Who stores the private key? An Exploratory Study about User Preferences of Key Management for Blockchain-based Applications

Clemens Brunner[1], Günther Eibl[1], Peter Fröhlich[2], Andreas Sackl[2] and Dominik Engel[1]
[1]*Center for Secure Energy Informatics, Salzburg University of Applied Sciences, Salzburg, Austria*
[2]*AIT, Austrian Institute of Technology, Vienna, Austria*
{*f_author, s_author*}*@en-trust.at,* {*f_author, s_author*}*@ait.ac.at*

Keywords: Blockchain, Trust, Key management, User study

Abstract: Applications based on blockchain technology have become popular. While these applications have clear benefits, users are not yet familiar with their usage, which could hinder further applications of this technology. In this paper, an online survey with 110 potential users, as a representative of an average citizen, was conducted. The focus of this survey is to explore their preferences concerning the interaction with blockchain-based applications by mainly focusing on how to handle private keys. To best of our knowledge this is the first study where average citizens are asked about the preferred management of a private key, which is necessary when interacting with blockchain-based applications. One of the main results was that about 80% of the participants would like to have the benefit of data sovereignty despite the cost of being fully responsible to backup their credentials.

1 INTRODUCTION

Blockchains or distributed ledgers create a globally unique decentralized database managed without a central trusted third party. Blockchains have been suggested as the underlying technology for a large number of use cases. However, while blockchain concepts (Nakamoto, 2008; Wood, 2017) and applications based on this technology (Madhusudan et al., 2019; Sovrin Foundation, 2018) have received a high level of attention in the scholarly, business and even political arena, users are not familiar with the optimal ways to interact with this technology.

Making blockchains accessible to users should essentially start with secure and understandable authentication mechanisms. In the context of blockchain technology, this is even more the case.

Despite its benefits of offering more sovereignty through less dependability from a central intermediary, a decentralized blockchain approach has a main drawback: The private key needs to be stored securely by the users themselves. Only the owner of the private key has full control over the attached data, e.g., cryptographic coins in case of bitcoin.

In case the private key gets lost or compromised, this can result in a full loss. So far, there is little knowledge on users attitudes towards the related benefits and risks. In this paper, conceptual variants of private key management and password backup for different decentralized, blockchain-based applications are explored and assessed regarding their acceptance by users. An online-survey was conducted in which participants can qualify different ways of key management which are varied by backup responsibility, application area, and key storage methods.

The rest of the paper is structured as follows. In Section 2 a description of (de)central controlled applications, possible use cases of decentralized applications and different variants to allocate the backup responsibility of the private key are described. The research questions are presented in Section 3. Section 4 describes the method for the questionnaire. The results are presented in Section 5 and discussed in Section 6. Finally, a conclusion is provided in Section 7.

2 BACKGROUND AND RELATED WORK

In this section, the process of authentication of users in centrally controlled application, blockchain-based applications, use cases for blockchain-based applications and different variants for managing the backup responsibility for the private key are described. Please note, that the management of a private

key is usually done by a software without the knowledge of the end users or by experts. In blockchain-based applications the management of the private key should be done by the end users, in order to be in full control of the data.

2.1 Centrally Controlled Applications

In this paper, "centrally controlled applications" are defined as all kinds of application, which are hosted or controlled by one organization. The identification of a user is done by, e.g., a username and/or and email address. The most common way to perform the authentication of users is by a password, which is linked to the user's identifier. A salted hash of the password is stored on the application controllers' server (Gauravaram, 2012). The verification of the password is performed by the application owner. If the user provides the correct password, the application owner enables the user to access the application. There are several ways to reset a password, e.g., confirming the email address or answering control questions. If the user provides enough evidence for the application owner to confirm the user's identity, the application owner replaces the old password with a new one. This password can then be used for future authentication attempts. One downside of centrally controlled applications is that the application owner can modify, delete or change information attached to users. Another downside is data availability, the application owner can decide to stop the application and delete all information, anytime. In this case, all users will lose their data.

2.2 Blockchain-based Applications

Blockchains, first proposed in (Nakamoto, 2008), are decentralized append-only databases with a globally consistent state. Transactions can be used to append data and valid transactions trigger a new state transition, which results in a new globally consistent state. Data in a blockchain is usually publicly and transparently available and replicated among all blockchain nodes. Numerous variants of blockchains, e.g., (Nakamoto, 2008; Wood, 2017; Larimer, 2017) and decentralized applications, e.g., (Fromknecht et al., 2014; Madhusudan et al., 2019; Ben-Sasson et al., 2014; Sovrin Foundation, 2018; Brunner et al., 2019; Gräther et al., 2018) based on blockchains exists. In the following, the difference of authentication and authorization between centrally controlled applications, e.g., online banking or Amazon, and decentralized applications is described.

In this paper, "decentralized applications" refers to applications based on blockchain technology. Such applications differ from centrally controlled applications in the way that the usage of passwords is not possible, because no application owner exists, who can confirm the correctness of the password. Decentralized applications which are based on blockchains use asymmetric cryptography and represent users by public/private key pairs. Users are then identified either by their public key or an address (which is usually derived from the public key using a hash function). The authentication of users and their transactions is done by a signature, created with the users' private keys. The private key for each user is linked to this user's identifier by cryptographic methods. The verification of the signature is done by all blockchain nodes during the block validation process. There is no way to reset a private key. If users lose their private keys, the users have no way to add new data to the blockchain or verify that data belongs to them.

Only the user has the possibility to prove ownership, add or update data attached to the user's identifier by providing a signature created with the user's private key.

In order to avoid that a user is locked out of the decentralized application, the user needs to backup the private key. Most blockchain clients force users to note a mnemonic[1] or the private key on paper before they can proceed. It is also possible to use, e.g., Shamir secret sharing (Shamir, 1979) for a secure distributed backup of the private key among trusted parties. For this purpose, the private key is split into n shares with a threshold $t \leq n$. To reconstruct the private key at least t of the shares are needed. There are several ways to backup the private key, but in all cases it requires additional (maybe unknown) action by the user.

2.3 Use Cases of Blockchain-based Applications

In the following, potential blockchain use cases, based on (Kshetri, 2017; Wüst and Gervais, 2017) are described. Due to the fact that there are always new ideas for potential blockchain use cases, this is not an exhaustive list, rather it should give the reader an idea of potential use cases.

Digital Cash: Cryptocurrencies, e.g., (Nakamoto, 2008; Wood, 2017; Ben-Sasson et al., 2014) are the most prominent use case for blockchain-based applications.

[1] A mnemonic is a sentence of a group of words and is used in combination with a Key Derivation Functions (KDF) to deterministically derive a private key.

Supply Chain Management: The possibility to avoid trusted third parties and the transparent and decentralized way of managing paperwork are reasons why blockchains were proposed to be used for supply chain management (Apte and Petrovsky, 2016; Francisco and Swanson, 2018).

Identity and Access Management: For authentication and authorization of identities, a blockchain can be beneficial in order to avoid a centrally trusted party which controls those identities. This can support users' self-sovereignty and trustful decentralized authentication mechanisms (Manohar and Briggs, 2018). uPort[2] and (Sovrin Foundation, 2018; Al-Bassam, 2017) provide blockchain-based solutions for this use case.

IoT Security: There are numerous contributions, e.g., (Christidis and Devetsikiotis, 2016; Won et al., 2018; Madhusudan et al., 2019; Orman, 2018; Faber et al., 2019), which aim to improve the security of IoT devices with blockchain-based applications. For example, using blockchains, a globally consistent and decentralized Public Key Infrastructure (PKI) can be established.

Proof of Ownership: Similar to PKI, a blockchain-based application can be used to proof ownership of digital data. Additionally, since this data is time-stamped, it is ordered. Verification that a file existed at a given point in time is used, e.g., by MIT[3] and by (Gräther et al., 2018; Brunner, 2017; Brunner et al., 2019) to issue diplomas.

Trading: Trading of all kind of assets is possible with blockchain-based applications, such as energy trading (Mengelkamp et al., 2018; Knirsch et al., 2019).

2.4 Assigning the Backup Responsibility

To use a decentralized application, the backup of the private key is essential. There are several ways to backup a private key, e.g., a note on paper, distribute shares to trusted parties or storing it on the computer. But in all these described cases, the user is responsible for the backup.

There also exist applications which provide an online wallet. An online wallet is provided by a central application owner to access decentralized applications. For example, Bitpanda[4] is an exchange service for digital assets, such as bitcoin. In order to provide a good user experience, they will backup and store the private keys of their users, which are represented by a username and a password. Bitpanda interacts with the decentralized application (bitcoin) in the name of their users. Hence, the users can use, e.g., bitcoin, without managing their own backup for their private keys.

In the following, and also for the questionnaire, two variants of using decentralized applications are differentiated:

Direct Variant: In the direct variant, the users are responsible for their own backup of their private key. The users interact direct without an intermediary with the decentralized application.

Hybrid Variant: In the hybrid variant users are not responsible for their backup, they will outsource it to a third party, named as *manager* which is contacted in the usual way using username and password. There would be no change in usability for users, when moving from the currently known centrally owned applications to decentralized applications. This includes the property that if the users forget their password to authenticate themselves to the manager, they can reset it. This variant is termed *hybrid*, because a user uses a blockchain-based application through a centrally-controlled application. The manager acts as proxy and interacts in the name of the user with the decentralized application. The private key is in the hand of the manager.

2.5 Related Work

Authentication has been a serious challenge for system usability and acceptance for a long time(Adams and Sasse, 1999; Bonneau et al., 2012). In (Bonneau et al., 2012) an overview of alternatives to a password for authentication are described and compared. Since authentication for blockchain-based applications is only possible with digital signatures this work focuses on the corresponding approach that uses a private key which needs to be stored. In (Adams and Sasse, 1999) a web-based questionnaire was conducted to evaluate the users' behaviors and perceptions relating to password systems. Password managers and specific bitcoin clients have shown to provide more comfort for customers, but still exhibit unsolved usability problems as described by (Chaudhary et al., 2019). A review of previous studies on public key authentication is presented in (Eskandari et al., 2015). They summarize that the metaphor and terminology behind public and private keys is often found confusing and that key migration between devices is regarded as difficult.

[2]https://uport.me [accessed: August 2020]

[3]https://blockcerts.org [accessed: August 2020]

[4]https://www.bitpanda.at [accessed: August 2020]

Online studies are increasingly used in HCI research. An overview of their advantages and disadvantages can be found in (Petralito, 2019). Advantages over co-located lab-based data gathering methods are the typically quicker and cheaper conduction, wider distribution of the participants sample, lower hurdles of participation and lower infrastructure costs, the convenient use of randomisation and features for embedding multimedia.

To best of our knowledge no related questionnaire exists which asks average users of blockchain-based applications about preferences of managing their private keys.

3 RESEARCH QUESTIONS

In this section, the research questions consisting of four (three main and one exploratory) questions are described in more detail.

RQ1: Which variant (direct, hybrid) for the key management will be preferred, depending on the application area?

The main research question asks who should have the responsibility for the backup of the private key and if people are willing to transfer the responsibility to an institution at the cost of also transferring control to this institution.

RQ2: How does the application area and the institution influence the possible preference of the hybrid variant?

It is expected that the application area would highly influence this preference. It was decided to choose only three use cases from the application areas, described in Section 2.3. The authors excluded cases which require technical expertise on the topic like IoT security and opted for the ones which are more common for the participants: proof of ownership (UC1: management of diplomas or certificates), identity and access management (UC2: management of passports) and digital cash (UC3: management of money). For the hybrid variant the possible managers were chosen based on the following expectations: for diplomas users might prefer the university, for ID document users trust the government and for financial transactions users trust a bank. Therefore, the questionnaire asks for the trust in all these institutions, plus an external company, for each of the three use cases above.

RQ3: Which method would users of the direct variant use to store their private key?

Here, it should be assessed, if users are aware of the need to backup the private key reliably using, e.g., technical solutions like password managers or not.

RQ4: Which personal characteristics influence these preferences?

The preference for different solutions is likely to depend on user characteristics, which is expected to be very diverse. Therefore, questions which address potential user influences, e.g., the current behavior or the technical knowledge are added. However, due to the sample size only a first, exploratory analysis was planned in order to generate hypotheses, which personal characteristics and behavior influence the decisions above.

4 METHOD

In order to address the research questions above, in March 2019 an online-study was conducted[5].

However, online studies also have often revealed limitations, amongst which sample distortion, carelessness and insufficient control of attention are the most severe ones for the purpose of studies in user-centered security (Petralito, 2019; Fahl, 2016). In order to alleviate these drawbacks in our study, we controlled the sample selection by commissioning a nationally leading online panel provider. Furthermore, we used a set of control questions in order to assess the participants understanding of the presented concepts, see Section 4.2.

Candidates from the whole country (Austria, Europe) were selected with regard to their compliance with the main target participant profile characteristics representative spread of regions, age between 18 and 65, gender, and educational level. Overall, 110 participants completed the survey. To generate representative results, the following characteristics were balanced among the participants: sex, age, residence in federal state and education level. The time to complete the survey was targeted to be under ten minutes, in order to ensure a high attention and answering quality level. Consequently, the questionnaire structure was kept comparably simple and short. The study was executed using the online tool limesurvey[6]. The participants received 5 EUR from the panel provider, after the survey was successfully completed.

The survey was structured as follows: it starts with a video that explains the research question. Due to the rather restrictive time constraints, it was decided that it should take no longer than three minutes. The video is followed by four control questions that

[5]Both the video and the questions are available at https://www.en-trust.at/downloads/

[6]https://www.limesurvey.org/de/ [accessed: August 2020]

check whether the participants understand the problem. Since the attention and memory of the video content is expected to decrease with time, the main research questions were asked directly after the control questions. Finally, possible influence factors were assessed ending with the simplest ones, i.e., the demographic data.

4.1 Explanation Video

The questionnaire starts with a video that was produced specifically for this study using a professional speaker and a video producer. In this video both the direct and the hybrid approaches are described. Their advantages and their disadvantages are explained both visually and acoustically.

The video is created in a way that participants do not need any prior knowledge. If possible, technical concepts like asymmetric cryptography and the use of key pairs are hidden or simplified. The private key was named "access code" and illustrated with a key symbol, see Figure 1, in order clearly distinguish it from the username and password approach. Therefore, in the following, the terms "private key" and "access code" are used interchangeably. The language of the video is German, English subtitles were added later on[7].

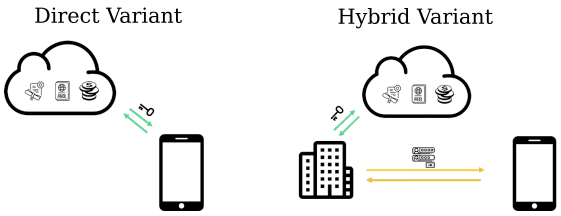

Figure 1: Screenshot of the explanation video for the questionnaire showing the differences between the direct and the hybrid variant[8].

The video is divided into six sections. The first section is an *introduction* about authentication and the differences between an access code and a password. The advantages and disadvantages of using a password are explained in the section *centrally controlled applications*. After that, the use cases (UC1, UC2, UC3) of decentralized applications are explained in the *decentralized applications* section. The advantages and disadvantages of the *direct variant* and the *hybrid variant* for the backup responsibility of the access code are explained in two sections. Finally, the last part of the video provides a short *summary* in order to increase the comprehensibility of the content. The video ends with an invitation to imagine a future where only decentralized applications exist that manage diplomas, passports and money.

[7]Video can be downloaded here https://www.entrust.at/downloads/

[8]Icons made by Freepik, Smashicons, SimpleIcons, Wanicon, Payungked, Geotatah, Recep Kutuk, Surang, Smalllikeart, Zlatko Najdenovski from www.flaticon.com

4.2 Control Questions

The survey starts with the following four control questions:

CQ1: Which variant lets me reset my password? (direct - **hybrid** - don't know)

CQ2: Which variant leaves me in full control of my data? (**direct** - hybrid - don't know)

CQ3: If I forget my access code using the direct variant, can I reset it? (yes - **no** - don't know)

CQ4: Which variant leaves full control over my data at my manager? (direct - **hybrid** - don't know)

Since the chance to answer two questions correctly is 25%, redundancy of the questions was introduced. Therefore, the questionnaire includes questions about the advantage of the variants (CQ1 and CQ2) and redundantly at the corresponding disadvantages (CQ3 and CQ4).

4.3 Main Research Questions

To determine the favored storage mode for the private key when using the direct variant (RQ3), the following options were proposed: note - save on paper - password manager - computer - mobile phone - print - email - share on facebook - other. Multiple answers were possible. The analogous question was asked for the password using the hybrid variant only for sake of potentially needed sanity checks.

To answer RQ1 and RQ2, the participants had to rate the following five possibilities for interacting with a decentralized application for all three use cases.

D: Direct

HC: Hybrid with an external company as manager

HU: Hybrid with the university as manager

HG: Hybrid with the government as manager

HB: Hybrid with the bank as manager

These ratings allow a comparison of the direct and the hybrid variant (RQ1). For the hybrid variant one can also determine the preferred managing institution depending on the use case (RQ2).

4.4 Influence of Personal Characteristics

In the next step of the questionnaire, the influence factors were assessed.

4.4.1 Questions of the Survey

Beside the usual demographic questions at the end, five different groups of questions were considered to be potentially important for the answers.

- Password-related
- Online behavior
- Technical knowledge
- Affinity to mobile phone usage
- Demographic information

The first group of questions are about password usage: first we ask about the self-assessment, how likely the participants would lose their access code (surely not - rather not - maybe - rather yes - surely yes). As a comparison, questions about the behavior of the participants about electronically bought tickets (print - save on mobile phone - get by email - don't buy electronically) are asked. Later on follow questions about how often the participants reset a password using the password-forgotten method (never - rarely - occasionally - often), how many passwords they use (one - own passwords for important applications - separate password for each application) or if they use a password-manager (don't know what it is - know but not use it - use it).

Questions about the online behavior of the users are added, i.e., how frequently online-applications are used, is online-booking the first option (always - if possible - mostly - rarely - never) and also how often the participants use social media platforms (daily - weekly - monthly - never).

Technical knowledge is assessed by questions for knowledge about electronic signatures, cryptocurrencies, installation of programs or programming skills.

As a side investigation, simplified questions about the affinity for different mobile phone usages, e.g., if the participants would like to save different documents on the mobile phone or if they already do it are added.

Finally, demographic data were assessed: Age, sex, education, job type.

4.4.2 Methodological Approach

The final research question aims at generating hypotheses about which factors could influence the preferences of the variants. Due to the exploratory nature of this part, the found dependencies are described graphically. Here, filtering participants due to their understanding of the control questions would assume that the control questions have the highest impact on the answers. In order to avoid this assumption the control questions are treated the same way as other possible influence factors. A second benefit of this approach is the possibility to include all participants. Therefore, other influences such as age can be analyzed with a higher sample size.

For this analysis the three outcome variables considered are the rating of the direct variant (**D**) for all three use cases. While logistic regression could be used to predict the outcomes and assess the influence of variables, this way of analyzing the data would not correspond to the exploratory goal of this study which aims at generating hypotheses for future studies. Instead, pairwise correlations between all variables (both possible influence factors and the outcome variables) were investigated. This kind of analysis can not only show influences of variables on the outcome variables but also show dependencies between influencing variables.

In the first step, from the large amount of possible influencing factors only the most promising ones were pre-selected. This was done by performing univariate Chi-Square tests. Only variables that are significantly dependent to at least one of the three outcome variables (considering p-values less than 0.01 as significant) were kept. In order to get valid p-values all variables were binarized before the application of the test. The three outcome variables were binarized by merging the ratings "surely yes" and "rather yes" into "yes" and merging the answers "surely not" and "rather not" into "no" and named `moneyDirectBin`, `passDirectBin` and `certDirectBin`, respectively. Furthermore, all possible influence factors were suitably binarized based on the relative frequencies of their attributes.

In the second step the correlations between all pairs of 11 variables (8 influence variables were selected in the first step plus the 3 outcome variables, see Table 1) were assessed using Cramer's V.

In the last step all these correlation coefficients were visualized in a correlation plot, see Figure 5, where a larger and darker circle indicates a stronger correlation between the corresponding pair of variables. In order to make it easier to see a structure in the plot, the variables where rearranged using a hierarchical clustering algorithm.

5 RESULTS

In a first step, the users of interest were selected from the obtained data[9]. Although the video contained all needed information and was designed to be as comprehensible as possible, it cannot be guaranteed that all participants understood the content and gave a well-founded answer to these technically related questions. Thus, for the subsequent analysis only the users that answered all four control questions correctly are considered. It was not surprising that for an average citizen the technical concepts of asymmetric cryptography and the importance of a backup of the private key was hard to understand. While 108 of 110 participants stated that they understood the content, only 43 out of the 110 participants answered all four control questions correctly, as illustrated in Figure 2. For the subsequent analyses, only participants who answers all control questions correctly are selected.

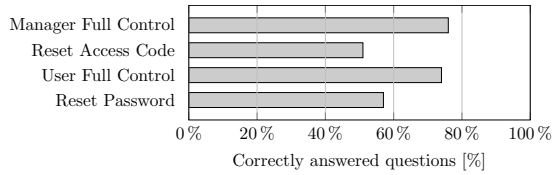

Figure 2: Percentage of correct answers for each of the four control questions.

5.1 RQ1: Preference of a Key Management Variant

The primary research question was whether the direct or the hybrid variant is preferred. For all three use cases the direct variant was preferred, see Figure 3. About 80% of the participants stated that they would surely or rather use the direct variant. In contrast, the most preferred hybrid variant reached only about 60%.

5.2 RQ2: Application Area and Institution Influences

As illustrated in Figure 3, the preferred parties for the hybrid variant are as follows: for UC1 (diplomas) the government and the university are equally preferred, for UC2 (passport) the government is preferred and for the UC3 (money) the bank is the most trusted manager. For none of the use cases a (not nearer specified) external company is preferred.

[9]The data and the Python-scripts are available at https://www.en-trust.at/downloads/

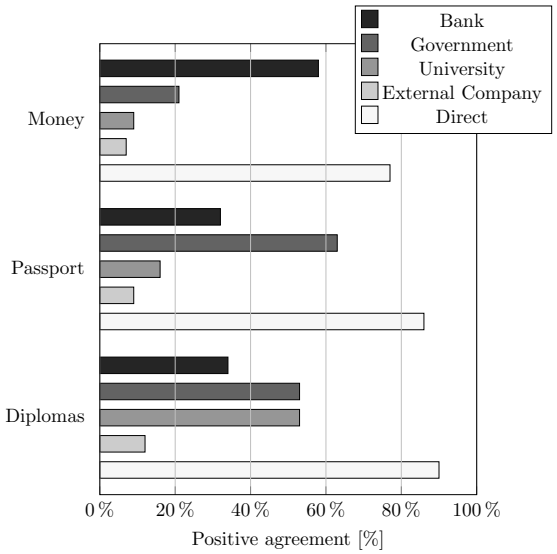

Figure 3: Comparison of the direct and the hybrid variant for three different use cases. In order to reduce complexity, for this graphic, the answers are binarized treating the attributes “surely yes” and “rather yes” as positive agreement.

There is only a small difference between the three use cases in the sense that for all of them the preferred institution reaches about 60% positive agreement.

5.3 RQ3: Key Storage Method

For this question multiple answers were possible. About half of the participants would note the access code on paper, followed by the equally frequent options storing on the mobile phone, in a password manager or printing it, as illustrated in Figure 4. Sending by email and sharing it with a trusted person are the two least favored options.

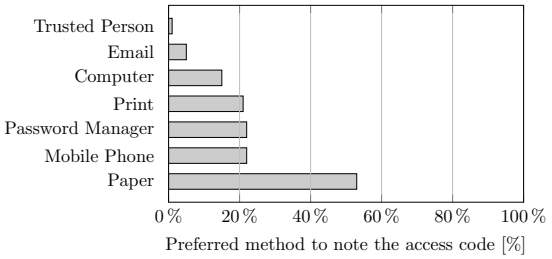

Figure 4: In this figure the preferred ways for the participants to backup a private key are illustrated.

5.4 RQ4: Influence of Personal Characteristics

For the exploitative part of dependencies all 110 participants of the questionnaire were selected. The set

Table 1: Description of significant influence factors.

Name	Short Description
CQ2	Which variant leaves me in full control of my data?
CQ3	If I forget my access code using the direct variant, can I reset it?
storePubTransp	I want to save transport tickets at my mobile phone.
usePubTransp	I buy transport tickets with my mobile phone.
CQ4	Which variant leaves full control over my data at my manager?
installMyself	I typically install programs myself.
useGovAlways	I always contact the government online.
ageBin	Binarized age: age $<$ 40 or age $\geq$ 40.
certDirectBin	I would use the direct variant for UC1.
passDirectBin	I would use the direct variant for UC2.
moneyDirectBin	I would use the direct variant for UC3.

of variables remaining after filtering with the Chi-squared test is described in Table 1. The ordering of variables is chosen such that highly correlated variables should be near each other.

The correlation plot in Figure 5 shows that the three outcome variables are grouped together (bottom up) since there is a large correlation between them (bottom right part). The variables (ageBin and useGovAlways) have a high correlation to the three outputs only, i.e., they are independent on the other influencing factors. Besides CQ4 these two factors are the only variables with a significant dependency to all three outcome questions. While control questions appear to correlate with the outcomes, these correlations are not clearly the biggest ones. In comparison, age unexpectedly appears to be an equally important influence factor. Furthermore, control question CQ1 does not have a significant dependency with any outcome and is therefore missing in the list. The remaining variables describe online behavior in general. From the variables that describe the skills only installMyself remains. It is remarkable that eleven out of twelve participants that use a password manager would use the direct variant for passports (although the influence of the variable itself was not significant).

Finally, in order to get an idea about the influence of different factors on the outcome, the effect of the factors with the highest correlation with the corresponding outcome variable is exemplified in Figure 6 for the passport use case UC2 and shows about 30% difference between the two possible attributes.

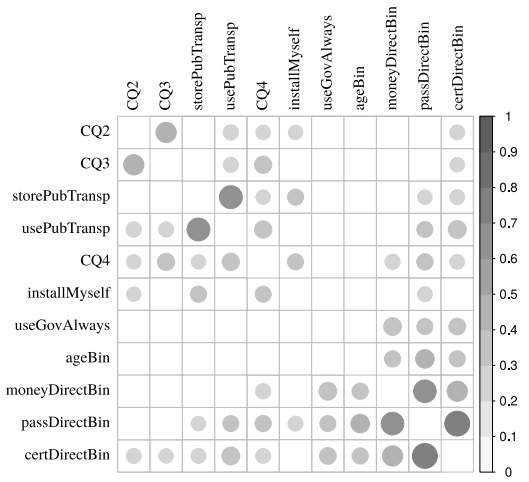

Figure 5: Correlation between influencing variables and the outcome variables. Higher correlations are visualized using larger and darker circles, only significant ($\mathbf{p < 0.01}$) correlations are shown.

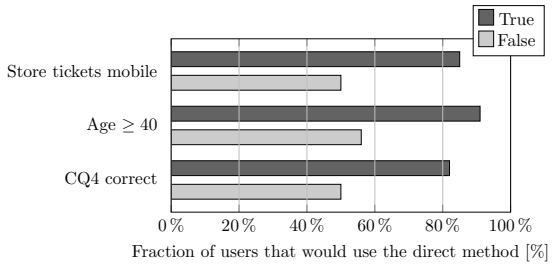

Figure 6: The influence for the question, if a participant will prefer the direct variant or not, for the passport use case (UC2) are illustrated. For three influence factors: control question 4 (CQ4) answered correctly, participant age is $\geq$ 40 and if participants store their public transport ticket on their mobile phone the results are split into two groups.

6 DISCUSSION

In the following, we discuss all research questions separately.

RQ1: The preference of the direct variant was not expected as such. This holds especially for the money use case since this use case is already well-established and online-banking is already common practice (102

of 110 participants state to use online banking "always" or "often"). There is indication that this preference of the direct variant might be overestimated. Firstly, most participants likely have never experienced a situation where documents or money were completely lost due to forgetting an access code. Currently, in most systems a way to recover from the loss of access information exists. Therefore, participants might underestimate the disadvantage. Additionally, the number of correct answers to the questions regarding full control over the data (CQ2 and CQ4), exceeds the number of correct answers to the questions regarding the possibility to retain a forgotten password or access code (CQ1 and CQ3). This could indicate, that the advantage of the direct variant was more comprehensible to the participants than its disadvantage. Maybe this also holds for the participants that correctly answered all control questions. An indication for this assumptions is that the same analysis done with the remaining participants yielded rather similar results.

RQ2: For the hybrid solution, the preferred parties are as expected. Due to the short amount of time in this study the three use cases were simply treated as separate. In reality, one could also ask for a minimal set of institutions that may relate to more than one single use case. Based on the result above, see Figure 3, it would be natural to let the bank provide access to the money use cases and let the government provide access to both the certificate and the passport use cases.

RQ3: The preferred key storage method is "noting on paper" followed by "printing", "using a password manager" and "storing it on a mobile phone". This results shows that the average user prefers analog variants ("printing" and "noting on paper") over more technical, digital variants (like "using a password manager" and "storing it on a mobile phone"). The analog variants have the advantage that the users control and understand the whole storage process. In contrast, the digital backup methods are less transparent and the users additionally have to trust the application developers. While nearly every participant has a smart phone, only slightly more than 20% would use the smart phone for storing important access codes.

RQ4: When analyzing possible influence factors, surprisingly variables that ask for behavior that are more tightly connected to the research questions like the amount of passwords or the behavior of noting passwords show no significant influences. For example, participants who estimate themselves to lose the access code, were expected to prefer the hybrid variant. However, there was no evidence that this is the case. Figure 5 also shows a property already stated above: The control questions about the reset possibility in case of loss of password and access code (CQ1 and CQ3) correlate less with the outcomes than the questions about which party has full control over the data (CQ2 and especially CQ4). The variables `storePubTransp`, `usePubTransp` and `installMyself` have correlations both between themselves and to the control questions. One could try to interpret them as the variables describing affinity to usage of modern techniques.

7 CONCLUSION AND OUTLOOK

In this paper an online survey about the preferences of average citizens concerning the handling of private-keys of blockchain-based applications was conducted.

While about 80% of the participants of the conducted online survey would use the direct variant that enables users need to store the private key themselves, only about 60% would use a hybrid variant where an external manager is responsible for the backup. So the participants of the study seem to value data sovereignty higher than the risk of losing money, passports or certificates.

The fact that average citizens, represented by the survey participants', are not familiar with technological concepts like asymmetric cryptography it would be crucial for blockchain-based applications to provide additional explanations and an user-friendly interface for guidance about handling the private key.

Future work might develop easier overall solutions that educate people with awareness phases about their options and includes or supports several methods to backup the private key. For example, solutions could be accompanied by additional tools including easy to setup and use password managers and redundant saving strategies. The two variants considered here are rather extreme cases, other methods that decrease the needed amount of trust in the manager could also be developed. A survey about the current user preferences of actual users of blockchain-based applications would also be interesting as future work.

Acknowledgements

The financial support by the Federal State of Salzburg and by the Austrian Research Promotion Agency (FFG) under project number 865082 (ProChain) are gratefully acknowledged.

REFERENCES

Adams, A. and Sasse, M. A. (1999). Users are not the enemy. *Communications of the ACM*, 42(12):40–46.

Al-Bassam, M. (2017). SCPKI: A Smart Contract-based PKI and Identity System. In *Proceedings of the ACM Workshop on Blockchain, Cryptocurrencies and Contracts*, pages 35–40, Abu Dhabi, UAE. ACM.

Apte, S. and Petrovsky, N. (2016). Will blockchain technology revolutionize excipient supply chain management? *Journal of Excipients and Food Chemicals*, 7(3):76–78.

Ben-Sasson, E., Chiesa, A., Garman, C., Green, M., Miers, I., Tromer, E., and Virza, M. (2014). Zerocash: Decentralized Anonymous Payments from Bitcoin. In *Proceedings – IEEE Symposium on Security and Privacy*, pages 459–474, San Jose, CA, USA. IEEE.

Bonneau, J., Herley, C., van Oorschot, P. C., and Stajano, F. (2012). The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes. In *Security and Privacy, 2002. Proceedings. 2002 IEEE Symposium on*, pages 553–567, San Francisco Bay Area. IEEE.

Brunner, C. (2017). *Eduthereum: A System for Storing Educational Certificates in a Public Blockchain*. Master's thesis, Universität Innsbruck.

Brunner, C., Knirsch, F., and Engel, D. (2019). SPROOF: A platform for issuing and verifying documents in a public blockchain. In *Proceedings of the 5th International Conference on Information Systems and Privacy*, pages 15–25, Prague, Czech Republic. SciTePress.

Chaudhary, S., Schafeitel-Thtinen, T., Helenius, M., and Berki, E. (2019). Usability, security and trust in password managers: A quest for user-centric properties and features. *Computer Science Review*, 33:69 – 90.

Christidis, K. and Devetsikiotis, M. (2016). Blockchains and Smart Contracts for the Internet of Things. *IEEE Access*, 4:2292–2303.

Eskandari, S., Barrera, D., Stobert, E., and Clark, J. (2015). A First Look at the Usability of Bitcoin Key Management. In *NDSS Workshop on Usable Security (USEC) 2015*, number February, San Diego.

Faber, B., Michelet, G., Weidmann, N., Mukkamala, R. R., and Vatrapu, R. (2019). BPDIMS: A Blockchain-based Personal Data and Identity Management System. In *Proceedings of the 52nd Hawaii International Conference on System Sciences*, volume 6, pages 6855–6864, Grand Wailea.

Fahl, S. (2016). *On the importance of ecologically valid usable security research for end users and IT workers*. PhD thesis, Hannover: Gottfried Wilhelm Leibniz Universität Hannover.

Francisco, K. and Swanson, D. (2018). The Supply Chain Has No Clothes: Technology Adoption of Blockchain for Supply Chain Transparency. *Logistics*, 2(1):2.

Fromknecht, C., Velicanu, D., and Yakoubov, S. (2014). CertCoin: A NameCoin Based Decentralized Authentication System. Technical report, MIT, Cambridge, MA, USA.

Gauravaram, P. (2012). Security analysis of salt || password hashes. In *International Conference on Advanced Computer Science Applications and Technologies (ACSAT)*, pages 25–30, Kuala Lumpur, Malaysia. IEEE.

Gräther, W., Augustin, S., Schütte, J., Kolvenbach, S., Augustin, S., Ruland, R., Augustin, S., and Wendland, F. (2018). Blockchain for Education: Lifelong Learning Passport. In *Proceedings of 1st ERCIM Blockchain Workshop 2018*, Amsterdam. European Society for Socially Embedded Technologies (EUSSET).

Knirsch, F., Unterweger, A., and Engel, D. (2019). Implementing a Blockchain from Scratch: Why, How, and What We Learned. *EURASIP Journal on Information Security*, 2019(2):1–14.

Kshetri, N. (2017). Blockchain's roles in strengthening cybersecurity and protecting privacy. *Telecommunications Policy*, 41(10):1027–1038.

Larimer, D. (2017). EOS.IO Technical White Paper. Technical Report ii.

Madhusudan, A., Symeonidis, I., Mustafa, M. A., Zhang, R., and Preneel, B. (2019). SC2Share: Smart Contract for Secure Car Sharing. In *International Conference on Information Systems Security and Privacy*, Prague, Czech Republic.

Manohar, A. and Briggs, J. (2018). Identity Management in the Age of Blockchain 3.0. In *CHI 2018 workshop on HCI for Blockchain*.

Mengelkamp, E., Notheisen, B., Beer, C., Dauer, D., and Weinhardt, C. (2018). A blockchain-based smart grid: towards sustainable local energy markets. *Computer Science - Research and Development*, 33(1):207–214.

Nakamoto, S. (2008). Bitcoin: A Peer-to-Peer Electronic Cash System. Technical report.

Orman, H. (2018). Blockchain: The Emperor's New PKI? In *IEEE Internet Computing*, volume 22, pages 23–28. IEEE.

Petralito, S. (2019). *Current challenges in HCI-research: quantifying open experiences, warranting data quality, and developing standardized measures*. PhD thesis, University_of_Basel.

Shamir, A. (1979). How to Share a Secret. *Communications of the ACM (CACM)*, 22(11):612–613.

Sovrin Foundation (2018). Sovrin : A Protocol and Token for Self- Sovereign Identity and Decentralized Trust. Technical Report January.

Won, J., Singla, A., Bertino, E., and Bollella, G. (2018). Decentralized Public Key Infrastructure for Internet-of-Things. In *2018 IEEE Military Communications Conference (MILCOM)*, pages 907–913, LAX Marriott. IEEE.

Wood, G. (2017). Ethereum: A Secure Decentralised Generalised Transaction Ledger. Technical report, Ethereum.

Wüst, K. and Gervais, A. (2017). Do you need a Blockchain. Technical report, International Association for Cryptologic Research.