Master's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Architecture and Design
Department of Design

Daniel Fernando Arevalo Espinel

How Usable are Block Explorers for Attesting Agreements Registered on Blockchains?

Evaluating the Learnability of a Blockchain Product's User Interface for Verifying Records of Cryptographically Certified Documents on the Ethereum Blockchain

Master's thesis in Interaction Design
Supervisor: Mariusz Nowostawski and Miriam Begnum
June 2019

NTNU
Norwegian University of
Science and Technology

Daniel Fernando Arevalo Espinel

How Usable are Block Explorers for Attesting Agreements Registered on Blockchains?

Evaluating the Learnability of a Blockchain Product's User Interface for Verifying Records of Cryptographically Certified Documents on the Ethereum Blockchain

Master's thesis in Interaction Design
Supervisor: Mariusz Nowostawski and Miriam Begnum
June 2019

Norwegian University of Science and Technology
Faculty of Architecture and Design
Department of Design

Abstract

Blockchain is a technology that makes it extremely difficult for any single one of the participants in a decentralized network to change what has been recorded at the end of a distributed ledger as time passes. The unfamiliarity of people with products utilizing Blockchain and the complexity of their user interface increase the cognitive friction during a user's interaction and diminish the adoption of this technology. Researchers, designers, and developers lack insights about block explorers' usability for implementing them as tools that assist on the corroboration of cryptographically certified documents with hashing algorithms that are timestamped via transactions recorded in immutable and distributed ledgers. Grounded on these challenges and lack of research about the topic, this thesis applies a usability engineering approach to propose a framework for learnability evaluation and utilizes it in a methodology to examine such important aspect of usability for the user interface of Blockchain explorers. Specifically, evaluating the learnability of these products during the first-time experience of users with varying levels of expertise and domain knowledge about Blockchain. First, a summative expert assessment gives a rating of learnability to three block explorers available in the market. Then, formative empirical testing is conducted on testers verifying data recorded in a transaction of the Ethereum Blockchain, validating a contractual document's fingerprint obtained with the cryptographical hashing algorithm SHA-256. The outcome of this research allows concluding that the impact of neglecting the recommendations for a user interface's suitability for learning might be more significant for products that utilize Blockchain than it is for products that do not require it altogether. This also suggests that a focus on improving the learnability of UIs can increase the adoption of solutions that implement Blockchain in an attempt to cover real-world needs.

Preface

The following thesis is the outcome of my last semester as a student of the Master of Science in Interaction Design program at the Norwegian University of Science and Technology (NTNU). The literature review, planning, and research were conducted during the spring of 2019 for a workload corresponding to 30 ECTS.

I contemplate the possibility that this thesis contributes to the field of human-computer interaction by proposing a framework for the evaluation of learnability in user interfaces that display the information recorded on Blockchains. The framework fundaments an iterative design methodology to explore and assess such important aspect of usability on Blockchain Products, giving insights to designers and developers of these solutions that want to improve them, impacting the increase of their ease-of-use and adoption.

In general, the user interface of block explorers helps to verify all data recorded on a Blockchain and provides evidence of the utility of decentralized applications. Particularly, this research is about product user interfaces that display information of a transaction recorded on a Blockchain for attesting the existence and immutability of contractual data as established during an agreement. Researching the learnability issues and challenges that users have when interacting with them can contribute with insights that support designers and developers working to improve the usability of other blockchain products.

I want to thank the following people whose help, guidance, and support encouraged me to complete the work of this thesis project and finish the Master of Science program:

My supervisor, Mariusz Nowostawski and co-supervisor Miriam Begnum for their splendid advice and very knowledgeable input/feedback about the proper conduction of research, going all the way from the selection of the topic to the methods used and the collection and analysis of data. Also, for their support which encouraged me to examine a topic that was for me as complex as it was relevant, therefore representing an excellent opportunity to put my skills at use for the completion of the thesis.

Finally, my partner in life Susanne Wesner and our wonderful son Nathan whose emotional support and capacity to bring joy to my life motivate me always to give my best, and enjoy doing so, regardless of the hardships and pressure of any challenge I decide to take on. I dedicate this and many more future accomplishments to both of you and our family.

Table of Contents

List of Abbreviations (or Symbols) x
List of Figures xi
List of Tables xi
1 Introduction 1
1.1 Justification, Motivation and Benefits 1
1.2 Research Questions 2
1.3 Contributions 3
1.4 Thesis Outline 4
2 Theoretical Background 6
2.1 Distributed Ledger Technologies 7
2.1.1 Distributed Systems 8
2.1.2 Consensus 9
2.1.1 Cryptographic hashing 9
2.1.2 Decentralization 11
2.2 Blockchain 12
2.2.1 Accounts and Transactions 12
2.2.2 Blocks and Data Structure 13
2.2.3 Mining and Proof-of-Work 15
2.3 The Application of Blockchain 16
2.3.1 The Utility of Smart Contracts 16
2.3.2 Certifying Contractual Documents with Blockchain 17
2.3.2.1 MIT Academic Certificates 18
2.3.2.2 Stamp.io By Stampery 19
2.3.2.3 Proof.ink on the Steem Blockchain 20
2.4 Blockchain Explorers and Wallets 21
2.4.1 Verifying Certificates with a Block Explorer 21
2.4.1.1 EtherScan.io 22
2.4.1.2 BlockScout.com 22
2.4.1.3 EthStats.io (now Aleth.io) 23
2.5 Usability for Blockchain Products 23
2.5.1 Defining Usability 24
2.5.1.1 Usability as the Ease of Use 24
2.5.1.2 Usability Goals 25
2.5.2 Learnability 25
2.5.2.1 Measuring Initial Learnability 25

2.5.3 Learning is a Cognitive Process 26
2.5.3.1 The Gulfs of Evaluation and Execution 27
2.5.3.2 Crossing the Distance Between User and System 28
3 Methodology 29
3.1 Evaluating Learnability 30
3.1.1 Suitability for Learning of User Interfaces 31
3.1.1.1 Learnability Heuristics 31
3.1.1.2 User's Cognitive Activities 32
3.1.1.3 Cognitive Model for Learnability 33
3.2 Context of Use 34
3.2.1 Users 35
3.2.1.1 Sampling 35
3.2.1.2 Screening Questionnaire 36
3.2.1 Environment and Resources 37
3.2.2 Goal and Tasks 37
3.2.3 Product 38
3.3 Summative Expert Assessment 39
3.3.1 Usability Inspection Methods 39
3.3.2 Framework for Learnability Evaluation 40
3.4 Formative Empirical Testing 43
3.4.1 Cognitive Task Analysis 44
3.4.2 Post-test Semi-structured Interview 46
4 Analysis 47
4.1 Part A: Summative Expert Assessment 47
4.2 Part B: Formative Empirical Testing 48
4.2.1 Results of Screening Questionnaire 48
4.2.2 Cognitive Task Analysis Results 49
4.2.3 Post-Test Interview Results 54
5 Discussion 61
5.1 Implementing the Learnability Evaluation Framework 61
5.2 The importance of Learnability for Blockchain products 63
5.3 Confounding Variables 67
5.4 Research Limitations 68
5.5 Summary 69
6 Conclusion 72
6.1 Future Research 73
Reference List 75

Appendices 82

List of Abbreviations (or Symbols)

UI	User Interface
DLT	Distributed Ledger Technology
PDF	Portable Document Format
UIM	Usability Inspection Method
ISO	International Organization for Standardization
DAO	Decentralized Autonomous Organization
MIT	Massachusetts Institute of Technology
TRE	Technology Review Editors
dApps	Decentralized Applications
TA	Think Aloud
CTA	Cognitive Task Analysis
SEA	Summative Expert Assessment
NIST	National Institute of Standards and Technology

List of Figures

Figure 2.1: Centralized Versus Distributed Ledger Technologies. 7
Figure 2.2: Outputs of a cryptographic hashing function using SHA-256. 10
Figure 2.3: Blockchain Accounts and Transactions. 13
Figure 2.4: Blocks and Merkle Tree Data Structure. 14
Figure 2.5: Canonical Blockchain and Global States. 14
Figure 2.6: Examples of the Digital Certificates Project (from Nazare et al. 2016). 18
Figure 2.7: Architecture of Blockchain Certification System (from Nazare et al. 2016). . 18
Figure 2.8: Stamp.io service's screen for uploading files. 19
Figure 2.9: Digital Certificate and Download Screen of Stamp.io service. 19
Figure 2.10: Digital Fingerprint and Storing process on the Steem blockchain. 20
Figure 2.11: Signed Transaction with Steem Wallet and Proof.ink's page confirmation. . 20
Figure 2.12: EtherScan's User Interface. Home (left) and search result (right) pages. .. 22
Figure 2.13: BlockScout's User Interface. Home (left) and search result (right) pages. . 22
Figure 2.14: Aleth.io's User Interface. Home (left) and search result (right) pages. 23
Figure 2.15: Gulfs of Evaluation and Execution (Norman, 1986). 27
Figure 3.1: Seven user cognitive activities when performing of a task (Norman, 1986). 32
Figure 3.2: A Cognitive Model for Learnability of User Interfaces 34
Figure 3.3: Usability in a context of use (ISO, 2018, p. 7) 35
Figure 3.4: Criteria for Tester selection with Screening Questionnaire 36
Figure 3.5: Flow Diagram of the Cognitive Task Analysis and Sections of the Test 45
Figure 4.1: Results of Screening Questionnaire for the Selection of Testers 49
Figure 4.2: EthStats.io Block Explorer Home Page and Tutorial Pop-up 52
Figure 4.3: EthStats.io Block Explorer Home Page 52
Figure 4.4: EthStats.io Block Explorer Search Result Page 54

List of Tables

Table 3.1: Feature comparison between Blockchain certification services 38
Table 3.2: Framework for Learnability Evaluation from a Feature Analysis approach 41
Table 3.3: Scoring model for expert assessment according to the analysis of features .. 42
Table 3.4: Total scores per category 42
Table 3.5: Acceptance criteria for the results of Expert Assessment 43
Table 3.6: Post-test Questionnaire for Semi-structured Interview 46
Table 4.1: Expert Assessment of Block Explorers 47
Table 4.2: Total Scores of Expert Assessment per Category 48
Table 4.3: Results of Cognitive Task Analysis 50

1 Introduction

Blockchain is a technology with a protocol embedded in its code that makes it extremely difficult for any single one of the participants in a decentralized network to change what has been recorded at the end of a distributed ledger as time passes (Barber et al. 2012). The decentralized and immutable nature of the Blockchain technology provides the guarantees that the parties of an agreement require to register tamperproof information without the need of involving intermediaries. Moreover, cryptographic algorithms protect the data recorded inside transactions against the collusion and attacks from adversaries, with a mathematically-proven elevated probability of remaining unchanged (National Institute of Standards and Technology, 2012).

In 2009, Bitcoin became the first example of Blockchain used at a global scale to make people agree about the value of a shared asset that they can exchange (Nakamoto, 2008). In 2014, an implementation of Blockchain named Ethereum did not only let people exchange the value of an asset but also allowed to logically program data exchanged between participants of the network via transactions grouped in blocks (Wood & Buterin, 2014). From looking at these two valuable use cases of Blockchain, we can assess that the main application of this technology is ensuring with high certainty that information commonly accepted as the truth by the majority of participants in a network is copied, distributed and shared by them.

As a result, many companies are exploring the virtues of utilizing Blockchain to remove middlemen for diminishing transaction costs incurred when certifying agreements or exchanging items of value (Cisco Systems, 2018). However, independently of the latent capacity of Blockchain products to enable systemic changes that replace centralized solutions, the majority of companies might be obviating the inertia to change that customers have for established solutions. For this reason, designers and developers of these new Blockchain-based solutions are obligated to recognize and deal with the high cognitive load and even behavioural change required by end-users of this technology.

Currently, the most accepted use cases of Blockchain come in the form of cryptocurrency, but Blockchain technology has been gaining the attention from entrepreneurs, developers, and designers who in the past years have been creating other applications in which the characteristics of Blockchain may be useful (ConsenSys, 2019). However, as it happens with many new technologies, the applications of Blockchain in products that attempt to cover real-world needs are failing to become adopted by a critical mass of people.

1.1 Justification, Motivation and Benefits

Blockchain products are those digital solutions that utilize Blockchain as a part or the entirety of the back-end supporting the functionality offered by their user interface (UI). A potential explanation for their low adoption is that various Blockchain products have been released before ensuring that they create more value to their users than existing solutions. However, another problem could be that even if incorporating Blockchain to a solution creates value, the inherent complexity, and technicalities of the technology might cloud the perception from users concerning its value, consistently deterring them from adoption. After all, the functionality and value perceived in digital products have become dependent

on the cognitive friction that the user experiences when trying to achieve a goal at using them (Cooper, 2004).

Although some people argue that some friction might have a particular value in design (Weaver, 2019), Blockchain products have to be designed for compensating the high friction created by the complexity of concepts integrated from cryptography and distributed systems. Consequently, to increase the ease of use of a product, designers have to employ widely adopted concepts of usability in order to "fix" products that utilize Blockchain. They evaluate the impact of each element included in a user interface that might hinder the user and increase the cognitive effort at operating it. However, the focus of many companies is directed toward aspects of usability that established products need to improve in order to compete in the market such as Effectiveness, Efficiency, and Satisfaction (Rusu et al. 2015), obviating the importance of designing for the ease of use of a product, so it has the chance of being adopted in the first place. In this regard, this thesis takes Nielsen's definition of web usability (Nielsen & Loranger, 2006), which is comprised by five quality components—Learnability, Efficiency, Memorability, Errors, and Satisfaction— in function of attaining such needed ease-of-use (Nilsen, 2012).

Due to the novelty of the Blockchain and its application, there have been few considerations about evaluating the usability of blockchain products beyond the ISO 25010 standard definition (International Organization for Standardization, 2011, p. 12). In such a context, it becomes challenging to identify if there are differences between the impact of usability on products that utilize Blockchain and the one in those that do not require it altogether. The motivation behind this thesis is to help designers and developers with a methodology for evaluating the usability of Blockchain products and improving their ease-of-use during the design process. For that reason, the research focuses first on defining a framework for the evaluation of just one aspect of usability, learnability. Additionally, there is limited research about learnability despite being important for the adoption of information technology (Lazar, Feng, & Hochheiser, 2010, p. 44). Therefore, the outcome of this master's thesis can benefit designers or developers of products that utilize Blockchain and other professionals that want to assess the learnability of user interfaces that display cryptographically certified information published on decentralized networks such as Blockchains.

An instance of a product with user interfaces designed to access information recorded by Blockchain products are the so-called Blockchain explorers or block explorers. These tools are purpose-built search engines that have been developed to allow users to look up and verify the data of transactions registered on a determined Blockchain. Their user interface provides access to details such as the fingerprints of digital documents called hashes, which can be used to cryptographically prove the existence and data integrity of any contract. Since the records of a Blockchain are immutable, any hash saved as input data in a transaction remains timestamped at a certain point in time. Thanks to this form of certification, a user can verify the information written in a contract with a Blockchain explorer and also obtain proof about the date it entered into effect, making possible to enforce its fulfilment with the other party in case of a dispute about the details agreed.

1.2 Research Questions

In general, the user interface of block explorers helps to verify all data recorded on a Blockchain and provides evidence of the utility of decentralized applications. Particularly, this research focused on user interfaces that display information of a transaction recorded on a Blockchain in products used for attesting the existence and immutability of contractual

data as established during an agreement. Researching the learnability issues and challenges that users have when interacting with them can contribute with insights that support designers and developers working to improve the usability of other Blockchain products. The title and research question addressed in this master thesis project is:

How Usable are Block Explorers for Attesting Agreements Registered on Blockchains?

For this thesis, usability is narrowed down to the learnability aspect of a user interface since evaluating learnability covers one of 5 quality components that determine how easy is to use a user interface from the first time a user interacts with it (Nilsen, 2012). In addition, the title of the thesis is complemented by the subtitle:

Evaluating the Learnability of a Blockchain Product's User Interface for Verifying Records of Cryptographically Certified Documents on the Ethereum Blockchain

Taking into account the recommendations in the ISO 9241:110 (International Organization for Standardization, 2006) for a user interface to become suitable for learning by guiding the users to complete their goal, this research proposes a learnability evaluation framework and uses it alongside a methodology to evaluate the learnability of Blockchain explorers answering two research sub-questions:

a) Are the recommendations of the Suitability for Learning principle considered in the design of block explorers' user interfaces to assist users in verifying the certification of documents via transactions recorded on a Blockchain?

b) *Does the learnability of the selected block explorer's user interface adequately support the verification of cryptographically certified documents for first-time users with varying degrees of domain knowledge about Blockchain?*

1.3 Contributions

From the users' perspective, two significant problems are considered sources of friction by designers when conceiving the user interfaces of products that utilize Blockchain. On one side is the limited understanding of new users about the benefits of Blockchain and the concepts that make this technology useful. On the other side, there is the inherent complexity of the information displayed by the user interface of Blockchain products. For example, to show the information of a transaction, block explorers expose all the data recorded on a Blockchain about that specific transaction, then users can confirm the logging of the actions they have taken with a Blockchain product. When a block explorer or any other product display such records, poor usability can cause the increment of the cognitive friction that users experience and the effort they put at interacting with the UI.

Blockchain explorers are integral not only for the development of other Blockchain applications, but they also make possible for the parties entering an agreement to verify that they have successfully recorded data of a contract representing the exchange of goods, information, or the transfer of digital currency with determined value. This research takes into account the current landscape of Blockchain products, looking at services that utilize Blockchain for certifying the existence of contracts in a digital format. Then, evaluating the subsequent use of block explorers for the verification of certified information and the irrefutable attestation of the data in a contract, which enables users of these Blockchain products to enforce the terms and conditions of an agreement during any legal dispute.

Currently, designers and developers have limited knowledge of how usable the interfaces of block explorers are for supporting the task of verifying data registered on a blockchain. Notably, there are differences at supporting the accomplishment of this action between users that have varying degrees of understanding of concepts such as cryptography and distributed systems. Even if two people are technically savvy, their domain knowledge about Blockchain can significantly affect how usable and relevant is the user interface for validating records of data registered with this technology. This thesis studies the differences of new-users of block-explorers which have different levels of familiarity with Blockchain concepts, generating relevant data to address the improvement of the learnability of a block explorer's user interface overall.

1.4 Thesis Outline

Although some of the technical concepts of Blockchain technology are mentioned along with the content of the thesis, this research does not attempt to give a comprehensive and technical explanation of the foundations that constitute blockchain. The main goal of providing an approachable overview of such concepts is to present the necessary information that contextualizes the author's evaluation of the Blockchain products selected to be the target of this study. In total, the thesis consists of 6 chapters:

Chapter 1 introduces the utilization of Blockchain and the challenges that the user interface of block explorers and other Blockchain products in general face regarding one particular aspect of usability, learnability. The justification, motivations and benefits of writing this thesis were laid out as well as the corresponding research questions to be addressed by the thesis. Finally, the author explains that the contributions are aimed to provide insight to designers and developers interested in evaluating the usability of Blockchain products, specifically, the learnability of their user interface.

Chapter 2 gives an overview of the main concepts that are fundamental for distributed ledger technologies, such as Blockchain, to work. Defining distributed systems, consensus and cryptographic hashing gives the overarching conceptual framework to understand how decentralization generates the guarantees for recording data agreements without intermediaries. Relevant applications of Blockchain are described after explaining the internal components of the most recognized Blockchain architecture. Three different products that enable users to certificate digital documents on permission-less blockchains such as Bitcoin and Ethereum are presented. Subsequently, the importance of block explorers is rendered and the process of redirecting the user to verify the certification of documents explained. A review of three block explorers is given followed by a definition of the usability concepts that are employed to evaluate the learnability of these products' user interface.

Chapter 3 describes the methodology for a usability engineering approach to evaluating learnability divided in two parts. The first part consists of the explanation of the framework proposed for the evaluation of learnability and the summative expert assessment made of a combination of two usability inspection methods (Heuristic Evaluation and Cognitive Walkthrough) in a feature analysis approach. The second part includes the experiment used to evaluate the learnability of block explorers through Formative Empirical Testing. A sequence of two methods compose this part of the chapter, a Cognitive Task Analysis, and then a semi-structured post-test interview with the selected testers.

Chapter 4 presents and analyses the findings of the methods used during the expert assessment of the selected block explorers and the empirically based study of the suitability for learning of just one of them.

Chapter 5 involves a discussion about the framework used to evaluate the learnability of block explorers and its validity, as well as specific findings from the triangulation of results between the methods used. Also, the interpretation of the results from the previous chapter and the assessment of general patterns seen in the qualitative data analysed are discussed to establish the importance of evaluating learnability for block explorers and Blockchain products in general. Finally, the possible confounders and limitations of the methods used are presented in this chapter. A summary of the research is also offered.

Chapter 6 provides the final conclusion taken from addressing the research question and sub questions. This chapter also reflects about the possibilities for designers, developers, and researchers for employing the evaluation framework used throughout the research to assess the learnability and other usability aspects not only of block explorers, but also the user interfaces of Blockchain products in general. Recommendations for potential future research are also outlined.

2 Theoretical Background

What happens when a person agrees with another person might seem simple, but it requires multiple aspects to be accounted for before and after the agreement. For example, when two people agreed to exchange items of value, they must define a situation in which the ownership of the items will change of hands. Contracts are used to record the terms and conditions of the exchange to make sure that an agreement prevails according to a legal framework. The parties involved in the agreement have to acknowledge that the information introduced in a contract reflects the reality of the situation to which they agree, then proceed to accept it by including their signatures as a way to identify themselves entering the agreement. Thereinafter, such contract holds the "truth" of the situation, and each of the parties keeps physical or digital representations as proof of what was agreed.

For the minimal configuration in which just two parties are trying to agree, to avoid any conflict, usually, it is required a third party that holds an extra copy of the contract and also agrees about the information registered on it. A third party is expected to keep the contract intact and accessible in case the parties argue about specific details and have to verify them to solve a dispute. Also, such third party has to be trusted enough by the agreeing parties for them to rest assured that the information in the contract will not be modified against their interests. However, that there is just one extra copy controlled by a third party, represents a risk for any one of the agreeing parties since the "trusted" third party could collude with another to make a change in the agreement.

When there is a third party —often called middleman—involved in keeping a version of the contract in any trade or agreement, the agreeing parties have to go through it to verify the terms and conditions of the exchange written in the contract. The third party acting as an intermediary obtains leverage to request compensation for the service of staying impartial by securing the data integrity of the details recorded in the agreement. Such compensation does not guarantee the middleman will keep the contract untampered, but by receiving it, they are also entering into an exchange of value with each of the parties to protect the data, remain fair or otherwise be punished by the law. That is of course in the case the colluders are caught so the capacity for two of the three parties to cooperate and get away with it cannot be removed entirely. Ultimately, when people are using the same third party to keep a record of their contracts and expect them to guarantee that these cannot be changed, the system is centralized.

The probability of collusion between two-thirds of the parties involved in an agreement can only be decreased as more participants are involved in keeping a record of the contract. If all the parties form a network, making it open enables every participant joining to hold a copy of what has been agreed by the parties. Additionally, if all the parties use the same network of participants for keeping a copy of their contracts, then all of them will have an incentive to maintain the data that has been recorded intact and exclude anyone that is trying to change it for their benefit. In essence, such a system which enables replicating, sharing and synchronizing all the information between the participants in a network, so they have the most updated version of the contracts is considered a distributed system.

Furthermore, agreements can have various terms and conditions registered in contracts. Therefore, to make sure that multiple parties acknowledge their agreements as valid, they

need a solution to register and preserve the data in such contracts acting as a "system of record." As an example, banks use software to track every single transaction that happens with the money they manage (Chain, 2017), that means all the agreements involving the exchange of this money are registered. This software is nothing more than a bespoke ledger, which is the term used for accounting books or archives that collect data of a particular type changing sequentially, typically in chronological order.

2.1 Distributed Ledger Technologies

Ledgers that record monetary transactions, ownership of assets or similar data have existed since ancient times. There are findings of such records in paper, clay and even animal skin. In the late 20th century, computerization brought the only notable innovation that ledgers have had since their invention. All the records were able to be transferred to bytes, mirroring the functionality of their paper counterparts (Walport, 2016). However, throughout history, a central authority needed to validate the authenticity of the transactions recorded in the ledgers. Again, in the case of banks, their software needs to verify each financial transaction, charging a fee for it and becoming intermediaries in the process (Rouse, Troy & Pratt, 2017).

In the case of agreeing to record transactions of information about an exchange of goods or services, people rely on the creation of legal contracts. However, completing a simple agreement can become more intricate as more people (e.g. lawyers, witnesses and auditors) are involved in acknowledging its details and certifying its validity with a contract. Additionally, if the variables that control the exchange are multiple, the agreement grows in complexity and keeping track of all becomes costly to manage. For example, when a good or service not only has a price, but also has characteristics of valuation or devaluation according to certain conditions. Also, when contracts only become valid after a determined amount of people that need to sign it or the deal establishes that the exchange only happens after a specific date.

Businesses that enter into complex contract agreements with each other require systems that keep track of such information. However, most companies currently use centralized databases that become a single point of failure. Building a centralized ledger for these contracts can be dangerous since any data that becomes corrupted or lost signifies a critical problem that erodes trust between the parties affected. The use of Distributed Ledger Technologies (DLT) is considered an alternative to solve this problem. As Belin (2018) explains, DLTs are databases that exist across several locations managed by multiple participants. A distributed ledger is decentralized to eliminate the need for a central authority or intermediary to process, validate or authenticate the actions recorded since all of participants in the network collaborate to accomplish the same goal (Figure 2.1). In this way, any single participant can add new data on top of what has been registered in the ledger without needing to trust an intermediary to keep it secure.

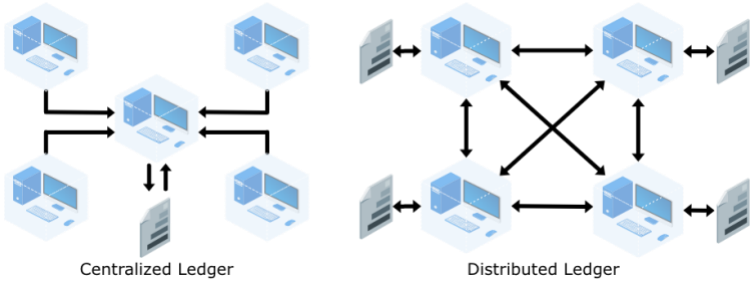


Figure 2.1: Centralized Versus Distributed Ledger Technologies.

However, distributed technologies that only replicate the data recorded in a ledger still have a problem. Even after having many participants in the network, all of them have to be sure that enough of them cannot easily cooperate to change the records. If collusion is still possible, people will fear that someone can change what has been agreed by the parties of a contract and share it with all the participants in the network as valid. After all, trusting strangers to keep records of data impartially is not something that rational human beings can do without guarantees that the contracts cannot be modified or tampered. Without guarantees of keeping the system honest and secure, the looming threat of adversaries accessing and potentially altering the records in their favor will always exist.

To solve this problem DLTs typically make a new record part of the ledger just after a consensus between the majority of participants, including the parties involved in the agreement, has been reached. Another step taken is that all the records and data that are part of the distributed ledger are then timestamped and given a unique cryptographic identification (Belin 2018). Having the data stored in multiple places and cryptographically hashed makes it possible to keep it secure and avoid that it can be changed easily. Concepts of consensus and cryptography determine essential features in all DLTs. Furthermore, the proprieties for maintaining a DLT performant and highly available relate to its architecture as a distributed system.

In general, distributed systems, consensus, and cryptographic hashing are together the main concepts that make Distributed Ledger Technologies possible. It is essential to understand these three characteristics of distributed ledger technologies because Blockchains are a type of DLT. Specifically, a Blockchain is an architecture of a distributed ledger with a distinct configuration of these three concepts. Thus, talking about any particular Blockchain, people would be referring to a DLT. However, talking about Distributed Ledger Technologies not always relates to the Blockchain technology. It is out of the scope of this research to look at multiple forms of DLTs, but the primary concepts that define these technologies are presented below.

2.1.1 Distributed Systems

At the core, the use of Blockchain is a distributed systems problem that involves several computers, identified as nodes, work together for agreeing on something while they all might be unknown and untrusting of each other. With blockchain, information is deployed across a network, where users who are potentially located on different sides of the world can interact with one another. As Takada (2013) explains, each of these computers accomplish two basic tasks, storing data and computing instructions. However, data needs to be copied around and computation tasks have to be coordinated, so adding a new machine does not increase the performance and capacity of the system linearly. Algorithms dictate the functionality of distributed systems and are in charge of dealing with the overhead that arises due to having separate computers.

There are two particularly relevant aspects when multiple computers work together. One is the performance, which is "characterized by the amount of useful work accomplished by a computer system compared to the time and resources used" (Oshana, 2013, p. 281). A key concept of performance is latency, which refers to the period between the initiation and the occurrence of an action taken by the system, determined by the speed at which new data "takes effect" in the system (Takada, 2013). In the case of Blockchains, latency describes the time it takes for a record to be included into the ledger and be copied to all the computers in a network.

The second aspect of a scalable distributed system is its availability, determined by the amount of time a system remains functional despite failures. A system made by a bunch of machines becomes unavailable if all of them fail, therefore, a notable characteristic of distributed systems is that they are fault-tolerant. Takada (2013) mentions in his book that even if the network is made of unreliable components, distributed systems have the ability to still be available and capable of behaving in expected manners once faults occur. Despite that the probability of components presenting a failure augments with their number, the system should be able to compensate and become more reliable as such number increases.

2.1.2 Consensus

All distributed systems are constrained by the number of nodes and the distance between them. In a Blockchain, a network of multiple nodes placed on different parts of the world interact with each other and agree on a common truth without trusting a single machine or authority. In other words, the network has to be able to reach consensus about the data they share. Instead of trusting the execution of individual processes or the reliability of any single machine, the nodes trust the general protocol and the math behind it (Blockchain at Berkeley, 2018). For example, as a use case of blockchain, the Bitcoin protocol makes sure several parties agree–or come to consensus–about who owns what amounts of the crypto asset.

For any distributed system to be correct, it must come to some form of consensus on the correct status of the system. Leslie Lamport, known for his valuable research on distributed systems called "Time, Clocks, and the Ordering of Events in a Distributed System" (Lamport, 1978), defined that the correctness of a system requires achieving its intended goal by establishing rules for what cannot happen called safety, and what must happen called liveness (Lamport, 1977). Despite, it is difficult that both properties of safety and liveness fulfil simultaneously, a protocol with strict rules of validity, agreement, and termination can make any distributed system to come to consensus.

A consensus-achieving protocol is the fundamental building block of Blockchain technology and is commonly known as consensus mechanism. In a blockchain, the consensus mechanism implemented ensures which node can write data on the ledger and the coordination of this prevents data corruption while all the nodes agree on some state of the system. Recording the final state of the system is determined by a majority rule, such as the 51% of all the nodes agreeing for the Bitcoin and Ethereum Blockchains, or with more specified majority of nodes based on the Byzantine Generals Problem (Lamport, Shostak & Pease, 1982). In the latter case, it is only possible to establish consensus if at least two thirds of the nodes agree, described as a system that has 3n+1 nodes where n is the total of faulty or byzantine nodes in a distributed system (Castro & Liskov, 2002).

2.1.1 Cryptographic hashing

Entering into an exchange agreement depends on the information that defines what is being exchanged. In the simplest of cases, two parties exchanging items of value between them, establish the characteristics of the item and their perceived value to each other, registering all these details in a contract. Such perception of value generally remains equal until the exchange happens. At this point, the transaction is considered fair even if the value of the items changes from that moment.

An exchange can be disputed, and the corresponding contract can be rendered invalid if at the moment of the exchange any of the parties encounter that what they received does not match the form or value of the items as described in the contract. However, the party that benefits from the agreement, might be able to change the contract to convince anyone that the exchange occurred as it was agreed. With malicious intentions, such a party could effectively tamper with any digital representation of the contract without leaving a trace. As such, there will not be irregularities that the affected party can establish as unfair.

Making sure of the existence and preservation of data are use cases of cryptography that are relevant to this scenario. Modern cryptography includes the creation and improvement of algorithms for data integrity and non-repudiation (Katz, 2018). Applying cryptography for the integrity of data in contracts involves complex mathematical algorithms whose output are extremely difficult to falsify. Their purpose is to associate files and any changes made to them with a particular individual, having evidence of the tampering so the malicious parties in an agreement cannot claim that they were not part of the agreement by modifying it or repudiating the authenticity of their signature (Zhou & Gollman, 2002).

Generally, obtaining a proof of data integrity can be accomplished by employing a cryptographic hash function. Such mathematical function encrypts files by mapping data of an arbitrary size with bit strings of a fixed size called digests, or hashes (Katz, 2018). For example, you could use the Standard Hashing Algorithm-256 (SHA-256) to process a small set of characters such as "Norway" or the entire PDF file of a book about Norway (Figure 2.2). Both inputs will return different fixed-size outputs, ranging in length from 160 to 512 bits, depending on the algorithm that is being used (NIST, 2012). Cryptographic hashing functions are efficiently computable because the output can be produced quickly and with few resources independently of the input size (Pearson, 2018).

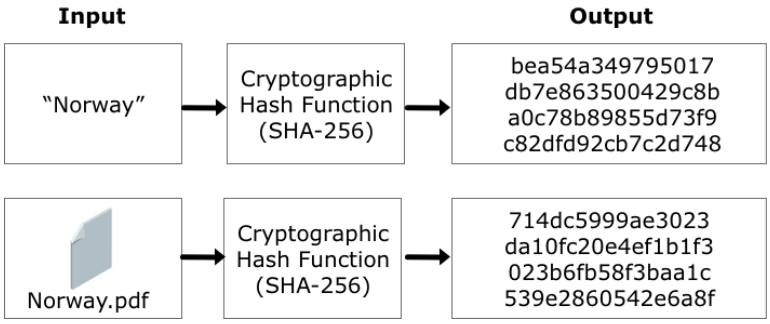


Figure 2.2: Outputs of a cryptographic hashing function using SHA-256.

Other three properties of cryptographic hash function are important for their implementation to keep the data integrity of Blockchains:

First, hash functions are considered to be extremely collision resistant. A collision can occur if two different inputs product the same output. This can happen considering that there are infinite inputs and a finite amount of ways to organize a fixed output, however the odds for it to happen are astronomical, making infeasible for data that is similar to generate the same hash (Pearson, 2018).

Second, hash functions are deterministic, so the same message always results in the same hash. Hashes are reproducible as long as parties have the same input (NIST, 2012). This property allows anyone to verify the data integrity of a file, such a contract. One party can verify that other party also has the same version of the contract by comparing the hash values produced by the same cryptographical algorithm.

The third property of hash functions is their irreversibility. This property ensures that an adversary party that holds the output has no feasible way to figure out what the input was, because the hashing function is designed to be one way and therefore, impossible to reverse. The only way to recreate the input data from a cryptographic hashing function's output is to convert all the possible inputs until finding a match (Menezes *et al.* 1996).

These properties make cryptographic hashing functions fundamentally valuable for Blockchain products that certify data integrity by generating a tamperproof record of hashes. Although, these algorithms convert any Blockchain into a tamper-evident ledger, people can become the target of hackers interested in modifying the data before getting processed. It is possible they tamper with data in transit, either through a man-in-the-middle attack or phishing. However, forging a record in a Blockchain of an agreement that has been hashed should be impossible without other participants in the network to notice.

2.1.2 Decentralization

The author of this thesis has defined distributed systems, outlined the requirements to achieve consensus, and described the properties of using hashing algorithms for the security and integrity of data. These three characteristics of Distributed Ledger Technologies allow them to maintain networks that are decentralized and functional. In such computational systems, decentralization provides three arguments (Buterin, 2017) that can act as the guarantees for people to agree with each other using a Distributed Ledger Technology such as Blockchain, without relying on witnesses or intermediaries.

1) **Distributed systems provide fault tolerance:** A distributed system is fault tolerant when it supports a network flexible and resilient enough to recover from faults. Decentralized systems are less likely to fail because they rely on many separate components (Buterin, 2017). Even if a large number of nodes in the network are separated from the majority of nodes in the distributed system, they would be able to keep operating independently.

2) **Consensus mechanisms make them collusion resistant:** "It is much harder for participants in decentralized systems to collude and act in ways that benefit them at the expense of other participants" (Buterin, 2017). The consensus protocol defines the process in which multiple honest parties rely on to solve conflicts in any agreement by supporting the correct version of the respective contract. In this context, a consensus mechanism enforces that a representative majority of the nodes involved in recording the data of an agreement in a distributed system are working in the best interest of the whole group and their common goal

3) **Cryptographic algorithms increase the attack resistance**: A conflict can arise if any of the parties can delete or modify data shared within the distributed system in the majority of its copies without leaving a trace of the tampering. Since a small change of the data results in a different hash value, an attacker cannot credibly establish that tampered data prevails over any other versions. Additionally, decentralized systems are more expensive to destroy, alter or manipulate because "they lack sensitive central points that can be attacked at a much lower cost than the economic size of the surrounding system" (Buterin, 2017).

In essence, these arguments for decentralization provide the guarantees that the parties of an agreement require to register tamperproof information without the need for

intermediaries. On one side, distributed ledger technologies employ algorithms with a mathematically-proven elevated probability to preserve data integrity (National Institute of Standards and Technology, 2012). On the other, a DLT's consensus mechanism protects the data that participants of a network record inside transactions against the collusion of adversaries that would benefit from altering the ledger. Moreover, their distributed architecture ensures the availability of all the records in the distributed ledger to the participants in the network, providing a verifiable and auditable history of all information stored or the possibility for querying a particular dataset.

2.2 Blockchain

Because Blockchain is a type of Distributed Ledger Technology, the guarantees for decentralization that apply to DLTs also apply to Blockchain. However, these guarantees are not present in the majority of information systems currently employed in the world, which have a high degree of architectural centralization (Kol, 2019). Despite that such guarantees for decentralization are not cheap, nor easy to enforce and maintain, the development of Blockchain technology is advancing to overcome its limitations. Advocates of Blockchain are trying to prove that decentralization provides more value and give competitive advantages to companies (Mamoria, 2017), but in reality, the adoption of decentralized applications has been slow.

A Blockchain is a type of distributed ledger technology that uses cryptography and economic incentives to record data in a tamper-evident way (Pearson, 2018). According to the MIT Technology Review, "A Blockchain's mathematical structure allows to store all kinds of valuable data in a way that is nearly impossible to fake" (MIT Technology Review Editors (MIT-TRE), 2018). As a DLT, the cryptographic database of a Blockchain is maintained by a network of computers called nodes, "each of which stores a copy of the most up-to-date version" (Orcutt, 2019). Different blockchains have different protocols that dictate how nodes verify new **transactions** and add them to the ledger grouping them in **blocks**. The consensus mechanism currently used by both Bitcoin and Ethereum to record each block in a chain is called **Proof-of-Work**.

2.2.1 Accounts and Transactions

A Blockchain can be defined as "a transaction-based state machine that reads a series of inputs and, based on those inputs, transitions to a new state" (Kasireddy, 2017). Each account has a state, and the aggregation of the individual states from all the accounts determines the global shared-state of the blockchain. When initialized, each of these accounts takes a unique address generated by public-key cryptography. This kind of cryptographic system utilizes a pair of keys, one public key that can be shared and used for identification, and a private key known only by the user, employed for authentication and encryption of data (Rivest, Shamir & Adleman, 1978). Usually, an account address is the hash of the public key, and the public key is the hash of the private key, all of them displayed as strings of seemingly random alphanumeric characters.

Transactions are at the core of everything that happens in a Blockchain. In general, transactions occurring between different accounts are what move the global state of blockchains from one state to the next (Kasireddy, 2017). They are the result of programming methods and functions in the Blockchain's code that allow accounts to interact with each other by passing messages. In a strict sense, a transaction is simply a cryptographically signed instruction that is executed by an account to communicate with

another (Kasireddy, 2017). Using a cryptographic signature method provides verifiability and unforgeability, confirming the ownership of a determined account and securing that the transaction cannot be tampered with (Pearson, 2018). Users of Blockchain accounts, employ their private key to produce such digital signature for each transaction.

Sharing an account address allows other people to use their accounts and target the recipient of a transaction. In Bitcoin, a transaction represents the transfer of the cryptocurrency from one account to another. After a Bitcoin payment is ordered, the transaction message is sent to the network and passed around all the network of participants, remaining in an 'unconfirmed' state (Lewis, 2015). Similarly, Ethereum includes a Turing complete programming language (Sipser, 1996) that supports the programming of all kinds of methods, functions and modifiers to create automatic transactions, also known as smart contracts. These transactions occur between two different kinds of accounts, external and contract accounts (Figure 2.3).

An externally owned account can send messages to other external accounts, merely executing a value transfer. However, these accounts can also send messages to contract accounts, invoking the execution of a method or function (Goldberg, 2018). Contract accounts are controlled by their code and are fired by external accounts when they sign a transaction using its private key, simply performing actions such as transferring tokens, writing to internal storage, performing some calculation or even creating new contracts (Kasireddy, 2017). As such, smart contracts are just computing programs executed after certain conditions encoded are met, emitting transactions in response to other transactions they have received (MIT-TRE, 2018).

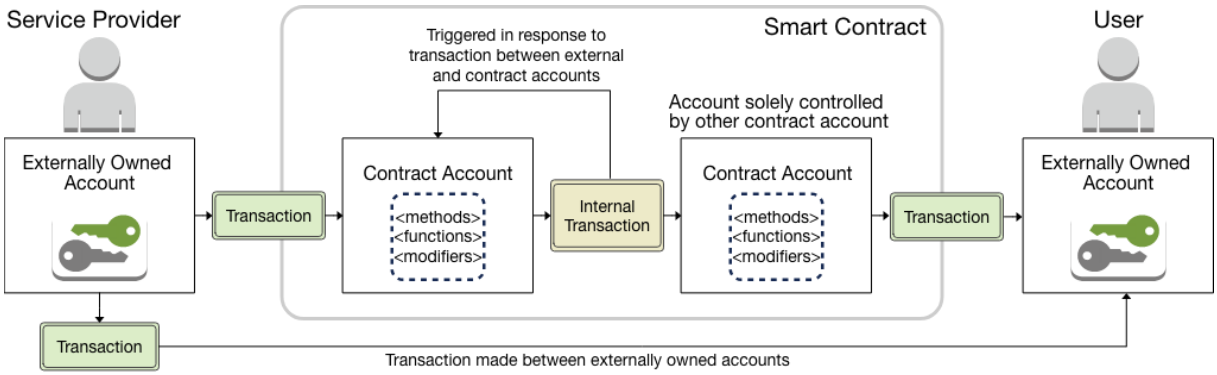


Figure 2.3: Blockchain Accounts and Transactions.

2.2.2 Blocks and Data Structure

After the transactions are generated, they are broadcasted to the nodes in the network. According to specific rules of each blockchain, the nodes validate the methods initiated from specific accounts by creating a list of transactions, however, the transactions are not written into the ledger yet (Lewis, 2015). A subset of nodes, called miners, organize valid transactions bundling them together with other data into blocks. A block contains data in a particular structure of recent valid transactions and a cryptographic reference to the previous block. Such structure and reference come from employing a cryptographic hashing function, such as SHA-256, used to create hash pointers which represent and help to identify specific data in the block.

Datasets belonging to the transactions are hashed together into a structure that looks like a tree branch with leaves, called Merkle tree (Merkle, 1988). Every leaf is labelled with the hash pointer of a transaction's dataset, and every non-leaf is labelled with the cryptographic hash of the children below, repeating the process until the total number of hash pointers is only one, called the root hash (Kasireddy, 2017). Internally, Merkle trees organize the hash pointers to each transaction listed and bundled in a block with this structure, allowing blockchains to efficiently and immutably store information (Figure 2.4). Consequently, each block contains a set of transactions, and the blocks are chained together, forming a cryptographic summary of the blockchain's state (Bansal, 2017).

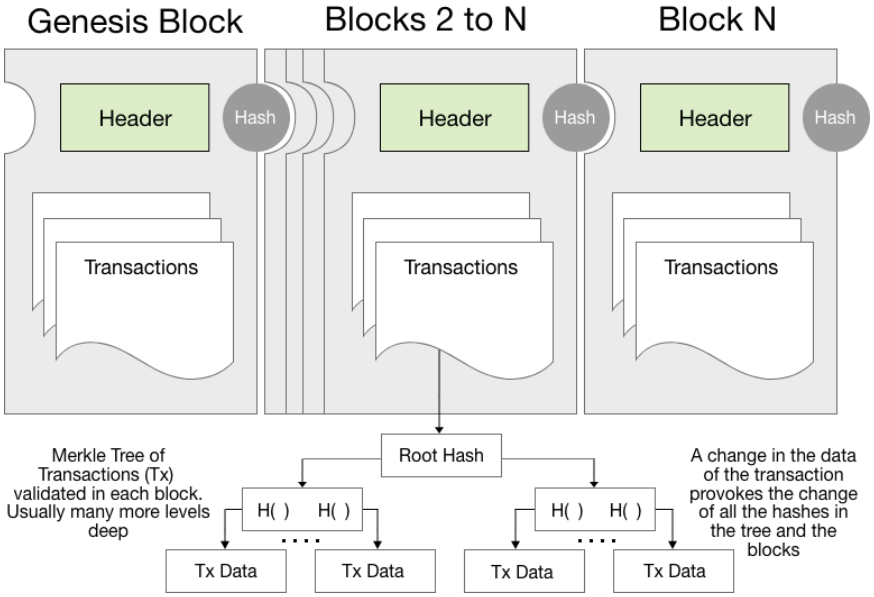


Figure 2.4: Blocks and Merkle Tree Data Structure.

Blockchains are often called triple entry accounting systems because the hash pointers in the Merkle tree secures the integrity of the ledger (Pearson, 2018). In a blockchain's structure, the root hash of every block is cryptographically dependent on the data stored in the tree (Kasireddy, 2017). Therefore, anyone can check the hash pointing at the "root" of the tree to verify that the data of the transactions remains the same, and its integrity is intact. Additionally, listing different transactions or modifying a single bit of data in them would modify the hash pointer of that block and all the subsequent blocks, generating an alternative global state or making easy to identify when someone tampers with the data of the ledger (Pearson, 2018). Hence, a blockchain's "triple-entry" is extremely difficult to forge, making sure the ledger only reflects the current and canonical state of the transactions.

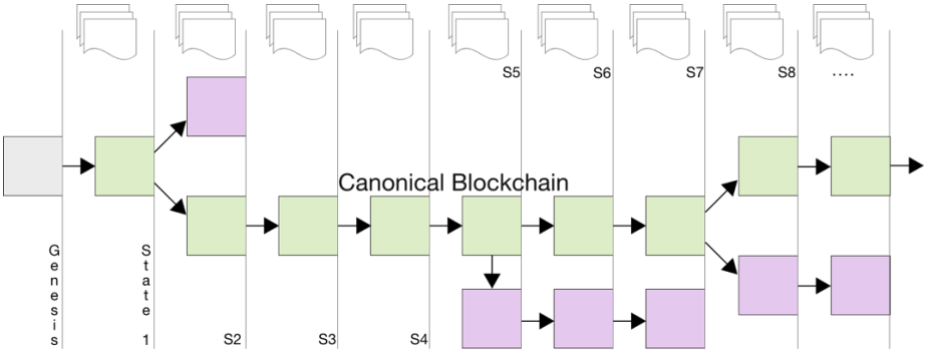


Figure 2.5: Canonical Blockchain and Global States.

2.2.3 Mining and Proof-of-Work

In Blockchains such as Bitcoin and Ethereum, mining is the last step required before recording a transaction into the blockchain. During this step, mining nodes race to validate a block that can extend the chain in a process that "requires solving a labour-intensive mathematical puzzle, which is unique to each new block" (MIT-TRE, 2018). The puzzle involves randomly selecting a number called a "nonce". The nonce is combined with other data in the block and then transformed into a hash by a specific cryptographic algorithm depending on the Blockchain (Hamilton, 2018).

The output of a cryptographic hashing function must meet specific criteria for the block to be considered valid (Lewis, 2015). In the case of Bitcoin, miners compete scanning for a value that when hashed using SHA-256, the resulting string begins with a certain number of zero bits. The miners keep trying nonce numbers and calculating hashes, but once the mining node has spent the CPU effort that satisfies the proof-of-work, "the block cannot be changed without redoing the work. As later blocks are chained after it, the work needed to change data in the block would include redoing all the blocks after it" (Nakamoto, 2008). The Ethereum's mining process also involves nodes competing against each other to complete a mathematical equation, but this time defined by the 'Ethash' mining algorithm (Wood & Buterin, 2014).

The proof-of-work challenge is a computationally expensive process in which miners employ a lot of "hash power" trying to guess a number that satisfies the difficulty set by the algorithm. As the editors of the MIT Technology Review (2018) explain, "mining difficulty is encoded in the blockchain's protocol; Bitcoin and Ethereum are designed to make it increasingly hard to solve a block over time." The first miner to discover a suitable solution to the problem earns a reward with the own economically valuable token of the blockchain, 12 BTC for Blockchain and around 3,5 ETH for Ethereum (Hamilton, 2018). Although it takes an enormous amount of tries to find a valid hash, the vast number of computers that are now participating in both networks reduces the time for the mining process. Hamilton (2018) explains that Blocks are mined in Bitcoin every ten minutes, and approximately 14-16 seconds in Ethereum to mine a newly added block.

The mining processes of both the Bitcoin and Ethereum blockchains use the proof-of-work consensus mechanism. As a combination of cryptography and distributed systems, consensus is what allow Blockchains to become a distributed consensus technology (Bansal, 2017). The Bitcoin protocol permits the nodes to agree about the block that defines who owns what amount of the crypto asset according to the recent transactions. In the case of the Ethereum network, nodes validate the block containing both the correct execution of transactions and the global state of the network before adding it to the end of the chain. In this way, as Hamilton (2018) explains: "Every miner node on the Blockchain works together to ensure the longest chain of transactions is the valid chain."

It is worth noting that, the extensive computer processing power required for mining both cryptocurrencies represent energy consumed by the mining nodes. Considering the number of mining nodes in these networks, the process equates to enormous amounts of electricity. Although there are valid critics, the lack of efficiency in the process is what makes Proof-of-Work Blockchains extremely expensive to tamper with or add reverse transactions that could facilitate double spending (Nakamoto 2008). It is prohibitive for any single participant even to try beating the network building the sufficient hash power. Thus, a Proof-of-Work consensus mechanism deters hackers from modifying the ledger altogether.

2.3 The Application of Blockchain

Blockchain technology is nascent, and the full realization of its potential might require it adopts multiple forms and keeps evolving. Some blockchains implement different configurations for securing their networks, protecting them from attacks by employing alternative consensus mechanisms to make the nodes agree on the status of the ledger (Zhelezov, 2018). However, "the proof-of-work consensus mechanism is the most thoroughly battle-tested." (MIT Technology Review Editors, 2018). Additionally, as a back-end technology, blockchains can adopt different characteristics for each context of use and application. Therefore, many argue that the future of Blockchain is becoming more exciting and promising as time passes and new implementations emerge.

Although Blockchain is usually labelled as a distributed database, it has many limitations in comparison to other technologies that do the same job (Bansal, 2017). During the design of Blockchain-based products, inevitable trade-offs have to be made in order to achieve the goal of decentralization. Ludwin (2017) argues that for now, "on almost every dimension, decentralized services are worse than their centralized counterparts." Blockchain implementations are "slower, more expensive, less scalable and have worse user experience" (Ludwin, 2017) than established solutions implemented in most industries. However, there is a single utility dimension of blockchains that excels over other technologies, their immutability.

The massive peer-to-peer network of nodes in the Bitcoin and Ethereum blockchains create public, immutable, decentralized and easily verifiable ledgers. The people and organizations behind the nodes supporting the mining process are motivated by economic incentives, making sure that the chain remains unchanged as long as these "honest nodes collectively control more CPU power than any cooperation group of attacker nodes" (Nakamoto, 2008). Consequently, Blockchain implementations can be considered the "right engineering implementations only for those applications that require a fully distributed system in an adversarial environment" (Bansal, 2017).

2.3.1 The Utility of Smart Contracts

When doing businesses, companies that are transferring or exchanging value need a third party that ensures they can trust each other. This trust is recorded on the contracts that both parties sign, after agreeing with all the terms and conditions of their engagement. Every single exchange of value requires extensive paperwork, especially for processes that have to be done routinely, resulting in expensive and timely processes that are prone to human mistakes. The first proposition of automating these contracts and making them "smart" dates to the 1990s when Nick Szabo published the papers "Smart Contracts: Building Blocks for Digital Markets" (Szabo, 1976) and "Formalizing and Securing Relationships on Public Networks" (Szabo, 1977). His research gives an excellent overview of how computerized transaction protocols in smart contracts could epitomize an efficient analysis of complex term structures in standardized digital contracts and record them with low transaction costs. However, the correct implementation of smart contracts without intermediaries was not feasible until the full potential of the concept was unlocked by the possibility to implement them on a Blockchain programming language that Ethereum provides (Wood & Buterin, 2014).

Blockchain protocols make sure that the terms and conditions programmed in a smart contract cannot be modified after they are deployed. Therefore, parties involved in an

agreement using smart contracts will only need to trust that a piece of code running on the Blockchain will continue to behave as designed (Dixon, 2019). Additionally, the data introduced as part of each transaction made by contract accounts is completely inalterable once they have been triggered (Mamoria, 2017). The capacity of running smart contracts on a Blockchain allows deploying fully decentralized applications that "can build a foundation for trust in the enterprise through the digitization of business processes, tokenization of assets, and codification of complex contracts." (Cisco Systems, 2018).

Besides recording value in crypto assets that can be exchanged in the form of digital currency, the ledger entries in a Blockchain can record other data, such as the unique fingerprints of ownership titles, identity records, and legal contracts. In this sense, the utility of blockchains has been stretched to support applications that involve processing and coordination of data or the digitization of physical assets by the creation of trusted and immutable records (Kumar, 2018). As Dixon (2019) states, "the benefit is that, unlike a traditional computer, a Blockchain can offer strong trust guarantees, rooted in the mathematical and game-theoretic properties of the system." Using Blockchain technology, the terms and conditions of any agreement can be converted into variables programmed into smart contracts. And unlike conventional contracts, that involve humans reviewing the process, "smart contracts consist only of software code and are executed by the Blockchain network" (Mamoria, 2017).

In general, smart contracts implemented on Blockchains and other Distributed Ledger technology could enable consumers, enterprises, and government institutions to replace their current processes with new digital solutions that promise to embed high trust and transparency in the transaction and management of data. Examples of Blockchain applications include recording data obtained by sensors or other devices in order to support smart cities (Cisco Systems, 2018). Also, the technology can be used for creating secure voting mechanisms for governance in companies or even governments, helping them to make decisions as a Decentralized Autonomous Organization or DAO (Orcutt, 2019). Finally, Blockchain permits the unequivocal digitalization of physical goods to be tracked and traced across the value chain, from their production, passing through transportation and finally their commerce and consumption (Kumar, 2018; Partz, 2019).

2.3.2 Certifying Contractual Documents with Blockchain

Two parties that need guarantees for trusting each other or leave a permanent record of their agreements for auditing purposes, can use Blockchain certification services to record the information of an agreement on the distributed ledger. Businesses and individuals interact with the user interface of these services to directly upload the files or raw data of their contractual documents. The service then makes a transaction or triggers a smart contract on behalf of the users, inserting the cryptographic fingerprint of the documents as input data. From the moment the data makes part of a mined block, users can verify that the hash of the contract has been recorded into the blockchain and can refer to it as a source of truth, instead of involving intermediaries. The use of these services can facilitate the operations of large enterprises by securely certifying legal contracts and commercial agreements without the need for human intervention. In this way, Blockchain "removes the need for an established third party to create a trusted relationship." (Cisco Systems 2018). Three examples of Blockchain certification services are presented below:

2.3.2.1 MIT Academic Certificates

The Massachusetts Institute of Technology (MIT) offers a service for creating, sharing, and verifying blockchain-based educational certificates that comply with the institution's vetting and verification processes.[1] Tamper-proof digital certificates are cryptographically signed and registered on the Bitcoin blockchain. Later called Blockcerts[2], the code of the product was released under an open-source license and can be used within any institution's technical infrastructure to reliably secure the unique versions of their academic certificates on the blockchain. The MIT has used this technology to create digital certificates to MIT Media Lab alumni, Learning Machine Employees, MIT Global Entrepreneurship Bootcamp alumni and the participants of the "Laboratorio para la Ciudad" workshop in Mexico City (Nazare, Hamilton & Schmidt, 2016).

Figure 2.6: Examples of the Digital Certificates Project (from Nazare et al. 2016).

The solution does not store the file on the Blockchain, but it saves the proof that an institution has signed a digital document. The educational certification contains information of the program as well as the name of the recipient, the name of the issue, an issue date. In the case of the MIT, the institution's signature, created using a private key only accessed by the Media Lab, is appended to the certificate itself. The certificate is hashed with SHA-256 to validate that nobody tampers with its content. Finally, the hash is recorded as data in a transaction of the Bitcoin Blockchain, proving that the academic certificate was issued "to a certain person on a certain date" (Schmidt, 2015). This system allows the recipient of the degree and other third parties to verify when the certificate was issued, to whom and by whom, corroborating the content of the certificate itself.

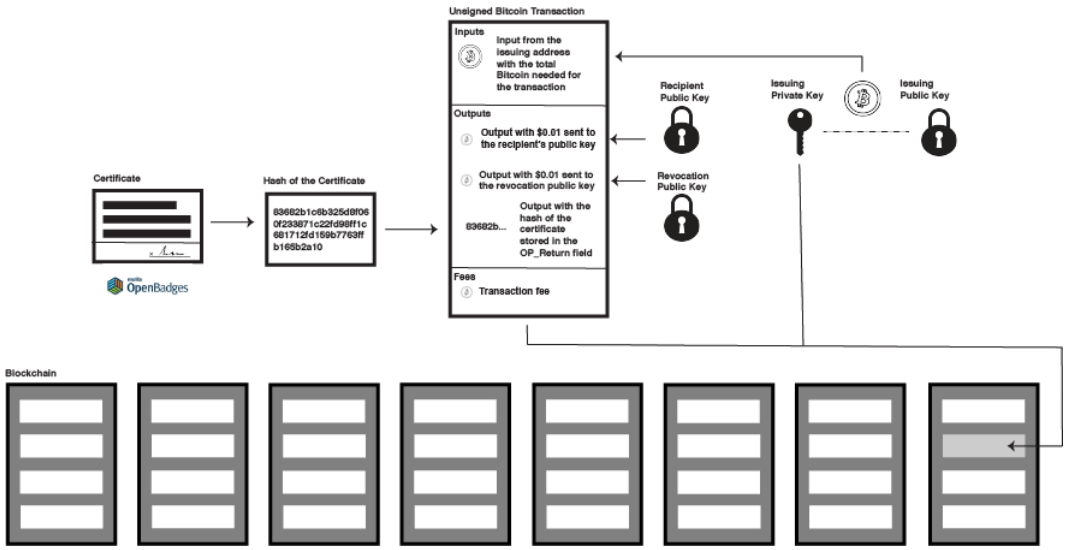


Figure 2.7: Architecture of Blockchain Certification System (from Nazare et al. 2016).

[1] Access the service and the certificates generated at http://certificates.media.mit.edu/
[2] Visit the corporate page of the certification solution at https://www.blockcerts.org/

2.3.2.2 Stamp.io By Stampery

Stampery is a company that uses Blockchain technology to certify who created, accessed or modified documents or datasets in enterprise settings. Their website mentions that "users can seamlessly and irrefutably prove the exact moment data existed for legal, compliance and business purposes".[3] Besides an API that can be used by enterprise customers, they offer a free tool called Stamp.io[4] for timestamping and certifying that individual documents, such as contracts, existed at a certain point in time.

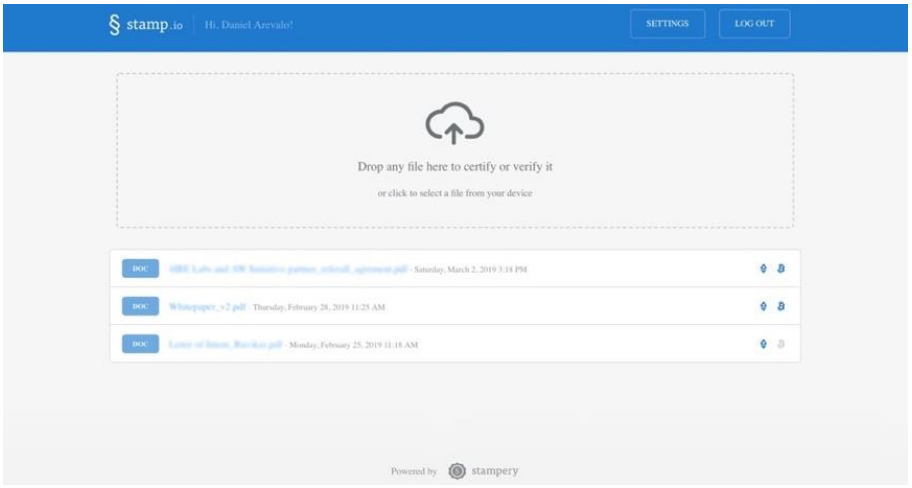


Figure 2.8: Stamp.io service's screen for uploading files.

The Blockchain service let users upload any kind of file and create a digital proof using cryptographic identifiers that irrefutably represent the data. The service not only creates a hash of the document uploaded using SHA-256, but it also provides the data of a Merkle tree structure containing the hashes of other datasets that are processed by their platform. The service then executes a transaction on both the Ethereum and Bitcoin Blockchains registering the Merkle root concatenated to a predefined prefix. After the transaction has been confirmed in the Blockchain, users can download a certificate with the timestamp and other technical details of the transaction. Anyone can share that certificate as a proof of existence, integrity, and ownership for the documents and files that have been uploaded.

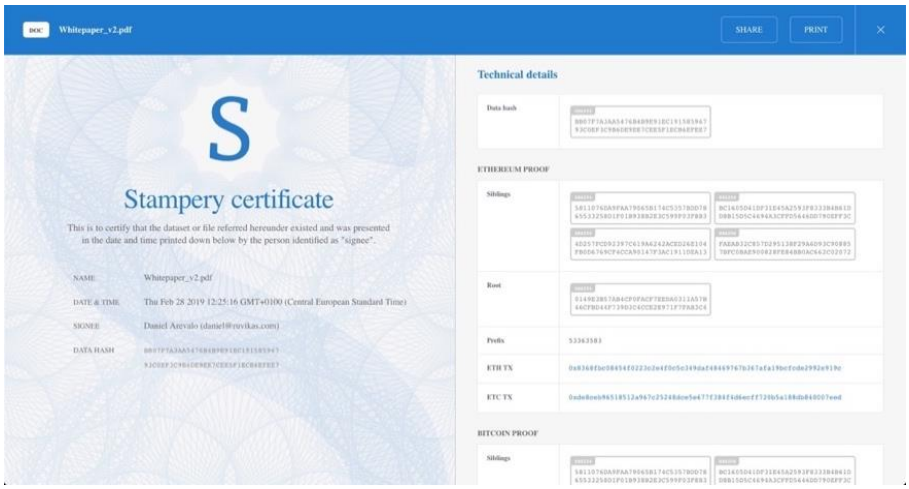


Figure 2.9: Digital Certificate and Download Screen of Stamp.io service.

[3] Visit the corporate page of the certification solution at https://stampery.com/
[4] Access the service and generate certificates at https://stamp.io/

2.3.2.3 Proof.ink on the Steem Blockchain

Proof.ink[5] is an online product that uses the Steem Blockchain to guarantee the integrity of files (Douma, 2019). Unlike the Bitcoin and Ethereum Blockchain that employ Proof-of-Work, the Steem Blockchain uses delegated Proof-of-Stake as consensus mechanism that is more similar to the consensus achieved by a Byzantine Fault Tolerant protocol (Larimer, 2017a). This allows the Blockchain to have block-times of approximately three seconds, require no fees and offer quick confirmations to its users (Larimer, 2017b). Currently, the Steem Blockchain remains amongst the five blockchains with the highest level of activity in the world, largely used as a social media and content-focused platform powered by an internal cryptocurrency named STEEM (Mathis, 2018).

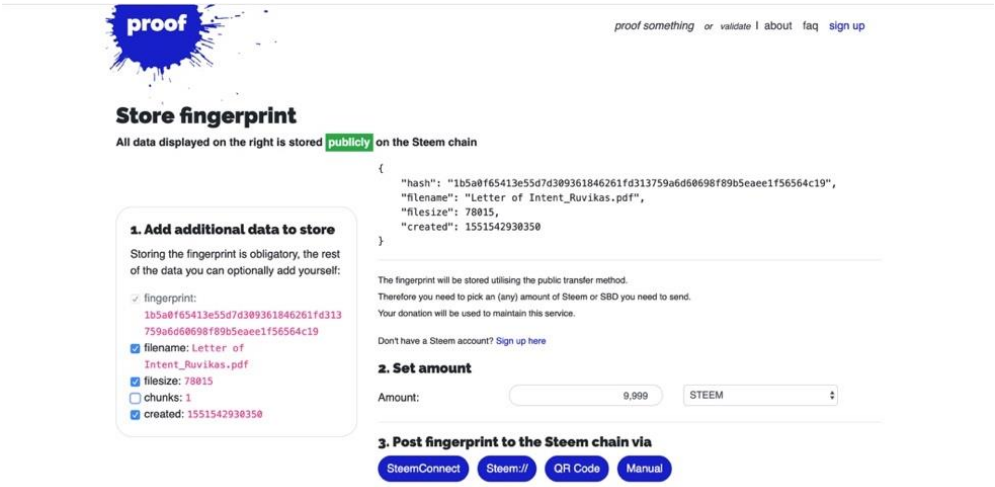


Figure 2.10: Digital Fingerprint and Storing process on the Steem blockchain.

Anyone can use Proof.ink's service to upload files of any size and format (word, excel, pdf, image, etc.) calculating a checksum of the digital data. The unique SHA-256 hash acts as a fingerprint of the document, just like the "unique" fingerprint of any human being. This hash gives a high degree of certainty that the file's content has not been tampered or edited in any form. The fingerprint is then attached to a transaction originated from the website, in which the user needs to provide a minimum amount of STEEM. The user obtains a prompt to sign the transaction from their respective account, making the transfer to Proof.ink's account and recording the data into the Steem Blockchain (Douma, 2019).

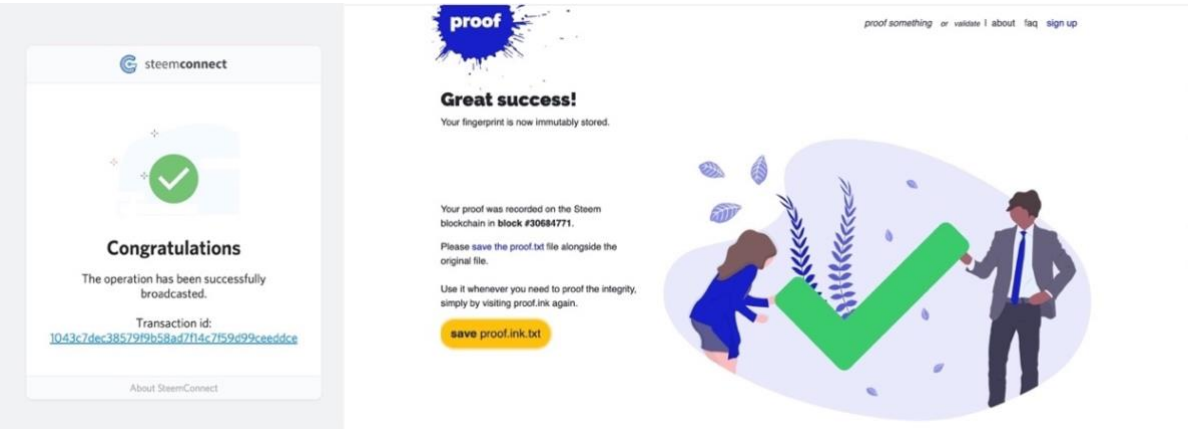


Figure 2.11: Signed Transaction with Steem Wallet and Proof.ink's page confirmation.

[5] Access the service and generate certificates at https://proof.ink/

2.4 Blockchain Explorers and Wallets

The Bitcoin and Ethereum blockchains are fully distributed permission-less and public, which means that their records can be accessed by anyone participating as a node in the network. Blockchain explorers, also called Block explorers, are purpose-built search engines that have been developed to allow users to look up, compare, and verify all the transactions and metadata registered on a determined Blockchain. As Barinov (2018) explains in his article about the importance of block explorers:

> To organize and make sense of this immense amount of data, we use a block explorer. A block explorer allows us to view information contained in any block in the chain, from the genesis block all the way up to the most recently created block. We can check transactions from any address, look at transaction history, and view and verify the contents of smart contracts. It is an essential part of the Blockchain ecosystem.

On the other hand, wallets are Blockchain applications that keep a record of how many crypto assets are assigned to a particular account. Usually, each account is controlled by an RSA Public-Private-Key-Pair (Rivest, Shamir & Adleman, 1978) that the user has to generate and manage properly in order to receive or sign transactions that go through his account. A wallet can be considered a form of block explorer limited to one account because it displays a searchable record of all transactions with the cryptocurrency that have come from or to the user's account address.

Users can verify the transactions made from their wallet and ask for a confirmation to the recipient who can verify that the transaction was successful from their respective client application. Any person aside from the owner should not have access to these wallets and see the transactions. Therefore, third parties that need to verify transactions and the data included inside them have to use a Blockchain explorer. Similarly, other cases obligate users of a wallet or any Blockchain application to distrust these programs and additionally verify the status of a transaction with a block explorer. Errors reported by a wallet or other Blockchain applications can vary from not displaying the information registered in a transaction, to glitches that make the transaction fail to be added in the next block mined. In some cases, these applications can take a long time before telling the user that transaction was erroneously broadcasted to the nodes and therefore never recorded.

Fortunately, end-users of Blockchain products have multiple options for verifying information separately from the program they use to record the data in the first place. Thanks to the benefits of decentralization, users can access a block explorer, or many of them at the same time, and have a valid alternative to revise the correct functioning of the Blockchain product used. They can use a blockchain explorer that can be either an open-sourced and public tool extracting data directly from the Blockchain, or a private custom-built tool for tracking the transactions relevant to specific users.

2.4.1 Verifying Certificates with a Block Explorer

A block explorer is a tool that provides attestation of any action that users take on a Blockchain product. Hence, they are essential to verify that transactions happened as expected and provide another source of unbiased "truth." By enabling users search and correlate the hash recorded as input data in a transaction with the hash obtained from the checksum of a digital contractual document, a block explorer provides attestation of a document's existence and data integrity. Users can use a Blockchain explorer search output to irrefutably verify a certified version of the agreement and request the fulfilment of the terms with the other part from the date its hash was recorded on the Blockchain.

Demonstrating that data in a contract has not been tampered is crucial for both individuals or businesses that require accountability, attribution, and auditability in their processes. Using a Blockchain explorer, anybody around the world can verify the authenticity of a particular digital document without having to rely on a third party or intermediary. However, in case of disputes, sometimes it would be necessary to take the proof to the court. Stampery's website[3] mentions that: "In United States Federal Courts, blockchain-based hashing and time stamping can make for exceedingly strong evidence of the authenticity of a document."

2.4.1.1 EtherScan.io

EtherScan[6] is the most popular Blockchain explorer for Ethereum, typically used to check if a transaction is confirmed, left an account or was received by another. Despite some critics (Barinov, 2018), EtherScan is a widely used tool by the Ethereum community to lookup, confirm and validate transactions. The block explorer has been developed by a small group of people, it's is close-sourced and independent of the Ethereum Foundation, the organization supporting the growth and development of the Ethereum blockchain.

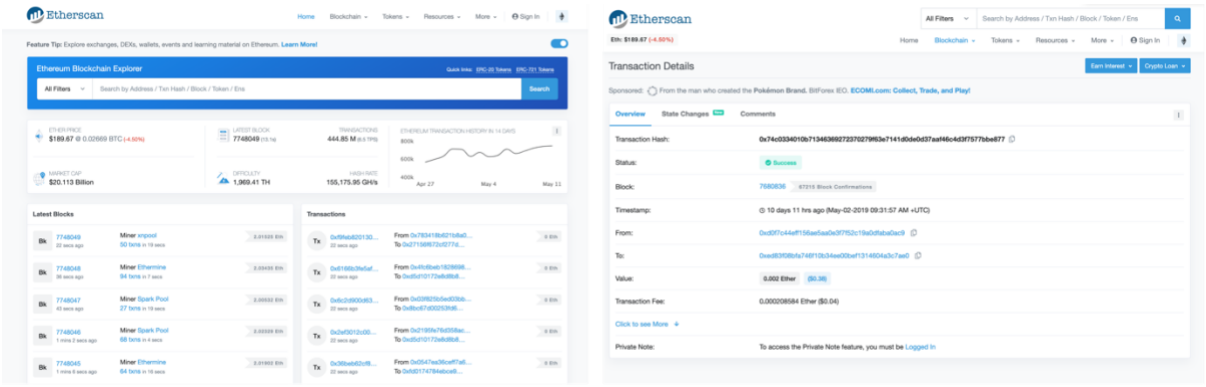


Figure 2.12: EtherScan's User Interface. Home (left) and search result (right) pages.

2.4.1.2 BlockScout.com

BlockScout[7] is an open-source block explorer developed by the POA Network that allows the inspection and analysis of Ethereum-based Networks. Being open-source is the major differentiator of this tool (POA Network, 2018). The Blockchain explorer offers transparency about its functioning, making it more trustworthy within the Ethereum ecosystem to analyse and validate transactions (Barinov, 2018). Unlike other block explorers, anyone can use its code to make a new version of it or improve it.

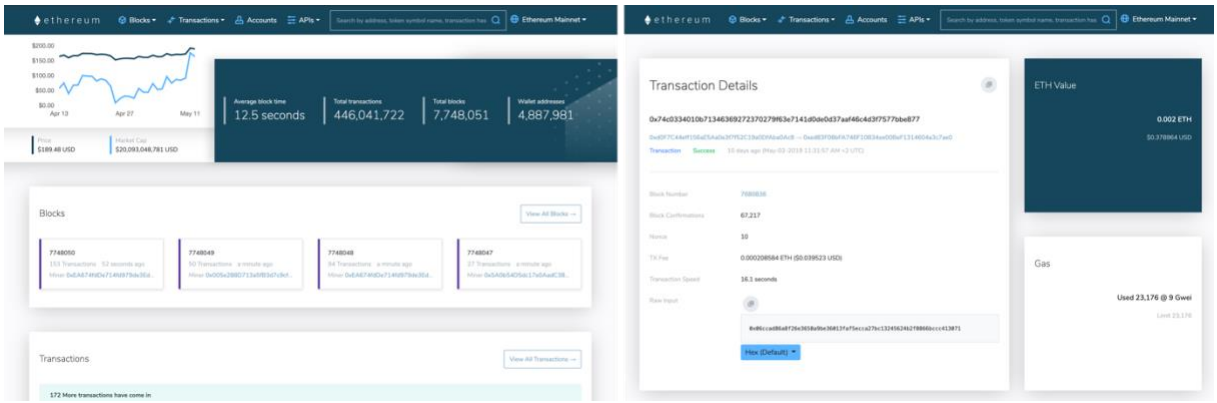


Figure 2.13: BlockScout's User Interface. Home (left) and search result (right) pages.

2.4.1.3 EthStats.io (now Aleth.io)

Recently rebranded as Alethio[8], it is an alternative to the most popular block explorers (Alethio, 2018). The company has been developed as a formation of ConsenSys, one of the largest technology companies focused on building infrastructure and applications for the Blockchain community (ConsenSys, 2019). Called EthStats while it was available in its Beta version and used for collecting user's interaction data presented in the results section of this research, the analytics platform helps users visualize, interpret, and evaluate their interactions with Blockchain accounts, smart contracts and decentralized apps commonly abbreviated as dApps (Crowley, 2019).

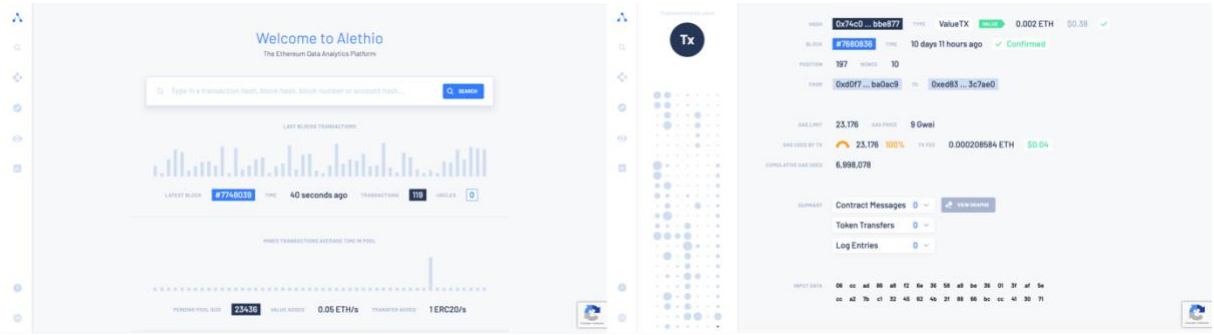


Figure 2.14: Aleth.io's User Interface. Home (left) and search result (right) pages.

2.5 Usability for Blockchain Products

Blockchain is an emerging technology being introduced as a backbone for managing information in new products, enabling a new set of software characteristics that are presented as features through user interfaces. Blockchain products are those digital solutions that utilize Blockchain as a part or the entirety of the back-end supporting the functionality offered by their user interface. A potential explanation for their low adoption is that various Blockchain products have been released before ensuring that they create more value to their users than existing solutions.

However, another problem could be that even if incorporating Blockchain to a solution creates value, the inherent complexity, and technicalities of the technology might cloud the perception from users concerning its value, consistently deterring them from adoption. After all, the functionality and value perceived in digital products have become dependent on the cognitive friction that the user experiences when trying to achieve a goal at using them (Cooper, 2004).

Interaction designers aim to reduce the cognitive friction of a product by improving its usability. For Blockchain products, designers are just starting to evaluate the usability according to the different concepts, methods and metrics available. Analysing how different definitions of usability converge before the evaluation of complex user interfaces help to confront the challenges of designing for the user's interaction with products utilizing Blockchain.

[6] Access the screens shown of the EtherScan Blockchain Explorer at https://etherscan.io/
[7] Access the screens shown of the BlockScout Blockchain Explorer at https://blockscout.com/
[8] Access the screens shown of the EthStats (now Alethio) Blockchain Explorer at https://aleth.io/

2.5.1 Defining Usability

Over the last three decades, usability has been defined by many authors to evaluate new interaction paradigms and software systems that have been developed (Rusu et al. 2015). One of the best known and widely used definitions of usability is proposed on the ISO 9241:11 (International Organization for Standardization, 2018, p.6):

> The extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use.

Despite this seems to be a converging understanding of usability, the ISO definition still gives ample scope for different interpretations. Although the terminology is expanded with many other well-recognized standards (International Organization for Standardization, 2006; 2011), the results of a usability study depend on the interpretation that the designer gives to this definition and the target user interface of the product evaluated during the study (Rusu et al. 2015).

Designers of Blockchain products have to increase usability and compensate for the friction created by the complexity of concepts integrated from cryptography and distributed systems. To increase the usability of a product, they need to evaluate the impact of each element included in a user interface, reflecting on the mentioned high-level factors of usability such as effectiveness, efficiency, and satisfaction.

2.5.1.1 Usability as the Ease of Use

According to Nielsen (Nielsen, 2012), besides usability, user interfaces also need to have the key quality attribute of utility. Utility refers to the design's functionality, meaning that the user interface does or provide the features that the users need to accomplish a goal. On the other hand, usability refers to how easy and pleasant any product and its features are to use. A product that hypothetically does what users want would suffer a lack of adoption if it cannot enable them to be easily operated because the user interface is complicated or confusing. Likewise, any product that could be perceived as easy-to-use would create frustration and will not be used again if the users do not accomplish the task. For Nielsen, "usability and utility are equally important and together determine whether something is useful" (Nielsen, 2012).

User research methods help to improve usefulness by identifying usability and utility issues, making a product as easy to use as possible. Designers employ such methods intending to detect complexity and reduce friction during the interaction with the user interface, hence contributing to increasing the adoption of a product. In this context, designers need to consider usability as a quality attribute that assesses not only the effectiveness, efficiency, and satisfaction but also the ease-of-use of a UI and the methods for improving such characteristics during the design process (Nielsen, 2012).

Nielsen's definition of utility is analogue to the effectiveness factor of a user interface as defined by the ISO standard. On the other hand, Nielsen's definition of usability comprises five quality components: Learnability, Efficiency, Memorability, Errors, and Satisfaction (Nielsen & Loranger, 2006). Considering that the attributes of efficiency and satisfaction included in Nielsen's usability appear in the ISO's standard definition, it is adequate to explore how a product’s ease of use can be tightly related to the aspects of learnability, memorability, and errors.

2.5.1.2 Usability Goals

Employing the ISO definition of usability would be more related to assessing the quality of a product (International Organization for Standardization, 2011), which is useful for comparing effectiveness, efficiency, and satisfaction between competing solutions. However, designers should not only focus on these aspects for products that have yet to be adopted and need to become easy to use. Taking into account Nielsen's definition of usability may become a better strategy for designing Blockchain products because it requires involving users to gain insights about their perspectives. In this regard, Sharp, Rogers, and Preece (2015) also provide a definition of usability that is more aligned to Nielsen's and which makes the convergence of the previous definitions clearer. For them, usability refers to "ensuring that interactive products are easy to learn, effective to use and enjoyable from the user's perspective" (Sharp, Rogers, & Preece, 2015).

Approaching to the definition of usability considering the user's perspective, Sharp, Rogers and Preece (2015) define all the aspects of Nielsen's usefulness (utility+usability) in terms of usability goals, typically operationalized as questions. Considering that the attributes of effectiveness and efficiency are presented as two of such usability goals, and the satisfaction factor–or that an interface is enjoyable–is a common denominator for the three definitions, determining how easy to learn is a user interface becomes prominent in a modern definition of usability. This attribute of a user interface can be evaluated as a usability goal, or component of usability, employing the term learnability (Nielsen, 2012).

2.5.2 Learnability

According to Grossman, Fitzmaurice, and Attar (2009), learnability is an important and well-accepted aspect of usability that has little agreement as to how should be defined, measured, and evaluated. In their paper, they present a survey of the previous definitions, metrics, and evaluation methodologies for software learnability. For them, the most agreed definition of learnability is: "The system should be easy to learn by the class of users for whom it is intended" (Michelsen et al. 1980). However, this definition can be broadly used, so the characteristics related to the interaction and the users' needs have to be specified before the definition can be used for evaluation.

There is a large number of dimensions upon which learnability can be considered; therefore, it is necessary to outline the concept of learnability that will be evaluated in the various aspects of a user interface. The most important consideration is concerning the learnability scope. Designers need to decide whether the evaluation focuses on initial learnability, and thus the interaction between user and system happening on a single time frame, or the evaluation focuses on the interaction over a long timeframe, and thus involving the study of extended learnability (Grossman, Fitzmaurice, & Attar, 2009).

Certainly, there are some systems where users have to be trained in order to overcome a hard-to-learn interface, but in most cases, systems need to be easy to learn. Learnability is sometimes considered the most fundamental usability attribute, because "most systems need to be easy to learn and because the first experience most people have with a new system is that of learning to use it" (Wilson, 2010).

2.5.2.1 Measuring Initial Learnability

Grossman, Fitzmaurice, and Attar (2009) extensively surveyed literature related to the measurement of learnability and found that various metrics exist, but "they are scattered

across various research papers over the last two decades." Measuring a system's learnability during an initial interaction falls in two possibilities for designers. On one hand, designers can focus on measuring the time it takes for a user "to rapidly begin to work with the system" (Holzinger, 2005) or "learn how to use the commands relevant to a set of tasks" (Shneiderman, 1997). On the other, it can require analysing how the users become competent at carrying out tasks without much effort (Sharp, Rogers, & Preece, 2013), as Santos and Badre (1995) put it: "the effort required for a typical user to be able to perform a set of tasks using an interactive system with a predefined level of proficiency."

According to Nielsen, learnability is addressed with the question: "How easy is it for users to accomplish basic tasks the first time they encounter the design?" (Nielsen, 2012). Notably, such ease of learning just considers the experiences of a novice user during the initial part of the learning curve enabled by the system (Nielsen, 1993). Almost all user interfaces have steep learning curves that start out with the user having zero efficiency when using them for the first time (Wilson, 2010). Cognitive processes of learning are triggered by learnable systems during such initial interaction. The goal of the user interface according to Nielsen is to enable users that have not interacted with the system before "to reach a reasonable level of usage proficiency within a short time" (Nielsen, 1993).

In every study about the learnability of a user interface there are discrepancies in what a "short time" and "reasonable level of proficiency" would be for each user (Grossman, Fitzmaurice, & Attar, 2009). These two characteristics are constantly included in many collections of learnability metrics. But the fact that there is not one single collection of learnability metrics that can be applied to all user interfaces makes clear that evaluating learnability is a complicated task and that designing user interfaces with a high degree of learnability will always require compromises according to the context of use for which the interactive system is designed. Despite this potential shortcoming, defining learnability based on initial user experiences is common to identify major flaws in user interfaces. As designers lack a set of well-accepted metrics for learnability, methodology to evaluate this aspect of usability can always be adapted from research (Grossman, Fitzmaurice, & Attar, 2009), allowing designers to obtain good qualitative data from the user’s interaction with new products and provide solutions to further improve their user interface.

2.5.3 Learning is a Cognitive Process

Independently of the approach that the designer takes to measuring learnability, they need to start by recognizing the cognitive processes that users endure when using a system. Cognition is described in terms of specific kind of processes that function in an interdependent manner in the mind of the user, and one of these processes is Learning. (Sharp, Rogers, & Preece, 2013). Furthermore, Norman (1993) defines cognition in two modes: experiential and reflective. The first one being the state of mind in which users perceive, act, and react to events effectively and effortlessly. The second kind of cognition involves thinking, comparing, and making decisions (Sharp, Rogers, & Preece, 2013, p.66).

Even more, Cognitive Engineering was invented to reflect upon the application of cognitive science to the design and construction of machines. The term refers to the process of analysing why many things, not only those associated with computers, are difficult to use. The goal of cognitive engineering is identifying significant difficulties in understanding and using the most complex devices, or as Donald Norman (1986, p.32) puts it:

> To come to understand the issues, to show how to make better choices when they exist, and to show what the trade-offs are when, as is the usual case, an improvement in one domain leads to deficits in another.

Furthermore, the work that Donald Norman did with Jim Hollan and Ed Hutchins titled "Direct Manipulation Interfaces" (1986) establishes that the discrepancies between users and systems can be defined under the concept of directness. Directness is the feeling or impression that a user interface gives when being used to accomplish a task and results from the commitment of less cognitive resources. Therefore, the need from a user to commit additional cognitive resources in the use of an interface leads to the feeling of indirectness. They recognize that the sensation of directness is always relative to the conditions of the interaction and the user's characteristics. Furthermore, they clarify that even if there might not be a way to directly measure the trade-off values in a cognitive process, establishing a framework to analyse what is being traded off against what qualitatively can help designers to understand the dynamics of interactions between users and systems.

2.5.3.1 The Gulfs of Evaluation and Execution

Applying Cognitive engineering over the concept of directness has led to a cognition model that encompasses how people actually interact with things. Norman (1986, p.38) describes a theory of action that models the cognitive efforts of a person interacting with a computing system. He explains that:

> The person's goals are expressed in terms relevant to the person–in psychological terms–and the system's mechanisms and states are expressed in terms relative to it–in physical terms. The discrepancy between psychological and physical variables creates the major issues that must be addressed in the design, analysis, and use of systems.

Hutchins, Hollan and Norman (1986) advocate for the representation of such discrepancies as two gulfs that must be bridged, these are called the *Gulf of Execution* and the *Gulf of Evaluation* (Figure 2.15).

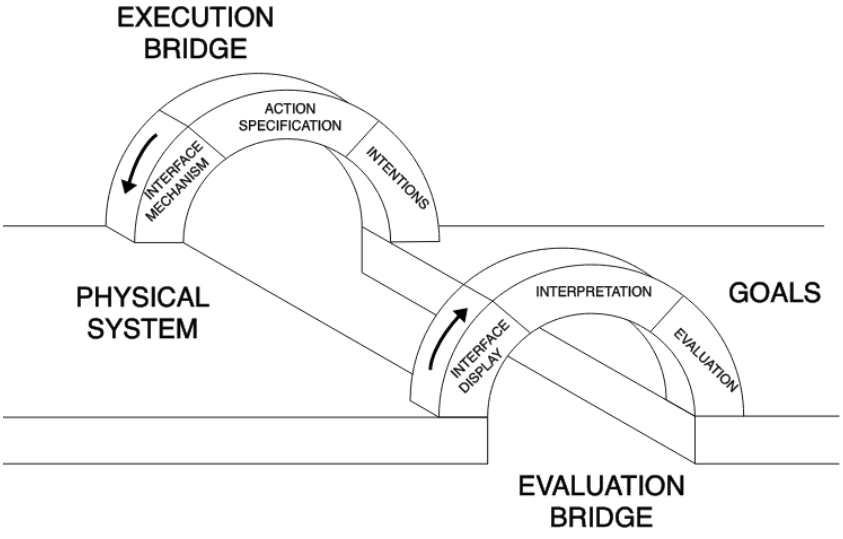


Figure 2.15: Gulfs of Evaluation and Execution (Norman, 1986).

The gulfs provide a representation of a distance between the system and the user. Such distance is bridged by introducing a user interface. Bridges covering the distances between the gulfs need to be built in both directions before the interface is presented to the user. It is the designer's job to construct the input and output characteristics of the interface to make better matches to the psychological needs of the user.

The *Gulf of Execution* is bridged from the system side when the designer builds the input characteristics of the interface, such as the commands and mechanisms of the interface that match the thoughts and goals of the user. The *Gulf of Evaluation* is bridged when the designer builds the output characteristics of the interface, such as visual representations and information presented on the interface or other assistant material.

2.5.3.2 Crossing the Distance Between User and System

Designers wishing to increase a product's ease-of-use have to bridge the gulfs in order to minimize the cognitive effort required from a user to understand the system and perform a task on the user interface. The execution part of a task means that the user crosses the distance between the gulfs in just one direction, going back to the other direction by evaluating the status of the system. In other words, for every interaction with a system, designers have bridged the gulfs and users cross them (Hutchins, Hollan & Norman, 1986). The job of designers requires the construction of physical characteristics of the interaction bridges from the assumptions they have made about the speed and direction in which the users will cross them, attempting to match the user's psychological characteristics perfectly.

In this sense, the distance that users need to cross by carrying out tasks usually is not covered linearly and efficiently. If users are not familiar with the system, the language used or the task, they need to learn how to use the interface in the way intended by the designer. Usually, when the bridges do not support the users adequately, it is likely that they will end up with an incorrect understanding of the system. Trying to understand the user interface or complete the task, they will employ resources for the cognitive process of learning but end up misusing the system, making errors, or even dropping its use entirely.

In the next chapter, this thesis presents the framework necessary to evaluate both sides of the interaction focusing on the aspect of learnability. On one side, there are features related to the cognitive process of learning in the mind of the user by defining the user's cognitive activities while performing a task (Hutchins, Hollan & Norman, 1986). On the other, the author presents the suitability for learning of a UI as the characteristics that digital products should have to support a learning interaction adequately (International Organization for Standardization, 2006). The user interface target user interface is that of Blockchain explorers supporting users to complete the task of verifying the certification of a contract on the Ethereum Blockchain.

3 Methodology

Due to the novelty of the Blockchain technology and its applications, few designers have considered evaluating individual aspects of a Blockchain product's usability such as Learnability. In such a context, designers cannot be sure if there are differences between the usability of products that utilize Blockchain and those that do not require it altogether. Grossman, Fitzmaurice and Attar (2009) assert the discrepancies in research about the learnability of user interfaces, quoting Santos and Badre (1995) statement that: "despite the consensus that learnability is an important issue in usability, few of those authors discuss at length the issue of learnability evaluation." Additionally, there is limited research about learnability despite being important for the adoption of information technology (Lazar, Feng, & Hochheiser, 2010, p. 44). Consequently, research about learnability is even more scarce for products that display information recorded on Blockchains.

The author of the thesis gives a response to the research question "**How Usable are Block Explorers for Attesting Agreements Registered on a Blockchain?**" by evaluating the learnability aspect of a block explorer's user interface. On this chapter, the methodology addresses the process to answer two sub-questions:

a) Are the recommendations of the Suitability for Learning principle considered in the design of block explorers' user interfaces to assist users in verifying the certification of documents via transactions recorded on a Blockchain?

b) *Does the learnability of the selected block explorer's user interface adequately support the verification of cryptographically certified documents for first-time users with varying degrees of domain knowledge about Blockchain?*

For the first part of the chapter, three block explorers go through a formative expert assessment methodology to answer the first sub-question. The expert assessment is structured as a combination of the heuristic evaluation and cognitive walkthrough methods under a feature analysis approach. A *Framework for Learnability Evaluation* that combines the effects of these usability inspection methods is constructed based on proposed learnability heuristics and the cognitive activities of users performing a task. The features and categories of the evaluation framework answering the first sub-question are described. They were established by mapping the ISO recommendations of a user interface's suitability for learning on the user's cognitive activities during the completion of a task.

The next part of this chapter explains a user-based usability test starting by the conduction of a cognitive task analysis. With this method, the author collects qualitative observational data from end-users about the challenges they have at interacting with the user interface to find an answer to the second sub-question. After the test is finished, the researcher conducted semi-structured post-test interviews with each of the participants. The questions in the interview seek information related to the seven recommendations that serve as a guide for the suitability for learning of a user interface (International Organization for Standardization, 2006). The goal at considering the same heuristics that form the basis of the Expert Assessment is to give more context to the observations made during the empirical user test and complement the answer to the second research sub-question.

3.1 Evaluating Learnability

The definition of usability given by the ISO standard 25010 emphasizes the three key factors of effectiveness, efficiency and satisfaction (International Organization for Standardization, 2011, p. 6). However, the standard refers to learnability as a sub-characteristic of usability on a quality in use model extended to include other product quality sub-characteristics such as user error protection and user interface aesthetics. Likewise, Nielsen (2012) stablishes that the kind of evaluations that should be made for each of the six components of usability might be distinct. For example, effectiveness, efficiency and satisfaction are usually measured after the use of a system to accomplish a goal. While issues regarding learnability, memorability and errors made by the user can be observed and measured during the use of a product even before completing a task.

Additionally, according to the ISO specifications: "In practice, within design situations for an interactive system, compromises will be made" (International Organization for Standardization, 2006, p.4). The applicability, priority and relative importance of each aspect varies with each usability research because of the context in which the user interface is evaluated. Such *context of use* and other design requirements are generated by the specific application, target user groups selected, the environment and the form of interactions to be studied.

The importance of learnability within the study of usability can be clearly seen in the IEEE's definition of usability as "the ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component" (IEEE, 1990). Such definition also evidences that other aspects of usability might be overlapping with learnability, and the evaluation of this one component might give results that also apply to evaluate the others. Since mapping the learning cognitive process of the user when interacting with an interface can be done by evaluating its learnability aspect, the methodology described in this chapter focuses on evaluating this one component of usability. More specifically, initial learnability.

After identifying a phenomenon in the interaction between the user and system, through secondary research designers create or select a cognitive model that helps them frame the particular aspect of usability they want to evaluate (Sharp, Rogers, & Preece, 2015). Next, they can employ qualitative and quantitative research methods to gather insights into user needs and challenges in relation to the system's user interface. Lewis (2014) explains that these methods can fundament an iterative usability evaluation. The iterative methodology is characterized by the empirical process of acting upon insights from a usability research to generate a new or comparable solution which can be again researched repeating the cycle (Al-Awar, Chapanis & Ford, 1981).

Ideally, during the iterative design process, practitioners should use two conceptualizations of usability: Summative and Formative (Lewis 2014). In general terms, summative usability focuses on metrics while formative usability focuses on problems detection and associated design solutions (Rusu et al. 2015). This distinction can also be made for learnability evaluations. As Grossman, Fitzmaurice and Attar (2009) state, "formative evaluations should expose learnability issues, while summative evaluations should provide an overall assessment of the system's learnability." While the later methods are based on measurements of the interaction, the former specialize on the diagnostic based on user observation.

For this thesis, these two types of evaluations are utilized in a methodology that combines traditional usability research methods to evaluate the learnability of block explorers,

demonstrating the process for designers, developers and researchers that wish to evaluate the same or other usability aspects of Blockchain products. The outcomes of using this methodology are considered the first part of one design iteration, which can be completed only after redesigning the product with the insights obtained from an analysis of results.

3.1.1 Suitability for Learning of User Interfaces

The visualization of the distance between user and system through the analogy of the gulfs allows a designer to better understand what is happening in the user's mind. For each interaction, defining the steps that occur during the accomplishment of the task help to localize the cognitive processes of learning. Since this thesis is concerned with evaluating the cognitive process of learning, it is necessary to describe a set of rules (heuristics) that let designers assess whether the concept of learnability is well supported by the user interface. The ISO 25010 (2011) also states that learnability can be defined in terms of a *Suitability for Learning*, one of seven Dialogue Principles. A user interface enables the occurrence of dialogues defined as:

> A dialogue is the interaction between a user and an interactive system as a sequence of user actions (inputs) and system responses (outputs) in order to achieve a goal, where user actions include not only entry of data but also navigation and other(control) actions of the user. (ISO, 2006, p.2)

A part of ISO 9241 (9241:110) sets forth these ergonomic design principles, "presented without reference to situations of use, application, environment or technology" (ISO, 2006). The fourth of the dialogue principles described in the standard, titled *Suitability for Learning*, state that "a dialogue is suitable for learning when it supports and guides the user in learning to use the system" (ISO, 2006).

3.1.1.1 Learnability Heuristics

Suitability for Learning as a principle contains 7 recommendations that can be considered the base of a framework for the development of dialog requirements or the design of specific solutions with high learnability. The ISO 9241:110 standard (2006) establishes that dialogues designed in accordance with the provided recommendations help to prevent users experiencing typical usability problems related with the ease of learning. Therefore, for this research the recommendations are taken as a set of heuristics (**H#**) that allow designers to evaluate the learnability of a user interface:

> **H1:** Rules and Underlying concepts which are useful for learning should be made available to the user.
>
> **H2:** If infrequent use or user characteristics require relearning of the dialogue, then appropriate support should be provided.
>
> **H3:** Appropriate support should be provided to assist the user in becoming familiar with the dialogue.
>
> **H4:** Feedback or explanations should assist the user in building a conceptual understanding of the interactive system.
>
> **H5:** The dialogue should provide sufficient feedback about the intermediary and final results of an activity so that the user learns from successfully accomplished activities.
>
> **H6:** If appropriate to the tasks and learning goals, the interactive system should allow the user to explore ("Try out") dialogue steps without negative consequences.
>
> **H7:** The interactive system should enable the user to perform the tasks with minimal learning by entering only the minimum amount of information required in the dialogue, with the system supplying additional information on request.

3.1.1.2 User's Cognitive Activities

Norman (1986) proposes that the process of performing and evaluating a task can be approximated by seven stages of user activity (Figure 3.1):

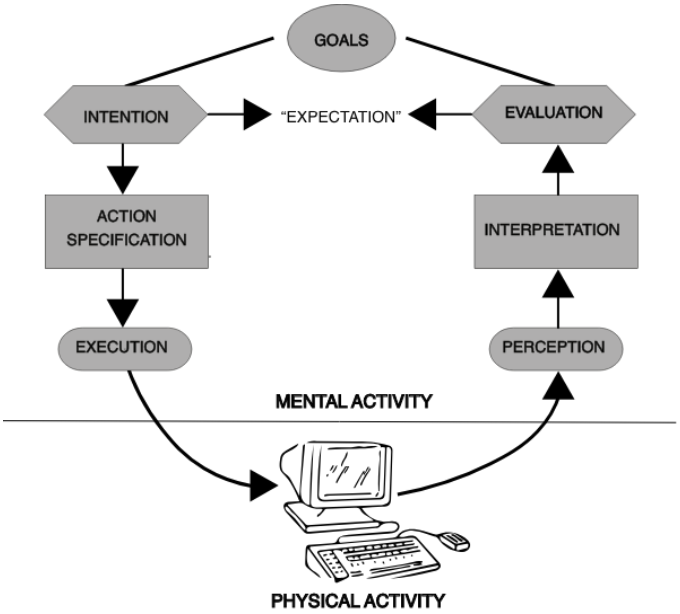


Figure 3.1: Seven user cognitive activities when performing of a task (Norman, 1986).

The stages can occur without any order, skipped, or even repeated. Moreover, users may not progress in the same pattern sequence all the time. The **A1) establishment of the goal** is the only required stage, happening before or after crossing one of the bridges. Users could have established a goal before having not even seen the system or they could start by evaluating the status of the system before deciding what to do. As such, sometimes the establishment of the goal comes from an interface (i.e. an event triggers a notification that makes the user react) and others it comes from the users (i.e. the user receives instructions to follow on a new interface). The user must diagnose the situation displayed on the interface or interiorize the instructions, appropriately establishing the goal and the emergent tasks to achieve it.

The *Gulf of Execution* is crossed by the user through three stages: **A2) forming the intention, A3) specifying the action sequence**, and **A4) executing the action by making contact with the input mechanisms of the interface.**

The intention is the first step at crossing the *Gulf of Execution*, it occurs when the language of the system enable thoughts by the person (Hutchins, Hollan & Norman, 1986). The user forms plans, action sequences, and interpretations moved by the goal that they are trying to accomplish with the system. During the second step, the user decides about just one specific action sequence. These two events occur mentally, right before the user physically interacts with the user interface. The third step involves the actual execution of an action. Crossing this gulf, users are bound to do something, whether it is to select information on the interface, communicate through it with written or verbal commands or to perform a complex motor sequence that is registered as an input.

The *Gulf of Evaluation* is crossed in other three stages: **A5) perceiving the system state, A6) interpreting the state,** and **A7) evaluating the interpreted state with respect to the original goals and intentions.**

When crossing the bridge in this direction, the first step for users is a perceptual processing to determine what is the system state by acquiring information directly from the user interface. The step of interpreting this information relies on complex relationships to the psychological variables in the mind of the user. Finally, reaching the other side of the gulf is accomplished after evaluation. This last step is just a thought process for checking that the intended actions have been executed properly by comparing the interpretation of system state with the original goals and intentions. If satisfied, the user diagnosis leads to either finish the interaction or the establishment of a new goal followed by the formation of a new intention, so the user starts crossing in the opposite direction.

3.1.1.3 Cognitive Model for Learnability

Models are typically abstracted from a theory coming from a contributing discipline that can be applied to interaction design (Sharp, Rogers, & Preece, 2015). First, the suitability for learning recommendations are presented again as a set of heuristics (**H#**) in the same order as proposed by the ISO 9241:110 standard (2006).

> **H1:** Rules and Underlying concepts which are useful for learning should be made available to the user.
>
> **H2:** If infrequent use or user characteristics require relearning of the dialogue, then appropriate support should be provided
>
> **H3:** Appropriate support should be provided to assist the user in becoming familiar with the dialogue.
>
> **H4:** Feedback or explanations should assist the user in building a conceptual understanding of the interactive system.
>
> **H5:** The dialogue should provide sufficient feedback about the intermediary and final results of an activity so that the user learns from successfully accomplished activities.
>
> **H6:** If appropriate to the tasks and learning goals, the interactive system should allow the user to explore ("Try out") dialogue steps without negative consequences.
>
> **H7:** The interactive system should enable the user to perform the tasks with minimal learning by entering only the minimum amount of information required in the dialogue, with the system supplying additional information on request.

Second, each of these heuristics is localized as part of a user cognitive activity happening in either the *Gulf of Execution* or the *Gulf of Evaluation*. This is achieved by examining the relationship between each learnability recommendation and the user activities (**A#**) that occur during the interaction between user and system.

> **A1:** Establishment of the goal (Before or after crossing any of the two Gulfs)
>
> **A2:** Forming the intention (Gulf of Execution)
>
> **A3:** Specifying the action sequence (Gulf of Execution)
>
> **A4:** Executing the action (Gulf of Execution)
>
> **A5:** Perceiving the system state (Gulf of Evaluation)
>
> **A6:** Interpreting the state (Gulf of Evaluation)
>
> **A7:** Evaluating the interpreted state (Gulf of Evaluation)

Third, the cognitive model for learnability is created (Figure 3.2) combining the recommendations from the dialogue principle corresponding to the *Suitability for Learning* principle and the cognitive activities of the interaction between user and system.

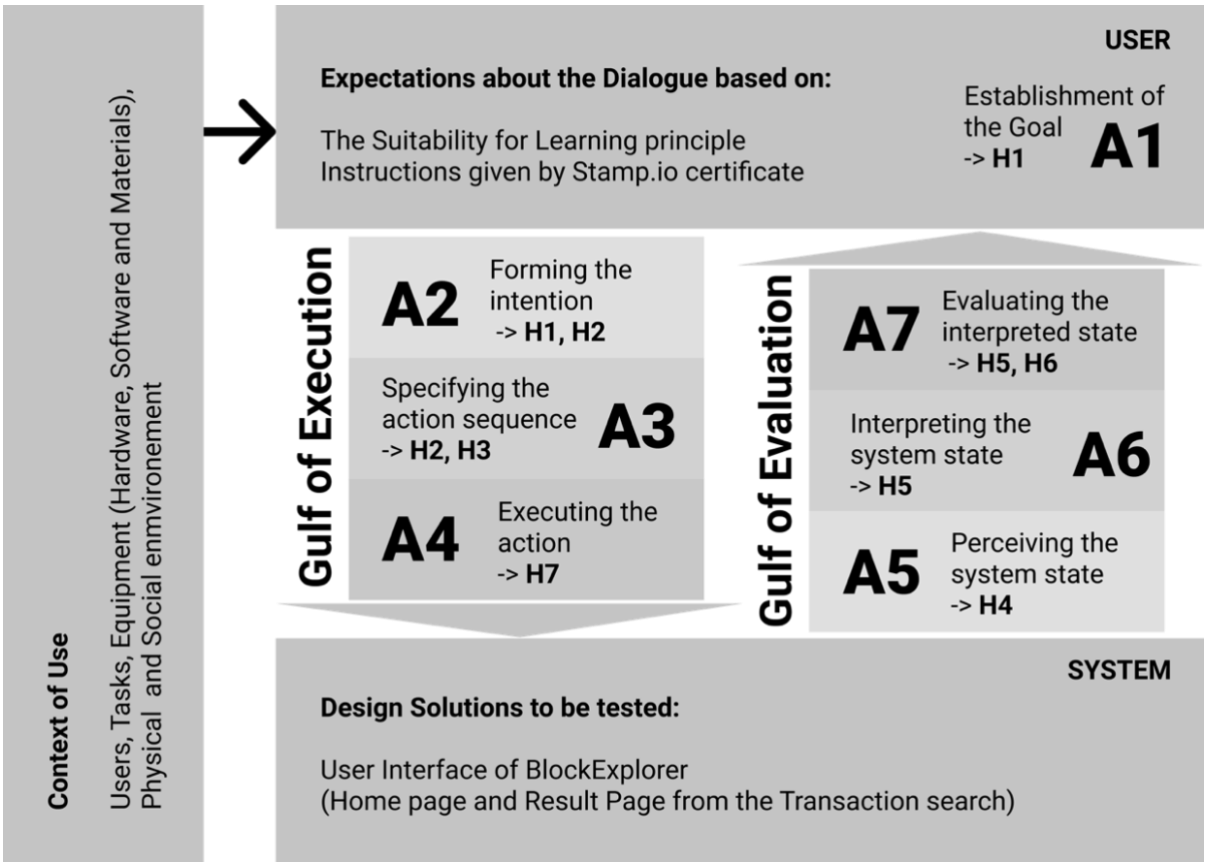


Figure 3.2: A Cognitive Model for Learnability of User Interfaces

The **H#** represents the Heuristic for evaluation of learnability and the **A#** represents the User's Activity during the interaction. A convention in the model can be read as:

> **In [the System], [H1]** Rules and Underlying concepts which are useful for learning should be made available to the user **when [the User] is [A2]** Forming the intention **for [a task].**

Similarly to the recommendations proposed by the ISO standard (2006, p.1), this model does not consider the impact of additional components of design presented by the user interface such as marketing, aesthetics, or corporate design. User interface designers can use this model and create a framework for the analysis, design and evaluation of the learnability of interactive systems.

3.2 Context of Use

For evaluating the learnability of any a Blockchain product's user interface, it is essential to describe the context of use as comprehensively as possible. The ISO 9241:11 (2018, p.4) defines context of use as the "combination of users, goals and tasks, resources and environment." Since the "use of the same system, product or service can result in significantly different level of usability depending on the goals, the types of users and other components of the context of use" (ISO, 2018, p. 7), explicitly specifying these elements allow other designers, developers and researchers to characterize the exact manifestation of usability that this evaluation of learnability conveys and how the outcomes are related to other concepts of usability (Figure 3.3).

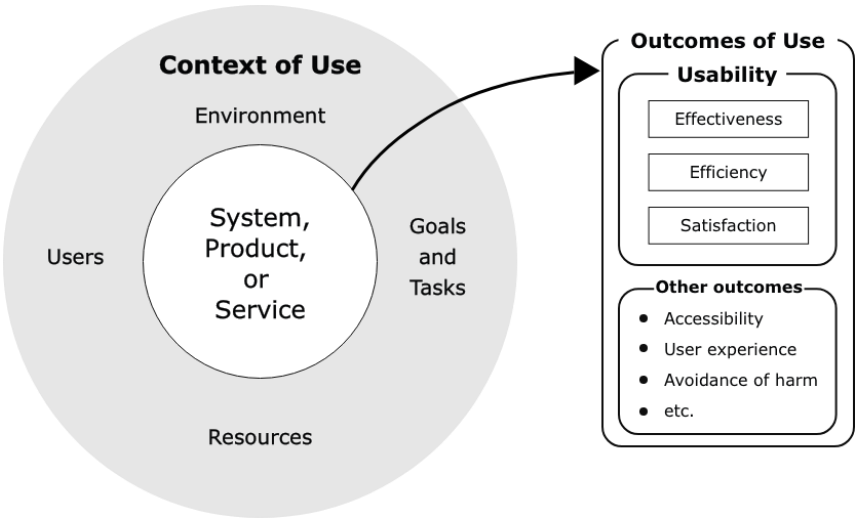


Figure 3.3: Usability in a context of use (ISO, 2018, p. 7)

3.2.1 Users

According to the ISO 9241:11 (2018, p. 3), a user is the person who interacts with a system, product or service. In business settings, the word "users" refer to the target group of the product, the people for which the product was designed and expected to be used by in the real-world. Alternatively, users can be just a subset of people selected as the participants involved in a controlled usability study. Recognizing how they are differentiated is highly important since each kind of user may perceive a product completely differently. For example, real users, when evaluated in a real-world setting, are "more likely unbiased compared to participants in a usability study" (Speicher, 2015).

3.2.1.1 Sampling

For this research, purposeful sampling is used (Palinkas et al. 2015). This non-probabilistic sampling technique, also known as selective sampling or purposive sampling, relies on the judgement of the researcher when it comes to selecting the users that are part of the study. The main goal of this kind of sampling is to focus on particular characteristics of a population which will best enable the researcher to answer the research questions (Lund Research Ltd, 2012). This sampling method is chosen because this research addresses sub-questions that attain to a group of people with specific characteristics in terms of technical knowledge, but also must present variability in terms of domain knowledge about one particular concept, Blockchain. Additionally, employing a purposeful sampling technique becomes particularly useful for exploratory qualitative research that counts with limited resources and focus on a single user interface.

The users that shall be selected for the test (testers) are individuals that have different levels of domain knowledge in regard to Blockchain and its applications. Though, they hold a basic, or above, expertise interacting with digital user interfaces of applications for managing data, ranging from accounting software to Microsoft Excel or similar spreadsheet tools. This particular combination of expertise and domain knowledge form the basis of this research's testing group (Figure 3.4), thus only users that are considered technically savvy were recruited for an exploratory assessment of the learnability of interfaces that display information related to Blockchain.

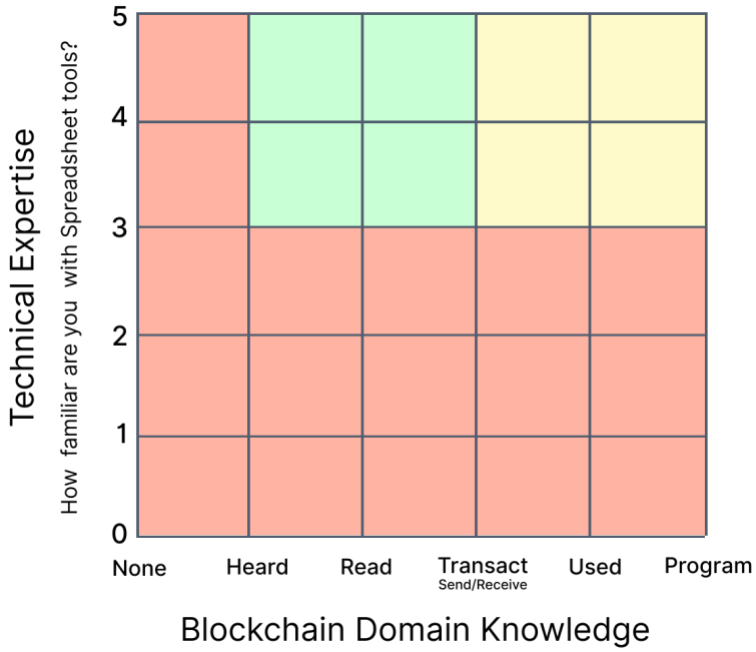


Figure 3.4: Criteria for Tester selection with Screening Questionnaire

The green area in the chart shows the criteria to select the users after a screening questionnaire. The yellow area represents the criteria for users that also qualify to become part of the research, but they are testers for which an assessment of learnability about the presented user interface might be affected by preconceptions generated from the interaction they have had with other Blockchain products. The red area represents the rejection criteria according to the people's results in the screening questionnaire.

3.2.1.2 Screening Questionnaire

Screening questions are prepared in order to assist in the qualification of users that meet specific characteristics to become part of the test. This selection process is done to omit the individuals that are not the most likely intended target group to interact with a block explorer's user interface. A simple questionnaire is provided to all candidates. The answers generate the thresholds beyond which people are considered fitting to become part of the study as testers.

1) Have you signed a contract with a digital tool or added your signature to a PDF file? (Y/N)

2) How familiar are you with digital systems to edit and save data such as spreadsheet tools? E.g. excel or accounting software (Ordinal scale from 1 to 5 with explanation)

3) Have you heard about the term blockchain? (Y/N)

4) Have you read/researched about Blockchain technology or its application? (Y/N)

5) Have you made or told other person to make a transaction on a blockchain? i.e sent or received cryptocurrency. (Y/N)

6) Have you used a decentralized application or dApp? (Y/N)

7) Have you programmed on a blockchain? e.g. smart contract or token issuing (Y/N)

8) Would you be willing to test a user interface that makes use of Blockchain and displays information recorded on it to assess the value of a product. (The movement of the cursor will be recorded together with your voice and utterances at the moment of interacting with the system. Data that personally identifies you will not be collected) (Y/N)

The first two questions assess the user's expertise signing contracts on the computer and the use of digital user interfaces for managing data in spreadsheet. The next five questions have binary answers regarding the candidate's domain knowledge about Blockchain and its applications. The last question is added as a formality to receive the consent of the user before becoming testers in this research. Aside from these questions, there are two questions that pertain to demographic data (gender and age) which had no effect on filtering out people that did not make part of the test. A copy of the screening questionnaire provided to candidates is added as Appendix 1.

3.2.1 Environment and Resources

As part of a context of use, the environment describes the surrounding conditions in which an interaction between user and system takes place during a test, "included the technical, physical, social, cultural and organizational elements" (International Organization for Standardization, 2018, p. 16). In essence, the environment is the setting in which the evaluation of the usability of a product occurs and should be described as precisely as possible. On the other hand, the resources include the equipment such as the hardware, software and physical materials with information that are provided to the tester in order to accomplish the tasks (International Organization for Standardization, 2018, p16).

Any aspect of the usability of a web interface differs if the user interacts with it from a desktop computer or a handheld mobile device, in consequence it is established that for this research the user interface is accessed from a browser application of a laptop computer. Additionally, the user receives two printed papers, one describing the scenario and other being a certificate which contains the information necessary to start the interaction with a block explorer web application running on the laptop computer. Furthermore, the environment is that of an office, in which the general characteristics of the setting are similar to a space that can be regarded by the testers as a place to do their everyday work activities. The presence of an observer is the most obvious indication that the test is conducted as a "lab study" and the environment might not necessarily be the same in the real world.

3.2.2 Goal and Tasks

The ISO 9241:11 (2018, p. 3) describes a goal as the intended outcome of an interaction between user and system, while the tasks can be a series of activities undertaken by a user in order to achieve a specific goal. The goal for this research can be established by the user from a scenario provided in a sheet of paper (Appendix 2). A part of the scenario description reads:

> "John offers to create a certificate of the agreement using a tool called "*Steempery.*" Such tool generates a PDF file that acts as certification. He has sent the certificate to you via email saying that you can trust he won't change the contract because the version you signed is impossible to change."

The scenario requires an interaction between the facilitator and testers for introducing the certificate "received by mail" and then offering the laptop with the target user interface to be tested. The second part of such facilitation only happens if the tester establishes the goal of using the product to be tested after reading the certificate. The goal is hinted as:

> "You will need to confirm that you have a proof that the contract won't be able to be tampered with. You should feel capable of successfully take legal measures in case that he doesn't perform or live up to his part of the contract."

3.2.3 Product

The hypothetical certification service called "Steempery" in the scenario, is the responsible for generating the certificate that is given to the testers. Some clarifications about its functionality are given in Table 1.1. The table shows the features that "Steempery" shares with the certification tools reviewed in the previous chapter. Despite the differences in regard to features, all these services provide the same value to their users: Timestamping the proof of data integrity of a document by publicly storing a file's fingerprint as data input of a specific Blockchain's transaction, delivering a certificate with the data necessary to find such proof on the distributed ledger.

Feature	*MIT Certificates*	*Stamp.io*	*Proof.Ink*	*Steempery*
Fingerprint with SHA-256	✓	✓	✓	✓
Simple hash-storage structure	×	×	✓	✓
Transaction between two Accounts	✓	×	✓	✓
Store file in Service's Database	✓	×	×	✓
Use of the Ethereum Blockchain	×	✓	×	✓
Link to verify with Block Explorer	×	✓	✓	✓

Table 3.1: Feature comparison between Blockchain certification services

After comparing the different Blockchain certification services, "Steempery Certificate" was chosen as the brand for the printed certificate (see Appendix 3), honouring the inclusion of features that the other services have. The name and logo are a combination of Stampery, the company that developed Stamp.io, and Steem, the Blockchain on which proof.ink operates. The hypothetical service creates certificates for files of any size and format, backing up a copy of them in the company's server as the MIT does for their alumni. The three Blockchain-based certification services reviewed require that their users or other parties visit block explorers for verifying the validity of their certificates. However, just two of them provide a direct link to a Blockchain explorer so their users can validate that the transaction made was successful. Similarly, the certificate given to testers in this research instructs them to validate the certification using a specific block explorer.

All the Blockchain explorers share the same search functionality despite aesthetics differences of their user interface. The test requires that users click on interactive elements of just two pages of the product, the home page and a page with the results from searching the ID of a specific transaction. While on the first page, users consult the certificate to copy a string of text and paste in the search input field of the block explorer, clicking search subsequently. On the second page, the users can see the details of the transaction, and search for the fingerprints of the contract mentioned in the scenario description. The fingerprint is just the hash digest calculated as a SHA-256 checksum of the contract's PDF file. The presence of that hash on a Blockchain serves as a form of timestamping and certification, valid from the moment the data was broadcasted to the majority of the nodes in the network. Therefore, after accessing the results page of a block explorer, the testers complete the goal by seeing the status of the transaction, checking the date in which it happened, and comparing the fixed-size string of characters in the input data field of the transaction with the data hash that has been given on the certificate.

According to the ISO 9241:11 (2018, p. 2) a system can be considered the product or service it provides, so the definition of product or service are analogue to the system's definition: "combination of interacting elements organized to achieve one or more stated purposes." For this research, the target product is the user interface of a chosen Blockchain explorer and not the service where the certification is generated.

3.3 Summative Expert Assessment

Lewis (2014, p. 665) recommends that "before conducting a test with users, part of the preparation of a study should include an inspection method such as heuristic or expert evaluation." Often, summative evaluations are conducted like experiments and should not have changes to the system or product during the study (Lewis, 2014). The summative methodology described in this section allows the author of this thesis to answer the first research sub-question:

> a) Are the recommendations of the Suitability for Learning principle considered in the design of block explorers' user interfaces to assist users in verifying the certification of documents via transactions recorded on a Blockchain?

3.3.1 Usability Inspection Methods

Furthermore, Lewis research (2014, as cited in Rusu et al. 2015) highlights that "the concept of summative usability led to ISO usability standard, which interpretation resembles to methods and metrics of experimental psychology, instantiated in human factors engineering." Also known as usability inspection methods or UIMs (Cockton et al. 2012), they are analytical evaluation methods that "only require availability of a designed artifact and trained analysts." In general, UIMs are based on experts' judgment and have become "cost effective ways of evaluating user interfaces to find usability problems" (Nielsen, 1994).

For example, the seven heuristics defined from the dialogue principle of *Suitability for Learning* can be applied in a heuristic evaluation, one of the most common usability inspection methods. Such expert evaluation involves a review of a product or system by specialists on usability or human factors who analyse every interactive element and dialog of a user interface (Rubin & Chisnell, 2008). However, a heuristic evaluation is considered the most informal method to judge the usability of a user interface, therefore, it is common to combine it with a "more explicitly detailed procedure that simulates a user's problem-solving process at each step through the dialogue" (Nielsen, 1994).

A cognitive walkthrough is such usability inspection method in which "evaluators work through a series of tasks and ask a set of questions from the perspective of the user" (Blandford et al. 2011). This method is commonly employed in the evaluation of the first experience of users with complex systems and software development tools. A cognitive walkthrough is an appropriate method to use in this research because it focuses on evaluating the learnability of a product or service (Wilson, 2014, p.65-79), and ensures that users can easily learn to perform tasks that the system is intended to support (Polson et al. 1992). According to Begnum and Foss-Pedersen (2018), cognitive walkthroughs are "more systematic than the trial-and-error approach of heuristic evaluations and may increase user empathy and needs sensitivity as well as focus on specific tasks and functionality aspects."

Despite the set of learnability heuristics defined can serve as guidelines to evaluate if a dialogue is suitable for learning through a simple heuristic evaluation, this research takes a step further to establish how the heuristics are relevant for each of the cognitive activities of the user while performing a task. The goal is to compound the effect of a heuristic evaluation with that of a cognitive walkthrough into one usability inspection method focused on the cognitive model for learnability already laid out. An expert evaluation utilizing usability inspection methods and specifically looking at the support or guides offered to users in learning to use a system can be called a ''learnability evaluation.''

3.3.2 Framework for Learnability Evaluation

A framework is an entity between a model and a method, containing a structure for the detection and comprehension of a defined result (Verbrugge, 2016). While a model helps to convey an idea by simplifying a phenomenon, a framework can be comprised of various models and offers the guidelines for their application. As Sharp, Rogers, and Preece (2015, p.57) elaborate the definition stating that frameworks "can come in a variety of forms, including steps, questions, concepts, challenges, principles, tactics, and dimensions." Also, compared with a method, a framework gives more freedom regarding the applicability of the models that it contains (Verbrugge, 2016).

The strength of a heuristic evaluation is its precise approach, allowing to check specific aspects of a solution's learnability against a specific set of tailored criteria. The strength of a cognitive walkthrough is to enable experts to frame the evaluation of learnability at each step that users take to make progress towards their goal. This complementary effect is also achieved in the cognitive model for learnability by the combination of the cognitive engineering concepts involved in the user's mental activity while performing a task and the ergonomic design principles that the ISO 9241:110 (2016) standard sets forth regarding the suitability for learning of a user interface.

The amalgamation of methods and model is achieved by approaching the construction of the framework from a feature analysis perspective, explained by Marshall and Brereton (2013) as:

> A qualitative form of evaluation involving a subjective assessment of the relative importance of different features and how well features are implemented. It is based on the requirements that users have for a particular task/activity and mapping those requirements to features that a tool aimed at supporting the task/activity should possess.

Feature analysis enables to conduct an expert inspection review as objective and non-biased as possible (Begnum & Foss-Pedersen, 2018). Additionally, the approach is used for comparing software (Kitchenham & Jones, 1997), perfectly suiting as a summative evaluation which is commonly used to "assess the overall usability of a system, in an effort to either compare to another competing system, or to determine if it meets requirements which have been set out" (Grossman, Fitzmaurice, & Attar, 2009).

Since this research has stablished well-defined references for the context of use (product, users, goals, task and environments) of the Blockchain technology, it is possible to take the cognitive model for learnability (figure 3.2) and generate a *Framework for Learnability Evaluation* combining the strengths of traditional usability research methods. The list of features is based on the learnability heuristics and cognitive user activities so the user interface of different block explorers can be analysed.

The block explorers that are subject to the feature analysis are the same three reviewed in the previous chapter. The features relate to aspects that are either present, partially present or totally absent from the expert's perspective and are assessed by a simple YES/PARTLY/NO nominal scale (Kitchenham & Jones, 1997). Table 3.2 presents the framework that experts need to take as assessment criteria for completing the evaluation of learnability through a feature analysis perspective.

Gap of Execution			Assessment		
Category	Feature	Description	Yes	Partly	No
A2: Forming the Intention	H1	**Rules and Underlying concepts which are useful for learning should be made available to the user:** Does the dialogue give an alternative to learn more about the elements shown in the interface offering an explanation about the information and terms displayed? (e.g.: What are transactions, blocks and on which kind of Blockchain it is all recorded)	Assessment		
	H2	**If infrequent use or user characteristics require relearning of the dialogue, then appropriate support should be provided:** Does the dialogue conforms with the user expectations of a search engine? Do conventions allow users to build up patterns and strategies for memorizing the actions that will be taken?	Assessment		
A3: Specifying the Action Sequence	H2	**If infrequent use or user characteristics require relearning of the dialogue, then appropriate support should be provided:** Are the interactive elements prominent, and the terminology used regarding taking action self-explained?	Assessment		
	H3	**Appropriate support should be provided to assist the user in becoming familiar with the dialogue:** Are tooltips, tutorials, links to support pages (FAQ and documentation) or live-chat assistance provided so the user can clarify doubts about the user interface?	Assessment		
A4: Executing the Action	H7	**The interactive system should enable the user to perform the tasks with minimal learning by entering only the minimum amount of information required in the dialogue, with the system supplying additional information on request:** Is the user prompted to introduce information they have? Does the dialogue give an example of what kind of information is allowed to introduce?	Assessment		
Gap of Evaluation			Assessment		
Category	Feature	Description	Yes	Partly	No
A5: Perceiving the System State	H4	**Feedback or explanations should assist the user in building a conceptual understanding of the interactive system:** Does the user receives feedback or explanations to understand the result of an action? If a wrong action is taken (incorrect input), feedback and explanations are provided in a recognizable manner? If taken to other dialogue, does the new user interface has a consistent appearance with the previous one?	Assessment		
A6: Interpreting the System State	H5	**The dialogue should provide sufficient feedback about the intermediary and final results of an activity so that the user learns from successfully accomplished activities:** Can the user interpret that the information provided by the user interface allows to complete the task?	Assessment		
A7: Evaluating the Interpreted State	H5	**The dialogue should provide sufficient feedback about the intermediary and final results of an activity so that the user learns from successfully accomplished activities:** Does the user realize when a task is completed? Does the user think that is in an intermediary step to receive the final results?	Assessment		
	H6	**If appropriate to the tasks and learning goals, the interactive system should allow the user to explore ("Try out") dialogue steps without negative consequences:** Can the user recover from mistakes if has introduced the wrong information? Can the users redo the task or inform about problems if the information they need isn't there or seems wrong?	Assessment		

Table 3.2: Framework for Learnability Evaluation from a Feature Analysis approach

Each of the categories listed in the framework correspond to the specific cognitive user activity when crossing the gulfs of execution and evaluation. A category is containing one of two features which correspond to the learnability heuristics that are relevant to the activity being analysed. The description of a feature is the recommendation of suitability for learning taken from the ISO 9241:110 (2006) and is complemented by a set of questions that the experts can use as a guide for making their assessment about the learnability of Blockchain explorers. The description can also give information to experts that are not familiar with the product or have limited domain knowledge about Blockchain.

This research's summative expert assessment is based on a feature analysis, which "increases the validity and reliability of expert evaluations by taking a transparent and structured approach" (Begnum & Foss-Pedersen, 2018; Lazar, Feng, & Hochheiser, 2010). Part of such reliability and transparency is reflected in the scoring defined pre-evaluation (Table 3.3), which provides a description of how the features are analysed and should be qualified by the expert, therefore promoting more objective evaluations.

Level	Description	Score
Yes	Feature is well implemented and satisfactorily complies with the learnability heuristic	3
Partly	Feature is implemented with shortcomings and partly complies with the learnability heuristic	1
No	Feature is not well implemented and does not comply with the heuristic.	0

Table 3.3: Scoring model for expert assessment according to the analysis of features

There are three levels in the scoring model used by the expert to assess a target user interface. The lower level indicates that a feature is not successfully implemented, probably hindering users from usage and therefore offering a low level of learnability for that particular cognitive user activity. At the medium level, the feature is implemented with only minor shortcomings, supporting the learnability of the user interface to some extent. The highest level indicates that the learnability heuristic has been well implemented as a feature of the user interface and satisfactorily complies with the recommendation of the ISO standard for the particular application and context of use.

After the expert has gone through task analysing each of the features and giving them a score, a total score is calculated by adding the scores of all the features (Kitchenham & Jones, 1997). The overall score resulting from the feature analysis can be also calculated as the sum of all the category scores (Begnum & Foss-Pedersen, 2018), which can result useful if the experts want to compare at a category level how different user interfaces support learnability for each cognitive activity of the user at performing a task (Table 3.4).

Category	Maximum Score	Minimum Score
A1	3 x 2 = 6	0 x 2 = 0
A2	3 x 2 = 6	0 x 2 = 0
A3	3 x 1 = 3	0 x 1 = 0
A4	3 x 1 = 3	0 x 1 = 0
A5	3 x 1 = 3	0 x 1 = 0
A6	3 x 2 = 6	0 x 2 = 0
Total Score	**27**	**0**

Table 3.4: Total scores per category

To properly employ the *Framework for Learnability Evaluation* as feature analysis and ensure validity in the results, all the features and category to be scored were prescribed before doing the expert assessment. This also applied to the acceptance thresholds which were separated in three possible rating results for the expert assessment of the user interface (Table 3.5). If the score of a user interface falls below one third of the total score, the user interface is rejected as it has poor learnability. The minimum threshold of learnability that is acceptable starts when all of the nine features receive at least a medium level (PARTLY) assessment from the expert. This minimum rating also indicates that the user interface has shortcomings in successfully meeting three or more of the learnability heuristics and a more exhaustive study should be done involving users in a real-world scenario. If the user interface has a score above two thirds of the total score, it is considered as having high learnability and therefore obtains a satisfactory rating.

Rating	Description	Threshold
Non-Acceptance	Doesn't comply with the heuristics or fulfils them below a level of learnability that is acceptable.	<33,33%
Minimum	Partially meets the requirements of the heuristics for an acceptable level of learnability to make part of a user research.	<66,66%
Satisfactory	Satisfactorily complies with the majority of heuristic recommendations and guidelines for a high learnability.	>66,66%

Table 3.5: Acceptance criteria for the results of Expert Assessment

In this thesis, the assessment criteria and acceptance thresholds of a feature analysis were established for evaluating Blockchain explorers with the commonly implemented functionalities expected from these solutions. According to Begnum & Foss-Pedersen (2018):

> Like other expert evaluations, a feature analysis may be viewed as a qualitative approach even if the feature assessments are quantified. This is due to the subjective aspects present in the approach, related to both determining feature assessment criteria and procedure at conducting the evaluation of feature implementation.

Therefore, this approach for expert assessment does not pretend to stablish a generalizable method for evaluating learnability. Although, it could be taken as a test of the learnability of a user interface if the assessments produced are peer reviewed by more than one specialist. Instead, the *Framework for Learnability Evaluation* suggested in this chapter serves to compare between user interfaces that receive minimum rating and select the product with highest score to take part of a formative empirical testing involving users.

3.4 Formative Empirical Testing

The second research sub-question is answered through formative empirical testing, where end-users are given a task to execute on a series of interfaces. A test is used to collect qualitative data about whether or not the user completes the task and the cognitive effort it takes to achieve the proposed goal. The data collected assists in the assessment of how the learnability of a block explorer's interface helps the tester to complete the assigned task answering the question:

> b) Does the learnability of the selected block explorer's user interface adequately support the verification of cryptographically certified documents for first-time users with varying degrees of domain knowledge about Blockchain?

The concept of formative evaluation is used to "learn about the usability problems associated with the system, in an effort to improve the interface" (Grossman, Fitzmaurice, & Attar, 2009). According to Lewis (2014), "formative studies permit a wide variation in

technique, which can be formal or informal, silent or talk aloud; participants can work solo or in pairs; use low- or high-fidelity prototypes; or use current, future, or competitive products." The formative usability evaluation methodology used during this research to assess the learnability of Blockchain products is essentially comprised by two empirical testing methods based on users' participation (Dumas & Fox, 2008): A cognitive task analysis and a semi-structured post-test interview.

3.4.1 Cognitive Task Analysis

Cognitive task analysis (CTA) is a usability testing method in which a task is broken down into steps that a tester must go through while a person conducting the test uses different methods to acquire qualitative and quantitative data about the interaction. The set of steps must relate to each other for examining how the user makes decisions taking each of them. The goal is to identify how much cognitive effort is involved in each step, and how the process might differ depending on the experience and knowledge level of the user (User Interviews Inc, 2019).

In this research, the test setup allowed the facilitator to observe users conducting the tasks and annotate their interaction with the components of a block explorer's UI. The tester is observed in-person, only requiring the laptop computer as equipment. The laptop is set to record the interaction within the screen and both the voice of the participant and the facilitator. Although, collecting data about of when and how the tester is accessing the information displayed is important, it is not necessary to run the study with an eye tracker. Acquiring a high level of detail about the interaction with the layout of the web application is not the goal of the test, therefore users are told to simply move the cursor following their eyesight and try to keep it wherever they are looking at the screen.

Instead, the goal of the test is identifying the cognitive process users have at trying to accomplish the goal, so they are instructed to use the Think Aloud (TA) technique (Lewis 2012, p.1274) by expressing themselves about what they are seeing and thinking while interacting with the user interface of the Blockchain explorer. They are also allowed to make questions but are warned that maybe they will not have an answer right away. Before the test starts, the users are welcomed by the facilitator reading a script (Appendix 4) that introduces them to what the process will be without getting into details. The script used was written based on the one provided in the book "Don't Make me Think" by Steve Krug (2006, p. 146).

During a Cognitive Task Analysis, designers can focus on the users while they are accomplishing the tasks and manifesting their goal-directed behaviours involving the product. This method is particularly helpful at conceiving, redesigning or diagnosing a segment of the user-flow (User Interviews Inc, 2019). The main tasks can be broken up into sub-tasks that can go from simple such as clicking a button to complex like memorizing and finding a long string of text. The user-flow is composed of three major sections for the test (figure 3.5), comprehending all the actions required for each task and sub-tasks, their order and how they relate to each other. For this research, determining what counts as a sub-task and how it is related to the main goal depends on the section of the test in which the task is done, a breakdown is shown in Appendix 5.

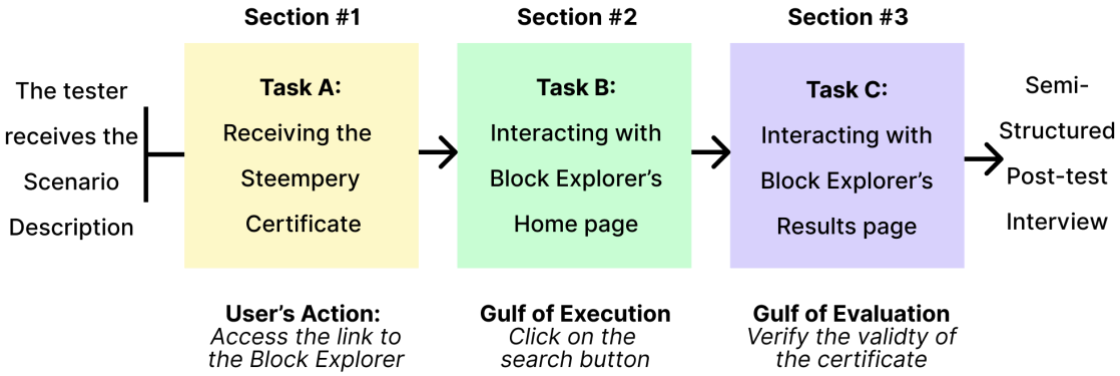


Figure 3.5: Flow Diagram of the Cognitive Task Analysis and Sections of the Test

During first section's sub-task, the participants carry out a revision of the certificate generated by the hypothetical certification service called 'Steempery' for the experiment. The document (see Appendix 3) is inspired in a version of the actual certificate that another certification service (Stamp.io) provides. The test certificate includes extra information explaining the data that is provided and a reference to the Blockchain explorer used during the next sections of the test. state

During this step, the user interprets the document in their capacity and establish the goal of using a block explorer to verify the information in the certificate and validate the data integrity of the contract signed. After reading the certificate, the testers are asked if they have understood its content and to explain how they would take the next step or mention if they have any doubts before moving to the next section. The facilitator is allowed to clear doubts related to the Blockchain-based service that created the certificate as testers never interact with its user interface. This might require briefly explaining the service features, branding or demonstrating the veracity of the contract's SHA-256 checksum, by using an online converter.

After this step, the tester would be accessing the Blockchain explorer, and they are encouraged to think aloud about the action they are making and the information they are taking from the user interface. The facilitator lets them progress alone through the tasks following the think aloud protocol and avoids leading the users to take any specific action, effectively reducing his interaction with the tester while focusing mostly on observation. In case that the testers become stuck after trying possible solutions or made questions that signaled problems to continue the task, the facilitator remained neutral during the interactions but made limited questions according to the Question-Suggestion protocol (Grossman, Fitzmaurice, & Attar, 2009) as explained in the next chapter.

The main goal of the test can only be achieved in the last section when the tester confirms that the immutable record in the form of fingerprint generated by the certification tool has been recorded as input data of a transaction of the Ethereum Blockchain. The test concludes when they acknowledge that the data displayed by the Blockchain explorer cryptographically proves the existence and data integrity of their contract from a certain point in time. This stage is the most critical of the test for the facilitator. Acting as observer, he has to focus on the user and the task, not on thinking how the user interface product assist them to complete the goal since it would be recorded and can be reviewed later. The purpose of this is annotating any interaction that deviates from the sub-tasks and that could affect the resolution of the test or the completion of its main goal. The task of finding the fingerprint labelled as input data in the UI might seem relatively easy, however the cognitive effort that each of the selected testers take to complete it might be different.

3.4.2 Post-test Semi-structured Interview

After finishing the test, a post-test interview takes place with each of the participants. The questions (Table 3.6) are prepared ahead of time based on the *Framework for Learnability Evaluation*. Each of the questions is pertinent to extract more information about how the learnability heuristics, which help to assess the *Suitability for Learning* of a Blockchain explorer, supported the user's cognitive activities when performing the task of verifying cryptographically certified data. Grounding the questions to the same heuristics that form the expert assessment helps to connect the results with the summative evaluation, but most importantly, gives more context to observations made during the CTA to complement the answer to the second research sub-question.

The interview is administered in a semi-structured way (Galletta & Cross, 2013), using a guide (Appendix 6) for post-test data collection instrument to obtain the user's comments on the learnability of a block explorer's user interface. The questions were made at discretion of the facilitator encouraging the tester to elaborate on the answer by asking why or how in some of their statements. Some questions were skipped if the tester had already answered after extending themselves on answering a previous question in the questionnaire. (Dumas & Fox, 2009).

Gap of Execution		
Category	Heuristic	Interview Questions
A2: Forming the Intention	H1	Do the terms shown by the interface helped you to know what to do?
	H2	What were you first thoughts of the Home page of the block explorer? Did you think about similarities with anything you know?
A3: Specifying the Action Sequence	H2	How did you realize you could make a search? What were your expectations of the result page?
	H3	Did you think about searching or entering the tutorial for the tool? Why you interacted or not with it?
A4: Executing the Action	H7	Did you need to learn a lot from the interface to make the search? Do you think you had to remember a lot of information?
Gap of Evaluation		
Category	Heuristic	Interview Questions
A5: Perceiving the System State	H4	What were you first thought when entered the Results page of the search? Did you receive Feedback or explanations that assist you understanding where the Input Data field was?
A6: Interpreting the System State	H5	How did you start confirming the Input data was the Data Hash? Did you think you were in an intermediary step or expected to do more?
A7: Evaluating the Interpreted State	H5	What did you realize after verifying the Hash given to you on the certificate? Would you able to verify another certificate if it's given to you in six months?
	H6	What would you do if the information wasn't shown / shown wrong? How would you start over if you introduced the information wrong?

Table 3.6: Post-test Questionnaire for Semi-structured Interview

4 Analysis

This chapter presents an analysis of the results from data collection by applying both the A) summative expert assessment of the suitability for learning of three selected Blockchain explorers, and the B) formative empirical testing for evaluating the learnability of the user interface of one of them.

4.1 Part A: Summative Expert Assessment

The expert in this research–the author of the thesis act as a ''double'' specialist in both usability principles as well as in the domain of cryptographic certification with Blockchain technology–performed a feature analysis review according to the proposed learnability heuristics extracted from literature. As an aspect of usability, learnability can be examined by an expert according to their previous professional experience. The set of heuristics to be used are specific enough to detect the most relevant usability issues related to the learnability of a user interface.

Having a broad definition of what learnability represents for Blockchain products, the expert examined the user interface to discover usability issues while taking the same viewpoint of users by walking through the steps for accomplishing the task. Taking a feature analysis approach, the expert gave a score to each feature in the framework according to the qualitative scoring system provided. Table 4.1 presents the scores for the three user interfaces of Blockchain explorers reviewed: EtherScan, BlockScout, and EthStats (now branded Alethio[8]).

Gap of Execution				
Category	Heuristic	A) EtherScan	B) BlockScout	C) EthStats
A2	H1	0	0	1
	H2	1	1	3
A3	H2	0	0	1
	H3	1	1	3
A4	H7	1	0	3
Gap of Evaluation				
Category	Heuristic	A) EtherScan	B) BlockScout	C) EthStats
A5	H4	3	1	1
A6	H5	0	1	1
A7	H5	0	0	0
	H6	3	1	3

Table 4.1: Expert Assessment of Block Explorers

The total scores are reordered in Table 4.2 according to each category corresponding to the user cognitive activities at performing the task. The categories have different weight because of the amount of heuristics that are relevant to assess during that specific cognitive activity in which the typical user engages with the interface. The Category A1: "Establishment of the goal" does not make part of the results because that cognitive activity does not occur on the interface being tested (Block explorer). Instead, the formation of the goal is achieved by reading the certificate generated with the hypothetical certification service named "Steempery" (Appendix 3).

Category	Expert 1 Totals		
User Cognitive Activities	Block Explorer A	Block Explorer B	Block Explorer C
A2	1/6	1/6	4/6
A3	1/6	1/6	4/6
A4	1/3	0/3	3/3
A5	3/3	1/3	1/3
A6	0/3	1/3	1/3
A7	3/6	1/6	3/6
Total	9/27	4/27	16/27
Acceptance	33,33%	18,52%	59,26%

Table 4.2: Total Scores of Expert Assessment per Category

The results show that two of the three block explorers (A and C) have a minimum rating of acceptance in terms of learnability while the other (B) falls below one-third of the total score and therefore is rated as having poor learnability. The minimum rating of A and C indicates that their user interfaces have shortcomings in successfully meeting the learnability heuristics and a more exhaustive study should be done involving users in a real-world scenario. In general, the recommendations for the *Suitability for Learning* principle are not sufficiently considered in the design of block explorers to assist users in verifying the certification of documents via transactions recorded on a Blockchain.

In the case that the methodology developed in this thesis is used as an iterative method to design user interfaces of a Blockchain explorer with high learnability, the results of a summative expert assessment based on the *Framework for Learnability Evaluation* (Table 3.2) would serve as a stopping rule to prevent more iterations than necessary (Lewis, 2012) detecting "when user performance and preference meet predefined summative goals" (Lewis, 2014, p. 665). For this research, that covers the first part of one iteration and employs only one specialist, the results allow identifying which of the block Explorers does a better job at following the recommendations of suitability for learning. Alethio's block explorer, EthStats, is chosen to become part of a formative empirical testing.

4.2 Part B: Formative Empirical Testing

Evaluating the user's interaction based on a set of heuristics allows designers to obtain an internal assessment of the system's learnability from an expert's perspective. However, the heuristics used within the framework for evaluating learnability can also be used to assess this important aspect of a product's usability while it is in use. This requires involving real users that interact with the product's UI in a specific context of use, after giving them resources and environment to complete the task and achieve a goal.

4.2.1 Results of Screening Questionnaire

The profile of users that became part of this research as tester fell into the criteria described by the sampling technique used. It is worth noting that typical users that designers might consider a good fit or ideal can be perceived as generators of a great deal of information, but just selecting them misses out on opportunities to obtain broader insights about the learnability of the user interface. Therefore, the survey-takers that had the potential to be included in this research, besides having at least a basic expertise with data management software or spreadsheet tools (minimum criteria), also are in a span of different levels of

domain knowledge about Blockchain. There are answers from both people with more than average knowledge about the application of Blockchain technology and people who have not interacted with Blockchain products because they are less familiar with the terms and concepts of the technology, despite at least have heard about it (minimum criteria). Figure 4.1 presents the distribution of people that took the screening questionnaire.

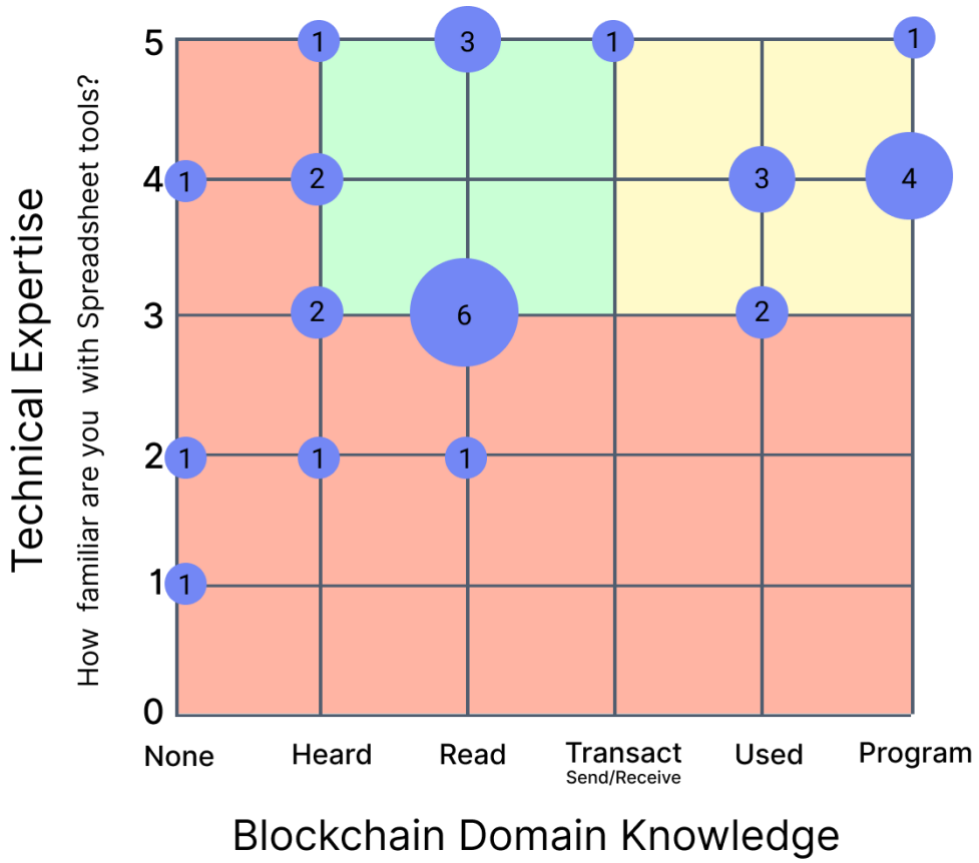


Figure 4.1: Results of Screening Questionnaire for the Selection of Testers

The total amount of people filling out the screening questionnaire was 30, from which 33% were female and 67% male. Additional demographic information such as age and the percentage distribution of all responses are found in Appendix 7.

4.2.2 Cognitive Task Analysis Results

Before the results of the empirically based study are presented, the researcher offers an overview of the users that made part of the research as testers. From the 30 people that completed the questionnaire, just 25 qualified as potential testers. From that sample, only 18 were capable of meeting in person, and 3 of them failed to establish the goal after receiving the certificate (Appendix 3) at the start of the test; therefore, they were disqualified before being introduced to the Blockchain explorer's user interface.

Independently of their domain knowledge about Blockchain, the testers were non-users of Blockchain certification services and have never seen the user interface of the block explorer that is part of the test. These characteristics that all share allow to collect data about the first experience of a new-user and impact of the UI's initial learnability. The results of the cognitive task analysis for the final 15 testers are in Table 4.3.

Tester	Expertise - Domain Knowledge	Establishment of Goal	Gulf of Execution	Gulf of Evaluation	Goal Accomplished	Other Actions Detected
#1	5 - 4	yes	yes	yes	yes	(T)(X)
#15	5 - 4	Yes	yes	yes	yes	(H)(D)
#2	5 - 4	yes	yes	partially	no	(X)
#3	5 - 4	yes	yes	yes	yes	none
#14	4 - 4	yes	yes	yes	yes	(D)
#6	4 - 4	yes	yes	partially	no	(H)(B)(X)
#7	4 - 3	yes	yes	no	no	(T)(H)(X)
#5	3 - 5	yes	yes	no	no	(X)(B)(H)
#11	2 - 5	yes	yes	yes	yes	none
#12	2 - 5	yes	yes	no	no	(T)(H)(B)
#4	2 - 5	yes	yes	yes	yes	(D)(B)
#13	2 - 3	yes	yes	no	no	(H)(B)(X)
#8	2 - 3	yes	no	no	no	(D)(X)
#9	2 - 3	yes	yes	no	no	(T)
#10	1 - 3	yes	yes	no	no	(D)(X)

Table 4.3: Results of Cognitive Task Analysis

The participants were numbered according to the turn in which they were part of the test and ordered in the table according to the level of domain knowledge about blockchain, followed by their technical expertise, defined as their familiarity with data managing software such as spreadsheet tools. Although some users that made part of this empirical testing of a block explorer had interacted with Blockchain products and had already developed preconceptions from the interaction with similar user interfaces, they were able to provide insightful information about the selected Blockchain explorer's learnability. The participants that have at least made or received a transaction on a Blockchain and are presented first (yellow), while people with limited experience and interaction with applications based on the Blockchain technology are positioned after (green).

The completion rate is defined by people that accomplished the test's goal, which results being just 40% of all the participants. Despite a good understanding of the Blockchain of some participants, they could not accomplish the goal. This partially answers the second sub-question of the research indicating that block explorers do not adequately support the verification of cryptographically certified data for first time users with different degrees of domain knowledge. This result can be explained in part because block explorers are generalist tools developed to solve many problems regarding the verification of data in transactions and not only the data of contractual documents that have been hashed and recorded on the Blockchain. However, this quantifiable metric does not give any information about the influence of the User Interface's learnability on the results.

The think-aloud method allowed to collect data from the testers regarding the specific needs or problems they were having at interacting with the block explorer's user interface. During the test, the testers are allowed to explore all the options that the user interface provides, even if they do not help to keep advancing in the task. While recording the data of the cursor's movement and observing the user's behaviour at interacting with the product can provide indications of how they learn to solve a problem, their comments and thoughts will stop being relevant after a certain point for systems that have yet to resolve their learnability issues.

When the thinking-aloud method becomes inappropriate to identify existing learnability issues, the researcher utilized the "question-suggestion" evaluation protocol that is comparable to traditional think-aloud evaluation but extends it by enabling the interaction of the tester with an expert present during the test. The expert was allowed to make questions to the tester so they could rethink what they were doing before getting stuck with the user interface. The protocol proposed by Grossman, Fitzmaurice, and Attar (2009) is based on the question-asking protocol (Kato, 1986) mentioned in Nielsen's work titled "Usability Engineering" (1993) and states that:

> This would replicate a scenario, where a user is performing the task next to a colleague, and the colleague notices a usage behavior which could be improved upon. This type of informal, or "over the shoulder" learning has been shown to be a common way for users to learn (Twidale, 2005). Including suggestions into the protocol would allow the system evaluators to identify causes for suboptimal performance, indicating barriers to *extended learnability*. Furthermore, allowing a coach to provide suggestions may allow the user to progress further through a task, which in turn could expose a larger set of learnability issues.

The many benefits laid out by Grossman, Fitzmaurice, and Attar (2009) at utilizing this protocol make it particularly suitable for highlighting learnability issues and contributing in a great way to evaluating the learnability of user interfaces of Blockchain products. First, the presence of the Blockchain expert makes possible to identify learnability challenges and inadequate behaviour from the user, which an evaluator less educated in the system might simply ignore. Second, by instructing the tester to make questions, but trying to find an answer first as they would in the real world, the participants are encouraged to verbalize their thoughts. Finally, the protocol generates valuable opportunities for testers to learn about the system they are using (Mack & Robinson, 1992). This last benefit lets them answer post-test questions comparing their thought process with the correct action, recognize which concepts they lack understanding of and help to identify what is required on the user interface to help them progress.

Testers only received support from the expert when conducting the tasks if they had already interacted with all other functions of the website, asked a question, or mentioned they were stuck. The questions and interaction with the expert helped them to progress, but the completion of that task was marked as a failure in that section of the test. Besides marking with "NO" in the table when the user did not accomplish the final goal of the test or crossed the gulfs of execution and evaluation appropriately, other actions were annotated helping to flag potential learnability issues of the user interface during a specific cognitive activity of the tester.

The tester starts to cross the *Gulf of Execution* after opening the block explorer's interface. According to Norman (1986), it is the designer's job to construct the input and output characteristics of the interface to make better matches to the psychological needs of the user. The designers of Alethio's block explorer provided an optional tutorial for users entering the website for the first time (Figure 4.2). Interacting with the tutorial is an action that is detected and marked as **(T) Found Tutorial and Interacted with the Guided Tour**. While some participants decided to interact with the tutorial, the majority decided to simply close the pop-up and start figuring out how to complete the task by themselves. A comment made by a participant while thinking aloud when interacting with the guided tour was:

> **#12:** "I click next on the tutorial since I am unfamiliar with this website."

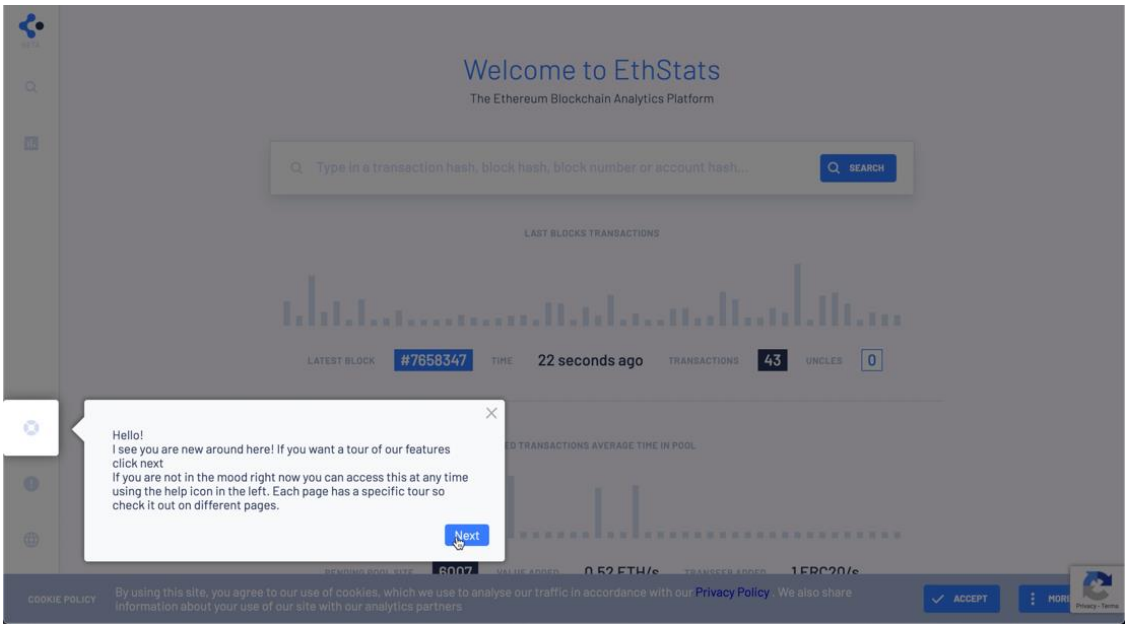


Figure 4.2: EthStats.io Block Explorer Home Page and Tutorial Pop-up

To bridge the *Gulf of Execution*, the EthStats block explorer's home page (Figure 4.3) contains a noticeable input field that can be recognized by the majority of the testers as a search box. Furthermore, it contains placeholder text establishing what kind of input is acceptable. Between the options, the first one is a transaction hash, which corresponds to data that the user can find in the certificate.

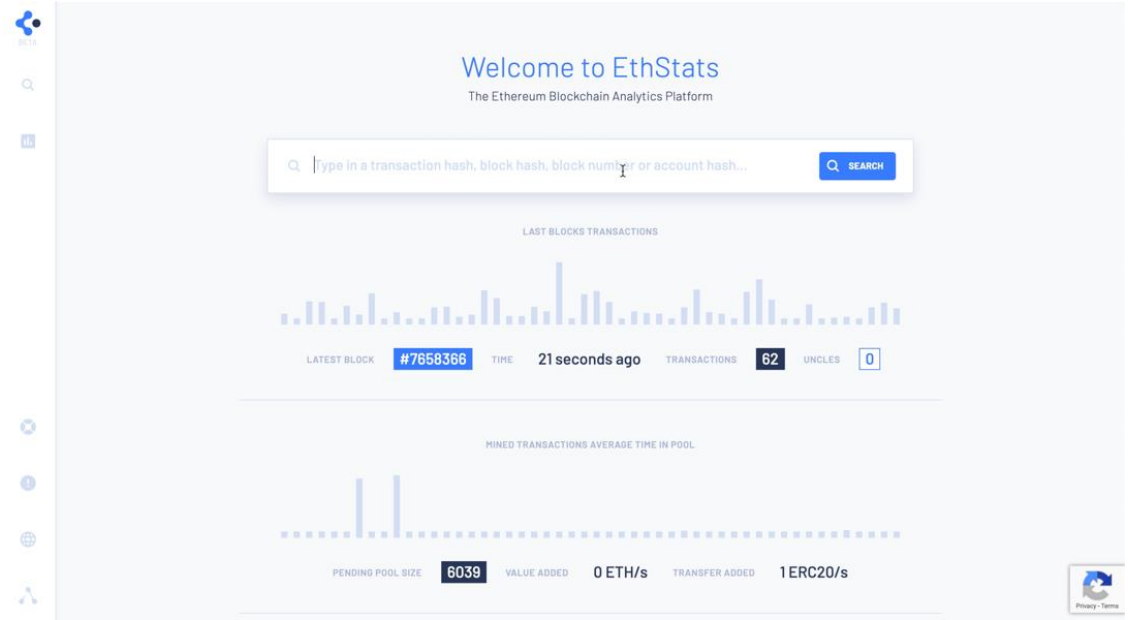


Figure 4.3: EthStats.io Block Explorer Home Page

Another action that both people with and without domain knowledge about Blockchain made was marked **(H) Used the data hash in the Search input**. This action refers to employing other data given in the certificate as input for the search, which results in the block explorer response "SORRY, YOUR QUERY DIDN'T RETURN ANY RESULTS." A comment recorded from one of the testers:

> **#13:** "Oh, I couldn't use the data hash here."

In general, the Gap of Execution for this particular task is straightforward and requires little cognitive effort from the user in terms of interacting with the block explorer's UI. Part of the cognitive effort may have already spent when the user is establishing the goal, which occurs after immersing in the Scenario (Appendix 2) and reading the certificate (Appendix 3). Participants with two different levels of domain knowledge expressed the goal as:

> **#9:** "I suppose I would need to retrieve the contract and make sure that it hasn't changed."
>
> **#15: "**Probably what I would need is some way to verify that I can find one of these hashes within the Blockchain data structure and it has to be during this particular time."

After the first two sections of the test, the tester starts crossing the *Gulf of Evaluation* when the search result page shows, and they interact with the output characteristics of the block explorer's UI (Figure 4.4). The output of the interface has a visual representation of the transactions in a block to the left and the information contained in the Blockchain transaction to the right. Of relevance for the tester is to identify the difference between the data of the transaction and data in the transaction, differentiating the details that characterize the transaction itself, and the data introduced by a person or a service to be recorded on the Blockchain.

> **#14:** "This shows us that at certain point in time existed a document with this data hash"
>
> **#15:** "It could be very tricky if you are not familiar with cryptography, know what a hash is, or you don't understand what the difference is between a data hash and a transaction hash, and how a transaction works."

The observations revealed that bridging the *Gulf of Evaluation*, designers should pay attention to the complexity of the elements displayed on the interface and give clarity to the presentation of information. In case of being difficult to understand, they should also provide assistant material. Some other actions executed by the testers are marked as **(X)Tried to compare From and To Account hash addresses**, **(D)Checked the date of certification in the transaction**, and **(B)Looked for information about the Block.** Comments about each of the actions are respectively:

> **#1:** "I want to check if this from and to is my fingerprint or has any relation."
>
> **#2:** "I know the contract has the hash and I want to make sure it has been timestamped."
>
> **#8:** "So this happened two and a half hours ago"
>
> **#13:** "Is there a way to find the document in here?"

Other comments that let the observer see if the user interface affected the trust that testers had on the process were also annotated:

> **#9:** "For all that I know this could be a fake site giving me the data that I have."
>
> **#13:** "I would use other software to see if I get the same result."
>
> **#15:** "It would be nice to know which of these accounts belongs to who."

The topic of trust is recurrent in the comments that users made along with the test. Despite seeking trust in the product could be taken as a sign of curiosity to learn more about the technology and complete the task, the topic leads to information that gets out of the scope of this research focused on learnability.

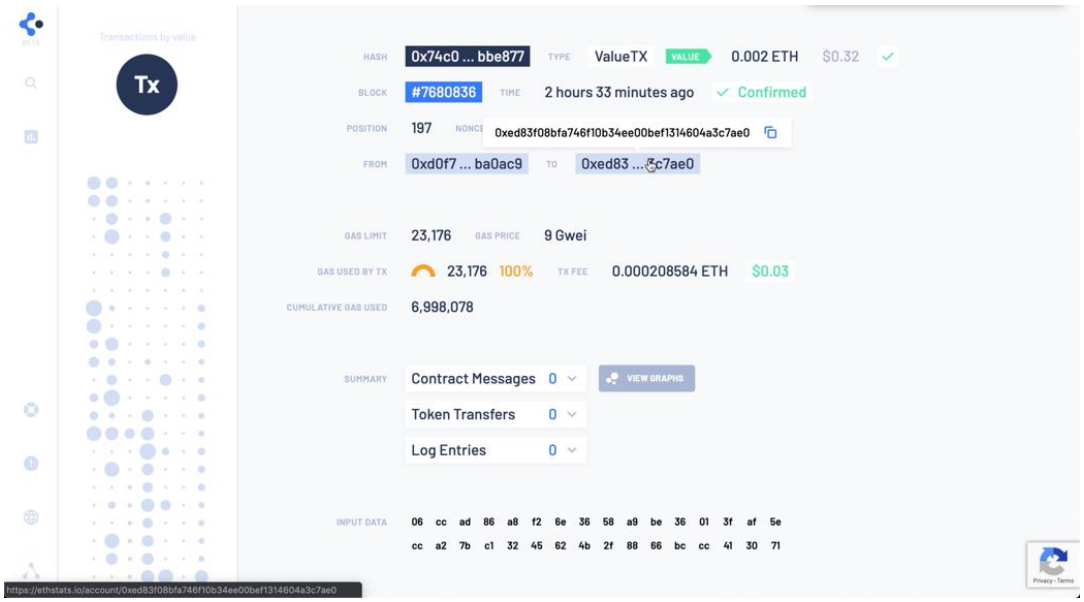


Figure 4.4: EthStats.io Block Explorer Search Result Page

4.2.3 Post-Test Interview Results

The guidelines for conducting the semi-structured questions made during a post-test interview are in Appendix 6. The answers from the testers conveyed more in-depth and valuable information about the functionality and learnability of a Blockchain product such as EthStats from the user's perspective. Questionnaires and interviews are some of the most common methods to collect data after empirical testing (Dumas and Fox, 2009). They generate a lot of qualitative data that needs to be interpreted individually requiring an effort from the evaluator to analyse it in an integrated way. All the comments extracted from the recordings of the post-test interview were ordered and analysed by category and heuristic according to the *Framework for Learnability Evaluation*. The score obtained by the Summative Expert Assessment (SEA) is also presented.

Category A2: Forming the Intention

Regarding H1: *Rules and underlying concepts which are useful for learning should be made available to the user*. (SEA: Partly)

EthStats gave a good first impression on the majority of testers. According to them, the interface is clean and straightforward. Many agree that the layout supports them to understand what is possible to do.

> **#6:** "The website looks really nice. The homepage helped a lot because I sort of intuitively knew I could paste something here"; "Looks like a huge search bar and there was only one thing I could do. I was forced to do the right thing."

Besides finding easy to recognize the search box, they mentioned that the placeholder text let them understand the rules they have to follow to take action.

> **#6:** "The placeholder text helped me to understand that I could paste that kind of information"; "The other parts were changing, so I understood they weren't applicable for the task and that the search bar was the only thing to use";
>
> **#15:** "It's very straightforward"; "It's very easy to Use"; "The first thing you see is the search tool and it says type the transaction hash and the block hash."

However, other participants commented about the difficulty of understanding the terminology and how its general meaning could confuse anyone out of context.

> **#3:** "I don't think that the terms helped me necessarily. I think I understood because I know the terms and I am used to work with them. I think there are a lot of foreign terms that are hard to understand."
>
> **#12:** "I think that the terms are too general to understand the context."

Regarding H2: *If infrequent use or user characteristics require relearning of the dialogue, then appropriate support should be provided.* (SEA: Yes)

Utilizing conventions and predefined mental models that testers have about other products with similar functions is a great way to bridge the distance between them and the system corresponding to the *Gulf of Execution.* As many recognizable search engines do, EthStats simplifies the main function of the tool and makes it prominent. This allows users, either new or occasional, to quickly recall what is possible to achieve and let them focus on learning what they still do not know about operating the product.

> **#1:** "It sort of looks like Google."
>
> **#3**: "I thought it was a clean looking page and understood immediately that the main function would be to look for transaction hash or some sort of hash of accounts."
>
> **#9:** "The user interface reminds me of Google or Google Analytics, so it is a familiar interface. It also reminds me of a stock market interface where I can see if my stocks are plummeting or not. This also reminds me of an information board at the airport. Or a QR code, or a statistics graph."

Category A3: Specifying the Action Sequence

Regarding H2: *If infrequent use or user characteristics require relearning of the dialogue, then appropriate support should be provided.* (SEA: Partly)

The expectations that users have about what they would get from a search should be met once they interact with it. However, many of these expectations are created by the goal users have already focused on achieving (i.e. verifying the validity of a certificate). A block explorer is a tool used for many purposes, and despite certifying that hashes have been introduced as input data in transactions can be one of them, the tool takes a more generalist approach, covering the needs of all the potential users. Adding the tutorial to introduce some of the functionality can help users to set expectations more aligned with what the product can do.

> **#1:** "I thought I would get information about the transaction, but I wasn't sure exactly what I would get from the search."
>
> **#12:** "Being greeted by the tutorial was useful for me to find where to search for the transaction"; "I was expecting something technical and it was confirmed when I did the search. In the context I was, it can be a little bit confusing, too general for me."

They were also expecting functionalities more typical of centralized applications were holding the information of the users gives more value to companies, instead of trusting the software as an intermediary between the people agreeing. Some testers were expecting to find out the actual contract again on the interface.

> **#2:** "I was thinking it would be really nice if this just shows me that the transaction is related to this contract and I can click on the contract and then I can see if its stored."

> **#6:** "I was expecting to see more of a legal sort of page. Basically, having the transaction information but also the contract attached to the Blockchain as a PDF."
>
> **#7:** "First input the data hash and I thought maybe that would lead me to the contract. the pdf, but it didn't happen."

Regarding H3: *Appropriate support should be provided to assist the user in becoming familiar with the dialogue.* (SEA: Yes)

This learnability heuristic enforces the need for supporting new and occasional users to become familiar with the dialog, the tutorial on EthStats was designed to that. Observations about testers that closed the tutorial popup had been made during the cognitive task analysis. Asking them about their reasoning produced opposing opinions about tutorials.

> **#1:** "I actually didn't think about looking for the tutorial, because when I get to a site, I always like to search for the thing I am looking for, and then if needed, I would look for help."
>
> **#3:** "I didn't look for the tutorial because I usually don't do that. I Usually just look around and try to figure out the things myself."
>
> **#4:** "I don't like tutorials. Personally, I would like to fiddle with it until I figure it out."

It might be necessary to change the way a tutorial is delivered from a walkthrough to something available when needed, like tooltips marked by question marks that open if the users click for more information about each element of the interface.

Some testers expressed a need to have guidance, and they felt they were left alone in the step they most needed help. Designers of Blockchain explorers could take a better approach at supporting the learning processes in the right moment and with elements that people would prefer to interact.

> **#6:** "I didn't think about interacting with the tutorial for the first page, but then this second page was a bit more confusing."
>
> **#12:** "However, there was not any content interesting for me in that context, so I didn't evaluate it more than I needed. I totally forgot about the tutorial tool in the second part of the page. I was too focused on finding the information myself. I wouldn't have noticed the tutorial either, if it wasn't for the notification on the first page. So, if you could get any kind of notification on the transaction hash page as well, that could be brilliant because there is where I needed to understand more."

Category A4: Executing the Action

Regarding H7: *The interactive system should enable the user to perform the tasks with minimal learning by entering only the minimum amount of information required in the dialogue, with the system supplying additional information on request.* (SEA: Yes)

Removing the need to learn new concepts might be a complicated endeavour for products which technology requires a certain understanding of complex topics. Still, users could be introduced to the concepts gradually or depending on the application, designers could hide all the complexity away and avoid jargon. However, the benefit of Blockchain in a solution is relying on it instead of an organization acting as an intermediary. This may require users to learn a little bit about the technology's properties and potential. Despite the complexity of the concepts, EthStats does try to moderate the amount of information that gives and requires just relevant information from the testers executing an action.

#1: "I didn't need to know much about Blockchain, but I think I needed to learn about the data hash."

#3: "Using the transaction hash is a little bit clumsy because you need to copy and paste some long string from somewhere and if you don't have in your computer then it would be a pain."

#10: "Since the hash is so complicated you don't bother on memorizing it."

#13: "I didn't need to remember information. It's very good that I just needed to copy and paste the transaction hash."

Category A5: Perceiving the System State

Regarding H4: *Feedback or explanations should assist the user in building a conceptual understanding of the interactive system.* (SEA: Partly)

At providing the results of the user's search and therefore bridging the distance on the *Gulf of Evaluation*, EthStats focuses on the information about the transaction, giving a structure and colours intended to differentiate the elements and content displayed. However, it does a poor job at helping the user understand this layout, and despite the tutorial could be triggered again, it remains hidden from the users, who forgot they could click on the icon again and receive help.

#6: "There was a lot information that is not interesting for the task, so I had information overload."

#10: "It's a bit too much information coming towards me, but it would change if I would be familiar with this and I would recognize the numbers."

The testers of EthStats were also interested in finding a particular set of data, and they commented about how the layout and the terminology to facilitate the process of identifying the information:

#7: "Shouldn't the label Input Data be called data hash? It should be the same so it would be easier to identify."

#13: "I didn't know why this is blue and why this is black. I don't understand why the input data doesn't have any colors. It should look like a result and adding colors would be enough."

They also portrayed conflicting opinions about why it makes sense to prioritize where the Input Data field should be displayed in the layout. Some suggested it should be at the top where it would make it quicker to identify and accomplish the task, but others mentioned that would find logic to be positioned at the bottom because they would be able to understand it as the affordance of paper contracts is to sign at the end.

#4: "I was looking for the Data hash at the beginning on the top, and I saw the transaction hash but not the data hash."

#8: "It was the most important information, but it should be bigger and with color, but here feels just extra at the bottom."

#10: "I didn't recognize the data hash looking like that. But it was at the bottom where it should be because its where you sign on papers to validate an agreement." "I was looking for a signature and obviously a signature in a contract should be at the bottom."

Category A6: Interpreting the System State

Regarding *H5: The dialogue should provide sufficient feedback about the intermediary and final results of an activity so that the user learns from successfully accomplished activities.* (SEA: Partly)

To accomplish the task, users need to compare two different strings of data, one given in the certificate and one displayed by the block explorer's UI. The comparison happens in the mind of the user when interpreting the data, and this makes the feedback that the interface can provide about fulfilling the last step more difficult. As a result, some users think that there is another step they need to take before completing the task.

> **#4:** "I was taking one by one to be 100% sure, I was expecting to confirm something, like confirm that I have confirmed it. Like a verification checkmark."; "I would expect to Log in somewhere and approve it has been checked from this party."

> **#9:** "The input data corresponds with the data hash, so I should copy it and compare"; "It took me quite a while to understand that this was the input data. I would think that there was another step, to get another verification like a green pop up doing Bing."

Also, just completing the process of comparing the hashes can be cognitively demanding and prone to mistakes. Some of the testers left the comparison process incomplete, and others thought it would be time-consuming. Seeing other elements in the interface affected their perception of how long the process was going to take.

> **#1:** "I started to compare them, I saw the first two numbers and then compared all of it overall and saw it was similar."

> **#10:** "I was looking at the date and the accounts. Especially I stopped thinking that it was confirmed, because the format of the time made me thing that it was time left before confirming"

> **#14:** "I looked at the three first and three last characters, because with hashes, even if you make a small change, they would change completely so I usually don't bother comparing the whole hash, I usually often look at the three and a last byte and I would be satisfied. However, if you would be at a court case you will show that every byte is the same."

There were some discrepancies in the way the interface displayed the data. It is typical to display a hash in full length as a string of text. EthStats has chosen to present the information broken in pairs of characters. For some testers, this was clearer to see and also to accomplish the task of comparing the values. However, others expressed that they did not find relevant to interact with that data because it looked entirely different from what they were expecting to find.

Designers of block Explorers should have in mind that even though structuring text in a particular format could make it easier to read, it does not necessarily make it easier to detect and interpret. The proper way to present complex information (ISO, 2017) should be addressed in future research.

> **#3:** "It says Input Data and I assumed it was it. But because it was in the same format but the structure it was made it a little bit hard to see that it was actually the Data Hash"

> **#6:** "I immediately though this wasn't the right thing to see. Because Hash are normally displayed as a string."; "When comparing, it was easier to read it this way because I could check specific numbers and combinations."

> **#12:** "I think that this is a confirmation that the transaction has been incorporated into the Blockchain. I was a little bit unclear because of the formatting of the hash"

> **#15:** "It was quite difficult to find the Data Hash because the format is a bit different. I look at the two strings and just compared them looking at the first 6 characters"

Category A7: Evaluating the Interpreted State

Regarding *H5: The dialogue should provide sufficient feedback about the intermediary and final results of an activity so that the user learns from successfully accomplished activities.* (SEA: No)

To complete the evaluation of the interpreted state, some testers felt the need to use additional software or do extra processing on the data they had already verified. This extra step could be avoided by implementing a comparison feature in the interface or giving a clear explanation that satisfies the user about completing the task.

> **#3:** "I would need to verify [the input data] against the one [data hash] that I do have. I would use an online tool to do a string comparison."
>
> **#5:** "I would put the Input Data and the Data hash on excel and use the function of Equal."

Adding more features might, of course, require complicating the interface. For example, letting the user change in which format the information is presented, as other block explorers do, requires giving more information about how the data hash is converted to a human-readable form. Solving this issue needs to be taken by designers of Blockchain explorers carefully, analysing the impact of trade-offs between functionality and complexity of an already complicated product.

> **#3:** "I think I would need to decode the Input Data."
>
> **#14:** "This data is Hexadecimal. So, you cannot put a string here because it wouldn't be readable still. I think it's all I need, since I confirmed that the transaction has been recorded on the Blockchain so I know that it is not a pending transaction that might be cancelled.**"**

Regarding H6: *If appropriate to the tasks and learning goals, the interactive system should allow the user to explore ("Try out") dialogue steps without negative consequences.* (SEA: Yes)

The last learnability heuristic overlaps with other concepts of usability, such as user error protection or freedom from risk (ISO, 2011). EthStats is a web application that can be accessed from any web browser, and as such it can be operated with the same principles of navigation than all other websites, allowing the user to go back to the previous page by clicking the back button of the browser. It also informs the user when they have introduced information that is not acceptable, displaying an error that allows correcting the mistake. In case of freezing the user can refresh the page.

> **#6:** "I would be able to refresh the page, but if it wouldn't load, then it would be difficult to search for another site."
>
> **#12:** "Entering the Page of the Block I was really not sure; however, I was able to go back to the transaction page by clicking on the back button of the browser.
>
> **#15:** "I did a wrong input of data and was able to search again for the correct hash."

That there is a Blockchain explorer that is failing at supporting the user to accomplish their goal does not mean that all other block explorers will also be operating wrong. The advantage that Blockchain gives with decentralizing the access to information is that users have many alternatives to accomplish their goal without relying on a single central authority or a third party to be available for them. As such, some testers indicated their intention to solve the problem by looking for alternative solutions and other block explorers before confronting the other party in the agreement.

> **#13:** "If there is an error in the data hash it would take away the credibility that nobody has tampered with the contract. To solve it a would look for other software and see if I get the same result, and if I get the same result everywhere, I would confirm with the person that everything is ok."
>
> **#14:** "I might for example use one of the other block explorers and try looking at the information with them instead. If I cannot find it then I would ask the person that claimed that has done it to explain where he did it. If they couldn't then I would say that I can't tell that is on the Blockchain and I cannot verify the document won't be changed."

In a block explorer, users can try many times to accomplish the same task of finding data on the Blockchain, and they wouldn't have consequences at making mistakes. However, when asked about the action they would take if the information was incorrect or inexistent, testers turned their comments toward the topic of trust. They indicated that errors in finding the information would make both the tool and the person that recommended its use to lose all credibility.

> **#4:** "I would be back to the contract party and ask him to send me a revised certificate. I would have checked three times with the interface before contacting the person."
>
> **#8:** "I would suspect that something was changed. I would try one more time before contacting the person I made the contract with. But that's the thing with Blockchain, that there is nobody to call."
>
> **#11:** "I wouldn't trust the contract; I would think the contract is not valid. Then, I would have contacted the sender."

Regarding TRUST: *Comments about the hashing algorithm, the certification service, the credibility of the person or the legitimacy of the block explorer's website.*

Testers were keen on giving their opinions about trusting this application of Blockchain as a default solution for them to verify the certification of documents. Their comments mostly focused on the measures they would take towards the person that sent the certificate, confronting him about their statement in the email, questioning either the legitimacy of the process that he chose to create the hash or his intention at telling them to confirm information that could have been wrong. However, the idea of using a block explorer to confirm the data seemed valid for them to assess the data integrity of the document as long as they could trust the company behind building the product.

> **#1:** "I am not sure if this is enough. I should look at the website to know if it is a real website"; "If I get to confirm it is a real company, that I can trust them, then I will confirm that the certificate is valid"
>
> **#2:** "I haven't gained any trust in that Steempery actually generate the hash. But if I can reproduce it and it has been timestamped, then I would be sure that I can say that this contract was presented at that time." "Because I haven't used EthStats before I would only confirm that the data is recorded on the Blockchain using other block explorer."
>
> **#6:** "I would probably contact the person that sent me this certificate. I wouldn't contact the tool because I feel I have the 'Google is always right' syndrome where I feel humans make mistakes, but Google doesn't.

After the analysis of the collected data from both the cognitive task analysis and the post-test interview, is possible to state that the learnability of the selected block explorer has deficiencies for supporting the verification of cryptographically certified documents for first time users regardless of their domain knowledge about Blockchain technology. The full set of excerpts are added as Appendix 8. Comments about trust and other topics are added as Appendix 9, but they are not considered in the discussion chapter as evaluating the impact of trust is out of the scope for this research.

5 Discussion

This chapter contains a discussion about the implementation of a learnability evaluation framework and its application to evaluating this aspect of usability on block explorers used to verify cryptographically certified documents on the Ethereum Blockchain. The discussion covers the components and design process towards the resulting framework and suggests some considerations that designers and developers should make when deciding to reutilize it for evaluating the learnability of other Blockchain products. After that, the insights obtained from analysing the data collected on a block explorer's UI are used to discuss the importance of learnability for the adoption of Blockchain products. Finally, the confounders and limitations that the researcher identified during the completion of the thesis are explained and pondered into the results of the research before offering final remarks in the last chapter.

5.1 Implementing the Learnability Evaluation Framework

Users learn how to use an interface when they become competent at carrying out tasks without much effort (Sharp, Rogers, & Preece 2015). Zooming in on the properties that make a user interface easy to learn, I found that choosing a single definition of learnability is needed in order to centre the focus of the research on one goal. Therefore, employing Grossman, Fitzmaurice, and Attar's (2009) taxonomy of learnability definitions, I established that the concept of learnability studied in this research was based on:

> The ability to perform well during an initial task for a user whose expertise with similar software is at least basic and domain knowledge about Blockchain can be varied.

Determining heuristics and constructing a framework on that single concept of learnability to analyse data collected fundaments a learnability evaluation and give proper internal validity to the results. Furthermore, establishing the concept based in terms of user aspects such as their expertise and domain knowledge, give the results of the methodology and nature of their analysis some degree of external validity. Particularly, I decided to seek for insights about how a Blockchain explorer's UI adequately support first-time users on verifying the existence, validity and even ownership of a digital contract's fingerprint so parties which signed an agreement in such a document can settle a dispute.

With the definition of learnability and the goal of the research clear, the next step was to construct the framework and methodology to evaluate learnability issues so they could be addressed during an iterative design process. Parting from the base that a set of heuristics can not only be used for evaluations, but also for guiding the design process (Begnum and Foss-Pedersen 2019), I felt there was room for establishing them and developing a framework to evaluate the learnability of Blockchain-based solutions. Creating a cognitive model for learnability and then adapting it as a framework for evaluation helped me to assess this aspect of usability during specific tasks performed by using a block explorer.

I decided to utilize both summative and formative conceptualizations of usability for proposing an integral learnability evaluation methodology of Blockchain products. But before completing a summative expert assessment or conducting a formative empirical test with users, I combined traditional usability inspections methods–heuristic evaluation and cognitive walkthrough–and took a feature analysis approach to construct the

Framework for Learnability Evaluation presented in this thesis. The substance of the framework is based on the recommendations in the ISO 9241:110 (2006) regarding the suitability for learning of user interfaces. This specific set of design guidelines are taken as learnability heuristics and then mapped in the cognitive process of learning occurring along with seven cognitive activities in which users normally engage when performing a task on a user interface. These user cognitive activities are identifiable while the user is performing a task on a user interface by crossing the *Gulfs of Execution* and *Evaluation* that define an interaction with the system.

I took the role of expert to investigate the user interface of different block explorers, record issues, and suggest improvements on their learnability aspect. By properly defining the context of use for the evaluation, I characterized the user's perspective during the interaction with the system. The context of use is comprised of user, goal and tasks, environment, resources, and the product, all of which are comprehensively defined for this research. The framework was used in a summative expert assessment, to evaluate whether potential testers would be affected by the presence or deficiency of features, conveyed by the learnability heuristics, in the selected user interfaces.

The user interface of three different Blockchain explorers with the capacity to support the same goal for the user was evaluated, all of them obtaining minimum or lower rating of learnability according to the acceptance thresholds. However, such an assessment is limited to my personal expertise and experience. Despite making an effort for augmenting its validity with a proper structuration of concepts and scoring, my only assessment, of course, lacks the empirical bases to be taken as a rigorous analysis of the learnability of these Blockchain products. Therefore, I employed the *Framework for Learnability Evaluation* to comparing between the UI of block explorers with a minimum rating of learnability and selected the product with the highest score to take part of formative empirical testing involving users.

Since formative evaluation methods have strong ties with the iterative design process (Lewis 2014) and can be used in the process of designing usable user interfaces, I employed two of these methods to explore and evaluate how the user interfaces of block explorers comply with the learnability component of usability. The formative evaluation of the block explorer selected was executed conducting a Cognitive Task Analysis immediately followed by and a semi-structured post-test interview. The goal with the succession of these exploratory methods was complementing my expert assessment of different Blockchain explorers with user-based qualitative data collected from empirically testing the learnability of just one of these Blockchain products.

The insights obtained from analysing the data collected with this methodology may not be possible to generalize for solving the learnability issues that users have when performing tasks on the UI of other Blockchain products. However, the *Framework for Learnability Evaluation* itself can be re-utilized by other designers, developers, and researchers to evaluate the learnability of other Blockchain-based solutions.

Re-utilizing the same framework and methodology proposed in this research to study the learnability of a different Blockchain product's UI, require that designers, developers, and researchers make changes to evaluate the impact of following the recommendation of suitability for learning in their particular context of use. First, they would need to select new testers and decide on the specific cognitive traits these users share and which they are distinct about. Also, specify the tasks and goal that users will complete, and finally, the environment and resources they need to perform those tasks interacting with the

Blockchain product. Once these specifications of the context of use have been considered, adapting the framework requires a reformulation of the questions that aid the expert during the evaluation of the interface, the actions that can be marked or flagged during the cognitive task analysis and the questions that will be made during the post-test interview.

In the formative part of the methodology, particularly during the conduction of cognitive task analysis, specific considerations have to be made. In case that accomplishing a goal within a particular blockchain product requires a multi-step process, compelling the user to perform multiple tasks, the framework could be employed the times necessary to cover all individual interactions in which there is a clear crossing of the *Gulf of Evaluation* and Execution by the user. Despite the framework can be re-utilizable, if the tasks to accomplish a goal make the duration of the cognitive task analysis too long–particularly when involving users that are less familiar with the application of the Blockchain technology–it would be better to use a method that accelerates the natural learning process of the tester. Therefore, after allocating the use of the *Framework for Learnability Evaluation* along with the multiple tasks, the facilitator of the cognitive task analysis may want to avoid using the think-aloud method altogether and embrace the question-suggestion evaluation protocol. According to Grossman, Fitzmaurice, and Attar (2009):

> Allowing a coach to provide suggestions may allow the user to progress further through a task, which in turn could expose a larger set of learnability issues. Thus, we believe the question-suggestion protocol could be a suitable methodology for learnability evaluation, since it captures both the initial and extended learnability dimensions of our taxonomy. By selecting appropriate users, the methodology can also capture the user dimension of the taxonomy.

Alternatively, similar efforts to extend the reach of this methodology could be made towards the evaluation of usability in general. Designers, developers, and researchers could modify the framework to evaluate another aspect of usability altogether. This would require changing and re-mapping the learnability heuristics in the framework for those of the component of usability evaluated. After adapting the framework, a methodology for evaluating the usability of Blockchain products could spawn from extending this thesis methodology, either by using the same methods to individually assess other aspects of usability apart from learnability or including the utilization of more methods according to the needs of the researcher.

All the data gathered during research, if obtained with an extended methodology and changes to the framework suggested, can provide a proper contextualization for evaluating the learnability or other usability aspects of Blockchain products. For this specific research, the utilization of the framework and methodology corresponds to an integral assessment of just one aspect of usability in block explorers used to verify cryptographically certified documents via transactions on the Ethereum Blockchain. After analysing and triangulating the results of the expert assessment, the cognitive task analysis, and the post-test interview, some considerations can be drawn and discussed not only about the importance of evaluating the learnability for block explorers but Blockchain products in general.

5.2 The importance of Learnability for Blockchain products

That there are different ways to define learnability makes the concept difficult to be interpreted and evaluated in just one way (Grossman, Fitzmaurice, & Attar, 2009). The results of a learnability study depend on the interpretation that the designer gives to the concept and the target user interface of the product evaluated during the study. While learnability can also be an attribute of everyday things such as appliances, information

booths or vending machines, in this thesis the concept is assumed in accordance with ISO 9241:110 (2006) towards the kind of user interface that let people access and interact with information online, particularly, data recorded on a Blockchain. Therefore, this thesis solely focuses on studying the learnability aspect of usability in block explorers that supports the goal of verifying the existence of an agreement by attesting that the data hash of a contractual document has been recorded via a transaction on a Blockchain.

Using block explorers is a necessary step for Blockchain-certification services because the record of a transaction is where lies the proof that a fingerprint of a document was successfully registered on the Blockchain, becoming extremely difficult to tamper as time passes. Thus, users can obtain a proof that the input data included in those transactions is correct, either to get guarantees for the data integrity of their document or to demonstrate the existence and validity of a contract from a certain date. While the content of the documents and contracts is not stored in the Blockchain, the proofs of the integrity, existence, and ownership of such content can be independently verified by anyone, including lawyers and courts. And due to the benefits of decentralization, these records will be available even if the service used to certify the document stops working or does not exist anymore.

Since transactions are the core of everything that happens on the Blockchain, the same tasks and cognitive processes of users explored on this research are usually sustained by someone that uses a block explorer to verify information about the transaction itself, or other data recorded as part of the execution of a smart contract. Therefore, a Blockchain explorer's user interface provides evidence of the utility of decentralized applications and studying the challenges that users have when interacting with them contribute with insights that can be considered by designers and developers to improve the usability of other Blockchain products. The outcomes of conducting this research made visible some of the learnability issues and shortcoming of current user interfaces that provide the visualization of transactions recorded on a Blockchain, a functionality that fundaments the majority of current Blockchain products and maybe those yet to be launched in the market.

I consider essential to study the learnability of Blockchain products because the outcomes of researching usability aspects of a low entry point for consumers to interact with Blockchain, as the use of block explorers, could become particularly useful for Blockchain products that need to improve their ease of use. After all, industries are bounded to adopt this technology (Columbus, 2019), but their reception in the market might be lagging due to the complexity and cognitive friction they present for all kinds of users. This opens the possibility to consider this research as an initial step to evaluate the not only the learnability but the general usability of other Blockchain product's UIs where users verify the data registered on an immutable and distributed ledger via transactions. Data recorded as a message in these transactions can be automated and displayed correctly through more usable user interfaces, therefore diminishing the dependence, intricacy, and cost of manual and centralized services.

The selected testers that participated in this research have expertise using software programs where they can edit or save information on spreadsheets. Therefore, the user interface introduced to them during the test might not have been new in terms of usage. However, accomplishing tasks on the user interface incorporated new concepts because of the utilization of Blockchain as the bases of the product's functionality and value proposition. Not all the participants had a high domain knowledge about Blockchain and its applications despite all of them have at least heard or read about the technology.

The most recurring learnability issues detected associated to the design of the User Interface were more related to the presentation of information. Despite the company's effort to simplify the UI and use colour to group or separate elements displayed, a feature that recognizably received more thought in comparison to competing products, users suffered from problems associated to low recognizability and information overload.

> **#4:** "It was not intuitive for me that I could press some elements. It looks like a pdf almost. It should visible that they can be pressed."
>
> **#6:** "The hashes were a little. It hard to see because they were shortened. So, I didn't think about them. Maybe if they would have been underneath each other and with the full length maybe they would make more sense"; "There was a lot information that is not interesting for the task, so I had information overload."

Testers failed to find the data hash displayed as input data due to the format in which was presented (e.g. #10, #12#15) and because the presence of other information (e.g. #2, #6, #8), making the task more time consuming and frustrating regardless of their domain knowledge about Blockchain.

> **#4:** "This information, for me since I don't know about Blockchain, would be irrelevant. And I was looking for the name in the contract and the rest of information was distracting, I would spend time looking at it."
>
> **#11:** "There is a lot of information that makes me think that I need to read though it several times. There is a lot of complexity and data that I don't know the vocabulary."

The certificate given to the testers explained that a hash is a long string of characters calculated by a mathematical function when it processes an input file, generating an output that is unique and equal every time the same input is introduced. According to their level of knowledge about the technology, they were more or less receptive to find the information, despite its low recognizability. The block explorer evaluated during the test made it difficult to identify this information changing the typical structure in which this information is displayed in other block explorers and Blockchain products in general.

Problems identified during the research associated to the lack of domain knowledge about distributed ledger technologies involve the users wanting to find a copy of the contract attached to the Blockchain transaction (e.g. testers #6, #7, #10, #12).

Blockchains limit the data-size in their transactions for the information stored by the nodes to be broadcasted, replicated, and validated efficiently across the network. Therefore, it is prohibitively expensive and not recommended to include more than few bytes of data, such as hashes, making the use of cryptographic hashing algorithms the perfect resource due being efficiently computable and their capability to secure the integrity of information. If the Blockchain certification service does not hold a copy of the content, when directing the user to use a Blockchain explorer or developing an integrated explorer to the service, they need provide the information and tools needed by the users to learn how to verify the information appropriately and in terms that they can comprehend.

For example, testers proposed to find information about the accounts that made the transaction by associating the account addresses to a form of identification. However, replacing the hashes corresponding to the participants in the transaction to human-readable information that is more familiar requires the involvement of established and trustworthy organizations that decide to associate the details they have about their users to the data registered on Blockchains. This abstraction, which might be the key to increase adoption of the technology, leads to centralization. Accordingly, the value of Blockchain products could be provided and appreciated by users just if they can point at a party as

responsible and accountable for managing the data. Two testers with different knowledge about Blockchain made relevant comments, mentioning solutions from organizations like banks and the post office as possible variables to involve in the equation.

> **#11:** "I don't know the service EthStats, so I would like to confirm it is legitimate." "I was comparing this to the same process using the Norwegian postal office solution, which I am using every day, signing contracts back and forth. Both for sending contracts and getting contracts, and all parts are basically the same except for the validation."
>
> **15:** "I would like to get some more proof that this [data hash] is actually the contract. So maybe there could be a mechanism to decrypt after providing some credentials like BankID for example."

Another issue that caught my attention was related to the terminology used, not for participants that were not knowledgeable in Blockchain, but for those that had a more than average domain knowledge about the technology. Apparently, the use of the word contract as part of the scenario described before conducting the test, affected participants that were familiar with the concept of "smart contract." These testers did not change the goal established when they read the certificate, but their knowledge made them think beyond a simple solution to complete the task. Trying to find the information related to the execution of a smart contract, three of them (#2, #5, #6) were not able to accomplish the goal of the test without receiving questions from the facilitator that "guided" them to take a simplistic perspective towards finding the information. They missed comparing the string of text in the input data field of the transaction despite that two of them noticed the data hash given in certificate on the user interface, partially crossing the *Gulf of Evaluation* (Table 4,3).

> **#2:** "I was hang up on it's probably stored in a smart contract and I didn't think it was going to be as data in the transaction."

This was identified as a learnability issue because is related to the concept of directness from a user interface, which requires matching the level of description required by the user interface's language to the level at which the person thinks of the task. According to Hutchins, Hollan, and Norman (1986), the problems associated to directness on the evaluation side are occasioned by the semantic distance between the system and the user:

> Such distance refers to the amount of processing structure that is required by the user to determine whether the goal has been achieved. If the terms of the output are not those of the user's intention, the user will be required to translate the output into terms that are compatible with the intention in order to make the evaluation.

The idea that the word contract has been widespread learned and applied in the vocabulary of Blockchain technology with the implementation of "smart contracts", may be the cause of such a discrepancy between the phycological terms of the person and the physical terms from the system. The disconnection between these variables create "the major issues that must be addressed in the design, analysis, and use of systems" (Norman, 1986). Finally, Hutchins, Hollan, and Norman (1986) argue that "some of the production of the feeling of directness is due to adaptation by the user, so the designer can neither completely control the process, nor take full credit for the feeling of directness that might be experienced by the user." In my view, such adaptation is the outcome of interacting with the interface while committing resources to the cognitive process of learning. This is the process that concerns learnability for which an evaluation framework of the interaction between users and systems was proposed.

5.3 Confounding Variables

Even though none of the testers had used the block explorer selected, those with higher domain knowledge about Blockchain may have already interacted with other block explorers to achieve similar goals, like checking information about the transaction itself. At first look, there was the possibility that including them in the research was going to give irrelevant data that would not be valuable for evaluating learnability. The reality was that they had not been in contact with an application of Blockchain for this purpose and collecting data about learnability issues they had was helpful to address problems that people with high expertise or but limited domain knowledge of Blockchain could also face when progressing across the learning curve of these products. In this regard, all the people included in the research were considered first-time users.

Another concern at recruiting first-time users of Blockchain explorers for this study was recognizing that participants who are not sufficiently familiar with technology could provide fewer insights and take a longer time to complete the test than people with knowledge about Blockchain. However, since the goal of this research was evaluating learnability, people with limited information about the concepts and technology supporting the functionality of the product were involved. They were useful for identifying learnability issues that people familiar with the product may have already overcome.

The most significant confounder identified was that testers in this, and maybe other learnability studies, may lack the urgency or commitment towards accomplishing the task. They acknowledged that they would feel a little bit more pressure if the contract was affecting them in real life. At least two testers expressed that they had difficulties internalizing the context and the situation of the scenario. They mentioned that they would have been paying more attention to all the details if it were a important contract for which they could be held accountable. Evaluating learnability is unequivocally affected by the level of interest that the user has in learning. Therefore, when conducting tests to evaluate learnability, researchers should try to make the user engage into the activity or task by making it as close as possible to a real-world scenario where the users would easily immerse in.

The research concerning the evaluating learnability of complex software solutions in which this research was largely inspired provide some clarifications and warnings about the use of the question-suggestion protocol in replacement to the use of the TA protocol for usability evaluations. Grossman, Fitzmaurice, and Attar's (2009) make remarks about the outcome at utilizing both protocols for observing how testers recover from errors and the time it takes for them to figure out things by themselves. They also state the complication of performing a study when it's needed to involve a software expert for conducting every test. Finally, they acknowledge that the question suggestion protocol accelerates the natural learning process. This is especially relevant for other researchers to consider if they do want to execute a study that is representative of how people naturally learn how to use specific software before it is released.

While pure classical experiments usually require large sample sizes, qualitative studies with intricate designs could have only a single participant (Rubin & Chisnell, 2008). Specifically, Grossman, Fitzmaurice, and Attar's research (2009) about initial learnability only involved ten people. This can be considered sufficient for an initial study, according to many authors. Rapid usability testing recommends a size of three to four people (Krug, 2006). While the authors of "eight is not enough" (Perfetti & Landesman, 2001) dispute the evidence

provided by Jakob Nielsen (2000) in "Why you only need to test with 5 users" that five is the rule of thumb to work, achieving 85% of problem discovery (Lewis, 2012; 2014).

If there is interest in adopting the results of this research, researchers should treat them with caution, since the summative evaluation counted with just one "double" specialist and the formative evaluation only involved 15 participants. Additionally, the learning issues identified with a user-based test have not been validated by researching more than one product. However, the goal of applying the methodology was not to achieve generalizable results but exploring unknown circumstances about the use of emergent and complex technology such as is Blockchain. Equally, the intention behind thoroughly describing the construction of the evaluation framework, was to provide internal validity so the results could be replicated and also facilitate the learnability evaluation of similar products in order to gain more generalizable insights.

Despite the small number of people making part of the test, selecting a purposive sample of users gives a framework of reference about the characteristics of people that would find similar limitations and challenges of user interfaces in Blockchain products. This is not to say that the cases of users selected for this research should be used to make statistical generalizations. However, it can be argued that they can help in making logical generalizations that pertain to the user interface and not to its users. If so, for future research, it should be considered looking beyond the sample selected for the test, stretching out the research by recruiting from a larger pool of potential users, expanding the opportunity to advance in the evaluation of learnability and usability of Blockchain products.

5.4 Research Limitations

Finding participants for the research was a reasonably easy task for the level of expertise required. Surprisingly, it was also possible to find participants with varied domain knowledge about Blockchain. The challenge was to test enough users from both extremes of the spectrum, trying to balance the number of participants with limited understanding of the technology with those that had already worked with it. In contrast, finding experts that could perform as a ''double'' specialist (Rubin & Chisnell, 2008) in both usability principles as well as in the domain area of cryptographic certification with Blockchain technology, was a difficult task. Alternatively, it was proposed to involve a second specialist with just a background in usability and universal design, but due to time constraints and scheduling, it was not possible. Since heuristic evaluations usually involve more than one specialist in order to confer validity to the results, it was decided that the summative expert assessment had to be used as a preselection method for this research.

Limitations in time also provoked to shorten the scope of the research to just the last part of the Blockchain certification process. Initially, it was proposed to evaluate the learnability of both the user interfaces of certification services and block explorers. Also, there was an intention to complete one design iteration, providing a redesign of the interfaces tested. These options were discarded due to the impossibility of achieving them as a master's program research in one academic semester.

5.5 Summary

This thesis started by defining distributed systems, outlining the requirements to achieve consensus, and describing the properties of using hashing algorithms for the security and integrity of data. These three fundamental components of Distributed Ledger Technologies allow preserving networks that are decentralized and functional. These networks are made of computers called nodes that communicate with each other validating the state of accounts, which function is sending and receiving transactions. Decentralization provides three arguments (Buterin, 2017) that have been defined in this thesis to act as guarantees for people to agree with each other without relying on witnesses or intermediaries. These guarantees are not present in the majority of information systems currently employed in the world, all of which have a high degree of architectural centralization. Because Blockchain is a type of Distributed Ledger Technology, the guarantees provided by decentralization that apply to DLTs also apply to Blockchain.

The first guarantee that is relevant from all Blockchains is that for transactions made to remain valid, there should be a consensus between the multiple parties that form the network. That means that all the participants should be able to agree on the transactions that took place, bundling them in a block that is validated and added (mined) to the end of a chain. Such consensus is probably the most crucial guarantee provided by a Blockchain's functionality. A protocol within its code attempts to create a reliable system from potentially unreliable parts, where the existence of malicious entities trying to subvert the network for economic gain is reduced because they can be excluded if don't align with the goals of the system as a whole. Depending on the strict rules of its consensus mechanism, each Blockchain implements a different protocol to incentivize and keep the participants honest. These incentives are responsible for avoiding collusion between the parties in any distributed system, and the novel variety of protocols has pushed forward innovation with different Blockchain networks of incredible potential.

The second guarantee and a significant characteristic of Blockchains is their capacity to be difficult to tamper, but extremely easy to demonstrate if there was a modification to the ledger or even an attempt at doing it. The clever application of cryptography achieves this. Blockchains employ cryptographic hash functions with a mathematically-proven immense probability to preserve data integrity, creating fingerprints of every single byte of data that is shared within the network. These functions can convert files of any size as input in a one-way process that creates a string of characters as output. The gibberish-looking identifiers are unique in a similar way that the fingerprints are to human beings. An important characteristic is their irreversibility, which means that it is infeasible to reproduce the output unless you try all possible inputs–which are infinite. This certifies that the only way a derived file's fingerprint is created is because the person that generated it has the right version of the data that was processed by the algorithm.

The third guarantee is that the distributed architecture of Blockchains ensures the availability of all the records in the distributed ledger to the participants in the network, providing a verifiable and auditable history of all information stored or the possibility for querying a particular dataset. To solve any dispute, agreeing parties can validate the information that established an exchange of valuable assets or information, finding the fingerprint of contractual documents recorded in a transaction. Since the data stored in any Blockchain offers immutable and independently verifiable records, any person can obtain proof of the data integrity of a contract agreement and can advocate for corrections in the real world. These proofs can legally obligate a person to compensate another for any

inequality or contradiction spotted in what was agreed. If a court wants to verify the legitimacy of a contract, they can check when its unequivocal fingerprint was timestamped by vetting the proof provided. Hence, Blockchains' role in solving all legal disputes about digital contracts or computer logs is to give unmodifiable public storage for cryptographic identifiers that irrefutable represent the data of agreements.

In a non-distributed system, a third party acting as an intermediary would be required to be present and keep fairness. However, in a distributed system, people can rely on the consensus about the form and value of every exchange, always having available proof in the case that there is a dispute. Along with the unfolding of this research, we explored the capacity for Blockchain-based products to record agreements on immutable and distributed ledgers that certify any contract's data integrity and timestamp their validity. Since the technology can make for convincing proof that a document is identical to one signed by many parties at a certain point in time, some companies have deployed online services for data and document certification. Additionally, the cryptographic hashing algorithms used are accepted by courts, having a legal precedent for their reliability on indicating the tampering of any data or digital document's fingerprint.

In general, scaled-up public and private Blockchains are also becoming accepted by academics, technology experts, and enterprises in many industries as reliable sources of truth for computer logs and digital documents. This is reflected in recent surveys, from firms such as Grant Thornton LLP, reporting that forty percent of senior finance executives expect their firms to invest in Blockchain technology over the next two years, in addition to the 22 percent who say their firms have already implemented the technology (Singer 2019). Likewise, Deloitte's Global Blockchain Survey (2019) reported that 86% of senior executives interviewed believe that Blockchain technology is broadly scalable and will eventually achieve mainstream adoption and that 83% of senior executives say their enterprises are seeing compelling use cases for Blockchain today (Columbus, 2019).

Decentralized applications of Blockchain technology may be on the cusp of providing more value and give competitive advantages to companies, but the reality is that their adoption until now has been slow. Despite that generating and maintaining guarantees for decentralization is not a cheap nor easy endeavour, the development of Blockchain technology is advancing to overcome its limitations. Three online certification services presented in this thesis use the guarantees given by a Blockchain's provable immutability and decentralization to timestamp and certify the data integrity of contractual documents without the need for intermediaries. These services use the Bitcoin, Ethereum and Steem Blockchain networks as they are permission-less and transparent ledgers that anyone can access from anywhere in the world. Using them to support Blockchain certification services allows people to create a permanent record of any digital document's data integrity embedding its corresponding fingerprint (hash) in a timestamped transaction of the Blockchain.

Block explorers are purpose-built search engines that allow users to look up, confirm, and verify information registered on the transactions of a Blockchain. These Blockchain products are valuable tools that provide an attestation for documents representing an agreement and therefore help individuals and organizations to enforce its fulfilment as accorded between the parties. However, it is evident that the records on Blockchain ledgers can be challenging to interpret. After all, these decentralized databases contain complex, encrypted data that is constantly added to the chain as new blocks containing validated transactions are mined. To adopt this technology, both cognitive and physical barriers to

access products using it should be lowered, democratizing its use and incorporating its value in new solutions. Designers can tackle this problem and help users with different levels of expertise and domain knowledge about Blockchain to use affordable Blockchain certification services that allow rebutting any objections to the admissibility of a cryptographic proof that a particular document exists and has not been tampered.

Designing products that utilize Blockchain requires to involve usability evaluation methods focused on the ease of use. This is to reduce the cognitive friction of interfaces that display information with high complexity and the multiple concepts that fundament Blockchain technology. Based on the challenges perceived by users, I applied a usability engineering approach to investigate the learnability of block explorers and study the degree to which the attestation of a contractual document's fingerprint recorded on a Blockchain is adequately supported. A set of learnability heuristics were extracted from the recommendations laid out in the ISO 9241:110 (2006, p.8) regarding the dialogue principle of *Suitability for Learning*. The learnability heuristics form the backbone of a *Framework for Learnability Evaluation* (Table 3.2) presented during the methodology chapter. The framework lets an expert adopt the cognitive model for learnability of user interfaces (Figure 3.2). The framework is used to assess whether or not a Blockchain explorer's user interface follows the recommendations of suitability for learning when they are most relevant for diminishing the cognitive effort of a user verifying the certification of data via transactions recorded on a Blockchain. Specifically, the methodology described in this thesis uses the framework as an alternative for evaluating the initial learnability of block explorers to verify information provided on a Blockchain certificate (Appendix 3).

The research primary used the *Framework for Learnability Evaluation* on three block explorers currently available in the marked as a preselection method by comparing their ratings of learnability according to subjective thresholds defined by the author of the thesis. The selected Blockchain product, called EthStats at the time, became the product subject of formative empirical testing. First-time end-users recruited after taking a screening survey went through a sequence of methods formed by cognitive task analysis and a post-test interview. These methods help the author identify learnability issues of the block explorer's UI supporting the attestation of a cryptographically certified document, more specifically, a data hash resulting from the SHA-256 checksum of a contract that was recorded as input data of a transaction on the Ethereum Blockchain.

6 Conclusion

Block explorers let end-users to navigate registries of immutable records of existence, integrity, and ownership of data recorded on a Blockchain for multiple purposes. Until now, when building one of these products, designers and developers couldn't base their decisions on any research about how usable the interfaces of block explorers are for supporting the task of manual attestation of data registered on a Blockchain. Especially in their role for verifying the fingerprints of contractual documents recorded via transactions. By analysing the data collected during the research, I elaborated suggestions that designers can consider when building a block explorer's UI that assists first-time users in learning how to interact with information displayed on products utilizing Blockchain.

The sample being investigated in this research is quite small, especially when compared with studies employing probabilistic sampling techniques that select random subjects to become testers and obtain a representative sample of the population. However, the use of purposeful sampling is a choice made for this thesis. Taking into account that the research methods employed are qualitative in nature and that the individual characterization of every participant portraits a variable context of use, which albeit defined and attempted to be controlled make the results obtained with this thesis tricky to generalize. Thus, the intention of this research is not to make a generalization about the learnability of all blockchain explorers but proposing one methodology for assessing this aspect of usability of them, with the hope that designers and developers use it to obtain insights about the importance of focusing on the learnability of Blockchain products in their design processes.

For that same reason, a discussion about the importance of the learnability of a Blockchain explorer's user interface in data verification processes took place together with a discussion about the learnability heuristics defined in this thesis to construct a learnability evaluation framework. In brief, using established methods in the field of interaction design, this thesis methodology provides some light over the evaluation of the aspect of usability most related to the user's cognitive process of learning. It is proposed then that the methodology that generated insights in this research could be adopted by designers, developers, and researchers to detect learnability issues associated with the perceived complexity of Blockchain products from first-time end-users.

Undoubtedly, the most significant contribution of this research corresponds to the *Framework for Learnability Evaluation* and a methodology that utilizes it. Both can be adapted and become useful for designers that wish to turn Blockchain products into innovative solutions by improving their learnability, for developers that would like to build products with high learnability to increase the adoption of Blockchain technology, and finally, for researchers that would like to obtain insights about the learnability of sophisticated software's UIs. The thorough description of its construction also enables them to replicate or reutilize the framework and the methodology, or even modify them, to obtain results that permit to improve the ease-of-use for other Blockchain products during design or usability evaluation processes.

Regarding the results obtained and their analysis, this research detected different usability issues of the selected block explorer's UI at supporting the accomplishment of the goal studied. Notably, the data collected from users that have varying degrees of expertise and

domain knowledge about Blockchain demonstrated how their understanding of concepts such as cryptography and distributed systems needs to be weighted in when designing Blockchain products. Determining the different context of use that people establish–as new users of a particular user interface regardless of their domain knowledge about Blockchain–generated relevant data that is helpful to address and improve the learnability aspect in a Blockchain explorer's user interface. The main take away concerning the UI of these Blockchain products is that even if people are technically savvy, failing to consider the learnability heuristics proposed in this thesis along with each of the cognitive activities when completing a task affects their performance at accomplishing the goal of verifying cryptographically certified information recorded in an immutable distributed ledger.

Involving people with varied domain knowledge of Blockchain and experience interacting with the technology can also become an excellent strategy to understand the different needs and challenges of a target audience that may benefit from using digital tools for accessing and verifying information recorded on a determined Blockchain. Designers can reutilize the same or extended versions of the framework and methodology proposed in this research in iterative design processes of this and other Blockchain explorers to keep addressing the limitations and challenges identified in this research related to the learnability of their user interface. Finally, the outcome of this research allows concluding that the impact of neglecting the recommendations for a user interface's suitability for learning might be more significant for products that utilize Blockchain than it is for products that don't require it altogether. This conclusion also suggests that a focus on improving the learnability of UIs can increase the adoption of solutions that implement Blockchain in an attempt to cover real-world needs.

6.1 Future Research

In every study about the learnability of a user interface, there are discrepancies in what a "short time" and "reasonable level of proficiency" would be for each user (Grossman, Fitzmaurice, & Attar, 2009). These two characteristics are constantly included in many collections of learnability metrics. But the fact that there is not one single collection of learnability metrics that can be applied to all user interfaces makes clear that evaluating learnability is a complicated endeavour and designing user interfaces with high learnability will always require compromises according to the context of use for which the interactive system is designed.

Despite this potential shortcoming, defining learnability based on the first user experience, as done in this research, is common to identify major flaws in user interfaces. Whilst designers and developers lack a set of well-accepted metrics for learnability, a methodology to evaluate this aspect of usability can always be adapted from research (Grossman, Fitzmaurice, & Attar, 2009), allowing to obtain qualitative data from the user's interaction with new products and provide solutions to improve their user interface further. From the seven categories of metrics that they proposed, the ease of learning how to use the product to complete a task can be decomposed in two major concepts, metrics based on task performance (Task Metrics) and metrics based on cognitive effort (Mental Metrics) (Grossman, Fitzmaurice, & Attar, 2009). In this research, the evaluation methodology focused on Metal metrics at assessing the Initial Learnability of the selected user interface.

For future research, the framework and methodology proposed in this thesis could be adapted to not only give an evaluation of the cognitive effort that users endure while performing a task but how helpful a redesign of the user interface is to increase the speed in which they accomplish their goals. Metrics of learnability that cover such evaluation need

to be proposed and the methods used should avoid the interaction of the facilitator or observers with the participants. Once integrated into the framework, these metrics would create a more precise evaluation of the natural learning process of users interacting with user interfaces of Blockchain products.

Additional research should be done to evaluate the learnability products and services that certify and verify data employing Blockchain. The impact of improving learnability can be useful for accomplishing tasks that require users to understand a lot more of the Blockchain technology, such as the creation of Blockchain accounts and the management of their corresponding combination of cryptographically paired private and public keys (Rivest, Shamir & Adleman 1978). Many companies are still figuring out the best way to facilitate this process (Baker, 2018), representing an opportunity to do more research employing other recommendations related to the guidance on visual presentation of information that can be extracted from the ISO 9241-125 (2017) or evaluating the impact of tutorials on increasing the learnability of Blockchain products.

Reference List

Al-Awar, J., Chapanis, A., & Ford, R. (1981), 'Tutorials for the first-time computer user'. *IEEE Transactions on Professional Communication*, *24*, 30–37.

Alethio (2018), 'Alethio: Lighting Up the Blockchain with Real Time Data'. ConsenSys Media. Available at: https://media.consensys.net/alethio-lighting-up-the-blockchain-with-real-time-stats-a80bb30576db (Accessed: 31 May 2019).

Baker, S. (2018), '*Designing for Blockchain: What's Different and What's at Stake'.* ConsenSys Media. Available at: https://media.consensys.net/designing-for-blockchain-whats-different-and-what-s-at-stake-b867eeade1c9 (Accessed: 31 May 2019).

Bansal, D. (2017), 'Why It's Hard to "Get" Bitcoin: The Blockchain Spectrum'. Unchained Capital Blog. Available at: https://blog.unchained-capital.com/blockchain-spectrum-806847e1c575 (Accessed: 31 May 2019).

Barber, S., Boyen, X., Shi, E., & Uzun, E. (2012), 'Bitter to Better — How to Make Bitcoin a Better Currency', in Keromytis, A. D. (ed.) *Financial Cryptography and Data Security*. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 399–414.

Barinov, I. (2018), 'Deploying your own customizable blockchain explorer for the Ethereum ecosystem'. Jaxenter. Available at: https://jaxenter.com/blockchain-block-explorer-152892.html (Accessed: 31 May 2019).

Begnum, M. E. N. & Foss-Pedersen, R. J. (2018), 'Digital assessment in higher education', in *Universal Access in the Information Society*. Springer Berlin Heidelberg, 17(4), pp. 791–810. doi: 10.1007/s10209-016-0513-9.

Belin, O. (2018), 'The Difference Between Blockchain and Distributed Ledger Technology'. TradeIX Blog. Available at: https://tradeix.com/distributed-ledger-technology/

Blandford, A., Bevan, N., Wilson, C., Werner, B., & Mascari, M. (2011), 'Cognitive Walktrough'. *Usability Body of Knowledge.* Available at: https://www.usabilitybok.org/cognitive-walkthrough

Blockchain at Berkeley (2018), 'BerkeleyX: CS198.2x Blockchain Technology'. edX Courses. Available at: https://www.edx.org/course/blockchain-advancing-decentralized-technology (Accessed: 14 May 2019).

Buterin, V. (2017), 'The Meaning of Decentralization'. Medium. Available at: https://medium.com/@VitalikButerin/the-meaning-of-decentralization-a0c92b76a274

Castro, M. and Liskov, B. (2002) 'Practical byzantine fault tolerance and proactive recovery', *ACM Transactions on Computer Systems*. doi: 10.1145/571637.571640.

Chain (2017), 'Introducing Sequence – Chain'. Medium. Available at: https://blog.chain.com/introducing-sequence-e14ff70b730. (Accessed: 14 May 2019).

Cisco Systems (2018), *Blockchain by Cisco, Build trust-based business networks for digital transformation*. Available at:

https://www.cisco.com/c/dam/en/us/solutions/collateral/digital-transformation/blockchain-whitepaper.pdf (Accessed: 14 May 2019).

Columbus, L. (2019), 'Blockchain Is Gaining Trust In The Enterprise'. Forbes. Available at: https://www.forbes.com/sites/louiscolumbus/2019/05/19/blockchain-is-gaining-trust-in-the-enterprise/#4c74eb8d3aa0 (Accessed: 30 May 2019).

Crowley, C. (2019), 'Our Next Step Toward an Integrated Blockchain Analytics Solution'. Medium. Available at: https://medium.com/alethio/our-next-step-toward-an-integrated-blockchain-analytics-solution-ee680c42f514 (Accessed: 30 May 2019).

Cockton, G., Woolrych, A., & Lavery,D (2012), 'Inspection-Based Evaluations', in *The Human – Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, Taylor & Francis, New York, pp. 1171–1189. doi: 10.1201/b11963-65.

ConsenSys (2019), '10 ConsenSys Announcements From Ethereal Summit NY: Kaleido, OpenLaw, Alethio, Pegasys'. ConsenSys Media. Available at: https://media.consensys.net/10-consensys-announcements-from-ethereal-summit-ny-kaleido-openlaw-alethio-pegasys-ee8cb5dbe2ac.

Cooper, A. (2004), *The inmates are running the asylum*. Sams. Available at: https://www.interaction-design.org/literature/topics/cognitive-friction

Deloitte Development LLC (2019), *Deloitte's 2019 Global Blockchain Survey.* Available at: https://www2.deloitte.com/content/dam/insights/us/articles/2019-global-blockchain-survey/DI_2019-global-blockchain-survey.pdf (Accessed: 30 May 2019).

Dixon, C. (2019), 'Blockchain Can Wrest the Internet From Corporations' Grasp'. Wired. Available at: https://www.wired.com/story/how-blockchain-can-wrest-the-internet-from-corporations/ (Accessed: 30 May 2019).

Douma, R. (2019), 'Proof.ink - Proven immutable data - File hashes stored on the Steem blockchain'. Busy.org. Available at: https://busy.org/@roelandp/proof-ink-proven-immutable-data-file-hashes-stored-on-the-steem-blockchain (Accessed: 30 May 2019).

Dumas, J. S. & Fox, J. E. (2008), 'Usability Testing: Current Practice and Future Directions', in *The Human-Computer Interaction Handbook*, Taylor & Francis, New York, pp. 1129–1149.

Galletta, A. & Cross, W. E. (2013), *Mastering the Semi-Structured Interview and Beyond*. doi: 10.18574/nyu/9780814732939.001.0001.

Goldberg, O. (2018), 'Building a Blockchain: The Grey Paper'. Hackernoon Blog. Available at: https://hackernoon.com/building-a-blockchain-the-grey-paper-5be456018040

Grossman, T., Fitzmaurice, G., & Attar, R. (2009), 'A survey of software learnability', in *Proceedings of the 27th international conference on Human factors in computing systems - CHI 09*. doi: 10.1145/1518701.1518803.

Hamilton, D. (2018), 'Ethereum vs. Bitcoin Mining: Why They Are Different'. Medium. Available at: https://coincentral.com/ethereum-mining-vs-bitcoin-mining-which-is-more-profitable/ (Accessed: 30 May 2019).

Holzinger, A. (2005), 'Usability engineering methods for software developers', in *Communications of the ACM*. 48(1):71-74. doi: 10.1145/1039539.1039541

Hutchins, E., L., Hollan, J. D. and Norman, D. A. (1986), 'Direct Manipulation Interfaces', in D. Norman and S.W. Draper (eds) *User Centered System Design*. Lawrence Earlbaum Associates, Hillsdale, NJ, pp.87-124.

IEEE (1990), '*IEEE Standard Glossary of Software Engineering Terminology'.* New York Std 610.12-1990, Office. doi: 10.1109/IEEESTD.1990.101064.

International Organization for Standardization (ISO) (2006), 'Ergonomics of human-system interaction - Part 110: Dialogue principles', *Iso/Iec*. ISO Standard No. 9241-110:2006. Available at: https://www.iso.org/standard/38009.html

International Organization for Standardization (ISO) (2011), 'Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — System and software quality models', *Iso/Iec*. ISO Standard No. 25010:2010(E). Available at: https://www.iso.org/standard/35733.html

International Organization for Standardization (ISO) (2017), 'Ergonomics of human-system interaction -- Part 125: Guidance on visual presentation of information', *Iso/Iec*. ISO Standard No. 9241-125:2017. Available at: https://www.iso.org/standard/64839.html

International Organization for Standardization (ISO) (2018), 'Ergonomics of human-system interaction - Part 11: Usability: Definitions and concepts', *Iso/Iec*. ISO Standard No. 9241-11:2018. Available at: https://www.iso.org/standard/63500.html

Kato, T. (1986), 'What "question-asking protocols" can say about the user interface', in *International Journal of Man-Machine Studies*. doi: 10.1016/S0020-7373(86)80080-3.

Kasireddy, P. (2017), 'How does Ethereum work, anyway?'. Medium. Available at: https://medium.com/@preethikasireddy/how-does-ethereum-work-anyway-22d1df506369 (Accessed: 30 May 2019).

Katz, J. (2018), *Introduction to Modern Cryptography*. doi: 10.1201/9781420010756.

Kitchenham, B. A. & Jones, L. (1997), 'Evaluating Software Engineering Methods and Tool Part 6: Identifying and Scoring Features', in *ACM SIGSOFT Software Engineering Notes*. doi: 10.1145/381790.381795.

Kol, T. (2019), 'Defining the Public Blockchain'. The Orbs Blog. Available at: https://medium.com/orbs-network/defining-the-public-blockchain-191834584ed5

Krug, S. (2006), *Don't Make Me Think! A Common Sense Approach to Web Usability.* NewsNature. doi: 10.1098/rspb.2009.1614.

Kumar, R. (2018), 'Blockchain vs. Distributed Ledger Technology'. Coinmonks Blog. Available at: https://medium.com/coinmonks/blockchain-vs-distributed-ledger-technology-b7b2e434093b (Accessed: 30 May 2019).

Lamport, L. (1977), 'Proving the Correctness of Multiprocess Programs', in *IEEE Transactions on Software Engineering*. doi: 10.1109/TSE.1977.229904.

Lamport, L. (1978), 'Time, clocks, and the ordering of events in a distributed system', in *Communications of the ACM*. doi: 10.1145/359545.359563.

Lamport, L., Shostak, R., & Pease, M. (2002), 'The Byzantine Generals Problem', in *ACM Transactions on Programming Languages and Systems*. doi: 10.1145/357172.357176.

Larimer, D. (2017a), 'The Problem with Byzantine Generals'. Steemit. Available at: https://steemit.com/blockchain/@dantheman/the-problem-with-byzantine-generals

Larimer, D. (2017b), 'STEEM Whitepaper: An incentivized, blockchain-based, public content platform'. Steem. Available at: https://steem.com/SteemWhitePaper.pdf

Lazar, J., Feng, J. H. J. & Hochheiser, H. (2010), *Research methods in human-computer interaction*. doi: 10.1002/asi.21187.

Lewis, J. R. (2012), 'Usability testing', in G. Salvendy (Ed.) *Handbook of human factors and ergonomics.* 4th ed., pp. 1267–1312. New York, NY: Wiley.

Lewis, J. R. (2014), 'Usability: Lessons Learned ... and Yet to Be Learned', in *International Journal of Human-Computer* Interaction. pp. 663-684, doi: 10.1080/10447318.2014.930311.

Lewis, A. (2015), 'A gentle introduction to bitcoin mining'. Bits and Blocks. Available at: https://bitsonblocks.net/2015/09/21/a-gentle-introduction-to-bitcoin-mining/

Ludwin, A. (2017), 'A Letter to Jamie Dimon'. Chain Blog. Available at: https://blog.chain.com/a-letter-to-jamie-dimon-de89d417cb80

Lund Research Ltd (2012), 'Purposive sampling'. Lærd Dissertation. URL http://dissertation.laerd.com/purposive-sampling.php.

Mack, R. & Robinson, J. B. (1992), 'When novices elicit knowledge: question asking in designing, evaluating, and learning to use software', in *The psychology of expertise: cognitive research and empirical AI*. Springer-Verlag, Inc. p. 245-268.

Mamoria, M. (2017), 'Your company will use blockchain in less than 10 years — here's how'. The Next Web. Available at: https://thenextweb.com/full-stack/2017/10/17/your-company-will-use-blockchain-in-less-than-10-years-heres-how/

Marshall, C. and Brereton, P. (2013), 'Tools to support systematic literature reviews in software engineering: A mapping study', in *International Symposium on Empirical Software Engineering and Measurement*. doi: 10.1109/ESEM.2013.32.

Mathis, J. (2018), 'Major Milestone: Steem Blockchain Hits 1 Million Accounts'. CCN Altcoin News. Available at: https://www.ccn.com/major-milestone-steem-blockchain-hits-1-million-accounts (Accessed: 30 May 2019).

Menezes, A., Katz, J., van Oorschot, P. C., Vanstone, S. A. (1996), 'Hash Functions and Data Integrity', in *Handbook of Applied Cryptography*. doi: 10.1201/9781439821916-9.

Merkle, R. C. (1988), 'A Digital Signature Based on a Conventional Encryption Function', in *Advances in Cryptology — CRYPTO '87, Lecture Notes in Computer Science*. pp. 369–378. doi:10.1007/3-540-48184-2_32.

Michelsen, C. D., Dominick, W. D., & Urban, J. E. (1980), 'A methodology for the objective evaluation of the user/system interfaces of the MADAM system using software engineering principles', in *ACM Southeast Regional Conference*. pp. 103-109.

MIT Technology Review Editors (MIT-TRE) (2018), 'Explainer: What is a blockchain?'. MIT Technology Review. Available at: https://www.technologyreview.com/s/610833/explainer-what-is-a-blockchain

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

Nazare, J., Hamilton, K., & Schmidt, P. (2016), 'What we learned from designing an academic certificates system on the blockchain'. MIT Media Lab. Available at: https://medium.com/mit-media-lab/what-we-learned-from-designing-an-academic-certificates-system-on-the-blockchain-34ba5874f196 (Accessed: 30 May 2019).

National Institute of Standards and Technology (NIST) (2012), 'Secure Hash Standard (SHS)', in *Federal Information Processing Standards Publication 180-4*. doi: 10.6028/NIST.FIPS.180-4.

Nielsen, J. (1993), *Usability Engineering*. Morgan Kaufmann Publishers Inc. doi: 10.1145/1508044.1508050.

Nielsen, J. (1994), 'Usability inspection methods', in *Conference companion on Human factors in computing systems - CHI '94*. doi: 10.1145/259963.260531.

Nielsen, J. (2000), *Why you only need to test with 5 users.* Jakob Nielsen's Alertbox.

Nielsen, J. (2012), *Usability 101: Introduction to Usability Why Usability is Important.* Available at: https://www.nngroup.com/articles/usability-101-introduction-to-usability/

Nielsen, J. & Loranger, H. (2006), *Prioritizing Web Usability*. Available at: https://books.google.no/books?id=YQsje6Ecl4UC (Accessed: 30 May 2019).

Norman, D. (1986), 'Cognitive engineering', in D. Norman and S.W. Draper (eds) *User Centered System Design*. Lawrence Earlbaum Associates, Hillsdale, NJ, pp.31-62.

Norman, D. (1993), *Things That Make Us Smart*. Addison-Wesley, Reading, MA.

Orcutt, M. (2019), 'Once hailed as unhackable, blockchains are now getting hacked'. MIT Technology Review. Available at: https://www.technologyreview.com/s/612974/once-hailed-as-unhackable-blockchains-are-now-getting-hacked/ (Accessed: 30 May 2019).

Oshana, R. (2013), 'Software Performance Engineering for Embedded Systems', *Software Engineering for Embedded Systems*. Newnes, pp. 281–311. doi: 10.1016/B978-0-12-415917-4.00010-4.

Partz, H. (2019), 'Retail Giant Carrefour Applies Blockchain for Tracking Milk Product Supply Chain'. Cointelegraph. Available at: https://cointelegraph.com/news/retail-giant-carrefour-applies-blockchain-for-tracking-milk-product-supply-chain

Pearson, T. (2018), 'Blockchain Explained: How Does Blockchain Technology Work?'. Medium. Available at: https://taylorpearson.me/blockchain-explained/

Perfetti, C. & Landesman, L. (2001), *Eight is not enough*. Available at: https://articles.uie.com/eight_is_not_enough/ (Accessed: 30 May 2019).

Polson, P. G., Lewis, C., Rieman, J., & Wharton, C. (1992), 'Cognitive walkthroughs: a method for theory-based evaluation of user interfaces', in *International Journal of Man-Machine Studies*. doi: 10.1016/0020-7373(92)90039-N.

Rubin, J. & Chisnell, D. (2008), *Handbook of Usability Testing*. Indianapolis, IN: Wiley Pub. doi: 10.1007/s13398-014-0173-7.2.

Santos, P. J. & Badre, A. N. (1995), 'Discount learnability evaluation', in *GVU Technical Report GIT-GVU-95-30*. Georgia Institute of Technology.

Sharp, H., Rogers, Y., & Preece, J. (2015), *Interaction Design: beyond human-computer interaction.* 4th edition, John Wiley & Sons. Ltd. England.

Shneiderman, B. (1997), *Designing the User Interface: Strategies for Effective Human-Computer Interaction*. Addison-Wesley Longman Publishing Co., Inc.

Singer, A.W. (2019), 'Blockchain Hype Turns to Mainstream Adoption in Billion-Dollar Corporations'. CoinResponse. Available at: https://coinresponse.com/crypto-coin-news/crypto-news/blockchain-hype-turns-to-mainstream-adoption-in-billion-dollar-corporations/ (Accessed: 30 May 2019).

Speicher, M. (2015), 'What is Usability? A Characterization based on ISO 9241-11 and ISO/IEC 25010'. Available at: http://arxiv.org/abs/1502.06792

Palinkas, L. A., Horwitz, S. M., Green, C. A., Wisdom, J. P., Duan, N., & Hoagwood, K. (2015), 'Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research', in *Administration and Policy in Mental Health and Mental Health Services Research*. doi: 10.1007/s10488-013-0528-y.

POA Network (2018), 'Introducing BlockScout—The Ethereum Explorer'. Medium. Available at: https://medium.com/poa-network/introducing-blockscout-the-ethereum-explorer-86b4ddd9e8a4 (Accessed: 30 May 2019).

Rivest, R., Shamir, A., & Adleman, L. (1978), 'A method for obtaining digital signatures and public-key cryptosystems', in *Communications of the ACM*. doi: 10.1145/359340.359342.

Rouse, M., Troy, S., & Pratt, M.K. (2017), 'What is distributed ledger technology (DLT)?'. WhatIs.com. Available at: https://searchcio.techtarget.com/definition/distributed-ledger

Rusu, C., Rusu, V., Roncagliolo, S., Apablaza, J., & Rusu, V. Z. (2015), 'User experience evaluations: Challenges for newcomers', in *Design, User Experience, and Usability: Design Discourse Lecture Notes in Computer Science*, 237-246. doi: 10.1007/978-3-319-20886-2_23.

Schmidt, P. (2015), 'Certificates, Reputation and the Blockchain'. MIT Media Lab. Available at: https://medium.com/mit-media-lab/certificates-reputation-and-the-blockchain-aee03622426f (Accessed: 30 May 2019).

Sipser, M. (1996), 'Introduction to the Theory of Computation', in *ACM SIGACT News*, 27(1), pp. 137–159. doi: 10.1145/230514.571645.

Szabo, N. (1996), 'Smart Contracts: Building Blocks for Digital Free Markets', in *Extropy Journal of Transhuman Thought*. doi: 10.1200/JCO.2011.40.6546.

Szabo, N. (1997), 'Formalizing and Securing Relationships on Public Networks', in *First Monday*. doi: 10.5210/fm.v2i9.548. Available at: https://nakamotoinstitute.org/formalizing-securing-relationships/

Takada, M. (2013), *Distributed systems: for fun and profit.* Available at: http://book.mixu.net/distsys/ (Accessed: 30 May 2019).

Twidale, M. B. (2005), 'Over the shoulder learning: Supporting brief informal learning', in *Computer Supported Cooperative Work*. doi: 10.1007/s10606-005-9007-7.

User Interviews Inc (2019), *'How to Do a Task Analysis to Design Better User Flows'.* UX Research Guide. Available at: https://www.userinterviews.com/ux-research-field-guide-chapter/task-analysis (Accessed: 30 May 2019).

User Interviews Inc (2019), 'How to Write Screener Surveys to Capture the Right Participants'. UX Research Guide. Available at https://www.userinterviews.com/blog/how-to-write-screener-surveys-to-capture-the-right-participants (Accessed: 30 May 2019).

Verbrugge, B. (2016), 'Best Practice, Model, Framework, Method, Guidance, Standard: towards a consistent use of terminology'. Van Haren Publishing. Available at: https://www.vanharen.net/blog/general/best-practice-model-framework-method-guidance-standard-towards-consistent-use-terminology/ (Accessed: 30 May 2019).

Walport, M. (2016), 'Distributed Ledger Technology: beyond block chain'. UK Government Chief Scientific Adviser, Office for Science. Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/492972/gs-16-1-distributed-ledger-technology.pdf (Accessed: 30 May 2019).

Weaver, J. (2019), 'The Value of Inconvenient Design'. Medium. Available at: https://medium.com/s/user-friendly/the-power-of-inconvenience-f0ae1773dd77.

Wilson, C. (2010), *User Experience Re-Mastered*, *User Experience Re-Mastered*. doi: 10.1016/C2009-0-20682-9.

Wood, G. & Buterin, V. (2014), 'Ethereum: A Secure Decentralised Generalised Transaction Ledger', in *Ethereum Project Yellow Paper*. doi: 10.1017/CBO9781107415324.004.

Zhelezov, D. (2018), 'PoW, PoS and DAGs are NOT consensus protocols'. Coinmonks Blog. Available at: https://medium.com/coinmonks/a-primer-on-blockchain-design-89605b287a5a (Accessed: 1 May 2019).

Zhou, J. & Gollman, D. (2002), 'A fair non-repudiation protocol', in. doi: 10.1109/secpri.1996.502669.

Appendices

Appendix 1: Usability Test of a Blockchain Product - Screening Survey

Appendix 2: Scenario and Task Description for Cognitive Task Analysis

Appendix 3: Steempery Certificate for Cognitive Task Analysis

Appendix 4: Test Guide to read before Cognitive Task Analysis

Appendix 5: Breakdown of Tasks and Subtasks in the User-Flow Diagram

Appendix 6: Guide for Semi-structured Interview

Appendix 7: Results of Screening Questionnaire

Appendix 8: Excepts from Recordings of the CTA and Post-test Interview

Appendix 9: Excepts from Recordings about Trust and other Topics

Appendix 1: Usability Test of a Blockchain Product - Screening Survey

Please answer the following questions.

Gender: Female ( ) Male ( ) Other ( ) I prefer to not say ( )

Age: Below 20 ( ) 20-30 ( ) 30-40 ( ) 40-50 ( ) More than 50 ( )

1. Have you signed a contract with a digital tool or added your signature to a PDF file?

   Yes ( ) No ( )

2. How familiar are you with digital systems to edit and save data such as spreadsheet tools? (For example: Excel or accounting software)

   5 Master (I use the software regularly operating the majority of its features)

   4 Intermediate (I use them occasionally and use many of its features)

   3 Basic (I know the programs and I have used them)

   2 Totally Unfamiliar (Have never interacted with such computer programs)

   1 None of the above describes my level of familiarity with a spreadsheet tool

3. Have you heard about the term Blockchain?

   Yes ( ) No ( )

4. Have you read about Blockchain technology or its application?

   Yes ( ) No ( )

5. Have you made or told other person to make a transaction on a blockchain? e.i. sent or received cryptocurrency.

   Yes ( ) No ( ) Maybe ( )

6. Have you used a decentralized application or dApp?

   Yes ( ) No ( )

7. Have you programmed on a blockchain? e.g. smart contract or token issuing.

   Yes ( ) No ( )

8. Would you be willing to test a user interface that makes use of blockchain and displays information recorded on it to assess the value of a product. (The movement of the cursor will be recorded together with your voice and utterances at the moment of interacting with the system. Data that personally identifies you will not be collected) *

   Yes ( ) No ( )

Appendix 2: Scenario and Task Description for Cognitive Task Analysis

You and a person named John Doe have signed a contract. You both know what is in the agreement and you have a digital copy of it. Together, you decided to certify that the contract exists using the Ethereum Blockchain, so you don't involve any third person as witness.

John offers to create a certificate of the agreement using a tool called "Steempery." Such tool generates a PDF file that acts as certification. He has sent the certificate to you via email saying that you can trust he won't change the contract because the version you signed is impossible to change.

Read the certificate that verify that a fingerprint of the contract has been undeniably saved on the Blockchain. You will need to confirm that you have a proof that the contract won't be able to be tampered with. You should feel capable of successfully take legal measures in case that he doesn't perform or live up to his part of the contract.

Appendix 3: Steempery Certificate for Cognitive Task Analysis (Page 1 and 2)

Steempery Certificate

This document certifies that the file referred hereunder exists and was presented in the date and time printed down below by the person identified as "signee".

NAME	Important_Contract_Signed.pdf
DATE & TIME	April 23, 2019. 03:53:46 PM +UTC
SIGNEE	John Doe (john@doe.com)

Blockchain technology ensures and certifies that your file existed at a certain point in time. Timestamping with the Ethereum Blockchain is a way to irrefutably prove the exact moment a document exists for legal, compliance and business purposes.

DATA HASH

01ad39c64a2db50fa3c802e4159bb158a9b928f6b635cc4d2b5e4ed605421684

A DATA HASH is a fingerprint of your document. A *Standard Hashing Algorithm* (SHA256) has been used to produce this unique DATA HASH. This cryptographic identifier irrefutably represents your data because any change to the file is mathematically proven to generate a totally different string of characters when processed again by the algorithm.

TRANSACTION HASH (TxHash)

0xd7ca7358ce6668c22a82ee36be2c1f748d2ae630cdd3d277ff9437d6237e7993

The TRANSACTION HASH uniquely identifies a particular transaction. All blockchain transactions have a unique TxHash. This set of letters and numbers can be used to find all the details of the transaction using a Block Explorer.* The proof that your file exists will remain stored as data in this transaction of the Ethereum Blockchain even if *Steempery* disappears.

* *We recommend EthStats to see the details of the transaction at https://ethstats.io/*

↓ Screenshots of the content in the file are presented below ↓

Intensjonsavtale

Intensjonsavtale for samarbeid om bruk av CSR verktøy med blockchain

Mellom **BlockchainCompany** org.nr. 123 456 789 (heretter "BlockComp") og **ClientCompany** org.nr. 987 654 321 (heretter "Client", og "Partene") er det på dags dato inngått slik intensjonsavtale.

1. BlockComp er et selskap som utvikler og tilbyr en et system for å forbedre sertifiseringen av sosialutviklingsfinansiering som påvirker landsbygdssamfunn fra selskaper som utfører prosjekter for samfunnsansvar (CSR).

Client er en organisasjon med formål å bedre levevilkårene for barn i Colombia. Utdanning er grunnsteinen i arbeidet. Client arbeider samtidig for å gjøre lokalsamfunn i stand til å fremme egen bærekraftig utvikling.

2. Partene ønsker å sammen se på mulighetene for å benytte CSR-verktøy med blockchain teknologi for et felles pilotprosjekt for investering av bedrifters samfunnsansvarsmidler i Colombia. Teknologien er utviklet med det formål å sikre midlenes bruk i forhold til formål, samt redusere administrasjonskostnader og forhindre korrupsjon. Samarbeidet i pilotfasen skal foregå på følgende måte;

3. Partene skal sammen identifisere et utviklingsprosjekt der Client allerede har etablert virksomhet. Prosjektet velges av Client med innspill fra Client på egnethet for et pilotprosjekt. Målet er med hjelp at Clients lokale nettverk og eksisterende organisasjon på plass, i kombinasjon med BlockComp løsning for (beskriv) kan verktøyet kommersialiseres gjennom å tilbys både investorer og mottakere for en definert brukerkostnad.

3.2 Denne avtalen innebærer ingen økonomisk forpliktelse mellom partene. Hver av partene bærer sine egne kostnader ved prosjektet dersom ikke annet er skriftlig avtalt. Partene vil i felleskap søke ekstern prosjektfinansiering fra lokale og norske myndigheter og bedrifter. Partene utser en arbeidsgruppe med minimum 1 person fra hver part som tar seg av søknader og koordinerer kommunikasjon mellom partene og institusjoner og partnere som det søkes støtte fra.

4. Partene skal etter beste evne samarbeide i all ekstern kommunikasjonen rundt prosjektet. Dette kan gjøres blant annet gjennom felles pressemeldinger og poster i sosiale media. All offentlig kommunikasjon skal godkjennes av den andre part før det offentliggjøres eller publiseres offentlig.

5. Denne Intensjonsavtalen tar til å gjelde når den er undertegnet av begge partene.

BlockComp
John Doe
Daglig leder
Email: John @doe.no
Tlf.: +47 33445566

ClientCompany
Jane Doe
Daglig leder
Email: Jane@doe.no
Tlf.: +47 22334455

Appendix 4: Test Guide to read before Cognitive Task Analysis

Hi [name of participant]. My name is Daniel Arevalo, I am a student at NTNU and I am going to be accompanying you during this usability test that I doing as part of my research. I appreciate that you are taking the time to help me.

The thesis is about blockchain products so you will be testing a website that people use to verify data recorded on a blockchain. I want to make clear that we are testing the site, not you. I want to see how people actually use it. There is not a right way for doing it, so you don't have worry about making mistakes.

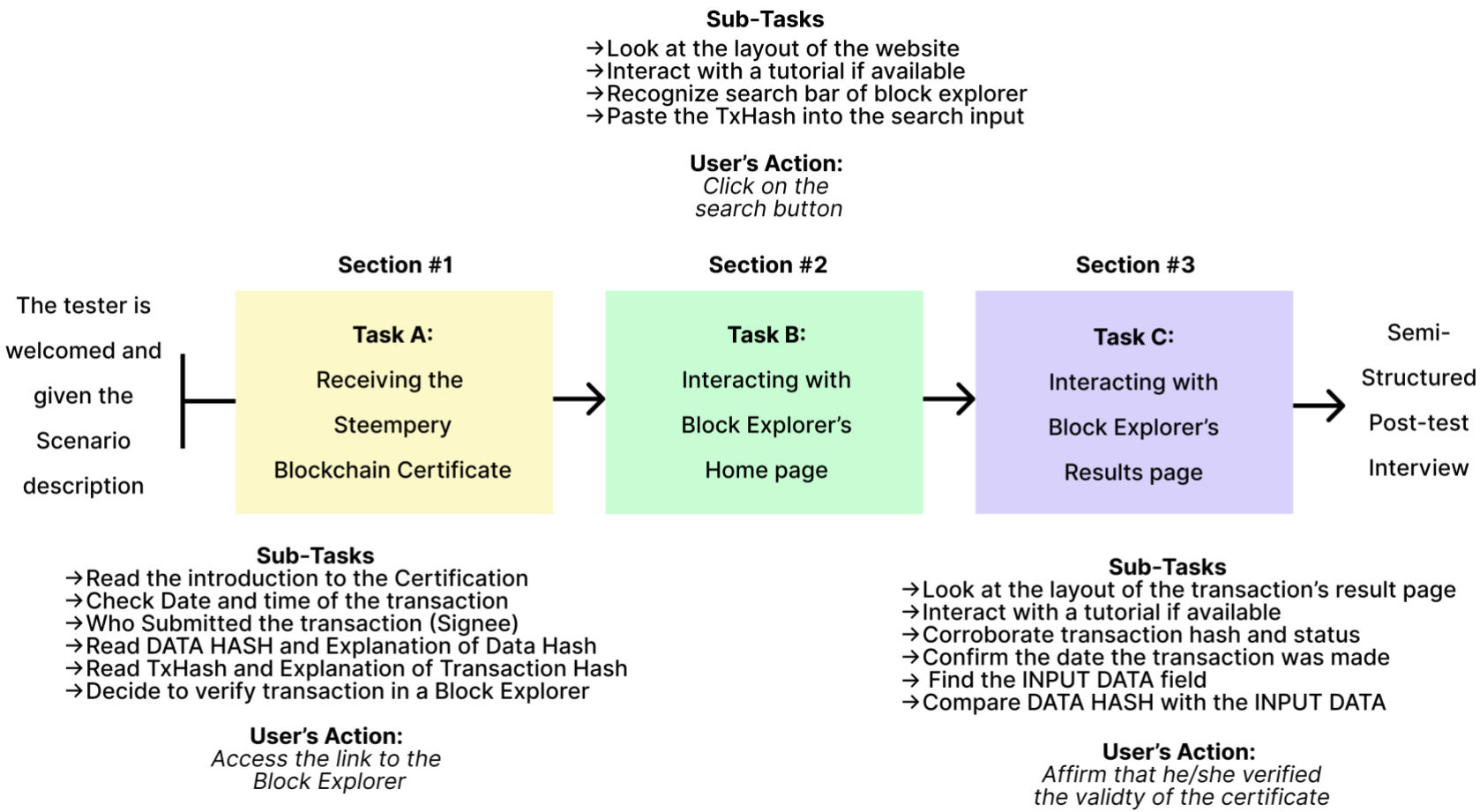
I want to clarify that with your permission I will record audio during the test, that means your voice, utterances and what you say about the interface. Also. I will be recording the movement of the cursor on the screen that you will be interacting with. The recording will only be part of the thesis to help analyse the usability of the tool and will only be seen by people working on this thesis. Data that personally identifies you will not be collected, and all will be deleted after the Thesis delivery.

When you are using the website, please say exactly what you are thinking. Read out loud and honestly mention what you have in mind when interacting with the website. If you have any questions, just ask, but try to find and answer first as you would in the real world. I also may ask you questions so you can identify what to do next.

While you are going through the website, try to follow your sight with the cursor so when I hear your thoughts I can know exactly where in the interface you are most likely looking at. For example, (show on the screen how to do this) when you are reading certain information that you consider important or have a question about, please move the cursor around wherever you are looking at on the screen.

Do you have any questions before we begin?

Appendix 5: Breakdown of Tasks and Subtasks in the User-Flow Diagram



Appendix 6: Guide for Semi-structured Interview

Before interacting with Block Explorer

A1: Defining the Goal

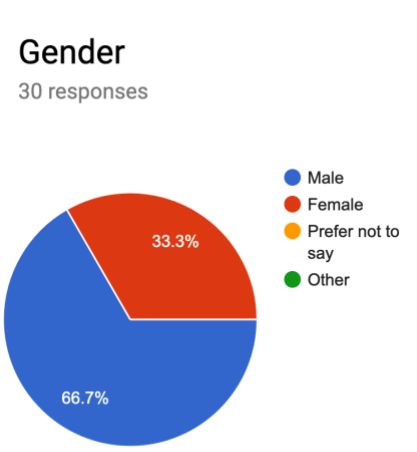
1. **What did you think was your goal after reading the Certificate?**
2. **Did you understand what DATA HASH and TRANSACTION HASH were?**

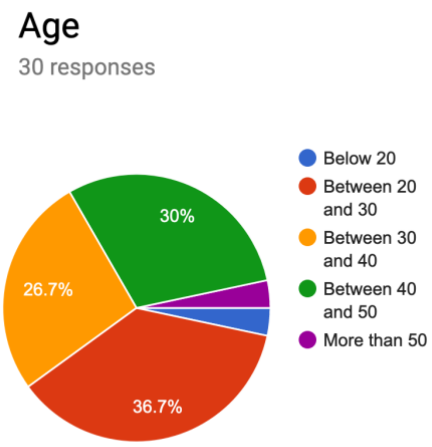
After interacting with Block Explorer

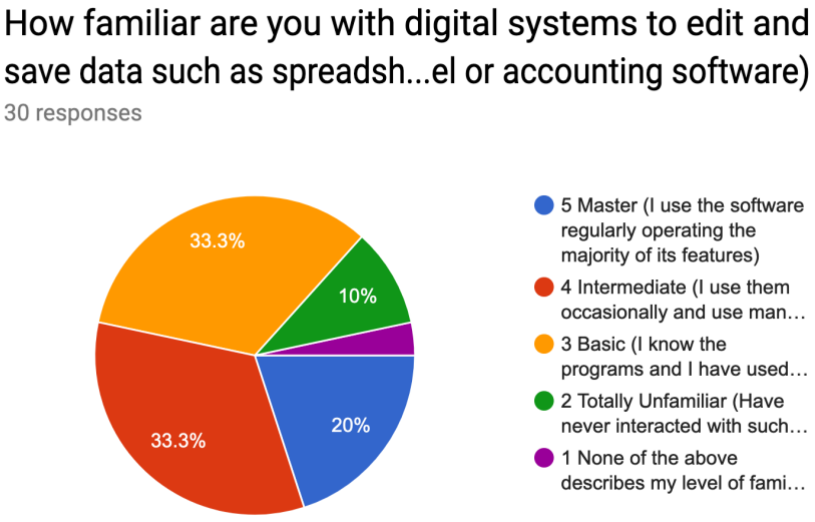
Walk me through what you though until you notice the Input Data section.

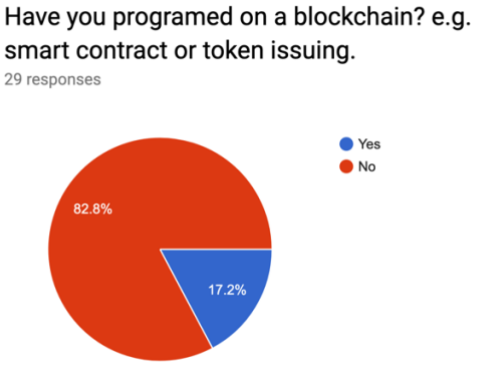
Gap of Execution		
Category	Heuristic	Interview Questions
A2: Forming the Intention	H1	Do you think that the terms shown in the interface helped you to know what to do?
	H2	What were you first thought when entered the Homepage of the block explorer? Did you think about similarities with anything you know?
A3: Specifying the Action Sequence	H2	How did you realize you could make a search? When you decided to make a search, what were your expectations of the result page?
	H3	Did you think about searching or entering the tutorial for the tool? Why you interacted or not with it?
A4: Executing the Action	H7	Did you need to learn a lot from the interface to make the search? Do you think you had to remember a lot of information?
Gap of Evaluation		
Category	Heuristic	Interview Questions
A5: Perceiving the System State	H4	What were you first thought when entered the Results page of the search? Did you receive Feedback or explanations that assist you understanding where the Input Data field was?
A6: Interpreting the System State	H5	How did you start confirming the Input data was the Data Hash? Did you think you were in an intermediary step or expecting to do something more?
A7: Evaluating the Interpreted State	H5	How did you realize you had verified the Data Hash given to you on the certificate? Would you able to verify another certificate if it's given to you in six months?
	H6	What would you do if the information wasn't there / was wrong? How would you start over if you introduced the information wrong?

Appendix 7: Results of Screening Questionnaire



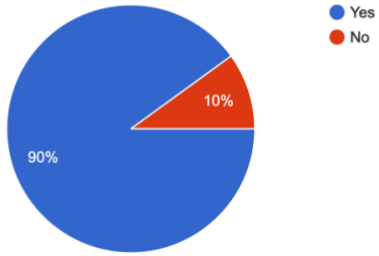






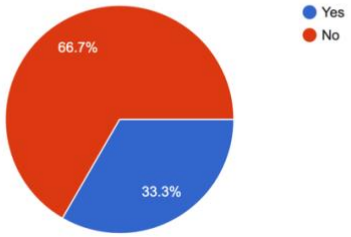
Have you heard about the term Blockchain?

30 responses



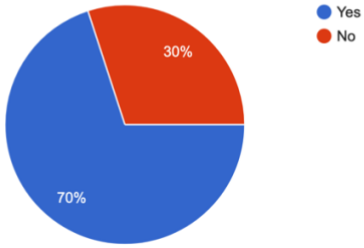
Have you used a decentralized application or dApp?

30 responses



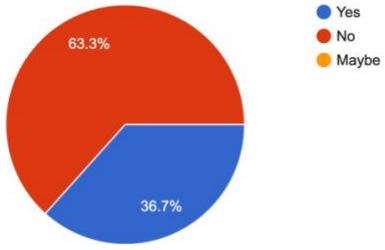
Have you read about Blockchain technology or its application?

30 responses



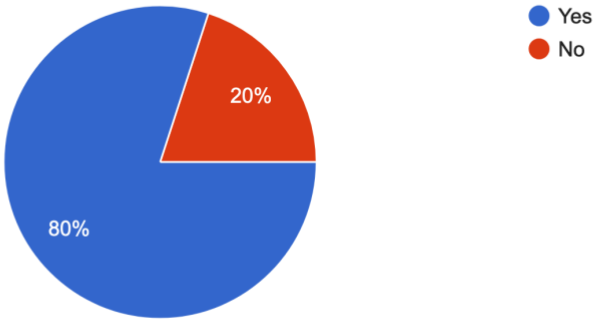
Have you made or told other person to make a transaction on a blockch...ceived cryptocurrency.

30 responses



Would you be willing to test a user interface that makes use of blockchain and displays information recorded on it ...ly identifies you will not be collected)

30 responses



Appendix 8: Excepts from Recordings of the CTA and Post-test Interview

Category A2: Forming the Intention

Regarding H1: *Rules and Underlying concepts which are useful for learning should be made available to the user.*

#1: "It was really easy, just typing in the transaction hash you can find the way. I just saw the welcome sign and then directly to the search so that it was the first thing I saw. It was useful"

#2: "It's very Straightforward" "It's very easy to Use" "The first thing you see is the search tool and the transaction hash the block hash"

#3: "I don't think that the terms helped me necessarily. I think I understood because I know the terms and I am used to work with them. I think there are a lot of foreign terms that are hard to understand."

#5: "I was a bit confused by the term block.

#6: "The website looks really nice. The homepage helped a lot because I sort of intuitively knew I could paste something here"; "Looks like a huge search bar and there was only one thing I could do. I was forced to do the right thing."

#7: "You have to be interested in blockchain technology to understand the terms."

#8: "This is Ethereum, but there are a lot of words that I don't understand."

#9: "Looks like something that a password generator spews out."

#10: "I didn't spend much time looking at the bars because I read the placeholder text right away."

#11: "The interface looks clean, and it was easy for me to link the transaction hash and look for it at first sight."

#12: "I think that the terms are too general to understand the context."

#13: "It literally says type the transaction hash, so it is very clear and the color contrast is good."

#14: "This is what I was looking for, so I didn't pay attention to the rest, because I was interested in searching for a transaction hash."

#15: ""It's very straightforward"'; "It's very easy to Use"'; "The first thing you see is the search tool and the transaction hash the block hash."

Regarding H2: *If infrequent use or user characteristics require relearning of the dialogue, then appropriate support should be provided.*

#1: "It sort of looks like Google."

#2: "The interface has clean design, minimalistic and good coloring."

#3: "I thought it was a clean looking page and understood immediately that the main function would be to look for transaction hash or some sort of hash of accounts."

#4: "I have used a registry where I need to input a certain code to get the data I am looking for."

6: "Looks nicer and cleaner that other search engines I have used. Normally block explorers would look very bad, clunky and hard to use, very like for engineering people. But this one looks very friendly."

#7: "The first think I saw from the site was the search bar and I knew right away that I needed to copy paste my data hash."

#8: "Reminds me of buying shares in a company or my bank account."

#9: "The user interface reminds me of Google or Google Analytics, so it is a familiar interface."

#11: "I think a Rainbow hash search engine to find about password hashes, perhaps 2 or three years ago."

#12: "The page focuses on the search tool which is quite important in this context. Also, it resembles many different search systems like Google and registries."

#13: "Is looks like a search engine."

#14: "It looks like a nice and professional website, and it's clear that it's for looking at data in the blockchain."

Category A3 Specifying the Action Sequence

Regarding H2: *If infrequent use or user characteristics require relearning of the dialogue, then appropriate support should be provided.*

#1: "I thought I would get information about the transaction, but I wasn't sure exactly what I would get from the search."

#2: "From the task, I understood I needed to make a search for getting to completion"; "I was thinking it would be really nice if this just shows me that the transaction is related to this contract and I can click on the contract and then I can see if its stored."

#3: "I was expecting to get all the information about the transaction in some structured form. It was a little bit messy what I did get. I was expecting to get all the low-level details of the transaction."

#4: "I would think that I would get details about the blockchain entry. That would be enough for me to verify the contract I have made with Mr. John Doe was entered into the blockchain."

#6: "The placeholder text helped me to understand that I could paste that kind of information"; "The other parts were changing, so I understood they weren't applicable for the task and that the search bar was the only thing to use"; "I was expecting to see more of a legal sort of page. Basically, having the transaction information but also the contract attached to the blockchain as a pdf."

#7: "First input the data hash and I thought maybe that would lead me to the contract. the pdf, but it didn't happen."

#9: "It also reminds me of a stock market interface where I can see if my stocks are plummeting or not. It reminds me of an information board at the airport. Or a QR code, or a statics graph. Interface to machine."

#10: "I was expecting to see a pdf or the contract that I have received. Some form of recognition of that contract."

#11: "I expected to see more neutral information, if it would have been just that the transaction is ok, I wouldn't have believed it. I am looking for pieces of data that would make me trust it"; "Although I haven't used the vocabulary, it gave me the impression that it was ok."

#12: "Being greeted by the tutorial was useful for me to find where to search for the transaction"; "I was expecting something technical and it was confirmed when I did the search. In the context I was, it can be a little bit confusing, too general for it."

13: "It felt really engineery, like someone said we need all this information so just put it there."

15: "One thing they could do is for example more information about how a transaction look like within the interface""

Regarding H3: *Appropriate support should be provided to assist the user in becoming familiar with the dialogue.*

> **#1:** I actually didn't think about looking for the tutorial, because when I get to a site I always like to search for the thing I am looking for, and then if needed, I would look for help
>
> **#2:** I always skip the tutorial because usually it is more hassle to read it than just try and figure it out. It is more efficient.
>
> **#3:** "I didn't look for the tutorial because I usually don't do that. I Usually just look around and try to figure out the things myself."
>
> **#3:** The data hash was stored in the transaction so I would check the transaction on a block explorer and enter the hash to find the transaction in question and look for the data that has been verified by the miners and by the network. And see if the data inside matches
>
> **#4:** "I don't like tutorials. Personally, I would like to fiddle with it until I figure it out."
>
> **#6:** "I didn't think about interacting with the tutorial for the first page, but then this second page was a bit more confusing."
>
> **#7:** "I got a walkthrough of the site."
>
> **#9:** "I would look for more information about transaction hash or how it would look like, trying to find out what does it mean for me."
>
> **#11:** "I would have just tried instead. If you need a tutorial, then the tool is not good enough."
>
> **#12:** "I click "next" on the tutorial since I am unfamiliar with this website. However, there was not any content interesting for me in that context, so I didn't evaluate it more than I needed. I totally forgot about the tutorial tool in the second part of the page. I was too focused on finding the information myself. I wouldn't have noticed the tutorial either, if it wasn't for the notification on the first page. So, if you could get any kind of notification on the transaction hash page as well, that could be brilliant because there is where I needed to understand more."
>
> **#13: "**I use the tutorial if I am lost, but I would like to have it pop up when you are in the results page.**"**
>
> **#14:** "I didn't want to look for the tutorial because the page was very easy to understand, and I was able to do what I needed without using the tutorial."
>
> **15:** "There was a pop up, that I just denied, so perhaps it would be more guiding"; "I have a pretty good understanding of the concepts, so I didn't feel it was necessary"

Category A4 Executing the Action

Regarding H7: *The interactive system should enable the user to perform the tasks with minimal learning by entering only the minimum amount of information required in the dialogue, with the system supplying additional information on request.*

> **#1:** "I didn't need to know much about blockchain, but I think I needed to learn about the data hash."
>
> **#2:** "If you know what a transaction hash is, I think you would be ok, but I didn't feel completely confident."
>
> **#3: "**I think the search part was easy. It was very clear where I needed to search, and what, and how"; " Using the transaction hash is a little bit clumsy because you need to copy and paste some long string from somewhere and if you don't have in your computer then it would be a pain."
>
> **#4:** "I didn't need to learn anything to make a search."

#6: "I didn't need to remember information, and I knew which to use first. That the hash would be for verification. I knew I was going to use that, and copy pasting was not difficult."

#9: "I suppose I would need to retrieve the contract and make sure that it hasn't changed."

#10: "Since the hash is so complicated you don't bother on memorizing it."

#11: "I didn't remember the data hash, but it was ok just to switching back and compare it."

#13: "I didn't need to remember information. It's very good that I just needed to copy and paste the transaction hash."

#14: "It was evident what to do, I didn't need to remember much information."

#15:"I didn't need to learn a lot right now or hold information in their minds as long as they understood the concepts"; "I was expecting to look up for the data hash and not just the transaction hash because the data hash actually has the contract while the transaction is more specific to the protocol."

Category A5 Perceiving the System State

Regarding H4: *Feedback or explanations should assist the user in building a conceptual understanding of the interactive system.*

#2: "I haven't used this interface before, there is a lot of information cluttered here."; "Much coloring, which is nice, but it doesn't help me find what I am looking for."

#3 : "There is a lot of information here, it was a little bit messy, there is a lot of information that is sort of structure, but there are also a lot of things happening here. Somehow it is difficult to see how things are connected with some terms that I am not used to"

#4: "Everything is new. The contract details, and all the descriptions of the contract I wouldn't have figure out"; "That the important field is labeled data hash and not input data. I would have found it in two seconds. Also, when the number is spread, suddenly it doesn't look right. It lacks similarity with the certificate."

#5: "I saw the label input data, but I was first looking for a lot of data, and in that a lot of data one would be my hash."

#6: "The hashes were a little. It hard to see because they were shortened. So, I didn't think about them. Maybe if they would have been underneath each other and with the full length maybe they would make more sense"; "There was a lot information that is not interesting for the task, so I had information overload."

#7: "Shouldn't the label Input Data be called data hash? It should be the same so it would be easier to identify."

#8: "It was the most important information, but it should be bigger and with color, but here feels just extra at the bottom."

#9: "This is the first time I see blockchain, so this green really tells me very little except that something is confirmed because my prior knowledge to the technology and what to expect is zero."

#10: "One of these numbers should stand that it is the fingerprint of the contract"; "I didn't recognize the data hash looking like that. But it was at the bottom where it should be"; "Since I am not familiar with the parts of the contract, it's a bit too much information coming towards me, but it would change if I would be familiar with this and I would recognize the numbers."

#11: "There is a lot of information that makes me think that I need to read though it several times. There is a lot of complexity and data that I don't know the vocabulary."

12: ” I was pretty confident I have found the hash of the transaction and it was in the right place as long as I have search for the hash, and it started with the same number.”

13: “The first thing I saw was the confirmed word, and since I don’t have much blockchain background at least it makes me feel that everything is good”; “I didn’t think there was an additional step. I thought that I was just going to copy and paste and then, here is the outcome, and say ok it’s good.”

14: “It was pretty much what I expected. I see the transaction hash and I see how much Gas was used to perform the transaction. Then I scrolled down for the hash of the document”; “Because the Input Data is very important it should be high up, because it is what’s important to me.”

#15: “When I search the Data [Hash] it didn’t find it, but then when I typed the transaction hash it immediately found it, so it was reassuring that it open automatically”; “There was a lot of information, basically the first I realized before thinking about finding the Data Hash, because that’s the information you get”; “I noticed [the transaction] had from and to. Also, that the date wasn’t specified, just how long ago the transaction had been done.”

Category A6 Interpreting the System State

Regarding *H5: The dialogue should provide sufficient feedback about the intermediary and final results of an activity so that the user learns from successfully accomplished activities.*

#1: “I started to compare them, I saw the first two numbers and then compared all of it overall and saw it was similar.”

#2: “I know the contract has the hash and I want to make sure it has been timestamped”; “I just read the first and last characters, but then I would have read all of them one by one, up to four pairs of numbers.”

#3: “There are sections that are just for smart contracts”; “It says Input Data and I assumed it was it. But because it was in the same format but the structure it was made it a little bit hard to see that it was actually the Data Hash”; “I think I would need to decode the Input Data.”

#4: “From the contract I check the time stamping and that I has been registered in the Ethereum blockchain.”; “I see that Ethereum is in the name of the whole place. So, for me it’s the right place.” “I was looking for the Data hash at the beginning on the top, and I saw the transaction hash but not the data hash.” “This information, for me since I don’t know about blockchain, would be irrelevant. And I was looking for the name in the contract and the rest of information was distracting, I would spend time looking at it.”

#5: “Somewhere in this block should be some metadata referring to my Data Hash, because I understand that my hash should be coded somewhere in this transaction.”

#6: “As long as I can find the contract on this transaction, it should be fine”; “I think there is like a contract on the Ethereum blockchain based on this, so we can both verify”; “Because these letters are spaced, I immediately though this wasn’t the right thing to see. Because Hash are normally displayed as a string.”; “When comparing, it was easier to read it this way because I could check specific numbers and combinations.”

#7: “I understand that the input data was a data hash, except that I didn’t understand it was the same one in the contract. So, I didn’t check it against the certificate.”

#9: “The input data corresponds with the data hash, so I should copy it and compare”; “It took me quite a while to understand that this was the input data. I would thing that there was another step, to get another verification like a green pop up doing Bing.”

#10: “I was looking at the date and the accounts. Especially I stopped thinking that it was confirmed, because the format of the time made me thing that it was time left before confirming”; “I was looking for a signature and obviously a signature in a contract should be at the bottom.”

#11: "I don't understand this, therefore I ignored it. Then I recognized the data hash because I remembered the first number"; "I would need a reminded of how to do it next time."

#12: "I think that this is a confirmation that the transaction has been incorporated into the blockchain. I was a little bit unclear because of the formatting of the hash"; "Being in the contract situation, I knew it had nothing to do with Gas"; "I saw the first four numbers of the transaction and then looked at the last digits, but I didn't look for the whole hash at first."

#13: "I didn't know why this is blue and why this is black. I don't understand why the input data doesn't have any colors. It should look like a result and adding colors would be enough."

#14: "I would like to see and explanation here saying 'we verified that the document with this hash existed at that point in time'"; "This data is Hexadecimal. So, you cannot put a string here because it wouldn't be readable still." "I looked at the three first and three last characters, because with hashes, even if you make a small change, they would change completely so I usually don't bother comparing the whole hash, I usually often look at the three and a last byte and I would be satisfied. However, if you would be at a court case you will show that every byte is the same."

#15: "It was quite difficult to find the Data Hash because the format is a bit different. I look at the two strings and just compared them looking at the first 6 characters"; "I feel that there is a lot more that I have to do. What if there are duplicate transactions? Then, I wouldn't be allowed to click and see if there is a duplicate one"; "That's why I think it would be very nice to search for input data, because if I search for input data, you could get automatically get the transaction that holds the contract and telling me if there is a duplicate."

Category A7 Evaluating the Interpreted State

Regarding *H5: The dialogue should provide sufficient feedback about the intermediary and final results of an activity so that the user learns from successfully accomplished activities.*

#1: "I am not sure if this is enough. I should look at the website to know if it is a real website."

#2: "If the information wasn't there and I really needed to confirm it, I would have also used etherscan."

#3: "I would need to verify [the input data] against the one [data hash] that I do have. I would use an online tool to do a string comparison."

#4: "I was takin one by one to be 100% sure"; "I was expecting to confirm something, like confirm that I have confirmed it. Like a verification checkmark."; "I would expect to Log in somewhere and approve it has been checked from this party."

#5: "I would put the Input Data and the Data hash on excel and use the function of Equal."

#6: "All the information of the transaction is being displayed right now, but it isn't connected to the contract."

#7: "I would contact the guy and he would be able to check how the hash is wrong and see that the contract is not valid. If he does it again he would have a new transaction hash with the right data hash and I would be able to confirm it."

#8: "After doing it one time I think I would be able to do it again."

#9: "I would be able to do it again in 24 hours, but I would be sure if I would be able in 6 months."

#10: "Since the contract in itself is unfamiliar, you kind of don't know what you are looking for."

#11: "If it's a service that let me verify that a contract has been signed, then perhaps it is too generic because I was trying to match the input data hash so I would think it would be more central information before the rest of information."

#12: "I didn't find the data hash stored in the blockchain and I couldn't find it, so I think the document is not part of the blockchain. If the digits are wrong, I would think that the contract is invalid."

#13: "If you know what you are looking for, it becomes really easy."

#14: "I think it's all I need, and also confirmed that the transaction has been recording on the blockchain so I know that it is not a pending transaction that might be canceled.**"**

#15: "I would like to get some more proof that [data hash] is actually the contract. So maybe there could be a mechanism to decrypt after providing some credentials like BankID for example."

Regarding H6: *If appropriate to the tasks and learning goals, the interactive system should allow the user to explore ("Try out") dialogue steps without negative consequences.*

#2: "I entered to the link at the hash of the external account and was able to go back to the transaction."

#4: "It was not intuitive for me that I could press some elements. It looks like a pdf almost. It should visible that they can be pressed"; "I would be back to the contract party and ask him to send me a revised certificate. I would have checked three times with the interface before contacting the person."

#6: "I would be able to refresh the page, but if it wouldn't don't load, then it would be difficult to search for another site."

#7: "I think it would be very natural to do it again."

#8: "I would suspect that something was changed. I would try one more time before contacting the person I made the contract with. But that's the thing with blockchain, that there is nobody to call."

#9: "I would be able to do it again in 24 hours, but I would be sure if I would be able in 6 months."

#10: "I would check with the person that sent me the email if it is correct. I would recognize that it was a human error, because I would think he did something wrong."

#11: "I wouldn't trust the contract; I would think the contract is not valid. Then, I would have contacted the sender."

#12: "Entering the Page of the Block I was really not sure, however I was able to go back to the transaction page by clicking on the back button of the browser.

#13: "If there is an error in the data hash it would take away the credibility that nobody has tampered with the contract. To solve it a would look for other software and see if I get the same result, and if I get the same result everywhere, I would confirm with the person that everything is ok."

#14: "I might for example use one of the other block explorer and try looking at the information with them instead. If I cannot find it then I would ask the person that claimed that has done it to explain where he did it. If they couldn't then I would say that I can't tell that is on the blockchain and I cannot verify the document won't be changed."

#15: "I did a wrong input of data and was able to search again for the correct hash." "If I find a mistake in the Data Hash, I wouldn't trust the algorithm. I would still trust the interface, but I would contact the business partner and tell them that I don't trust the certificate because if the contract had changed, I wouldn't understand what changed and how the hash was generated. Without those two pieces of information, I wouldn't be able to validate the certificate."

Appendix 9: Excepts from Recordings about Trust and other Topics

#1: "I am not sure if this is enough. I should look at the website to know if it is a real website" The participant then went through looking for the company that had created the website and reading more about it. "If I get to confirm it is a real company, that I can trust them, then I will confirm that the certificate is valid"; I would report him or first speak to him and ask him if he is sure if the certificate is real.

#2: "I was hang up on it's probably stored in a smart contract and I didn't think it was going to be as data in the transaction" I think I had a hard time getting myself into this situation, but before calling out the person, if it would be in a power position, I would have checked if there had been something wrong with the hashing"; "I think I had a hard time getting myself into this situation, but before calling out the person, if it would be in a power position, I would have checked if there had been something wrong with the hashing"; "I haven't gained any trust in that Steempery that actually generate the hash. But if I can reproduce it and it has been timestamped, then I would be sure that I can say that this contract was presented at that time." Because I haven't used Ethstats before I would only confirm that the data is recorded on the blockchain using other BlockExplorer."

#3: The user requested see the service Steempery "basically to hash the document just to make sure that it is the same, that he didn't in fact changed it"; "If the information was wrong, I would assume that the other person is trying to scam me, so I would contact that person and talk about that. But what I could is that I could put the data hash in a transaction myself so I would have a proof in the blockchain."

#4: "I would like to look for details that are in my contract the log entries would make me comfortable that my partner hasn't edit anything since last time, if he had it would be probably a number one showing here"; "And this is the fee for having the contract saved, this is the amount. This can confirm me that it has been registered and if I knew about blockchain these numbers would give me insight about where, and what is the security level of it."

#5: "I need to confirm that the contract generates that Data Hash."

#6: "I would probably contact the person that sent me this certificate. I wouldn't contact the tool because I feel I have the "Google is always right" syndrome where I feel humans make mistakes, but Google doesn't."

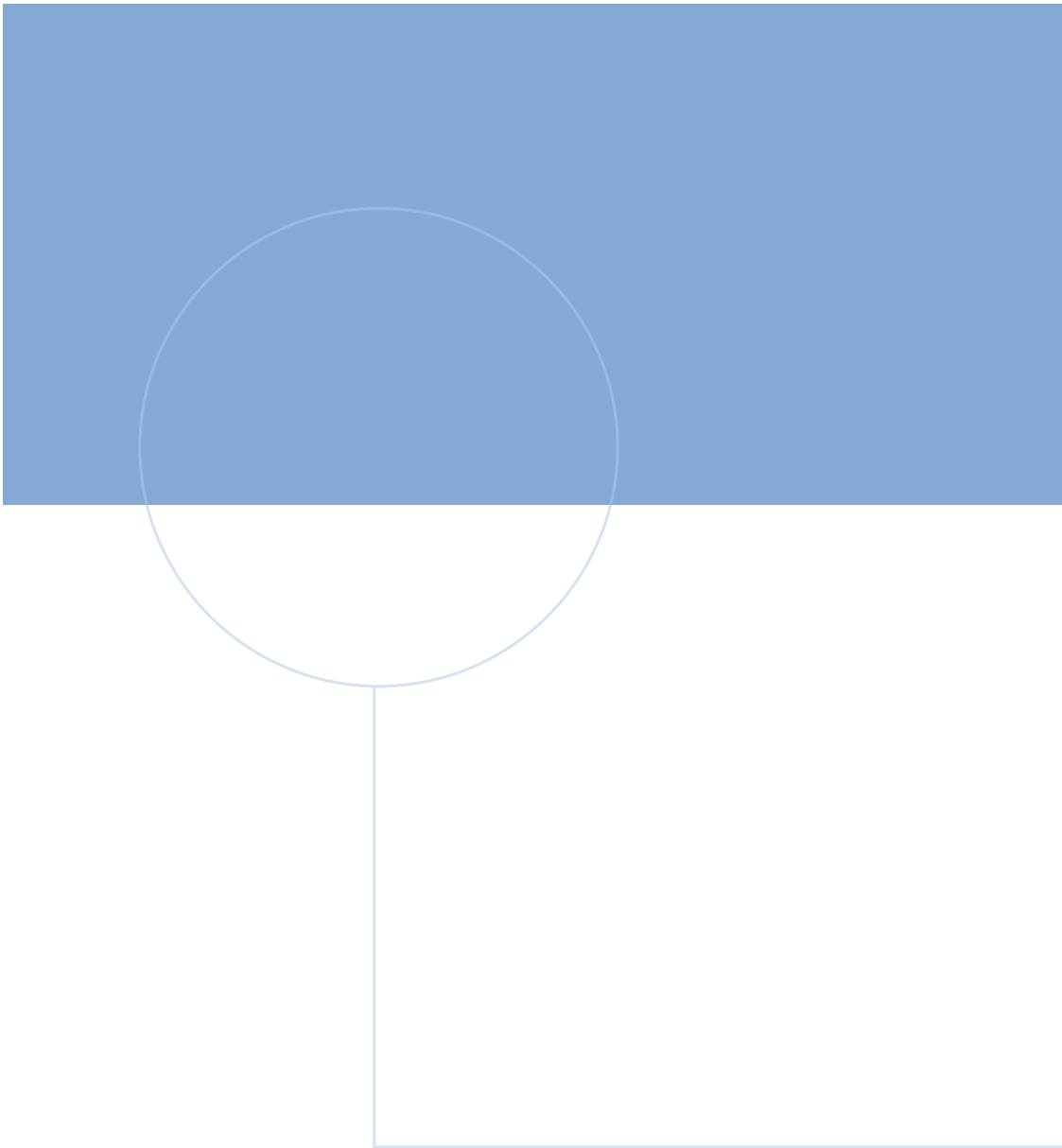
#9: "I wouldn't trust the person because I tend to trust just machines that have been instructed properly."

#11: "I don't know the service EthStats, so I would like to confirm it is legitimate." "I wouldn't trust the contract; I would think the contract is not valid. Then, I would have contacted the sender"; "I was comparing this to the same process using the Norwegian postal office solution, which I am using every day, signing contracts back and forth. Both for sending contracts and getting contracts, and all parts are basically the same except for the validation."

#12: "As a developer, I am unsure whether or not this is enough to ensure the data integrity with this algorithm." "Looked for information about the Hashing Algorithm on Google" "I think that this is a confirmation that the transaction has been incorporated into the blockchain" "I didn't find the data hash stored in the blockchain and I couldn't find it, so I think the document is not part of the blockchain." "Entering the Page of the Block: "I am really not sure" however he was able to go back to the transaction page by clicking on the back button of the browser.**"**

#13: "If there is an error in the data hash it would take away the credibility that nobody has tampered with the contract. To solve it a would look for other software and see if I get the same result, and if I get the same result everywhere, I would confirm with the person that everything is ok."

#15: "I would like to get some more proof that [data hash] is actually the contract. So maybe there could be a mechanism to decrypt after providing some credentials like BankID for example. "I would need information about the algorithm and knowing if the algorithm is certified because if someone ends up using a different algorithm and it delivers a different output then people would think it is not valid"; "If I find a mistake in the Data Hash, I wouldn't trust the algorithm. I would still trust the interface, but I would contact the business partner and tell them that I don't trust the certificate because if the contract had changed, I wouldn't understand what changed and how the hash was generated. Without those two pieces of information, I wouldn't be able to validate the certificate."



NTNU
Norwegian University of
Science and Technology