University of Zurich[UZH]

MASTER THESIS – Communication Systems Group, Prof. Dr. Burkhard Stiller

An Educational Blockchain for the University of Zurich (UZHBC)

Jerinas Gresch
Zurich, Switzerland
Student ID: 12-716-627

Supervisor: Bruno Rodrigues
Date of Submission: August 19, 2018

University of Zurich
Department of Informatics (IFI)
Binzmühlestrasse 14, CH-8050 Zurich, Switzerland

ifi

Master Thesis
Communication Systems Group (CSG)
Department of Informatics (IFI)
University of Zurich
Binzmühlestrasse 14, CH-8050 Zurich, Switzerland
URL: http://www.csg.uzh.ch/

Abstract

Diplome haben einen hohen Stellenwert in der Gesellschaft, da sie als offizieller Nachweis für die Ausbildung und die Fähigkeiten der Inhaber dienen. Daher ist es nicht verwunderlich, dass Fälschungen solcher Dokumente alltäglich geworden sind. Arbeitgeber und Headhunter sind sich dieses Problems bewusst, weshalb es oft erforderlich ist, dass die Diplome vom Aussteller überprüft werden müssen. Dieser Informationsaustausch ist derzeit manuell und damit zeitaufwendig. Die heutige Technologie eröffnet jedoch Möglichkeiten, diese Hindernisse zu überwinden. Insbesondere die Art und Weise, wie Menschen und Unternehmen miteinander interagieren, wurde durch die Blockchain revolutioniert. Die in der Blockchain gespeicherten Informationen werden verteilt, und für jeden transparent und gleichzeitig manipulationssicher verwaltet. Auf dieser Basis kann eine ganzheitliche Lösung realisiert werden, die die Ausstellung, Speicherung und Verifizierung von Diplomen beinhaltet. Diese Master Arbeit stellt die Konzeption und Implementierung eines End-to-End-Blockchain-basierten Systems zur Verwaltung von Diplomen mit dem Namen UZHBC (University of ZuricH BlockChain) mit der Universität Zürich als Anwedungsfall vor. Durch die Einbezug von Interessensgruppen wird aufgezeigt, welche Anforderung ein solches System erfüllen muss, und wie daraus schlussendlich ein lauffähiger Prototyp entsteht, der die nötigen Funktionalitäten aufweist.

Diplomas have a high importance in society since they serve as official proofs for the education and skills of the recipients. Therefore, it is not surprising that forgeries of such documents have become commonplace. Employers and headhunters are aware of this problem, which is why it is often required that diplomas have to be verified by the issuer. This exchange of information is currently manual and, therefore, time-consuming. However, today's technology opens up opportunities to overcome these obstacles. In particular, blockchain has revolutionized the way in which people and enterprises interact with each other. The information stored in the blockchain is managed in a distributed manner that is transparent for everyone, and at the same time tamper-proof. Based on this, a holistic solution that includes issuance, storage and verification of diplomas can be realized. This master thesis presents the design and implementation of an end-to-end blockchain based system for managing diplomas called UZHBC (University of ZuricH BlockChain), using the University of Zurich as the issuing instance. By involving stakeholders, it is shown which requirements such a system must fulfill and how, in the end, an executable prototype with the necessary functionalities is created.

Acknowledgments

I want to thank my supervisor Bruno Rodrigues for his support and guidance during my thesis. Furthermore, I would like to thank Professor Stiller, who motivated me to participate in the BIS workshop and who prepared me accordingly. I would also like to express my thanks to Eder Scheid and Professor Salil Kanhere, who gave me feedback on the paper submitted to the BIS workshop. Also, I would like to thank all involved stakeholders of the UZH who provided me with the necessary information for the thesis. Finally, I want to thank Michael Furrer for proofreading my thesis.

Contents

Chapter 1

Introduction

In an increasingly competitive market, a diploma from a higher education institution has major relevance in the labor market. Academic certificates are seen as a sign of capability, certifying the level of education and skills of individuals. Globally, enterprises are having difficulties in finding skilled professionals to fill up vacancies [31]. Unfortunately, this has led to an increase in *diploma fraud* which ranges from inflating academic grades to outright fake diplomas. There now exist several diploma mills, *i.e.,* unscrupulous organizations with the sole purpose of providing illegitimate academic degrees and diplomas. The number of individuals "owning" fake credentials globally is hard to estimate. In 2015 the Association of Certified Fraud Examiners [24] estimated that only in the US (United States) about 41% of job applicants presented falsified information about their education. In 2017, it is estimated that about 500 fake doctoral diplomas are sold monthly in the US [27].

Recognition and accreditation systems are commonly used to verify which institutions are recognized (*i.e.,* trusted or reputable) and authorized to award academic or professional qualifications. However, this system is not always effective in countries where the recognized higher education institutions could not meet the demand of certified professionals required by the labor market. This creates a fertile ground for these 'diploma mills' to sell fake credentials to unqualified individuals attempting to take advantage of this shortfall. In this regard, the digitalization of the processes of issuing and verifying diplomas including cryptography primitives to guarantee the identity of the diplomas becomes increasingly essential to ensure that enterprises are recruiting genuinely qualified individuals.

Currently, the majority of degrees is granted in a paper-based format, which can easily be faked and scanned into a digital representation. A survey has also shown that paper-based degrees are being used with decreasing frequency [19]. The documents are primarily used to prove the academic education to a potential employer in the form of an application. Today's labor market is predominantly digitized and applications are exchanged via digital channels [28]. Therefore, the paper-based certificates are scanned. The survey also reveals that graduates want to receive an equivalent digital diploma. However, digital documents can also be forged. The motivation for this work is to achieve that in the future the

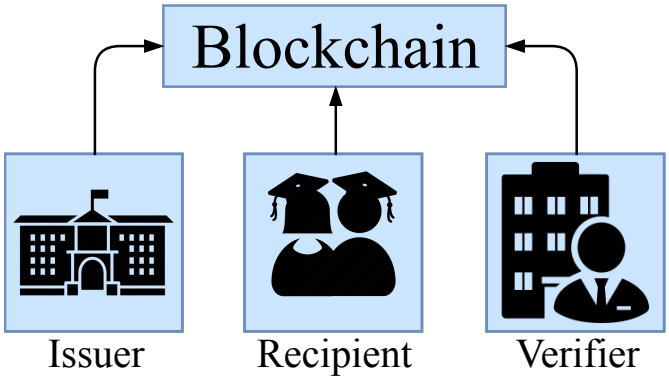

Figure 1.1: Stakeholders

university will hand out digital diplomas and be able to guarantee that these certificates are authentic.

As a counter-measure, many universities implement mechanisms [35] or use services [12] to issue and verify a digital representation of the paper-based diploma. The verification can be automated by including the identity of the diploma into a central database, which can be accessed by a company wishing to verify the credentials. However, this process is rather ad-hoc and there are no unified mechanisms or standards in place such as a public registry that is maintained by multiple institutions and accessible to everyone.

As mentioned in [18, 38], there is not a perfect type of diploma certification. While paper-based diplomas are still seen as the cheaper and safest form of accreditation, it has some drawbacks in contrast to digital-based diplomas. For example, paper-based certificates require more manual tasks for issuing and verifying diplomas than a digital one, and the security of these diplomas are as high as the level and expertise that one has to include security features such as watermarks or invisible fibers. In contrast, digital diplomas are more straightforward to be issued and verified against a central database maintaining these diplomas, and their security relies on available security cryptographic protocols.

Digital diplomas using a centralized database, however, have some drawbacks that blockchain-based approach can overcome. For example, centralized databases are a single point of failure and, using a blockchain (*c.f.,* Figure 1.1), issued diplomas cannot be tampered with as data stored in a block is replicated across the blockchain network. Once blocks are distributed, any party connected to the blockchain can access the stored diplomas, meaning that any verifier in possession of a diploma can quickly verify the authenticity of the diploma. Furthermore, hashes can create a link between the official digital diploma issued by the University and the same diploma held by the recipient to a verifier, which can then check whether the generated hash represents the original diploma.

Recently, there have been works ([3, 6, 11]) on the use of blockchain technology for creating a standardized platform for issuing and verifying diplomas. Thus, the infrastructure maintaining the information related to the diplomas is transparently replicated by the chain of nodes, so that it is not possible to change diplomas issued by previously authorized institutions. This way, only diplomas created by valid issuers are published and the falsification can be omitted. Based on these works, a blockchain based end to end system

is presented in this master thesis, implementing an approach to issue and verify diplomas at the Univerisity of Zurich (UZH).

1.1 Description of Work

This master's thesis demands, in an initial stage, to acquire the conceptual elements that are involved in the process of certification at the UZH. Also, it is expected in the initial stage to provide an overview of the expected system in contrast to the existing certification process and the originating requirements. The expected outcome is that the student understands the necessary background to accomplish the thesis objectives and obtain a global vision of the required activities in order to list and organize these activities in a timeline.

Based on the problem's overview acquired in the first stage, a questionnaire covering needs and requirements regarding desired system's characteristics and usability has to be created. This task will be given to crucial stakeholders from the UZH (i.e., people involved in the process of issuing certificates) and define basis to guide the selection of requirements of the prototype. Whereas this stage demands an interaction with UZH personnel to extract desired characteristics and features, the second stage focuses on the engineering aspect to design a prototype able to fulfill previously listed requirements. Therefore, the student should study and evaluate related works and propose a solution capable of satisfying these requirements. In this regard, design aspects shall be discussed through regular meetings with the advisor to examine the feasibility of the proposal.

The third stage involves the master's thesis development, evaluation through a pilot deployment and documentation. A pilot shall be run within the UZH to pave the path towards certificates digitalization in UZH. The working prototype shall consist of the integration of the designed and implemented elements into the system, producing results to be contrasted with the requirements listed on the questionnaire. Also, the feasibility of the pilot demonstration shall be discussed with the advisor. In this regard, depending on the development complexity of the system, partial requirements could be fulfilled leaving some requirements as future work.

1.2 Thesis Goals

Driven by the outlined work description, the following master's thesis goals are required:

1. **Questionnaire and requirements:** This goal aims to identify relevant aspects such as desired system functionalities and usability, with the personnel involved in the process of issuing certificates in UZH. In order to develop a prototype beneficial not only to educational institutions but also students involved in the process, it is necessary to raise requirements to check which are the essential and desired features of such system, as well as its applicability according to a legal point of view.

2. **Pilot prototype:** Based on listed requirements, a prototype should be developed aiming to pave the initial steps towards digitalization of certificates in UZH. In this regard, the system will serve as a basis for future contributions and improvements towards the process of certificates digitalization in the UZH. It is worth noting that the feasibility of the development shall be discussed with the supervisors so as the thesis be concluded by the deadline.

1.3 Methodology

This master thesis follows several steps towards its achievement. First of all, the relevant processes and the existing problem have to be identified. To achieve this, the most important stakeholders must be consulted. Questions should be asked on how the stakeholders relate to the academic certificates and how they work with them. This allows illustrating the process of issuance and verification of academic certificates at the UZH. From this, the most important requirements can be determined which a system must fulfill to solves the identified problem. With the identified requirements, a prototype with the necessary functionalities can then be developed. Several approaches try to solve the identified problem. Therefore, it should be checked whether elements of related works can be reused in this thesis.

1.4 Thesis Outline

- **Chapter 2** presents the related work that influences this thesis.
- **Chapter 3** describes how the processes at the UZH work and how the requirements for a verification system look like, as well as showing the architecture of such a system.
- **Chapter 4** builds on the previous one. Based on the architecture, it is explained in detail how the prototype is implemented. It introduces the different technologies used and also shows how to keep the cost factor low.
- **Chapter 5** documents a preliminary evaluation, which compares the requirements with the prototype based on it, and whether it can meet the criteria.
- The discussion of the results is presented in **Chapter 6**.
- In the last **chapter 7** a final consideration of the work is made and the possible future work is pointed out.
- **Appendix A and B** document the transcripts of the interviews with the stakeholders and the results of the survey with the students.

Chapter 2

Related Work

In this chapter the related works are listed which have a direct or indirect influence on the master thesis. It must be said, however, that there are other approaches that are not mentioned, as this would unnecessarily lengthen the scope of this chapter, as many related works are very similar. Section 2.1 presents the concept *Proof of Existence*, which serves as a foundation for this thesis.

When blockchain is used for the issuance of diplomas, there is an opportunity not just to verify a degree certificate, but to enrich and add value to the verification ecosystem. In its purest form, a blockchain acts like a shared, replicated, append-only database where participants can depending share, write, access and participate in the validation process [5, 30]. By providing a trustworthy, decentralized, and publicly available data storage, blockchain has become a disruptive technology that has seen interest from many application domains beyond the FinTech (Financial Technology) area. Although the application of blockchains in education is in its infancy, there are many interesting projects (blockchain-based or not) that have explored the possibility of digital diplomas as a countermeasure to faking degrees.

BADGR [8] and Mozilla Open Badges [23], both are existing unified solutions for managing the entire educational history of students by collating all digital certifications acquired by them at different academic institutes and associating it with a single identity. Although these solutions do not use blockchain, they demonstrate how to integrate multiple certifications into a student identity.

The goal of blockchain in the educational area is to create a digital certificate as an automatically verifiable piece of information that can be consulted by third parties through an immutable proof system. According to [18], blockchain can be implemented in two distinct ways in the area of education. While the first requires that diplomas be stored in plain text to create a publicly available database, the second requires that only the hash of a diploma be stored to secure the digital certificate awarded to the student. Therefore, published student data can be seen by anyone, as they do not contain any confidential information. As the diplomas are required to be tamper-proof, using a blockchain as decentralized storage is appropriate.

The first notable use case storing hashes of diplomas is Blockcerts [22], an initiative by the MIT (Massachusetts Institute of Technology) to create an open standard for issuing and verifying credentials on the Bitcoin blockchain. The stored diplomas are accessible via an App termed Blockerts wallet, which enables students to get a verifiable, tamper-proof version of their degree which they can share with employers, schools, family, and friends. Blockcert is seen as an enabler towards digital certificates in the blockchain.

Similarly to the approach of Blockcerts, the National Research and Education Network of Greece (GRNET) [6] are also storing the hashes of diplomas in a blockchain to protect the confidential student data. The goal is to create a system that can verify student diplomas on the Cardano blockchain reducing the manual verification process and cases of fake degrees. However, the GRNET project [6] differs from Blockcerts [22] in the sense that it can store not only hashes of diplomas, but also the entire verification process. Verification requests, successful or unsuccessful proof and the forwarding of the result to its requester are steps that will be stored.

BCDiploma [3], EduCTX [33] and UNIC (University of Nicosia) [34] have started their blockchain-based projects to issue and verify diplomas. BCDiploma and EduCTX share the same goal towards a global certification network of higher academic institutions. However, UNIC aims to digitize and decentralize their internal processes having issued their first academic certificates as a proof of concept.

SAP also provides a solution to guarantee authenticity for digital credentials. With TrueRec [32], any digital credentials - and not only academic certificates - can be verified through a blockchain based approach. Like many others, they also use hashes as digital fingerprints. However, the digital credentials are presented in a specific format that SAP created and so far, only openSAP certificates can be verified. In figure 2.1, the verification process of TrueRec is shown. In general, this approach can be found in all other related and blockchain-based work of academic certification. The institution writes the hash of the credentials, which in this context is often seen as a fingerprint, into the blockchain. The credentials are sent to the owner, who can share them with others, such as an employer. The credentials can then be used to create the same fingerprint that can be found in the blockchain.

The FIT (Fraunhofer Institution for applied information technique) [17] is also planning to realize an open platform for securely issuing and verifying academic credentials through blockchain. However, specific details have not yet been released.

My eQuals is a result of a collaborative initiative involving 46 participating universities in Australia. In contrast to the previous approaches, My eQuals stores the certificates in a central database. This can be associated with the problem that certificates are not guaranteed to be permanent or unchangeable. Also, all certificates from 46 universities can be found in one place and students are required to pay a fee to use this service [12]. Even though My eQuals ensures state of the art protection, a privacy issue might still exist. If an adversary can access the centralized database, all sensitive information is exposed. When it comes to performance, a central database might be superior to a blockchain. However, in the context of academic certification, performance is not necessarily a leading factor. My eQuals shows that there is a need to verify the authenticity of digital diplomas, but the technologies used will not affect this thesis.

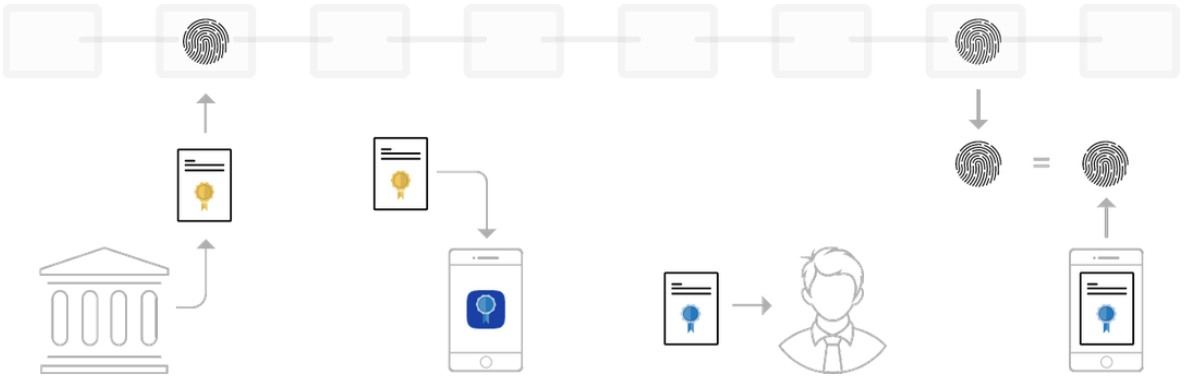

Figure 2.1: TrueRec Verification Process [32]

Although these approaches are already mature, they are either not meeting the requirements of the UZH or not easy to integrate into the structure of a university. Therefore, this thesis shows a prototype that, besides considering these works as starting points, takes into account specific requirements raised from the UZH - for instance, the ease of deployment into their existing IT infrastructure or extending the existing functionality to create diplomas. Nevertheless, the amount of related work that is tackling the problem underscores its necessity.

To guarantee the authenticity of a document, digital signatures can also be used. However, the UZH stated they will not to apply this solution, mainly due to financial reasons. Also, software exists that can bypass those protections and manipulate the content of a document [39].

2.1 Proof of Existence

The underlying concept of verification by blockchain technologies is called *Proof of Existence*. The principle behind this is that the proof of the existence of a document can be published anonymously and securely online. The service stores the cryptographic hash of the file. It is essential with this concept that the actual document is not stored or published anywhere under any circumstances. Therefore, the user does not have to worry about private matters to protect his information [9].

Chapter 3

System Design

This chapter presents the information obtained through interviews with the stakeholders of the UZH. Section 3.1 presents the relevant stakeholders. With the collected information, the current certification scenario is then explained in section 3.2. From this, the requirements can be derived, which are shown in section 3.3 and from which the desired architecture of the prototype can finally be presented in section 3.4.

3.1 Stakeholder Analysis

The University of Zurich consists of seven faculties:

- Economic scientific faculty
- Philosophical scientific faculty
- Law scientific faculty
- Medical scientific faculty
- Natural-mathematic scientific faculty
- Theological scientific faculty
- Veterinary scientific faculty

Each faculty is independent when it comes to the issuance of diplomas. The respective dean's office of the faculty is responsible for this task. Regarding issuance, each faculty can be seen as an independent entity that has its conditions for graduation, which is why an overarching department does not issue the diplomas. Therefore, the **Dean's Office** can be classified as an important stakeholder. The **Diploma Office** is also involved in the issuing process. Their responsibility is to print the documents.

Another relevant entity is the **student administration office**. It is primarily the admissions office for new students. Admission and exchange semesters are regulated by it. Furthermore, it has the task of verifying the authenticity of diplomas.

Since diplomas contain sensitive data, they must be treated with caution. In order not to violate any data security guidelines, the University of Zurich has set up a **data protection delegation**, which is also involved as a stakeholder.

All IT (Information Technology) relevant topics for the UZH are handled by the "Zentrale Informatik" (**ZI**). They offer IT-infrastructure and services for students, professors and all employees of the UZH. Therefore, they are a critical stakeholder in this project. A department at the ZI called BAP (Business Application) is maintaining the IT system in which the diplomas are created. The system generates the certificates as a PDF file that can be further used by the Dean's offices. Furthermore, the diplomas are archived on an internal database.

The verification requests received by the student administration office are mainly sent by headhunter companies. They want to have the diplomas checked for authenticity. For the sake of simplicity, these stakeholders are summarized as one and referred to as **verifiers**.

The last parties to be involved are the **students and graduates**. As these are the recipients of the diplomas, they are directly affected by any changes in the issuing process.

3.2 Scenario Analysis

Two processes have emerged from the discussions with the relevant stakeholders. On the one hand, there is the process of issuing the diplomas and on the other hand, the verification.

The issuance process involves several steps and different aspects must be considered. At each promotion deadline, the dean's office of the responsible faculty checks if the students fulfill all requirements for the promotion. The legacy system, provided by the ZI, delivers different kind of information related to the students. It is checked whether the necessary modules have been completed, whether the tuition fees have been paid, and other faculty-dependent factors. If the audit of the degree has been done, the dean's office creates the PDF files of the diplomas in the system. They can be seen as digital counterparts to the physical documents that are sent to the graduates. They are then handed over to the diploma office that prints and returns the documents to the dean's office. Afterwards, the PDF files are archived. If a previously issued diploma has to be replaced, *e.g.* in the event of its destruction, the archived document can be delivered for a fee of CHF 100. To underline the authenticity of the diploma, the UZH signs the document with a faculty-specific stamp. The issuance process is not only carried out for an individual but for all those who have registered to graduate. The diplomas are therefore handled in batches. The finance department of the UZH publishes annual statistics, including figures on graduates [2]. In 2017, there were 5777 graduates. The majority were students of the faculty of philosophy. With a share of about 36%, this faculty issues by far the most diplomas each year. The specific amount of graduates is depicted in figure 3.1. Before the

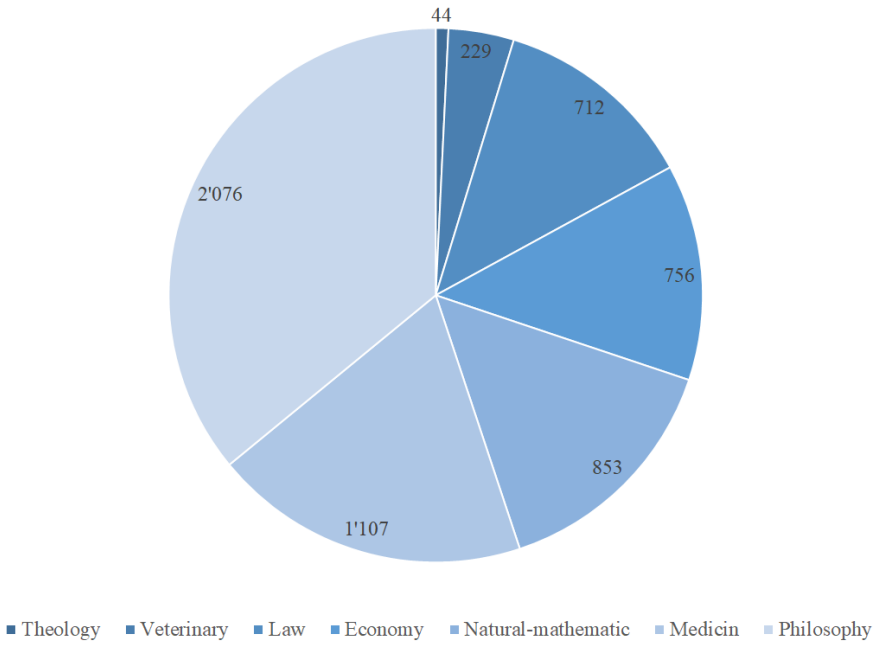

Figure 3.1: Graduates of the Faculties

diplomas can be handed over to the graduate, they are first rechecked in the dean's office and specially packaged. The activity diagram of the issuance process is shown in figure 3.2.

The student administration office verifies diplomas. Employers, headhunters and background check companies request the verification of degrees. To issue this information to the requester, the UZH needs to have the students' consent. If the requester will not deliver this authorization, it has to be obtained by the UZH. After the permission is granted, the UZH responsible has to look for the particular student in the database of the UZH. The requester will then be informed about the authenticity of the questioned certificate. Verifications are requested on an irregular basis. There are up to four requests per day but at least two per week. In 2017, 270 verification requests were answered. As stated in the interview with the UZH responsible, there have been several cases where the diplomas were fake. Either the grades were exaggerated, or the full diploma was manipulated. From the requesters' point of view, it is not explicitly described which department of the UZH they have to contact. This circumstance can lead to the problem that the request is not sent to the one responsible.

3.3 Questionnaire and Requirements

No specific questionnaire was prepared for the stakeholder survey. The fundamental question was how stakeholders relate to the diplomas and how they work with them. From this question, a conversation arose from which one got the necessary information. Table 3.1 presents the requirements derived from interviews with stakeholders which can be found in the appendix. This includes the student administration office that is responsible for

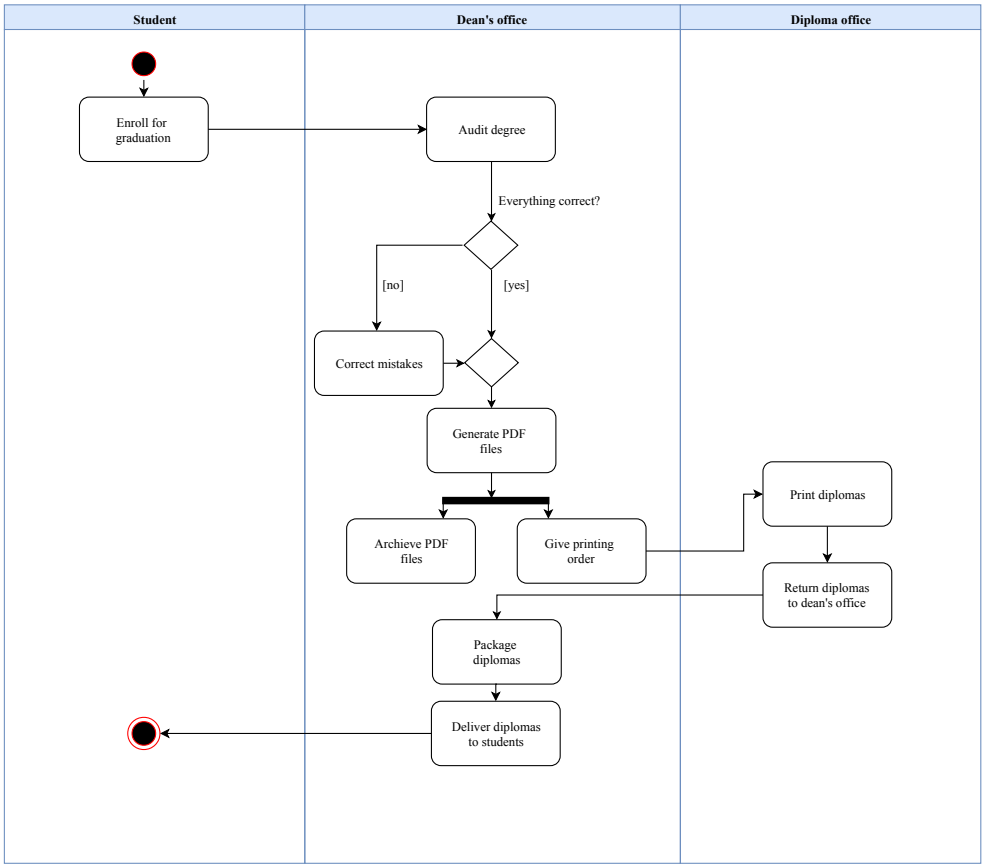

Figure 3.2: Current Issuance Process

the verification. Also, the faculty of economic science, which issues the diplomas for all economics students, was questioned. To not violate any legal aspects, the data protection delegation of the UZH was interviewed. For all IT relevant topics, the UZH employs the ZI, who provides IT infrastructure, software, and services for students and employees of the UZH (herein termed legacy system). While RQ (Requirements) 1-4 are related to the issuer, *i.e.,* conditions that UZH demands from the system, RQ5-6 are related to the requirements for a company that wants to verify diplomas. The most frequent requesters of verifications by the student administration office were background check companies. Finally, RQ7 is related to the delivery of the diploma in a digital form to the student.

- **RQ1**: related to the guarantee that diplomas can only be issued by authorized issuing instances, for example, UZH faculties. Thus, diploma mills are not able to fabricate any certificates. For the verifiers, it is important to be ensured that the diplomas can only be issued by an authorized institution.
- **RQ2**: addresses the confidentiality of student data, which should only be accessible by the student and potential verifiers. Also, the 'right to be forgotten' defined in the new GDPR (General Data Protection Regulation) declares that data of consumer (*i.e.,* students) cannot be permanently stored [29]. Hence, the diploma itself cannot be stored in the blockchain. The blockchain should therefore store a hash of the document in order to prove the authenticity of the digital diploma sent to the student.

Table 3.1: Requirements elicited during Interviews with Stakeholders

Issuer	
RQ1	Only authorized UZH departments are allowed to issue diplomas
RQ2	Diploma data should be confidential to its recipients
RQ3	Process of issuing and verifying diplomas should abstract technical complexities
RQ4	Multiple diplomas should be processable in batch
Verifier	
RQ5	Verification capabilities should be accessible to any company
RQ6	Diplomas should be verified autonomously
Recipient	
RQ7	Graduates should receive their diplomas in a digital format

- **RQ3**: defines that technical details involved in the process of issuing diplomas must remain transparent to involved users (issuers, verifiers and recipients). In this sense, the use of blockchain (or any other infrastructure) for issuing or verifying diplomas should not require technical know-how from the users (*e.g.,* extracting the hash of an academic certificate at the verification process). The system should simplify processes for all stakeholders and save effort in order to increase their acceptance. To achieve this, the new system must not require more resources that already exist.

- **RQ4**: relates to the system scalability concerning the ease to create and verify multiple diplomas at once, as in a bash service. The goal is to avoid the manual exchange of information between companies wishing to verify degrees and the university as an issuer instance. Since the dean's offices handle the diplomas in batch, this should also be addressed in the system.

- **RQ5**: allows anyone in possession of a diploma to verify its authenticity. As any company that receives a diploma from a graduate might want to verify its authenticity, this functionality has to be publicly accessible. The openness of the verification must not violate the integrity of the system.

- **RQ6**: describes an always available service with an automated response of the verification. If the diploma is authentic, the system has to recognize it, whereas tampered documents need to be rejected. The automated verification process is intended to replace the time-consuming manual exchange of information between issuer and requester and at the same time deliver reliable results.

- **RQ7**: graduates shall receive their diplomas in a digital format. Physical diplomas can get lost or damaged, whereas digital diplomas are not affected by these problems. In addition, forgery of physical documents is generally easier. An own survey has shown that graduates in Switzerland receive their diploma in paper format. They mainly use it for the application process, for which they need the document in a digital form to be able to share it through a digital channel. To do this, the graduates scan the paper. In the survey, respondents also indicated that a digital equivalent would be welcome. The specific figures can be found in appendix B.

3.4 Architecture

Figure 3.3 is divided into three different parts. The first covers issuer requirements and the second covers recipient (graduate student) requirements. The third is related to companies wishing to verify a diploma sent by recipients. On the UZH side, the issuing instance, the system is embedded into the legacy system, taking as input diplomas in a digital form (PDF files). Currently, these digital diplomas are not sent to students but used to print paper-based diplomas which are then granted to graduate students.

3.4.1 Hashes

A hash is the output of a hash function that expects an input value - in this case, PDF documents - and generates an output value in the form of a string of fixed length. The main feature of hash functions is that it is almost impossible to find two different input values that generate the same hash value [16]. The hash function used in this approach is SHA-3 with a length of 256 bits. SHA-3, unlike MD5, is considered collision resistant, which means that the chance that two different input values produce different output values is very high. Hashes can be used to prove the authenticity of software artifacts. In this case, one speaks of checksums. To inform a user about the authenticity of downloaded software, companies often highlight the checksum on their website. The software can also generate a checksum which has to match with the checksum from the website. The checksum functionality can also be used for diplomas. If someone makes even the smallest changes to the document, the hash will change completely. SHA-3 is a one-way function, which means that it is not possible to recreate the input from the output. This property and also the uniqueness of hashes make it possible to encrypt diplomas without revealing confidential information. No one can interpret the content of the diploma with the resulting hash, but it can be regarded as a unique link pointing to the official certificate.

3.4.2 System Overview

In the first step, the issuing institution has to create the digital diploma, which is part of the UZH legacy system workflow. The generated digital diploma is sent to the recipient, the graduates. Furthermore, the document is needed as input for the new system, because in the second step a unique hash is generated from it.

This hash will be stored in a smart contract that is deployed on the Ethereum blockchain. Since nobody should change or even delete the hashes of the diplomas, which are authentic and valid for life, the use of a blockchain is recommended. Due to its immutability, this requirement can be met.

A verifier company that receives the diploma from a student could then verify the authenticity of the document without contacting the university. Therefore, the verifier can use the UZHBC front-end that takes the digital diploma as an input, generates the corresponding hash and checks the authenticity of it. In order to do this, the generated hash

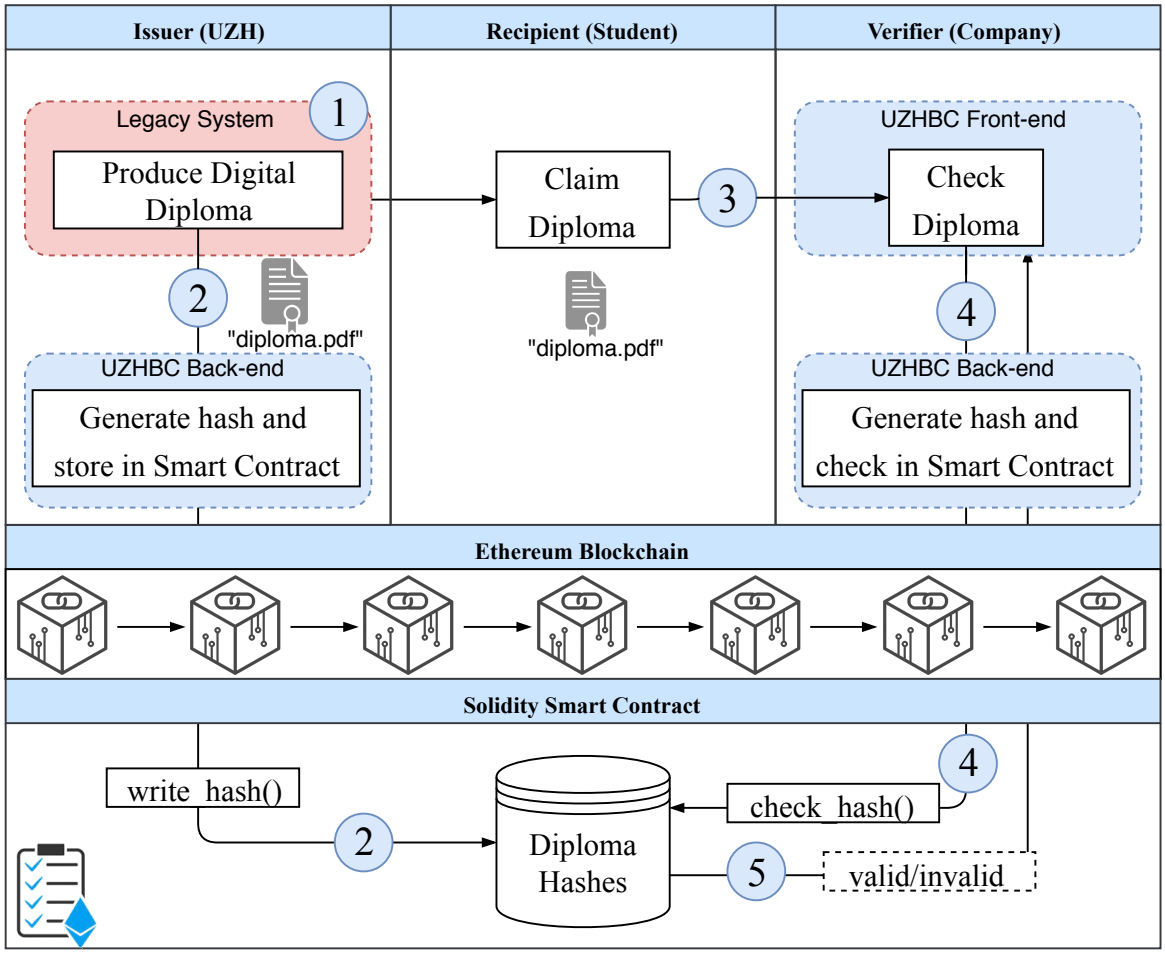

Figure 3.3: UZHBC System Architecture

will be compared with all hashes that are contained in the smart contract. If it exists, it means that the university has previously stored the same hash in the smart contract that belongs to the official diploma. In this case, the user is informed about the authenticity. If no match is found, this means that either the complete document has not been issued by the UZH or changes have been made to the document, resulting in a different hash. In both cases, it can be assumed that the document is not authentic.

Chapter 4

Implementation

This chapter shows the realization of the prototype. In section 4.1, essential implementation details are discussed, and insights into the interface of the prototype are shown. Section 4.2 explains step by step which methods can be used to achieve a cost-efficient interaction with the blockchain.

4.1 Implementation Details

4.1.1 Client

```
let txParams = {
  nonce: this.state.txCount + 1,
  gasLimit: this.state.minimumGas,
  to: this.state.contractAddress,
  from: this.state.owner,
  value: "0x0",
  data: rawTxData,
  gas: this.state.minimumGas,
  gasPrice: web3.utils.toHex(5000000000)
};
```

Listing 4.1: Raw Transaction Parameters

Listing 4.1 shows how the transaction is structured on the client side. Web3 prescribes the used parameters in transactions. The nonce indicates how many transactions have already been sent to this smart contract and must always be incremented by one. The gas limit must be specified, which describes the maximum amount of gas the transaction can consume. The recipient must also be defined, which in this case is always the smart contract. The sender's address is the owner, who is the only user that can write to the smart contract. In the value tag, the amount of Ether to be transferred would be specified. However, this is not necessary in this use case, since no funds have to be transferred and can therefore it can be left to zero. The hashes are specified in the data tag, which can be passed to the smart contract by calling the corresponding method, which is client-side

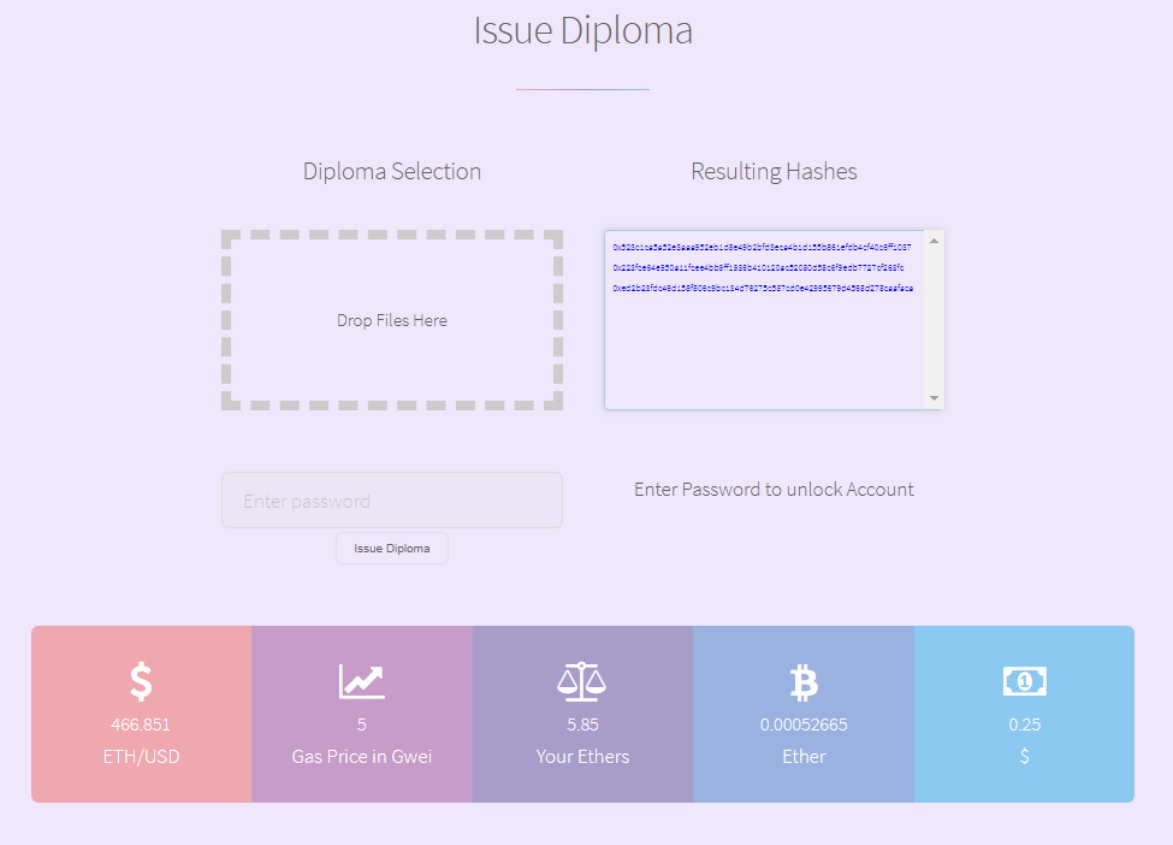

Figure 4.1: Interface to issue diplomas

encoded by the Application Binary Interface (ABI). The amount of gas that one wants to spend effectively is indicated in the gas tag. This value can be calculated and is the same as the gas limit. The last tag, the gas price, reveals how much one is willing to pay for a unit of gas. A higher gas price leads to faster execution of the transaction. More about costs is discussed in section 4.2.

The client is responsible for the interaction between user and system. A simple web-based application enables access to the corresponding functionalities. Only two input fields are needed to store the diplomas in the blockchain and to check if the document is authentic. For the first input field, which is used for the issuance process, the hashes resulting from the diplomas are displayed in a list. At the same time, a password is required to keep access to issuance functionality regulated. The interface is shown in figure 4.1. Furthermore, information regarding costs that would arise if transactions were made on the Ethereum network is presented. After loading the documents into the input field, the hashes with the size of 256 bits are generated [20]. These are then sent to the smart contract after the transaction has been successfully signed. To allow the client to interact with the blockchain, the library web3 [13] is used. Furthermore, the client must be connected to a synchronized Geth terminal, which is responsible for the communication between the different participants of the blockchain. An additional interface is provided for verification, shown in figure 4.2. This again provides an input field which generates the hash from the document and searches the blockchain for this hash. If the hash is

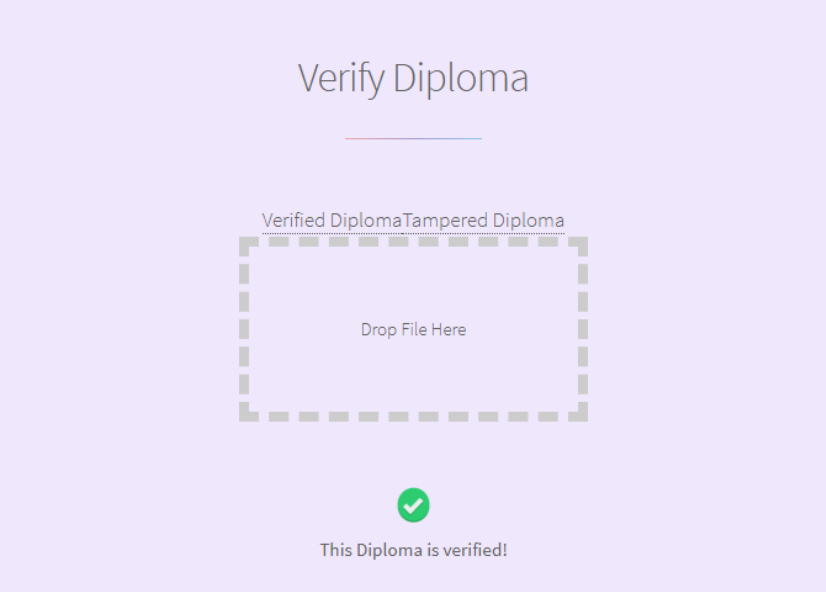

Figure 4.2: Interface to verify diplomas

found successfully, this is indicated accordingly via the user interface.

4.1.2 Blockchain Syncing

Since the blockchain has to be distributed over the entire network, suitable software is required. As mentioned previously, in this approach, Geth is used. Geth is the official implementation of the Ethereum protocol written in Go. To interact with the blockchain, transfer funds, create accounts or deploy smart contracts, the user needs a continuously updated version of the blockchain. Geth is used for this purpose. However, many users do not need the entire blockchain that contains years old data, which is why it is possible to get a shortened version through the fast sync mode. In addition to different synchronization modes, Geth offers many other parameters that allow a customized configuration on the individual clients. A disadvantage of Geth in the context of a distributed application like the UZHBC is that it has to synchronize constantly if someone wants to use the app. This requires a decent bandwidth on the one hand and considerable computing power on the other. To address this problem, services have been developed to outsource the syncing of the blockchain. One of the better-known providers of such a service is MetaMask [21], which serves as a wallet that can also send transactions to smart contracts. MetaMask is a browser extension that runs in the background of the application and automatically detects when a transaction is sent to the blockchain. It turns out, however, that it is not sufficiently developed yet to replace a Geth synchronization. For example, it is not possible to send multiple transactions at once, which is necessary for this approach. Furthermore, the use of a centralized service like MetaMask would be against the principle of a blockchain, which should provide advantages through its distributed nature. Therefore, Geth is used for this prototype. In the future, one could consider whether to solve this via a cloud service.

4.1.3 Smart Contract

At the moment, the prototype uses a public test network called Rinkeby. For test purposes and cost reasons, this is very suitable, since it behaves similar to the public Ethereum network but at the same time does not require any financial costs. The smart contract contains only two functions. The write function `issueCertificate` is responsible for storing the hashes in the smart contract. Also, it is only possible to call this function as an owner of the contract, which is the university. Transactions to the Ethereum network are associated with costs. How these costs are kept as low as possible is shown in the next section. It can also be addressed in the smart contract. In the first version of the smart contract, a single hash was passed as a string. Since operations on strings are relatively expensive, the hash was passed as a byte in the second version. In the last and most current version, a byte array of hashes is passed. Solidity offers different data types of bytes, from bytes to bytes32, which has a storage capacity of 32 bytes [14]. Since the SHA-3 is used with 256 bits hash function, a hash with 256 bits is generated, i.e., 32 bytes. Therefore, it is most efficient to use Solidity's bytes32 data type to avoid wasting storage space. Furthermore, it is only possible to pass bytes or integer arrays as parameters. Therefore, the use of strings in this approach is unfavorable and costly. The `issueCertificate` function can be found in listing 4.2 between line eight and 13.

```
1  pragma solidity ^0.4.18;
2  contract UZHBC {
3
4      address public owner = msg.sender;
5
6      bytes32 [] public diplomaHashes;
7
8    function issueCertificate(bytes32 [] byteHash) public {
9        if (msg.sender != owner)
10             revert();
11         for(uint i=0;i<byteHash.length;i++)
12           diplomaHashes.push(byteHash[i]);
13   }
14
15
16 function verifyCertificate(bytes32 byteHash)public constant returns(bool
       ){
17     uint counter = 0;
18     bool verified = false;
19     while(counter<diplomaHashes.length){
20         if(diplomaHashes[counter]==byteHash){
21             verified = true;
22             return verified;
23         }else{
24             counter++;
25         }
26     }
27     return verified;
28   }
29 }
```

Listing 4.2: UZHBC Smart Contract

The hashes are stored in an array of bytes32. When a verification request occurs, the `verifyCertificate` method will iterate through the array of hashes. If any hash in the array matches with the given hash from the parameter, the method returns `true` as an indicator of a verified diploma. The dedicated function is depicted in listing 4.2 between line 16 and 28. The university issues about 6000 diplomas annually which require a combined storage space of 192 kilobytes. This is repeated every year and the required storage space is constantly increasing. If the smart contract reaches a point where the storage space is used up and it can no longer save diplomas, a new smart contract has to be created, because a deletion in a blockchain is not possible. However, there is a storage capacity of 2^{261} bytes, which is such a huge amount that it would take the UZH thousands of years to use it up. It would be more likely that a hash collision would occur first [37]. The verification functionality in the smart contract is called whenever someone uses the verification interface. In figure 4.1, the associated costs are shown, but not in figure 4.2 for verification. The reason for this is that all write operations on the blockchain are associated with costs, but all read operations remain free.

4.2 Performance Enhancements

4.2.1 Transaction Fee

Costs in Ethereum are calculated with gas. Gas can be seen as a type of fuel required to perform transactions on the blockchain. Generally speaking, they are transaction fees that the miners receive as a reward. The costs of a transaction consist of low-level operations, so-called opcodes, which are weighted differently. A detailed overview of the opcode costs can be found under [10]. Beside the gas used by the opcodes, there is always a base fee of 21'000 gas, which is needed for the execution of the transaction. In other words, a transfer of funds for which there is no interaction with a smart contract costs 21'000 gas. If there is interaction with a contract, the total gas costs result from the basic fee and opcode costs. The total gas required for the transaction is called gas limit. With web3, the gas limit has to be specified to execute a transaction. To little specified gas leads to the rejection of the transaction. If the gas limit is above the required gas, the remaining gas will be refunded. To finally calculate the transaction fee, a second measure is required, the gas price. This specifies how much Ether a unit of gas is going to cost. Ether is the cryptocurrency resulting from the Ethereum network. The gas price is determined by the user who executes the transaction. It describes the amount of Ether per gas that someone is willing to pay. The gas price is normally given in Gwei, where 1 Gwei corresponds to 0.000000001 Ether. There are three different gas price categories: SafeLow, Standard and Fast. The higher the specified gas price, the faster the transaction will be mined into the blockchain. SafeLow indicates the minimum possible gas price. A transaction will be mined after approximately 30 minutes. If the gas price is in the fast category, a duration of two minutes can be expected. Through the gas price and the gas limit, the transaction fees can now be calculated by multiplying the two values.

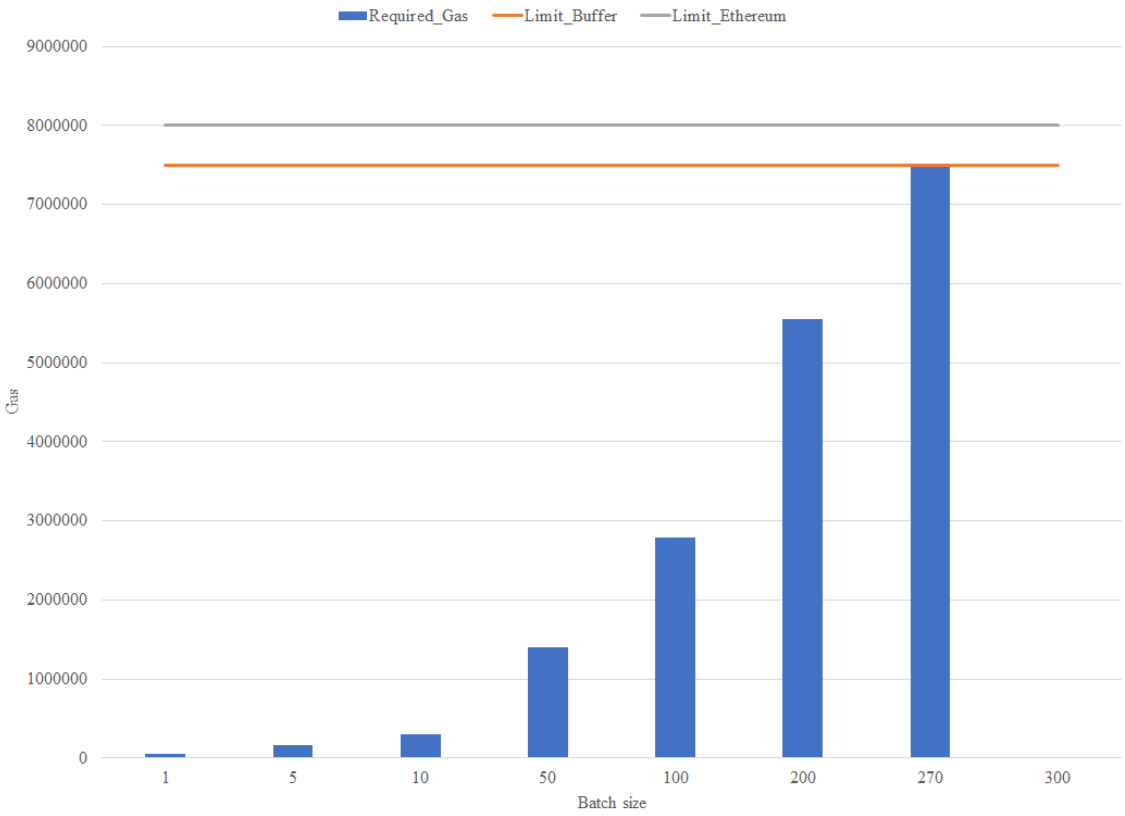

Figure 4.3: Number of diplomas per batch

4.2.2 Single Transactions vs. Batch

As mentioned in section 3.5.2, the smart contract expects an array containing multiple hashes of diplomas, not just a single hash per document. It is cheaper to have as few transactions as possible. The base fee previously mentioned will be credited to each transaction. One transaction would require 21'000 gas, 100 transactions would already require 2'100'000 gas as a base fee. However, since the UZH issues about 6000 diplomas annually, a huge base fee would arise if a transaction only contained one diploma hash. Therefore, as many hashes as possible are put into one transaction. This can significantly reduce the number of transactions. The size of the block limits the maximum amount of gas that a transaction can contain. It is possible that a block consists of only one large transaction, which consumes so much gas that there is no room for further transactions. This can be achieved by putting as many hashes as possible into a batch. First, however, the block size must be determined. In the Ethereum blockchain, the block size varies from block to block. However, there is a trend towards a growing block size. While in January 2017 a gas limit of 4 million gas per block was set, one year later, 8 million gas per block [15] was possible. In this approach, an attempt is made to get as close as possible to this limit. However, it should also be noted that it is possible that the limit may be below the 8 million mark, as it is only an average value. For safety reasons, a buffer of 500'000 gas is included, which can be removed without great effort. Figure 4.3 shows how to gradually reach this limit. A single hash transaction requires about 50'000 gas. A transaction with five hashes requires about 160'000 gas. One could assume that gas costs have to be five

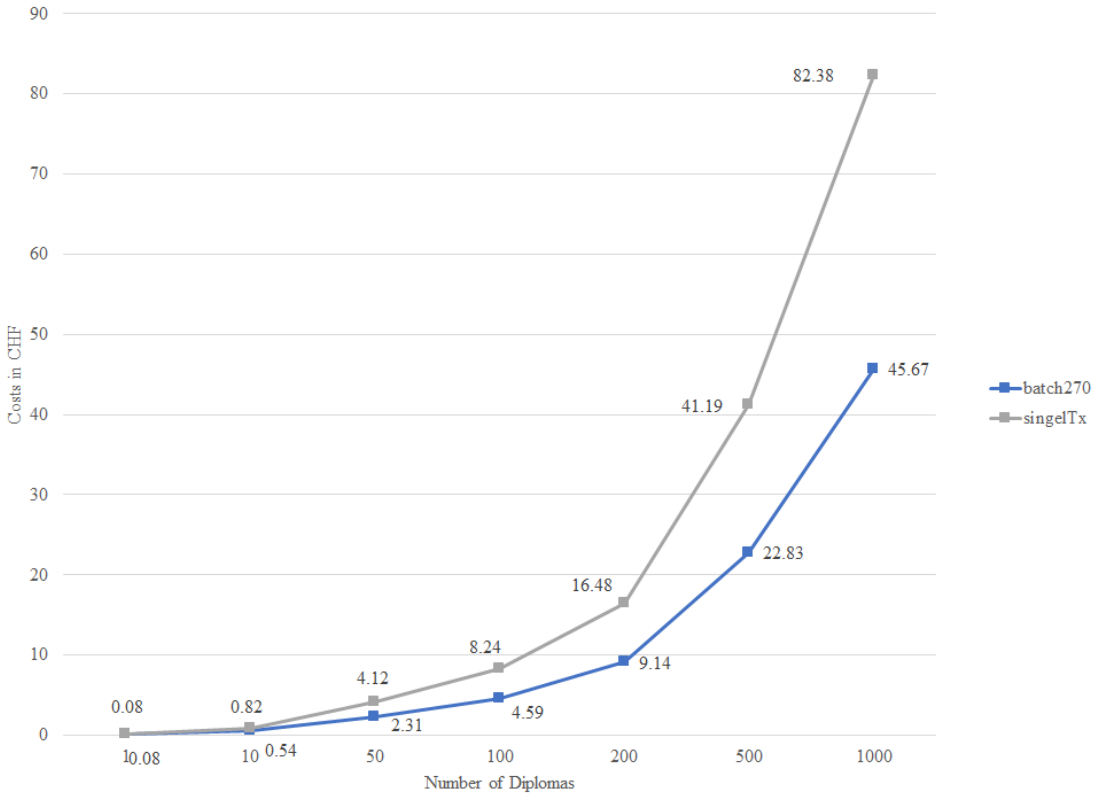

Figure 4.4: Total Transaction Costs

times higher, however, the base fee is not included five times but only once. By increasing the batch size continuously, the limit will be reached. This is the case for 270 diplomas, which require approximately this amount of gas. Surpassing this limit leads to a rejection of the transaction since the capacity of a block would be exceeded.

4.2.3 Costs

The sections discussed above make it possible to calculate the actual costs incurred. It was shown how costs per se arise, how these can be reduced in this approach and what quantity of gas is required. The first parameter to be set is the gas price. A higher gas price shortens the waiting time until a transaction is mined. In the environment of diplomas, the waiting time does not play a significant role. It takes about one to two months from the date of the promotion to the award of the diplomas. When it comes to mining, we are in a much shorter period. At the lowest possible gas price, the waiting time is around 30 minutes, which is very short compared to the waiting time for the physical document. As can be seen in [1], the SafeLow gas price varies from time to time. The average value is approximately 5 Gwei and is therefore used for this approach. Also, the gas required for the transactions can be calculated with the help of web3. By multiplying these two metrics, the gas costs are determined. To bring these in the context of fiat money, the conversion rates between cryptocurrency and fiat money must be used. There are several API's that offer this service, whereby CoinMarketCap [7] is used here.

However, the resulting currency is in dollars, which is why the exchange rate between dollar and Swiss franc must still be considered. The total costs in Swiss francs are shown in figure 4.4. The exchange rates used were taken up on 02.08.2018. For 1000 diplomas in a batch of 270, there will be costs of 45 CHF. Since the UZH has around 6'000 graduates per year, there would be costs in the area of CHF 270. Furthermore, the figure also shows the costs incurred if only one diploma was used per transaction. As is evident, the option with the batches is much more efficient. Also, it is possible that a batch may contain even more hashes in the future, as there is a trend towards growing blocks in the Ethereum network.

Chapter 5

Evaluation

This chapter discusses the evaluation of the UZHBC prototype. An analysis is conducted to verify whether the prototype can satisfy the requirements identified in section 3.3. Likewise, the fulfillment of the requirements by the related work was analyzed and compared against the prototype. This comparison is presented in Table 5.1.

Table 5.1: UZHBC and Related Work on Requirements

	Blockcerts	GRNET	EduCTX	UNIC	BCDiploma	BADGR	UZHBC
RQ1	✗	✓*	✓*	✓	✗	✗	✓
RQ2	✓	✓	✗	✓	✓*	✗	✓
RQ3	✗	✗	✗	✗	✗	✗	✓
RQ4	✓	✓	✗	✓	✓	✗	✓
RQ5	✓*	✓*	✓*	✓	✓	✓	✓
RQ6	✓	✓	✓	✓	✓	✓	✓
RQ7	✓	✓	✓	✓	✓	✓	✓

Note: * indicates that the requirements have been partially met.

The UZH consists of seven faculties whereas each faculty includes many departments. In UZHBC, each of these faculties would represent an issuing instance able to record diploma hashes into the blockchain. Other blockchain-based approaches such as Blockcerts [22] and BCDiploma [3] extended the number of issuers in their works. For example, new issuing institutions can register themselves on the platform which could work as a universal diploma verifier. However, at some point, new issuers would have to prove their ability to certify degrees to the developers of the platform. This dependency between developers and issuers cannot be neglected, and a fully automated process cannot be achieved. The most critical issue is that this prototype is intended to solve the falsification of diplomas through individuals or diploma mills. Therefore, granting issuing rights needs to be strictly regulated and the ability to add issuers is not desired. However, it is important to note that the respective UZH faculties are acting independently. The requirements for graduation, deadlines to be met and the entire process of graduating are different at each faculty. Thus, a faculty has to be considered as an autonomous entity with respect to the issuance of diplomas.

The provided UZHBC functionalities achieve the requirement (RQ1) as presented in listing 3.1, which shows that writing access is only granted to the actual owner of the smart contract, the UZH. Similarly to Blockcerts and BCdiploma, GRNET [6] and EduCTX [33] allow multiple issuers. Nevertheless, write permissions are not readily granted. While GRNET consists of a group of predefined universities as issuers, new universities at EduCTX should be selected by the existing participants.

Regarding the RQ2, a hash generated through a one-way function is used to represent the diploma. By only recording the hash, one is not able to identify confidential data about the actual content of the diploma. To verify the authenticity of a degree, a verifier needs an actual diploma document sent by a student (*e.g.*, in a job application). The provided functionality for verification generates another hash, and if this hash is already contained in the smart contract, it can be considered as authentic. Issued hashes are publicly available without compromising the confidentiality of its owner.

As depicted in table 5.1 many approaches also use cryptographic hashes. BCDiploma [3] stores encrypted diploma data and claims to solve the problem of the new GDPR "right to be forgotten" [29]. Diplomas can be decrypted through a persistence key, which is unique and kept by the owner of the degree. However, losing this key implies that the diploma cannot be retrieved anymore and encrypted data would remain on the blockchain.

Intense acceptance and usability test scores with the university and verifiers are required to gain more insights concerning system practicability. However, the amount of interaction with the system, which can be seen as the actual additional effort, requires fewer interactions in contrast to other approaches. This includes sending invitations or transaction addresses, registration, maintaining a hash list and more. Therefore, it must be stated that comparing different approaches is not straightforward since these are slightly different concerning their functionalities. As UZHBC currently offers two interaction possibilities (recording and verifying), complexity is reduced to a minimum (RQ3). For example, these functionalities are translated into a simple action at the front-end, such as dragging a file into a field.

The UZHBC can verify the authenticity of diplomas without relying on manual intervention by the university. However, it requires some additional steps to achieve this. At the moment where the paper-based diplomas are delivered to graduates, the digital equivalents have to be processed into the system. The extra effort can be limited since the prototype allows to prepare as many documents as desired. As the UZH handles diplomas in batches (for printing), it is also feasible to use the prototype and RQ4 can be met.

As confidential data is not disclosed in the verification process, the front-end interface can be publicly accessible (RQ5). Other approaches (*e.g.,* Blockcerts [22] and EduCTX [33]) use invitation mechanisms, where the graduate sends a link to his academic credentials. With UZHBC (and UNIC [34]), this invitation is handled by sending the digital diploma in a job application. The interface of verification is accessible to everyone, but without a diploma, it is useless. It is important that the awareness of such a system needs to be spread, so employers know where to verify the received diplomas.

Background-check companies, headhunters and also regular companies are the typical entities that need to verify student diplomas. This task is seen as rather time-consuming as

there is currently no automated verification system for diplomas in Switzerland. Thus, the process relies on the manual interaction between the employer, university, and graduate. Nonetheless, universities in Switzerland are not allowed to send any information without the graduate's consent. Thus, verification requests are rather time-consuming. Based on UZHBC, verifiers are only required to submit the received digital diploma to the front-end verification provided by UZH. Therefore, the hash will be generated again and checked at the blockchain whether it is authentic or not, fulfilling the initial requirement from the employers (RQ6).

From the perspective of the recipients, *i.e.,* graduate students, digital diplomas would be granted in addition to the conventional paper-based diplomas. At the moment, these are obtained by scanning the paper-based document to have a digital equivalent. To fulfill RQ7, the UZH will deliver, through the UZHBC system, these documents added to the paper-based diplomas. However, this requirement relies on the cooperation of the university concerning its internal regimentations. This is also a prerequisite for the other related work, as all the academic credentials are handled digitally.

Chapter 6

Discussion

This chapter first discusses the results of the evaluation and their implications for this master thesis, detailed in section 6.1. Afterwards, section 6.2 explains several limitations of the prototype.

6.1 Evaluation Results

The first requirement was the need for no unauthorized persons to have write access to the smart contract. As mentioned in the introduction, the blockchain is characterized by its immutability, which means that it is very resistant to attacks that attempt to change content. Therefore, the decision to use this technology has already partly led to the fulfillment of this requirement. To have write access, one must be in possession of the private key, which must be held locally. In the case of local storage, the access rights are theoretically secured, but the UZH is ultimately responsible for this. If it is decided that more institutions (apart from the UZH) must have access to the smart contract, the private key would only have to be passed on to the relevant parties.

The SHA3 hash function has so far proven to be safe, which means that no two different inputs resulted in the same output. At the same time, it is not possible to generate real information from a hash. By using this hash function, it is possible to meet the second requirement which states that no confidential information may be published.

To optimize processes, the new method cannot require more effort than already exists. However, the automation of the verification process should reduce the effort of all parties involved. The effort for verification by the UZH would be eliminated, and the verification requests from employers would also be significantly reduced. However, it should be noted that part of the work is transferred to the dean's office, which must still store the diplomas in the blockchain during the issuing process. Whether there will be less effort for all the parties involved is not clear at this stage. The architecture and implementation of the prototype aimed to provide a very pragmatic solution that requires as little user interaction as possible. To know whether the third requirement has been met, users and acceptance tests must be carried out in advance.

Processing diplomas in batches during the issuing process was not only a requirement on behalf of the UZH but also helps to save transaction costs on the blockchain. With the prototype, it is possible to load several hundred PDF files into the input field. The maximum number of files varies from device to device, as it depends on the browser's memory. Another possibility would be to allow compressed files, for example in ZIP format, as input. To do this, however, an unzipping algorithm would have to be implemented. There was also the question of what would happen to the diplomas that were issued and handed over to the graduates many years ago. The UZH archives all diplomas. If a diploma is lost or damaged, it can be ordered again. This archive can also be used as input to capture degrees in the blockchain. However, the digital diplomas would still have to be handed over to the graduates.

The presented interface in the prototype shows the issuance and verification process on the same page. Of course, within a running system, these two functionalities would be strictly separated, most likely on different URLs. For this master thesis, however, they were presented in the same system for simplification. The automated verification process is straightforward: only the diploma has to be handed over to the input, the hash is generated and searched for in the smart contract. Moreover, the read operations on the Ethereum blockchain do not produce any costs, which is beneficial to everyone. RQ5 and RQ6, which require that the verification functionalities must be publicly accessible and automated, can thus be fulfilled.

Diplomas should be made available to graduates in digital format. In the age of the paperless office, this makes sense. A small survey has also shown that this would be desirable for graduates in Switzerland. However, this cannot be achieved with the prototype, but it is a prerequisite for the approach to succeed.

6.2 Limitations

The interviews are, strictly speaking, unofficially. The UZH has never taken any measures to develop such a system. Therefore, the requirements found in this thesis are not necessarily justifiable by the UZH, which is why the interviewees remain anonymous. The requirements were derived from the interviews, whereby a certain degree of interpretation on the part of the author must also be considered.

The smart contract was deployed for the purpose of development on the test network Rinkeby. Rinkeby is based on the Proof-of-Authority (PoA) consensus mechanism [26] which promises advantages over the Proof-of-Work (PoW) consensus mechanism [25] regarding speed. The thesis talks about using the public Ethereum blockchain, which is based on PoW. If speed becomes a critical factor, it may be worth switching to PoA.

Chapter 7

Final Considerations

The digitalization of the processes within the UZH for issuing and verifying diplomas including cryptography primitives to ensure the identity of the diplomas becomes an increasing necessity. Graduates are expecting diplomas to be issued digitally in the future. At the same time, however, it must be ensured that they are authentic. This master thesis has achieved its general objective by presenting the elicitation of requirements, the design of an architecture and implementation of a prototype which is tailored to the needs of the UZH.

To set up such a system, the involved stakeholders first had to be questioned in order to understand how the processes currently function. The next step was to determine the requirements of the university's stakeholders to create an initial prototype which demonstrates the functionalities and highlights advantages. The requirements resulted from interviews with the stakeholders, who are dealing with diplomas in some way. They are rather high-level requirements and do not go into technical details. From this, the existing processes, as well as the bottleneck, could finally be identified and a prototype can be built upon.

However, the thesis was accompanied by certain difficulties. As already mentioned in the limitation, information from the interviews had to be partly interpreted by myself in order to receive a requirement. Although the interviewees always gave positive feedback on the thesis after the discussions, they had little incentive to show more participation. Understandably, they could not be asked to contribute more, since it was only a study project and not launched by the authorities of the UZH. This could also result in the fact that not all necessary requirements were identified.

The implementation turned out to be quite feasible, which is why many related works took similar approaches to protect the authenticity of diplomas and underline their necessity. The prerequisite for such a system to work would be that the UZH issued diplomas in a digital format. Firstly, this would require an adjustment of the regulations and as such, an amendment cannot be implemented immediately. Furthermore, the UZH would have to agree to use such a system at all. This would be a way to verify the authenticity of digital diplomas automatically. Naturally, the UZH could dispense with such a system, but the effort of verification will then continue.

Summarized, the thesis shows how an operating system can be created by targeted steps, which enables the verification of digital documents, and at the same time, can stand out from conventional systems through the use of blockchain technology. Due to the growing demand for such systems, it is only a matter of time before the UZH will make use of it. In this case, an initial basis would have been established with this thesis.

7.1 Future Work

Since the developed system is only a prototype, several improvements are foreseen for future work. For instance, the Master thesis started on my initiative, without a specific assignment from the UZH, which would be the primary stakeholder in this entire project. More precise requirements are needed to drive the project forward. These must inevitably come from the UZH, which has also understood that such a system is necessary. This can only happen if the UZH board of directors is involved. They may decide to press ahead with the digitization of diplomas in order to use such a system. Adjustments would have to be made to meet the new requirements and regulations of the university.

Since the falsification of diplomas affects not only the UZH but all institutions issuing an academic certificate, it must be considered whether further institutions besides the UZH should be included in this project. The necessary adjustments to the prototype would not require much effort. To keep track of the situation, one could run a smart contract for each institution.

If the project continues, user and acceptance tests would have to be carried out at the beginning. In my opinion, the chosen user interface was selected as pragmatically as possible, which is why even more effort and research would have to be invested in this area. Furthermore, the implementation could be improved in some places, since the prototype should mainly proof the concept, and for example, the application of design patterns was not a high priority.

Bibliography

[1] ETH Gas Station (2017), https://bit.ly/2iIVWE7, Accessed: 2018-06-13

[2] Abteilung Finanzen UZH: Abschlussstatistik (2017), https://bit.ly/2KIkcqi, Accessed: 2018-04-18

[3] BCDiploma: Degrees Certified on the Blockchain, https://bit.ly/2rp95qC, Accessed: 2018-03-15

[4] Blockgeeks: What is Ethereum Gas: Step-By-Step Guide, https://bit.ly/2LUXg7T, Accessed: 2018-07-31

[5] Bocek, T., Stiller, B.: Smart contracts - blockchains in the wings. In: Digital Marketplaces Unleashed, pp. 169–184. Springer (2018)

[6] Castor, A.: Cardano Blockchain's First Use Case: Proof of University Diplomas in Greece (January 2018), https://bit.ly/2DVsrYt, Accessed: 2018-03-20

[7] CoinMarketCap: Cryptocurrency Market Capitalizations, https://bit.ly/1dqX6ht, Accessed: 2018-06-13

[8] Concentric Sky: Make your badges meaningful with Badgr, https://bit.ly/2L271Ne, Accessed: 2018-03-09

[9] Crosby, M., Pattanayak, P., Verma, S., Kalyanaraman, V.: Blockchain technology: Beyond bitcoin. Applied Innovation **2**, 6–10 (2016)

[10] Danny Ryan: EVM OPCODE Gas Costs, https://bit.ly/2uEZc93, Accessed: 2018-07-28

[11] Elizabeth Durant, A.T.: Digital Diploma debuts at MIT (Oct 2017), https://bit.ly/2xPRWXC, Accessed: 2018-03-11

[12] eQuals, M.: The Official Platform of Australian and New Zealand Universities, https://bit.ly/2qjHtE9, Accessed: 2018-03-15

[13] Ethereum: Ethereum JavaScript API, https://bit.ly/2ujOLrj, Accessed: 2018-04-27

[14] Ethereum Foundation: Solidity Documentation, https://bit.ly/2uAXVQr, Accessed: 2018-04-19

[15] Etherscan: Ethereum GasLimit History, https://bit.ly/2uLNQBc, Accessed: 2018-06-13

[16] Fox, D.: Secure hash algorithm sha-3. Datenschutz und Datensicherheit-DuD **37**(2), 104–104 (2013)

[17] Fraunhofer Institut: Blockchain for Education - Lebenslanger Lernausweis, https://bit.ly/2Nq0pd9, Accessed: 2018-07-16

[18] Grech, A., Camilleri, A.F.: Blockchain in Education. Tech. rep. (2017)

[19] Gresch, J.: Survey about digital academic certificates (May 2018), https://bit.ly/2wLFXyP

[20] Mattias Andrée: SHA-3 and Keccak checksum utility, https://bit.ly/2NDRK6V, Accessed: 2018-05-03

[21] MetaMask: Brings Ethereum to your browser, https://bit.ly/2DIukHT, Accessed: 2018-04-24

[22] MIT Registrar's Office: Digital diploma pilot program faqs, https://bit.ly/2JYw4zT, Accessed: 2018-03-11

[23] Mozilla: Open Badges, https://bit.ly/2u7YS3n, Accessed: 2018-03-09

[24] Musee, N.M.: An Academic Certification Verification System Based on Cloud Computing Environment. PhD diss., University of Nairobi (2015)

[25] Nakamoto, S.: Bitcoin: A peer-to-peer electronic cash system (2008)

[26] Parity Ethereum - Wiki: Parity Ethereum Documentation, https://bit.ly/2v3DkW4, Accessed: 2018-07-23

[27] Park, H., Craddock, A.: Diploma Mills: 9 Strategies for Tackling One of Higher Education's Most Wicked Problems, https://bit.ly/2DoEeyu, Accessed: 2018-03-02

[28] Recruiting-Tools.ch: Recruiting-Statistiken Schweiz (Apr 2018), https://bit.ly/2u9k09m, Accessed: 2018-06-22

[29] Regulation, G.D.P.: Right to erasure ('right to be forgotten'), https://bit.ly/2zMT9Vl, Accessed: 2018-04-12

[30] Rodrigues, B., Bocek, T., Stiller, B.: The use of blockchains: Application-driven analysis of applicability. In: Pethuru Raj, G.D. (ed.) Blockchain Technology: Platforms, Tools and Use Cases, Advances in Computers, vol. 111, pp. –. Elsevier (2018), https://bit.ly/2zqkdtO

[31] Rutkowski, J.: From the shortage of jobs to the shortage of skilled workers: labor markets in the eu new member states (2007)

[32] SAP: TrueRec - Take the first step toward a trusted digital you, https://bit.ly/2tstjkx, Accessed: 2018-07-17

[33] Turkanović, M., Hölbl, M., Košič, K., Heričko, M., Kamišalić, A.: EduCTX: A Blockchain-based Higher Education Credit Platform. IEEE Access (2018)

[34] University of Nicosia: Academic Certificates on the Blockchain, `https://bit.ly/2I5G3mj`, Accessed: 2018-04-02

[35] University of Southern Denmark: The Digital Diploma, `https://bit.ly/2I3Bid5`, Accessed: 2018-04-21

[36] Varga, S., Guignon, C.: Authenticity. In: Zalta, E.N. (ed.) The Stanford Encyclopedia of Philosophy. Metaphysics Research Lab, Stanford University, fall 2017 edn. (2017)

[37] Vitalik Buterin: Is there a (theoretical) limit for amount of data that a contract can store? (2016), `https://bit.ly/2uAAEyd`, Accessed: 2018-07-10

[38] Warasart, M., Kuacharoen, P.: Paper-based Document Authentication using Digital Signature and QR Code. International Conference on Computer Engineering and Technology (ICCET 2012) (April 2012)

[39] Zeichick, A.: Can Blockchain Solve Your Document And Digital Signature Headaches? (Apr 2018), `https://bit.ly/2tstjkx`, Accessed: 2018-06-04

Abbreviations

ABI	Application Binary Interface
Eth	Ether
Geth	Go Ethereum
GDPR	General Data Protection Regulation
GRNET	Greek Research and Technology Network
Gwei	Giga Wei
MD5	Message-Digest Algorithm 5
Opcode	Operation code
PDF	Portable Document Format
PoA	Proof-of-Authority
PoW	Proof-of-Work
RQ	Requirement
SHA	Secure Hash Algorithm
UZH	University of Zurich
UZHBC	University of Zurich Blockchain

Glossary

Authenticity As stated in [36], authenticity refers to a subject being of undisputed origin or authorship. When one says that something is authentic, it means that it is what it claimed to be or what it aspires to be.

Digital Diploma In this work, a digital diploma is considered the digital equivalent of a real physical diploma. The type does not necessarily have to be a PDF, as in this thesis.

Gas In [4], gas is described "as a unit that measures the amount of computational effort that it will take to execute certain operations." To execute operations on the Ethereum blockchain, a certain fee needs to be paid to the network, which is measured in gas.

Issuance In the thesis, this term is used in two ways. Firstly, it means the physical creation and transmission of diplomas to the recipient. Secondly, it means the digital creation of hashes of the diplomas and saving them in the smart contract.

Opcode Opcode is an abbreviation and stands for operation code. Each low-level operation on the Ethereum blockchain is afflicted with costs, which are detailed in the opcode list [10].

Rinkeby Rinkeby is an Ethereum based blockchain which can be used free of charge for testing purposes. In contrast to the public Ethereum blockchain, which is based on the consensus algorithm Proof-of-Work, Rinkeby uses the Proof-of-Authority algorithm.

Smart Contract The smart contract is the extension of the traditional transactions on the blockchain, where funds are sent from A to B. The smart contract is a computer protocol designed to digitally simplify, verify or negotiate in a more performant way than conventional contracts [5].

Verification Like the term issuance, the verification is used in this thesis in two ways. The UZH receives requests from various companies as to whether a particular diploma corresponds to authenticity. The UZH verifies this and responds to it. The prototype also has verification functionality. In this case, however, the UZH no longer needs to intervene.

UZHBC The UZHBC is the product of this thesis. It allows to store digital diplomas in the blockchain and enable external parties to check the authenticity of the degrees.

web3 Web3 is a library which provides functionalities to send transactions from a client to the Ethereum blockchain [13].

List of Figures

List of Tables

Appendix A

Interview Transcripts

A.1 Dean's Office of the Economic Science Faculty

What is your responsibility regarding diplomas and how are you related to them?

We generate the diploma PDF files out of the system. At the same time, we check if everything is correct which means that the graduate fulfills the conditions and that there are no typos. This process happens at every promotion date. The diplomas are also digitally signed by the rector or dean, but some are signed in written form. Afterward, we create a print order and send the files to the diploma office. Once they are printed and returned to us, we send them to the graduates.

What about the print order, who is responsible for this?

There is a diploma office which solely responsible for the printing. They have special printers and paper types to print the diploma in the required format.

Why are the diplomas not handed to this office directly?

They do not know about the promotion date, because there are several faculties, and each dean's office of the faculty are sending the diplomas to them. The diploma office is independent of the faculty, it is an overlapping office of the university. Also, since each faculty has its condition to pass, the check has to be done from our side. We check if the student has all the required modules, paid his bills, no typos in his bachelor/master thesis title etc.

Are the PDF files used elsewhere?

They are archived here but not especially used anymore. Only for quality management, if we see a mistake. But then we create a new file out of the system and do not modify the original one.

What happens if the diploma will be destroyed or lost?

Every student only gets one diploma, the original. If it gets lost or damaged, you only receive a duplicate which is a bit different from the original, and it costs 100 CHF. You only get one original out of legal reasons. We were told that this could prevent diploma fraud. It is worth to mention that We had cases where we received fake diplomas.

A.2 Registry Office

How are you related to the diplomas?

Students who want to apply for the first time at the university, they have to deliver the original high school or bachelor diploma. Also, if they are coming from another university. We have to check if these diplomas are valid. We are the first contact instance when a student wants to join the University.

Are you verifying Diplomas?

We do not verify swiss diplomas. Sometimes, if we do not know the University of the received diploma, we contact the corresponding University. Verification requests are handled by the student administration office.

A.3 Managing Director of the Economic Science Faculty

What is your responsibility regarding diplomas and how are you related to them?

We, the dean's office, create them out of the SAP system, send them to the diploma office, receive them back end send it to the students. In our SAP CRM, each student will be archived, with its documents and all his records.

What do you think might be the reason we do not send the diplomas digitally?

Good question, we think it has to do with verification. We do not have an online verification system. Therefore, only the original paper is relevant. But here, we can recommend you to the student administration office. There was a project about a new design of the diplomas which was managed by them. There were a lot of legal aspects to be considered.

Since there is no online verification, who is verifying the diplomas?

Yes, we get many verification requests from companies. There is no transparent process behind this, companies do not know where to ask. We refer them to the student administration office, who is responsible for the verification. We need approval from the students that we are allowed to give this information out. Mostly, the companies are sending this approval together with the request. As a side note, there might be a project in this area. The ZI might give information about this. The project is called *"Announcement and Verification of titles academic ranks and degrees and scheduled for 2020".*

If the diplomas are sent in a digital format in addition to the paper-based format, who could make such a decision?

There are many steps for this. The certificate of performance, in which all modules, credits, and grades for a student are listed, is already digitized. You can access and download it at the student service page. To do this, we had to adjust the legal foundation. New is that this document can be offered through digital infrastructure. The decision would come from the direction of the UZH. For your project, we are not sure if the legal foundation is already adjusted, but since we plan a project in this direction, it will be necessary. By the way, there is a decree that the University has to publish all its degrees, but we are not doing it right now in a decent manner.

A.4 Student administration Office (Kanzlei)

As I was told you are responsible for the verification of diplomas. Is this true?

Yes, we get verification requests from employers or job hunting companies. They need to know if a student has a degree. However, we need the authorization from the student to do this. This process is not automated, the company has to contact us. We know this is an issue and that there is a desire to improve.

How often do you verify?

Very often. We print every request. Sometimes, We get 3-4 requests on a day. At least 2 per week.

Have there been any fake diplomas?

Yes. We get the copy from the companies, and we see that they are obviously fake. Sometimes, the grades are exaggerated, or the document is entirely wrong.

How does the verification look like?

We mainly verify the title of the degree, the date of completion, the duration of enrollment and the subject of study. In about one-third of all cases, the final grade is also verified. We practically never have to verify performance records. In 2017, we answered 270 verification requests.

Are there any particular companies that are requesting verifications?

Usually, the requests are spread across many different companies, some sending a little more frequently or regularly than others. We think the relatively most frequent requests come from the companies HireRight and First Advantage, although this - subjectively speaking - only affects every 7th or 8th request. Unfortunately, we do not have any statistics on the frequency of company requests, so we can only give you ad-hoc assessments. However, we also receive requests from many other companies, which gives the picture of a generally rather fragmented field of market participants in the background check.

A.5 Diploma Office

What are you doing with diplomas?

We have three different print-versions, bachelor, master and Ph.D. After the Bologna agreement, we were able to create the diplomas out of our SAP system. They are digitally stored there. Also, there is a long-term archive for diplomas before the SAP time, when we had the Liz-system. We are responsible for the correct print, and quality control, to check if the documents got dirty. Bachelor and master diplomas are sent to the dean's office, Ph.D. diplomas are sent directly but are signed manually before that. All others are signed digitally. We also put a seal of the corresponding faculty on each diploma.

What do you think is the crunch point with digital diplomas?

The UZH just made its first step in this direction. You can get your academic record digitally. The legal foundation to do this is key.

Since the university has to publish the degrees, who is responsible for this?

Yes. We are maintaining a promotion list for Ph.D.'s which is published monthly on the homepage. There are plans to do this for master and bachelor as well. For this topic, we need to work closely with the data security officer of the UZH. There is a specific department for data security guidelines at the UZH.

Since all faculties are sending the diplomas to you, is there a fixed print date for each faculty?

Your question is not easy to answer, as these processes are not harmonized, but are completely different depending on the faculty and type of degree. A basic distinction is made between Bachelor's, Master's, teaching diploma and doctoral degrees. One faculty has one doctorate per month, some have five doctoral appointments per year. Others print the Bachelor's and Master's degree documents at the end of each semester. In principle, a so-called Degree Audit is carried out after receipt of the grades. Upon successful completion, the Diploma Office receives the print order from the faculties.

A.6 Central Informatics (ZI), Business Application (BAP)

Why are students not receiving the diplomas in a digital format?

We think that the acceptance is relatively low. We asked around and found out, that people want to receive a diploma in paper format. We tried to do something with digital signature but realized that it is going to be too complicated. However, the PDF files are here, we just archive them or print them again if someone lost them.

We want use blockchain technology to store and verify the diplomas. What is your opinion on that?

There were cases when diplomas had to be delivered twice. Maybe if a gender changed, or there were mistakes. So the diploma is not valid for lifetime. In this case, you need to store it twice but cannot delete the old hash. On the other hand, the university has a publication duty.

What systems are used?

Basically, everything is done in SAP. It is called SAP Student Life-cycle Management. The whole management of the students is solved through this. All records, achievements etc. For the diploma, some rules will be applied, to check if the student fulfills all the requirements (degree audit). Those rules are also dependent on each faculty.

How much effort would it take to send the diploma PDF files?

Not much at all, we can upload the PDF files the same way as the digital record. Also, we would do it for all the students. What needs to be discussed are legal requirements. For your solution, we also need to integrate this into the workflow, for the prototype, this can be omitted at first.

How do you see the future of the diplomas?

We assume that we are going in this digital direction. Since we have already done this step with the digital records (which lists all modules, credits and grades for a student), the next step would be to make this verifiable somehow. Also, the whole development rises economic advantages. We could lessen our print and delivery costs by having a digital delivered diploma. Using blockchain makes sense in my opinion. Especially regarding the current progress in the industry. Also, the requirement for the verification was already there.

A.7 Data Security Department of the UZH

We want to store hashes of diplomas in a blockchain, is this possible with respects to the data security law of the UZH?

As hashes do not reveal any information about its origin, the diplomas, this will not be a problem. However, it is essential to know who is permitted to write. Also, the issuance must be traceable to the University. As you want to use blockchain, the right to be forgotten can be violated. However, as only hashes are stored, this can be ignored.

Since I talked with many relevant stakeholders now, how could we proceed at this point?

This project will mainly face political burdens. The acceptance of such an idea will not be high. Blockchain and hashes are not well known yet, especially in the upper management of the UZH, which will have to decide about such things. However, the idea is good and makes sense.

Appendix B

Survey

This survey was made to identify the need for digitized academic certificates. 32 different graduates from different Universities were conducted. A Google form was used to collect the answers. The four questions are depicted below in B.1 - B.4.

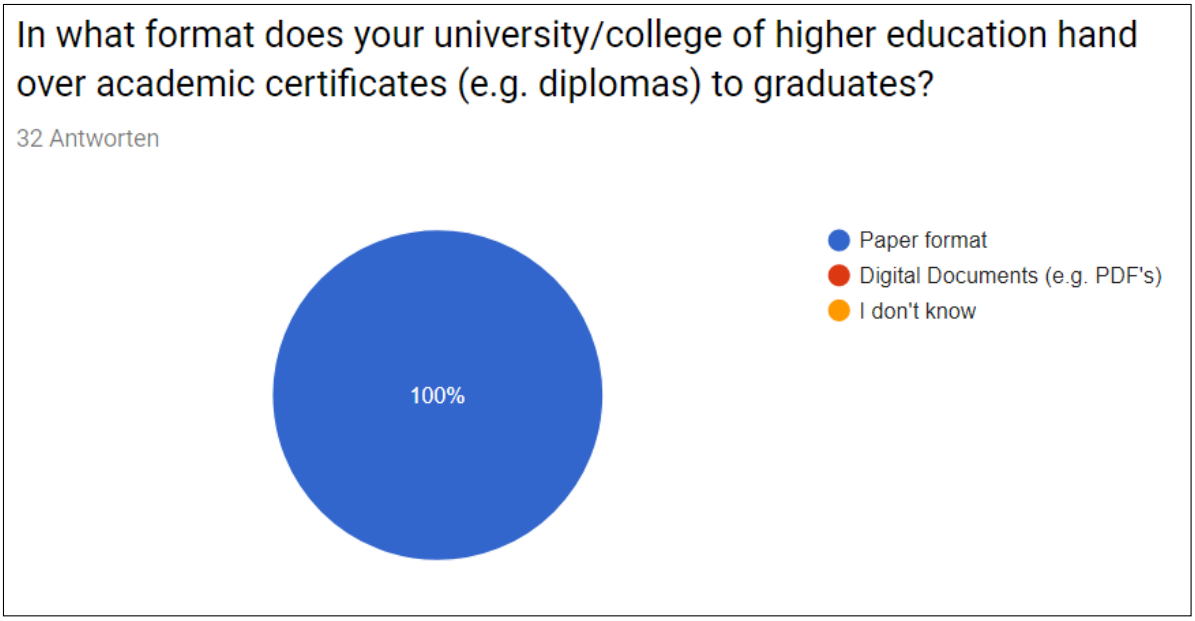

Figure B.1: Question 1

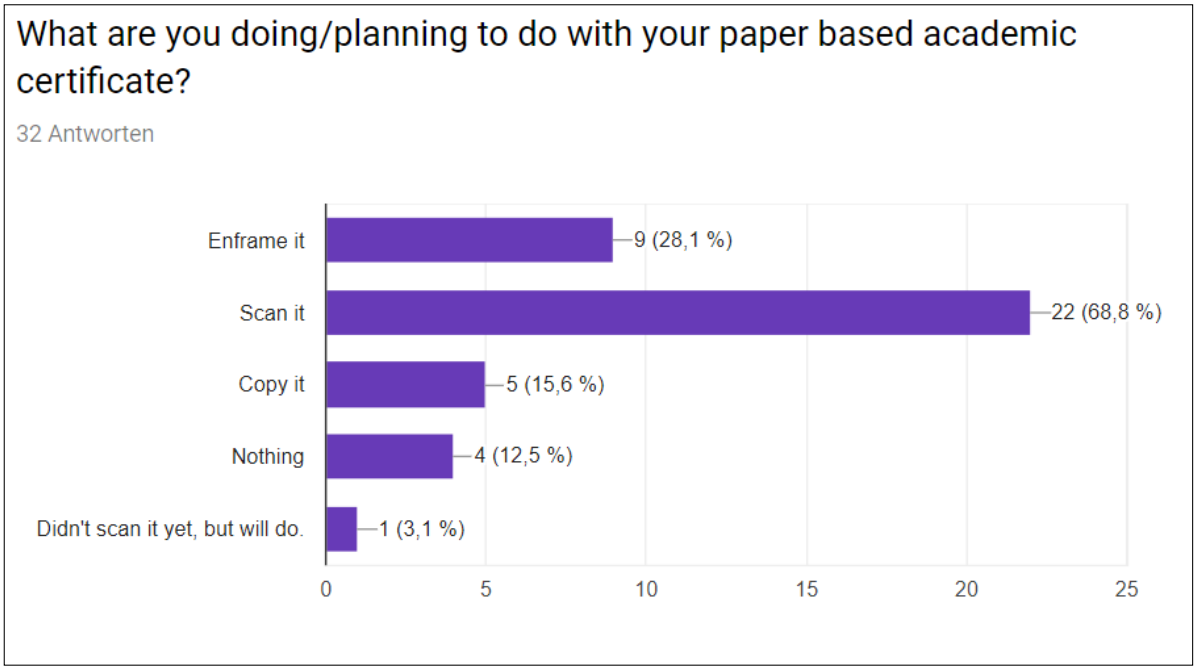

Figure B.2: Question 2

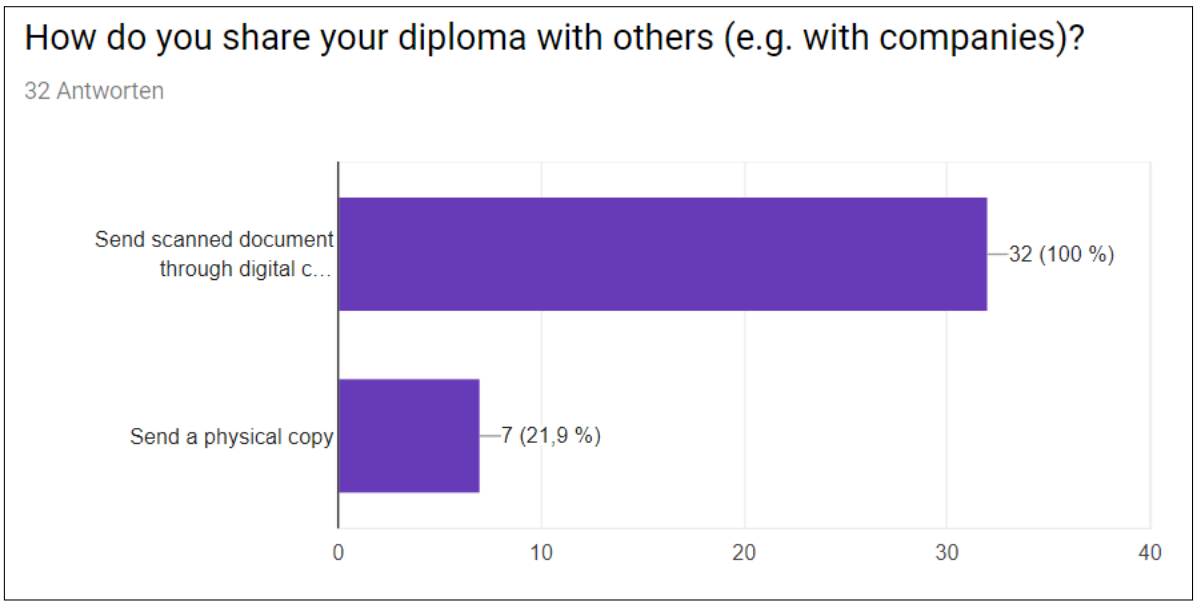

Figure B.3: Question 3

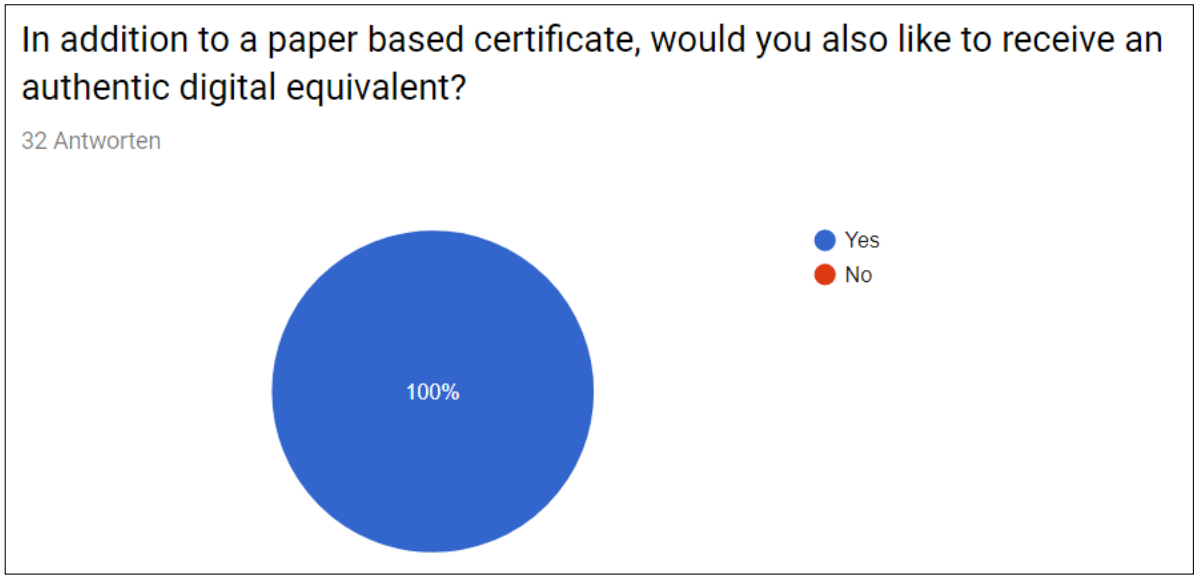

Figure B.4: Question 4

Appendix C

Installation Guidelines

As a prerequisite, Node.js and Geth needs to be installed in advance. After Geth is installed, start the syncing process by opening a terminal and run:

- `geth --rinkeby --fast --rpc --rpccorsdomain "*"--rpcapi "admin,eth,web3 ,personal,net"console`

This can take up to 20 minutes. The keystore file in the code folder at `public\keystore` has to be placed in your Geth-Keystore directory `..\User\AppData\Roaming\Ethereum \Keystore/`. Once the blockchain is synchronized, copy the *code* folder from the CD to your device, and locate into this directory. Open up a new console and run:

- `geth attach HTTP://localhost:8545`

This will connect the client with the running Geth instance. Lastly, start the server by opening another console and run:

- `nodemon`

The web application can be accessed on *http://localhost:3000/*. More details can be found in the README.md file which is stored in the code folder. The password to submit a transaction is `123456789`.

C.1 Code Documentation

The client of UZHBC is a web application implemented with React. The root file App.js, which contains the different React components, is located in the `\src` folder. The respective components footer, index, navbar, introduction, issuance and verification are located in `\src\components`. Issuance.js implements the issuance process and verification.js implements the verification process. The connection to the blockchain is implemented in

the web3.js file, and is located in the `\src\utils` directory. Changes to the code can be recompiled by opening a console in the root folder and run `webpack --w`, which creates a bundle.js that can be found in `\public\build`. All versions of the smart contract are located in the `\contracts` directory. However, these do not run on the client, but on the Rinkeby blockchain. All representation specific implementations can be found in the `\public\assets`. As mentioned above, the README.md describes further details about the prototype.

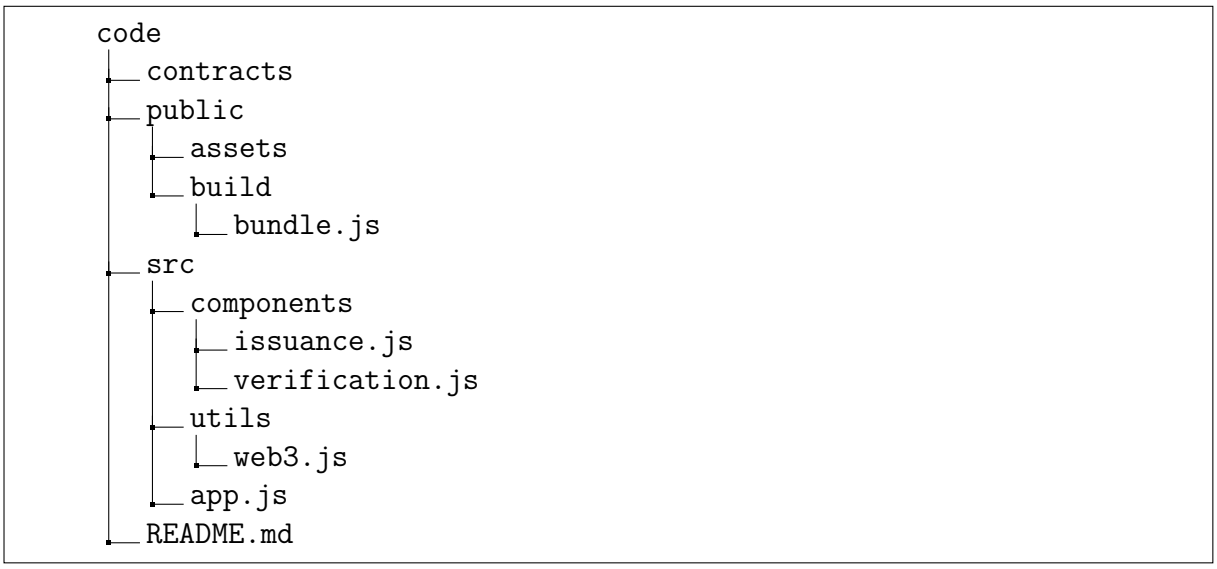

Figure C.1: Directory Tree of the code

Appendix D

Contents of the CD

The content of the CD is structured as follows:

- The **code** folder contains the final implementation of the prototype. The structure of this directory has been discussed in the previous chapter.
- The latex project of the master thesis can be found in the **thesis** folder. The `\figures` folder contains all used figures, and the `\tables` directory holds all used tables.
- The root directory also contains the proposal of the thesis, the final master thesis report, the proceeding for the BIS conference, the data collected for the costs of the smart contract, and the final presentation.

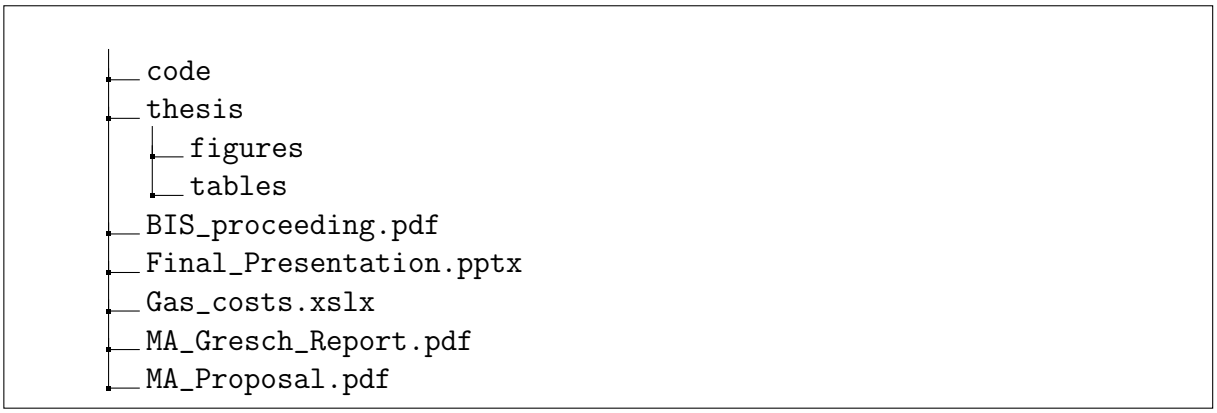

Figure D.1: Directory Tree of the enclosed CD