed as per the technology used. However, the actual number of packets to be sent over the network is affected by higher and lower latencies. Excessive latency prevents data from filling the network pipe, thus decreasing throughput, and limiting the maximum effective bandwidth of a connection. Therefore, I set our goal of the evaluation to retrieve the latency alongside memory consumption in each case.

To observe the effect of the multichain core daemon being stopped and reconnected into the network, I made the scripts that run with the gap of 1 minute for every new observation with the following Multichain commands:

multichain-cli model stop

multichaind model daemon

I observed the latency from the first node Node1, when it connected to another single-node N2 for scenario S1, next when it connected to the other two nodes N2 and N3 for scenario S2 and similarly with other seven nodes N2-N7 for scenario S3 in a total of 20 observations.

For S3, I first recorded the latency from Node1 to connect it with the other 7 nodes in the network and then finally took their average to find the mean latencies for connecting 7 different nodes from Node1.

Experimental Setup 2: Memory Consumption

I carried out another experiment to observe the memory consumption for the nodes when the corresponding multichain core daemon started on the given node. A total of five observations were carried out, one of which is shown in Figure V.1 and Figure V.2. The figures show the total memory usage during the pre- and post-activation of multichain Daemon.

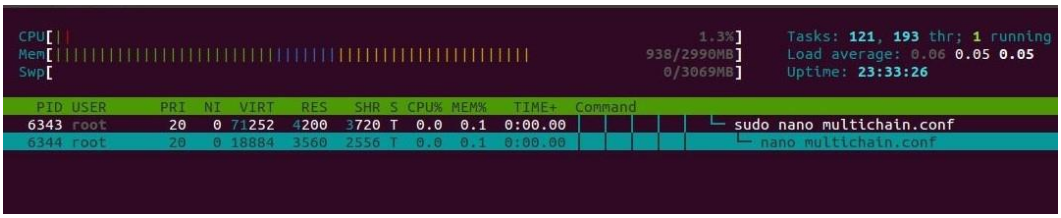

Figure V.1. Memory status at normal state

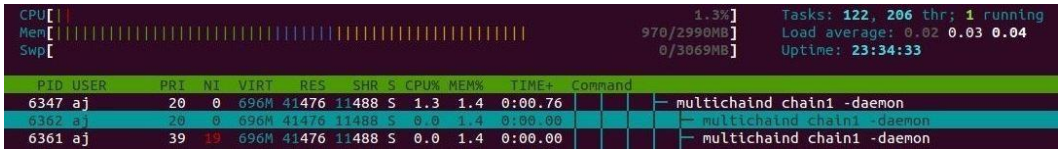

Figure V.2. Memory status after multichain daemon started

Experimental Setup 3: Transaction Cost

To achieve the last goal —evaluating the transactions' validation cost, the experiment involved deploying the smart contracts codes with the Remix IDE on the Ethereum Ropsten Testnet. For the given date, the smart contracts required excess gas fee to deploy the codes and commit the transactions

into the Ethereum blockchain. The transaction cost is paid in its cryptocurrency ether. ETH Gas Station[77] provides three categories of gas prices. They are SafeLow (less than 30 minutes), Standard (less than 5 minutes) and Fast (less than 2 minutes). The gas limit is helpful to optimize the gas used to provide a safety mechanism, as sometimes code with bugs might keep on consuming unnecessary gas for its execution. I used a gas price of 25 Gwei (2.5e-8 ether) which was the then price for achieving a faster transaction. In fact, the cost of a transaction always increases when the gas price goes higher. I have provided one of the instances of the contract creation in this section with the following transaction hashes:

H1: 0xe7b67820af62d3dc5c41357a6de1192de14f8c39cc0416a2830c57c28e3c32b6
H2: 0x8b38cc9b56f584ccec022678f61b16b166e32e581fd0329a394599000af8f226
H3: 0xd333d232646a571be438cda52548503f07047fa07dcce026f18b8df2c88870d0

In this case, a contract was first created at an address "0x2aadf80E4CE7Fc2Db5d57dD975e0337D373e1C50" with the transaction hash H1. Two *ethers* were transferred into the contract from a customer, which was at an address "0x5378fa11529725ccc491bb6708f9e2f06a1639d5". The data price was set as one *ether*. So, the customer must receive three *ethers* at the end of the transactions. The data consumer at an address "0x923c1eDfAdB6332254C83BCbAE85B2cA6b9Bb36e" with the transaction hash H2 transferred two *ethers* to the contract in which one *ether* was the deposit. Finally, the data consumer got the customer data, and then the smart contract transferred one ether to the data consumer and three *ethers* to the customer with the transaction hash of H3. The corresponding logs are given in Figure V.3. More of the details can be obtained using the corresponding transaction hashes on the Live Ropsten Ethereum explorer[78].

Figure V.3. Logs of successful transaction and payment

Result Analysis

The data on latency for the first part of the observation is shown in Table V.1 and Figure V.4. All the scenarios had the minimum and maximum latency around 100ms and 150ms, respectively, giving the average latency time of less than 125ms. Thus, there is no scalability limit in terms of

[77] https://ethgasstation.info/
[78] https://ropsten.etherscan.io/

node count because each node does not need to connect to every other to create a fully connected peer-to-peer network in the private blockchain. Moreover, before conducting the experiments on the real machines, I had initially conducted similar tests on many virtual machines, and I found that our results were similar in both cases: virtual machines and real machines. The only difference with the real physical nodes would be that it reflected the realistic scenario with no network factor influencing the experiments. However, for all the node catch-up time, new nodes joining the chain must replay all transactions from the beginning, and so it could take them significant time before they are up to date. The exact amount of time depends on how many blocks and transactions are in the chain.

Since the experiment had been carried out with only 10 streams having a total of 100 items, the total memory capacity was less than 100MB. This is because we were only concerned with the latency. In addition to that, since no smart contracts are committed into the Multichain, unlike in Ethereum, there is no execution of any automated program for every message on every blockchain node. That surely contributed to the low latency during our observations.

I also analyzed the memory consumption for the Multichain nodes when their core daemons were initiated. I carried out 20 successive observations this time to see the memory consumption, which changed from the initial memory usage of 938 MB to 970 MB after the multichain daemon began. So, it can be concluded that the memory usage was around 28 MB and is not so huge to operate the model with the Multichain blockchain. Moreover, it is also based on the number of unspent transactions. In fact, there are also around 300 bytes of memory already kept for each block in the chain. Therefore, if the node is subscribed to millions of streams, then that would increase memory usage. However, my model focused on storing the user profile data, and even 1 million of those data would have a size of just around 100 MB. So, this model is very effective in terms of a quick start, quick response and less memory consumption.

Table V.1. Latency (milliseconds) summary for three test scenarios

Scenarios	N	Min	Max	Avg.	SD
S1	20	85	159.5	122.57	19.32
S2	20	80	156	126.2	24.24
S3	20	106.86	144.7	127.22	11.01

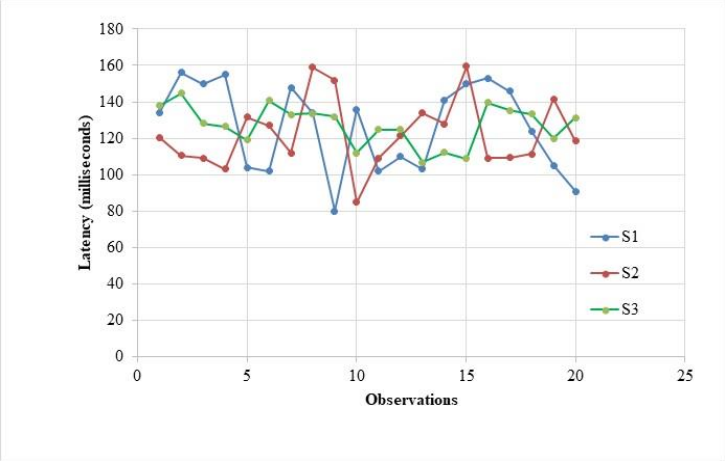

Figure V.4. Latency Test results in a chart for three test scenarios

Furthermore, I analyzed the transaction cost while deploying smart contracts and executing the associated functions. As can be seen from Figure V.5, for instance, the cost to execute the function for payment settlement (transferring ethers from deployed contract to both customer and consumer)

was 0.001247 ether when the higher gas fee for a faster transaction was chosen. This cost was 0.18 USD and is considered as acceptable as per the standard gas fee for Ethereum[79]. All the functions of the contract were tested successfully with the respective role confinement.

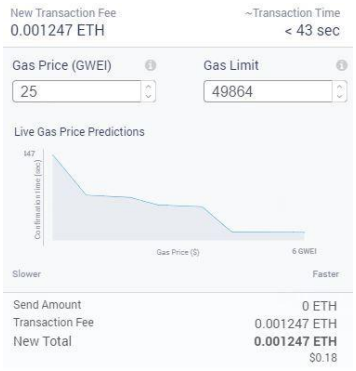

Figure V.5. Tx cost for executing a smart contract function

Table V.2 presents the transaction fees and time used for different functions (including constructors) to change the contract's states. The total cost associated with the contract was 0.022979 ether, which is the sum of the cost for executing constructors and other functions. The cost to execute the function to change the contract state from "Locked" to "Inactive" is higher than that to change the state from "Created" to "Locked". This increase is justified because the former method caused the contract to transfer ether to both data consumer and customer, but the latter made the data consumer transfer ether to the smart contract.

Table V.2. Transaction cost for a smart contract execution

Contract state	**Methods**	**Gas Price (**GWEI)	**Gas used**	**Tx fee (**ether)	**Tx time (**sec)
None to Created	contract_deploy	25	834625	0.020866	< 46
Created to Locked	consumer_buy	25	34639	0.000866	< 43
Locked to Inactive	payment_settle	25	47611	0.001247	< 43

[79] https://ethgasstation.info/

APPENDIX VI

Smart Contract: Customer_cart_share_trade.sol

```
pragma solidity >=0.4.22 <0.6.0;
contract Cart2 {
    uint public dataprice;
    address payable public dataowner;
    address payable public dataconsumer;
    enum State {
        Create, Lock, Release, Inactive
    }
    State public state;
    modifier condition(bool cndn) {
        require(cndn);
        _;
    }
    modifier onlydataconsumer() {
        require(msg.sender == dataconsumer,"Only dataconsumer can execure it!");
        _;
    }
    modifier onlydataowner() {
        require(msg.sender == dataowner,"Only dataowner can perform this operation!");
        _;
    }
    modifier inState(State s) {
        require(state == state, "Invalid State!");
        _;
    }
    event Cancelled();
    event PurchaseConfirmed();
    event ItemAccessed(string datawhat, string conditions, string purpose, address dataowneraddr, uint reward);
    event PayOwner();
    constructor() public payable {
        dataowner = msg.sender;
        dataprice = msg.value / 2;
        require((2 * dataprice) == msg.value, "Please provide even number!");
    }
    function cancel() public onlydataowner inState(State.Create) {
        emit Cancelled();
        state = State.Inactive;
        dataowner.transfer(address(this).balance);
    }

    function confirmPurchase() public inState(State.Create) condition(msg.value == (2 * dataprice)) payable {
        emit PurchaseConfirmed();
        dataconsumer = msg.sender;
        state = State.Lock;
    }
    function confirmAccessed (string memory datawhat, string memory conditions, string memory purpose) public onlydataconsumer
inState(State.Lock) {
        emit ItemAccessed(datawhat, conditions, purpose, dataowner, dataprice);
        state = State.Release;
        dataconsumer.transfer(dataprice);
    }
    function payDataowner() public onlydataowner inState(State.Release) {
        emit PayOwner();
        state = State.Inactive;
        dataowner.transfer(3 * dataprice);
```

```
    }
}
```

Smart Contract: Enterprise_cart_share_trade.sol

```
//for charity, sponsorship, fund raising.sol
pragma solidity >=0.4.22 <0.6.0;
contract Cart2new {
    uint public dataprice;
    address payable public dataowner;
    address payable public dataconsumer;
    address payable public miscellaneous;
    enum State {
        Create, Lock, Release, Inactive
    }
        State public state;

    modifier condition(bool cndn) {
        require(cndn);
        _;
    }
    modifier onlydataconsumer() {
        require(msg.sender == dataconsumer,"Only dataconsumer can execure it!");
        _;
    }
    modifier onlydataowner() {
        require(msg.sender == dataowner,"Only dataowner can perform this operation!");
        _;
    }
    modifier inState(State s) {
        require(state == state, "Invalid State!");
        _;
    }
    event Cancelled();
    event PurchaseConfirmed();
    event ItemAccessed(string datawhat, string conditions, string purpose, address dataowneraddr, uint reward);
    event PayOwner();
    constructor() public payable {
        dataowner = msg.sender;
        dataprice = msg.value / 2;
        require((2 * dataprice) == msg.value, "Please provide even number!");
          //address to be used for donation/sponsorship/charity
        miscellaneous = 0x71129A80fE77492D82DaFE55EE046c850E809006;
    }
    function cancel() public onlydataowner inState(State.Create) {
        emit Cancelled();
        state = State.Inactive;
        dataowner.transfer(address(this).balance);
    }
    function confirmPurchase() public inState(State.Create) condition(msg.value == (2 * dataprice)) payable {
        emit PurchaseConfirmed();
        dataconsumer = msg.sender;
        state = State.Lock;
    }
    function confirmAccessed (string memory datawhat, string memory conditions, string memory purpose) public onlydataconsumer
inState(State.Lock) {
        emit ItemAccessed(datawhat, conditions, purpose, dataowner, dataprice);
        state = State.Release;
        dataconsumer.transfer(dataprice);
        dataowner.transfer(2 * dataprice);
        miscellaneous.transfer(dataprice);
    }
}
```

APPENDIX VII

Consent Form II

UNIVERSITY OF SASKATCHEWAN

CONSENT FORM

You are invited to participate in this study entitled **"Evaluating user acceptance of the blockchain based data sharing framework"**. We intend to evaluate the user acceptance of the system (e.g., usability, satisfaction, privacy, security, trust) and if it depends on certain user characteristics (e.g., age, academic experience). Findings from this study will inform the design of the persuasive approach for blockchain-based data sharing systems, mainly but not limited to sharing of the user data. Please read the form carefully, and feel free to ask the researchers any questions you might have.

Title of Study: Evaluating user acceptance of the blockchain based data sharing system.

Ethics Application Number: 2106

Researchers:

- Julita Vassileva, Department of Computer Science, University of Saskatchewan, (306-966-2073), jiv@cs.usask.ca
- Ajay Kumar Shrestha, Department of Computer Science, University of Saskatchewan, (306-716-6726), ajay.shrestha@usask.ca
- Jennifer Just, Department of Computer Science, University of Saskatchewan, (jsj298@mail.usask.ca)

Procedure: In this study, your participation will be online for 30 minutes and is divided into two sessions:

1) You will be given a brief introduction of how distributed ledgers (blockchains) are used to store and share data. We will ask you to remotely use our system to
    a. Create an account in the system (link to the website), and
    b. Explore the functionalities offered by the system.
2) You will be asked to complete a questionnaire to evaluate the ease of use, usefulness, usability and satisfaction with the system, security, privacy and trust and your intention to use such a system.

Potential Benefits: You may find interesting studies related to your research, people working in areas similar to yours. Besides, you will have a chance to contribute to the design of the persuasive approach for the blockchain based systems. **You may send an email to ajay.shrestha@usask.ca if you wish to know about the study result and/or be contacted for future study.**

Potential Risks: There are no known risks in these **studies**.

Confidentiality: Your participation in the study will be anonymous. The analyzed data will be password protected and will be available only to the researchers involved. This survey is hosted by Survey Monkey. Your data will be stored in facilities hosted in Canada. Please see the following for more information on Survey Monkey's Privacy Policy.

Dissemination of Results: Results from this study will appear in a Ph.D. thesis and articles published in peer-reviewed conferences and scientific journals.

Right to Withdraw: Your participation in this survey is voluntary, and you can decide not to participate at any time or choose not to answer any questions you do not feel comfortable with. The data you share in the system can not be removed since it is stored on a Distributed Ledger technology which creates a permanent record. **You can decide not to participate at any time by closing your browser. Survey responses will remain anonymous. Since the survey is anonymous, once it is submitted it cannot be removed.**

Questions: If you have any questions regarding the study, please feel free to ask the researchers at any point, including at a later time. This research project has been approved on ethical grounds by the University of Saskatchewan Research Ethics Board. You could call (306) 966-2975 or **out of town participants may call toll free 1-888-966-2975**, or email Research Ethics Office at ethics.office@usask.ca regarding any questions on your rights as a participant.

Follow-Up: If you would like to know the results of this study, you can contact the researchers.

Consent to Participate: By completing and submitting this questionnaire, your free and informed consent is implied and indicates that you understand the above conditions of participation in this study.

APPENDIX VIII

Survey Questionnaire for Pretest

Survey Questionnaire for Pretest

* 1. Please provide a random ID. You may generate a random ID on this page.
Keep a note of this ID as you need to put the same ID in the next survey.

* 2. Please indicate your Gender
- Female
- Male
- Other

* 3. Please indicate your Age
- 18 to 24
- 25 to 34
- 35 to 44
- 45 to 54
- 55 to 64
- 65 to 74
- 75 or older

* 4. Please indicate your highest education level
- Graduated from high school
- Bachelors
- Masters
- PhD
- Other (please specify)

5. Which of the following best describes your current occupation/field of study?
- Business and Financial Operations Occupations
- Computer Science
- Engineering
- Life, Physical, and Social Science
- Arts, Design, Entertainment, Sports, and Media
- Education
- Other Sciences
- Other (please specify)

* 6. What continent are you from?
- Africa
- Asia
- Europe
- Oceania
- North America
- South America

* 7. I am familiar with research social networks (e.g. ResearchGate, LinkedIn, ORCID).

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 8. I am familiar with online shopping sites (e.g. Disney Store, Walmart).

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 9. I am familiar with blockchain and smart contracts.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

We want to introduce you to the advantages of blockchain and smart contract technologies, how they can improve the applications you know and are familiar with, namely online shopping. Here is a video that explains blockchain and smart contract technologies on an abstract level (please watch the video with HD quality in full screen).

Summary of the video: **Blockchain is a secure chain of digital ledger existing on multiple computers simultaneously such that no records on the ledger can be erased or edited. It is an unalterable distributed ledger.**
And, the smart contract is a self-executing contract that stores rules to verify and execute the agreed terms for executing some actions (e.g. accessing the records). Thus, these rules can define who gets access to the stored records, under what conditions, for example, for what declared purpose, in exchange of what (payment or virtual credit) and they allow every access to the data to be recorded. With the smart contracts and the blockchain, users can enjoy the increased transparency and protection of data from falling into the wrong hands.

Please answer all the pretest survey questions before using the system.

* 10. I believe that the information I provide to blockchain-based systems will be handled by appropriate processes.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 11. I believe that the information I provide to the blockchain-based systems will be stored securely.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 12. I believe that only legitimate organizations can view the information I provide to blockchain-based systems.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 13. I believe that the blockchain-based system is trustworthy.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 14. I am confident in the blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 15. The blockchain-based system protects my privacy.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 16. The blockchain-based system secures my information.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 17. The blockchain-based system has integrity.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 18. The blockchain-based system is dependable.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 19. The blockchain-based system can be relied on to keep its promises.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 20. I can trust a blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 21. I am familiar with a blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 22. I am aware of which organizations collect information I provide during the use of a blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 23. I am aware of the exact nature of the information that will be collected during the use of a blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 24. I believe that the information I put on a blockchain-based system can not be misused.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 25. I believe that the blockchain accounts I use can not be intercepted by someone else.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

This page measures your general behavior towards using online services.

* 26. I only register for web services that have a privacy policy.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 27. I read the privacy policy of a web service before I register.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 28. I look for a privacy certification of a web service before I register.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 29. I read license agreements fully before I agree to them.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

This page measures your general behavior towards using online services.

* 30. I regularly remove browser cookies.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 31. I regularly use a pop-up window blocker.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 32. I regularly check the computer for spyware.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 33. I regularly clear my browser history.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 34. I believe that using the blockchain-based system would be beneficial for me.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 35. In my opinion, it would be desirable for me to use the blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 36. It would be good for me to use the blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 37. Do you have any other comments, questions, or concerns?

APPENDIX IX

Survey Questionnaire for SCS

Survey Questionnaire for Shopping Cart System (SCS)

* 1. Please provide the same ID that you generated for the pretest survey

* 2. I am familiar with blockchain and smart contracts.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

After completing the pretest survey and testing the first part of the system, you may now complete this questionnaire. If you want to see the video on how to use the first part of the system, please click this video link (please watch it with HD quality in full-screen mode).

Summary of the video: **The proposed system enables customers to leverage the opportunity to receive features provided by the blockchain and smart contract technologies. They can obtain the provenance of every transaction, share their data as per their preferences and receive incentives for sharing it.**

Blockchain and smart contracts based system can support users in the following ways:
(1) Allow the users to specify the purposes of data sharing, which kinds of data that can be shared, and which applications or institutions can access the data;
(2) Give the users full transparency over who accesses their data, when and for what purpose;
(3) Provide an incentive to users for sharing their data (in terms of payment for the use of the data by applications, as specified by the contracts).

Furthermore, the second part of the system (which is not shown here) allows companies such as Online Shopping-Cart Enterprise to share user data among other companies in their consortium network via blockchain that incentivizes every customer with micro-payment for the use of their data.

Please answer the following questions to evaluate the first part of the proposed framework which is an example use case of the user-controlled privacy-preserving user data sharing framework based on blockchains and smart contracts.

* 3. Learning to operate this system is easy.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 4. I find it easy to get this system to do what I want it to do.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 5. My interaction with this system is clear and understandable.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 6. I find this system to be flexible to interact with.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 7. I feel it is easy to become skillful at using this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 8. I find this system easy to use.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 9. Using this system would improve performance in online transaction with transparency over privacy.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 10. Using this system would increase productivity in online transaction with more control over privacy.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 11. Using this system would increase effectiveness in privacy policy formulation.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 12. Using this system would make it easier for me to set data sharing preferences.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 13. Using this system would make it easier for me to receive incentives for sharing my data.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 14. I find this system useful for setting my data sharing preferences.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 15. I am satisfied with how the system lets me set the data sharing preferences.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 16. I am satisfied with how the system can create proof of data sharing choices.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 17. I am satisfied with how the system lets me decide how to share data with different companies/applications.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 18. I am satisfied with how the system gives incentives for sharing data.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 19. I would like to use this system to set data sharing preferences.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 20. I would like to use this system to receive incentives for sharing my data.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 21. I would enjoy using this system when I need to use it.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 22. It is worthwhile to use this system to set data sharing preferences.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 23. I will use this system to decide how my data is shared.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 24. I intend to use this system to share data and receive incentives for sharing it.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 25. I believe appropriate processes will handle the information I provide with blockchain.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 26. I believe that the information I provide will be stored securely.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 27. I believe that only legitimate organizations can view the information I provide to the blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 28. I believe that this blockchain-based system is trustworthy.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 29. This system can be relied on to keep its promises.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 30. This system is dependable.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 31. This system has integrity.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 32. This system protects my privacy.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 33. This system secures my information.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 34. I am familiar with this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 35. I am confident in this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 36. I can trust this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 37. I am aware of which organizations collect information I provide during the use of this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 38. I am aware of the exact nature of the information that will be collected during the use of this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 39. I believe that the information I put on this system can not be misused.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 40. I believe that the blockchain accounts that I use on this system can not be intercepted by someone else.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 41. I believe that using the blockchain-based system would be beneficial for me.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 42. In my opinion, it would be desirable for me to use the blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 43. It would be good for me to use the blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 44. Do you have any other comments, questions, or concerns?

APPENDIX X

Survey Questionnaire for DSS

Survey Questionnaire for Data Sharing System (DSS)

* 1. Please provide **the ID** that you generated for the pretest survey

* 2. I am familiar with blockchain and smart contracts.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

After completing the pretest survey and testing the second part of the system, you may now complete this questionnaire. If you want to see the video on how to use the second part of the system, please click this video link (please watch it with HD quality in full-screen mode).

Summary of the video: **The proposed system enables enterprises to leverage the opportunity to receive features provided by the blockchain and smart contract technologies. This is the second part of the system and it allows companies such as Online Shopping-Cart Enterprise to share user data among other companies in their consortium network via blockchain and incentivize every customer with micro-payment for the use of their data. They can also obtain the provenance of every transaction.**

Please answer the following questions to evaluate the second part of the proposed framework which is an example use case of the user-controlled privacy-preserving user data sharing framework based on blockchains and smart contracts.

* 3. Learning to operate this system is easy.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 4. I find it easy to get this system to do what I want it to do.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 5. My interaction with this system is clear and understandable.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 6. I find this system to be flexible to interact with.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 7. I feel it is easy to become skillful at using this system.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 8. I find this system easy to use.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 9. Using this system would enable me to share data more quickly in the consortium network.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 10. Using this system would improve performance with data sharing.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 11. Using this system would increase productivity with data sharing.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 12. Using this system would increase effectiveness with data sharing.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 13. Using this system would make it easier to share and access data by incentivizing the data owners.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 14. I find this system useful for sharing and accessing user data by incentivizing the data owners.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 15. I am satisfied with the feature of sharing user data in the consortium blockchain network.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 16. I am satisfied with the methods of accessing user data by providing incentives to the data owners.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 17. I am satisfied with how the system grants access to user data after fulfilling the terms of smart contracts.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 18. I am satisfied with how the system creates a proof of the existence of transactions via smart contracts.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 19. I would like to use this system to share user data.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 20. I would like to follow smart contracts for accessing user data.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 21. I would enjoy using this system when I need to use it.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 22. It is worthwhile to use this system to share user data and provide incentives to the data owners for accessing their data.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 23. I will use this system to share user data and provide incentives to the data owners for accessing their data

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 24. I intend to use this system to share user data and provide incentives to the data owners for accessing their data.

strongly disagree | moderately disagree | slightly disagree | neither | slightly agree | moderately agree | strongly agree

* 25. I believe appropriate processes will handle the information I provide with blockchain.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 26. I believe that the information I provide will be stored securely.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 27. I believe that only legitimate organizations can view the information I provide to this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 28. I believe that this blockchain-based system is trustworthy.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 29. This system can be relied on to keep its promises.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 30. This system is dependable.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 31. This system has integrity.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 32. This system protects my privacy.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 33. This system secures my information.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 34. I am familiar with this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 35. I am confident in this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 36. I can trust this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 37. I am aware of which organizations access the user data during the use of this blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 38. I am aware of the exact nature of the information that will be collected during the use of this system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 39. I believe that the information I put on this system can not be misused.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 40. I believe that the blockchain accounts that I use on this system can not be intercepted by someone else.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 41. I believe that using the blockchain-based system would be beneficial for me.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 42. In my opinion, it would be desirable for me to use the blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 43. It would be good for me to use the blockchain-based system.

strongly disagree	moderately disagree	slightly disagree	neither	slightly agree	moderately agree	strongly agree
○	○	○	○	○	○	○

* 44. Do you have any other comments, questions, or concerns?

APPENDIX XI

Demographics Data and Factor Analysis

Table XI.1. Demographics of participants

Criterion	Subgroup	Number (#)		Percentage (%)	
		SCS	DSS	SCS	DSS
Gender	Female	15	11	24	22
	Male	48	39	76	78
	Other	0	0	0	0
Age	18-24	6	4	9.5	8
	25-34	45	34	71.4	68
	35-44	11	9	17	18
	44-54	1	3	1.6	6
Highest Education Completed	High school	1	0	1.6	0
	Bachelors	13	12	20.6	24
	Masters	39	31	61.9	62
	PhD	10	7	15.9	14
Area (Expertise)	Business	2	1	3.2	2
	Comp Science	39	36	61.9	72
	Engineering	20	13	31.7	26
	Social Science	2	0	3.3	0
Continent	Africa	3	2	4.8	4
	Asia	30	20	47.6	40
	Europe	5	3	7.9	6
	N. America	19	18	30.2	36
	Oceania	6	7	9.5	14
	S. America	0	0	0	0
Familiarity with blockchain and smart contracts	Extremely	12	14	19	28
	Moderately	20	19	31.7	38
	Slightly	30	15	47.6	30
	Neither	0	1	0	2
	Slightly Not	1	0	1.6	0
	Moderately Not	0	0	0	0
	Extremely Not	0	1	0	2

Table XI.2. Exploratory factor analysis

Construct	Item	Factor Loading		
		Pretest	SCS	DSS
Attitudinal Privacy	AP1	0.835	0.812	0.864
	AP2	0.808	0.862	0.864
	AP3	0.88	0.774	0.685
	AP4	0.854	0.842	0.816
Attitudes Towards System	ATS1	0.964	0.938	0.948
	ATS2	0.971	0.955	0.943
	ATS3	0.962	0.942	0.954
Beh Privacy-General Caution	BP-GC1	0.775		
	BP-GC2	0.92		
	BP-GC3	0.888		
	BP-GC4	0.882		
Beh Privacy-Technical Protection	BP-TP1	0.605		
	BP-TP2	0.869		
	BP-TP3	0.775		
	BP-TP4	*0.39*		
Security	PS1	0.919	0.891	0.899
	PS2	0.922	0.858	0.934
	PS3	0.883	0.859	0.893
Trust	T1	0.854	0.821	0.88
	T2	0.877	0.792	0.899
	T3	0.858	0.721	0.809
	T4	0.82	0.765	0.82
	T5	0.766	0.753	0.783
	T6	0.727	0.786	0.8
	T7	0.735	0.696	0.734
	T8	0.924	0.806	0.817
	T9	0.673	0.819	0.931
Intention to Use	ITU1		0.89	0.804
	ITU2		0.917	0.936
	ITU3		0.854	0.926
Perceived Ease of Use	PEOU1		0.817	0.885
	PEOU2		0.818	0.888
	PEOU3		0.798	0.883
	PEOU4		0.805	0.809
	PEOU5		0.866	0.755
	PEOU6		0.876	0.865
Perceived Usefulness	PU1		0.743	0.857
	PU2		0.685	0.843
	PU3		0.661	0.883
	PU4		0.856	0.871
	PU5		0.725	0.811
	PU6		0.73	0.749
Quality of System	QOS1		0.88	0.896
	QOS2		0.901	0.943
	QOS3		0.83	0.872
	QOS4		0.71	0.868

APPENDIX XII

Summary of The Survey Data

Table XII.1. Survey data for pretest

Indicators	No.	Missing	Mean	Median	Min	Max	Standard Deviation	Excess Kurtosis	Skewness
PS1	1	0	5.88	6	3	7	0.986	0.708	-0.96
PS2	2	0	5.74	6	2	7	1.22	0.225	-0.904
PS3	3	0	5.19	6	1	7	1.513	-0.074	-0.807
T1	4	0	5.99	6	4	7	0.83	-0.505	-0.408
T2	5	0	5.74	6	4	7	0.945	-0.75	-0.34
T3	6	0	5.45	6	1	7	1.243	1.375	-0.952
T4	7	0	5.8	6	1	7	1.078	3.779	-1.301
T5	8	0	6	6	4	7	0.915	-0.612	-0.54
T6	9	0	5.64	6	3	7	1.06	-0.762	-0.468
T7	10	0	5.76	6	1	7	1.076	3.75	-1.357
T8	11	0	5.77	6	4	7	0.966	-0.724	-0.436
T9	12	0	5.2	5	1	7	1.345	0.405	-0.754
AP1	13	0	4.81	5	2	7	1.458	-1.006	-0.09
AP2	14	0	4.82	5	1	7	1.501	-0.3	-0.55
AP3	15	0	5.05	6	1	7	1.593	-0.454	-0.705
AP4	16	0	5.3	6	2	7	1.495	-0.807	-0.576
BP-GC1	17	0	5.32	6	1	7	1.603	0.588	-1.05
BP-GC2	18	0	4.22	4	1	7	1.773	-0.927	-0.306
BP-GC3	19	0	4.49	5	1	7	1.854	-0.909	-0.238
BP-GC4	20	0	3.5	3	1	7	1.803	-0.979	0.29
BP-TP1	21	0	5.01	5	1	7	1.681	0.094	-0.981
BP-TP2	22	0	5.93	6	2	7	1.082	1.655	-1.104
BP-TP3	23	0	5.3	6	1	7	1.625	0.062	-0.939
BP-TP4	24	0	5.14	6	1	7	1.765	0.129	-1.052
ATS1	25	0	5.68	6	2	7	1.092	0.292	-0.717
ATS2	26	0	5.57	6	2	7	1.128	0.036	-0.633
ATS3	27	0	5.61	6	3	7	1.076	-0.897	-0.353

Table XII.2. Survey data for SCS

	No.	Missing	Mean	Median	Min	Max	Standard Deviation	Excess Kurtosis	Skewness
PEOU1	1	0	6.254	6	5	7	0.689	-0.855	-0.389
PEOU2	2	0	6.063	6	5	7	0.774	-1.334	-0.112
PEOU3	3	0	6.19	6	3	7	0.94	4.156	-1.804
PEOU4	4	0	6.063	6	1	7	0.906	14.264	-2.75
PEOU5	5	0	6.365	6	4	7	0.741	1.525	-1.197
PEOU6	6	0	6.365	6	5	7	0.697	-0.726	-0.652
PU1	7	0	6.032	6	3	7	0.975	0.632	-1.011
PU2	8	0	6	6	3	7	0.926	0.644	-0.861
PU3	9	0	5.921	6	1	7	1.088	5.842	-1.883
PU4	10	0	6.365	6	4	7	0.674	1.057	-0.925
PU5	11	0	6.381	7	5	7	0.7	-0.692	-0.7
PU6	12	0	6.413	7	5	7	0.705	-0.6	-0.797
QOS1	13	0	6.19	6	4	7	0.794	0.092	-0.752
QOS2	14	0	6.238	6	4	7	0.728	0.052	-0.662
QOS3	15	0	6.381	6	4	7	0.677	1.088	-0.969
QOS4	16	0	6.397	7	4	7	0.746	0.439	-1.045
Enj1	17	0	6.079	6	3	7	0.803	2.421	-1.091
Enj2	18	0	6.302	6	4	7	0.769	0.805	-1.017
Enj3	19	0	6.19	6	4	7	0.814	-0.115	-0.733
ITU1	20	0	6.063	6	4	7	0.833	-0.131	-0.628
ITU2	21	0	6.063	6	4	7	0.889	-0.273	-0.684
ITU3	22	0	6.063	6	3	7	0.871	1.847	-1.161
PS1	23	0	5.73	6	3	7	0.84	0.596	-0.434
PS2	24	0	5.841	6	4	7	0.963	-0.75	-0.436
PS3	25	0	5.794	6	3	7	1.041	0.035	-0.694
T1	26	0	6.111	6	5	7	0.645	-0.582	-0.111
T2	27	0	6.016	6	3	7	0.917	1.133	-1.044
T3	28	0	5.81	6	3	7	0.99	-0.159	-0.609
T4	29	0	6.159	6	4	7	0.801	-0.077	-0.68
T5	30	0	5.937	6	3	7	0.871	0.894	-0.761
T6	31	0	5.905	6	4	7	0.849	-0.336	-0.451
T7	32	0	5.841	6	3	7	0.912	0.46	-0.703
T8	33	0	5.889	6	3	7	0.875	0.653	-0.652
T9	34	0	5.921	6	4	7	0.841	-0.216	-0.501
AP1	35	0	5.762	6	4	7	0.904	-0.82	-0.163
AP2	36	0	5.667	6	3	7	1.054	0.167	-0.704
AP3	37	0	5.397	5	3	7	1.134	-0.793	-0.239
AP4	38	0	5.492	6	3	7	1.344	-0.589	-0.767
ATS1	39	0	6.048	6	4	7	0.765	0.028	-0.519
ATS2	40	0	5.952	6	4	7	0.785	-0.406	-0.317
ATS3	41	0	6	6	4	7	0.797	-0.419	-0.386

Table XII.3. Survey data for DSS

	No.	Missing	Mean	Median	Min	Max	Standard Deviation	Excess Kurtosis	Skewness
PEOU1	1	0	5.96	6	3	7	0.916	0.63	-0.724
PEOU2	2	0	5.72	6	3	7	1.217	0.181	-0.945
PEOU3	3	0	5.86	6	4	7	0.895	-0.849	-0.232
PEOU4	4	0	5.48	6	2	7	1.22	0.019	-0.599
PEOU5	5	0	5.94	6	3	7	1.047	1.686	-1.278
PEOU6	6	0	5.72	6	2	7	1.201	1.163	-1.077
PU1	7	0	5.9	6	3	7	1.204	0.535	-1.148
PU2	8	0	5.62	6	1	7	1.384	1.078	-1.144
PU3	9	0	5.7	6	1	7	1.345	1.562	-1.204
PU4	10	0	5.88	6	4	7	0.993	-0.772	-0.51
PU5	11	0	6.28	6	4	7	0.776	1.143	-1.077
PU6	12	0	6.34	7	4	7	0.764	0.33	-0.959
QOS1	13	0	5.82	6	3	7	1.052	-0.263	-0.687
QOS2	14	0	6.16	6	4	7	0.784	-0.403	-0.555
QOS3	15	0	6.14	6	4	7	0.825	-0.776	-0.493
QOS4	16	0	6.42	7	4	7	0.666	1.98	-1.162
Enj1	17	0	5.9	6	3	7	0.985	0.356	-0.829
Enj2	18	0	6.1	6	4	7	0.831	-0.22	-0.626
Enj3	19	0	5.92	6	3	7	1.017	0.279	-0.894
ITU1	20	0	6.16	6	3	7	0.88	2.132	-1.234
ITU2	21	0	5.92	6	2	7	1.036	2.685	-1.281
ITU3	22	0	5.94	6	2	7	1.121	2.241	-1.372
PS1	23	0	5.94	6	3	7	0.925	0.751	-0.814
PS2	24	0	5.96	6	2	7	1.038	3.738	-1.575
PS3	25	0	5.9	6	3	7	1.1	0.878	-1.097
T1	26	0	6.3	6	4	7	0.755	0.271	-0.863
T2	27	0	6.22	6	4	7	0.782	-0.241	-0.678
T3	28	0	6.06	6	3	7	0.968	0.985	-1.08
T4	29	0	6.32	6	4	7	0.76	0.295	-0.91
T5	30	0	5.96	6	3	7	0.916	2.43	-1.208
T6	31	0	5.9	6	3	7	1.005	0.127	-0.768
T7	32	0	6.04	6	4	7	0.894	0.001	-0.775
T8	33	0	5.9	6	3	7	0.964	0.368	-0.759
T9	34	0	6.1	6	4	7	0.831	-0.22	-0.626
AP1	35	0	6.3	7	4	7	0.806	-0.206	-0.85
AP2	36	0	6.16	6	3	7	0.946	1.429	-1.211
AP3	37	0	5.32	6	1	7	1.378	0.58	-0.937
AP4	38	0	5.56	6	2	7	1.344	0.089	-0.82
ATS1	39	0	5.92	6	4	7	0.891	-0.713	-0.363
ATS2	40	0	5.76	6	3	7	1.05	-0.3	-0.67
ATS3	41	0	5.9	6	4	7	1.005	-0.603	-0.646