"I Can't Believe It's Not Custodial!" Usable Trustless Decentralized Key Management

Tanusree Sharma*
tsharma6@illinois.edu
Informatics, University of Illinois at Urbana-Champaign
USA

Vivek C Nair*
vcn@berkeley.edu
University of California, Berkeley
USA

Henry Wang
henryw4@illinois.edu
University of Illinois Laboratory High School
USA

Yang Wang
yvw@illinois.edu
Information Sciences, University of Illinois at Urbana-Champaign
USA

Dawn Song
dawnsong@berkeley.edu
University of California, Berkeley
USA

ABSTRACT

Key management has long remained a difficult unsolved problem in the field of usable security. While password-based key derivation functions (PBKDFs) are widely used to solve this problem in centralized applications, their low entropy and lack of a recovery mechanism make them unsuitable for use in decentralized contexts. The multi-factor key derivation function (MFKDF) is a recently proposed cryptographic primitive that aims to address these deficiencies by incorporating commonly used authentication factors into the key derivation process. In this paper, we implement an MFKDF-based Ethereum wallet and perform a user study with 27 participants to directly compare its usability against traditional cryptocurrency wallet architectures. Our results show that MFKDF-based applications outperform conventional key management approaches on both subjective and objective metrics, with a 37% higher average SUS score ($p < 0.0001$) and 71% faster task completion times ($p < 0.0001$) for the MFKDF-based wallet.

CCS CONCEPTS

• **Security and privacy → Usability in security and privacy**.

ACM Reference Format:
Tanusree Sharma, Vivek C Nair, Henry Wang, Yang Wang, and Dawn Song. 2024. "I Can't Believe It's Not Custodial!" Usable Trustless Decentralized Key Management. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3613904.3642464

*Two authors contributed equally to this research.

CHI '24, May 11–16, 2024, Honolulu, HI, USA

ACM ISBN 979-8-4007-0330-0/24/05
https://doi.org/10.1145/3613904.3642464

> "Cryptography turns a security problem into a key management problem."
>
> – Murphy's first law of cryptography [16]

1 INTRODUCTION

For decades, key management has been a known hard problem in the field of usable security, with classic studies repeatedly demonstrating the difficulty users face with understanding and securely handling cryptographic keys [57, 60, 68]. In most centralized systems, practitioners turn to password-based key derivation functions (PBKDFs) as an imperfect but widely accepted key management solution. Today, PBKDFs are used in a wide variety of popular operating systems [15, 20], network protocols [43, 44], and applications [2, 25].

Recently, the rise of blockchain-based cryptocurrencies has created a new class of decentralized key management challenges that PBKDFs cannot readily address. Passwords are highly insecure as a sole authentication factor[3, 28, 38], and while centralized platforms can supplement passwords with multi-factor authentication (MFA), decentralized applications, such as non-custodial cryptocurrency wallets, usually lack a trusted authority to verify credentials as required by the most common forms of MFA. Moreover, the lack of a central recovery mechanism leaves users no way to recover their key if they forget their password, which they often do [28].

Given the infeasibility of using PBKDF for decentralized key management, most cryptocurrency wallets today have reverted to a more traditional form of key management in which users are tasked with manually storing and managing their key in the form of a key file or mnemonic phrase. The cumbersome and unfamiliar nature of these mechanisms has increasingly driven users to instead adopt centralized custodial wallets, contributing to the centralization of cryptocurrency holdings [70] and reducing public trust in blockchain technology when large custodians collapse [13].

The multi-factor key derivation function (MFKDF) [51] is a novel cryptographic primitive that builds upon password-based key derivation by incorporating the entropy of existing, commonly used authentication factors into the key derivation process. MFKDF has specifically been proposed for use in non-custodial cryptocurrency wallets [50], providing the recovery, portability, and resilience of a

custodial wallet while remaining secure, trustless, and decentralized.

In this work, we implement and test the first functional Ethereum wallet based entirely on MFKDF via a user study of novice and advanced cryptocurrency users. We do so by designing and implementing two further reference wallets, based on the typical key management functionality of existing custodial and non-custodial wallets respectively. These wallets otherwise present an identical UI and UX to the MFKDF-based implementation, isolating the key management mechanism from confounding factors like branding and design. We aim to answer two key research questions:

(1) Does MFKDF address the main usability concerns associated with non-custodial wallet key management?
    Does MFKDF improve the wallet creation experience?
    Does MFKDF improve the wallet recovery experience?
    Does MFKDF improve the wallet portability experience?
(2) Does MFKDF improve the risk, security, and trust perception of non-custodial cryptocurrency wallets?

To answer RQ1, we conducted a usability study with 27 participants (23 had used crypto wallets before the study and the other four had not). We asked them to use cryptocurrency wallets with three distinct architectures. We found that the MFKDF-based wallet design significantly outperformed the non-custodial control wallet, and performed on par with the custodial control wallet, according to several objective and subjective measures of usability. For example, we observed a 37% higher SUS score ($p < 0.0001$) and 71% faster task completion times ($p < 0.0001$) for the MFKDF-based wallet. Regarding RQ2, we asked participants to think aloud while performing different various tasks with the three wallets. Participants expressed a subjective preference for the MFKDF-based wallet, which effectively addresses many of the prominent usability issues in conventional wallets, and provides a *"best of both worlds"* key management solution that is trustless, decentralized, and user-friendly.

Contributions:

- We provide open-source reference implementations for custodial and non-custodial Ethereum wallets (§3.1).
- We implement a fully functional open-source Ethereum wallet based on multi-factor key derivation (§3.2).
- We present a user study evaluating a variety of wallet architectures in a head-to-head comparison (§3.4).
- The MFKDF wallet demonstrates favorable performance on several objective and subjective metrics with rich qualitative insights on usability, security, and trust perceptions (§4).

2 BACKGROUND & RELATED WORK

2.1 Key Management

Usable key management has been a persistent challenge in the field of Human-Computer Interaction (HCI) and usable security research. Several studies have investigated the usability aspects of key management, aiming to improve the user-friendliness and effectiveness of cryptographic systems.

Perhaps the most well-known study in this area is "Why Johnny Can't Encrypt" (1999) [68], in which 12 users were tasked with using PGP 5.0 to send encrypted emails to each other. The paper identifies key management as a significant obstacle; most users failed to grasp the public-key cryptography model and thus committed errors that gravely undermined their own security. While other aspects of usable security have greatly improved, key management has continued to be a notable difficulty in several subsequent studies [57, 60]. Recently, the key management problem has been particularly pronounced in decentralized systems, such as non-custodial cryptocurrency wallets. Unlike centralized systems, which can take advantage of trusted servers for key management, decentralized systems largely still rely on a more hands-on key management that many users may find cumbersome.

When using asymmetric cryptography, the key management problem can be divided into public key management and private or secret key management. The problem of public key management essentially involves securely linking cryptographic keys to known user identities and often involves highly application-specific solutions. For example, in centralized applications, public keys are often managed by a public key infrastructure (PKI) with one or more certificate authorities [14]. In cryptocurrencies, public keys are managed by the use of deterministic wallet addresses. By contrast, private or secret key management essentially involves securely storing and accessing a series of bits constituting a cryptographic key, with a handful of techniques being used across nearly all applications (e.g., key derivation, see §2.3).

In this paper, we aim to address the problem of private key management in the context of decentralized systems, and therefore will use the terms "key management" and "private key management" synonymously moving forward. While cryptocurrency wallets are a timely and representative example of a consequential decentralized private key management task, the findings of this work may be generally applicable to private key management in other decentralized applications, as discussed further in §5.5. However, public key management is not a focus of this paper and continues to require highly distinct, application-specific considerations at the time of implementation.

2.2 Cryptocurrency Wallet Usability

2.2.1 Wallet Architectures. Currently, there are three main paradigms for cryptocurrency asset management: custodial wallets, non-custodial hardware wallets, and non-custodial software wallets. Using any of these approaches implicates a variety of security and usability consequences.

Custodial cryptocurrency wallets, in which a third-party service provider is trusted to store and manage private keys on behalf of users, remain one of the most popular ways for novice users to hold crypto assets due to their relative ease of use. The three largest centralized platforms, Binance [18], Coinbase [22], and Kraken [46], together account for nearly $20 billion in daily trading volume [8], orders of magnitude larger than the largest decentralized exchanges [4].

Committee-based wallets have been proposed as an alternative to custodial wallets, whereby a private key is secret shared with a committee of nodes, at least some threshold of which are presumed to be honest [41, 73]. In practice, however, these solutions do not provide the security properties of a fully decentralized approach.

The nodes constituting a trusted committee are often homogeneous in design and control, and thus subject to common vulnerabilities or influences. Thus, we consider these wallets, in their current form, to be an instance of custodial wallets, whereby the custodian is a joint entity controlled by a committee of nodes rather than a single party.

Non-custodial hardware wallets, such as those offered by Ledger [48] and Trezor [9], use a purpose-built chip for storing and managing keys, making them a highly secure mode of crypto asset management, resistant to most software vulnerabilities. However, they are also amongst the least user-friendly options, due to their relatively high up-front cost, cumbersome physical interface, and intrinsic limitations on the number and types of supported cryptocurrencies [62].

Non-custodial software wallets, such as MetaMask [5], allow users to directly manage and store their keys, usually in the form of a key file or BIP39 seed phrase. These keys are then used on the client side to sign and approve cryptocurrency transactions, without the use of specialized hardware to ensure their security. Thus, when correctly implemented, their security and usability properties essentially reduce to an instance of the classic usable key management problem.

As such, software-based non-custodial wallets are a perfect vehicle for understanding and evaluating potential improvements in the field of usable key management, and are the main focus of this work. We begin by discussing a number of prior works that highlight the known wallet usability challenges.

2.2.2 User Studies. Non-custodial cryptocurrency wallets have generally been difficult for novice users to grasp, with intimidating features such as manual key management and seed phrase recovery [27]. There has been extensive literature on understanding the usability challenges of crypto wallets and transactions from various angles and approaches. In many of these studies, inexperienced users were unable to perform basic transactions due to a lack of technical knowledge [35]. With a qualitative analysis of 6,859 user reviews, Voskobojnikov et al. [66] identified both general and domain-specific UX issues of mobile crypto wallets. The wallet initialization process was found particularly challenging by some users. Other issues included a lack of guidance on gas fees, transactions, etc. These usability issues also exacerbated users' misconceptions about crypto wallets and assets, e.g., the assumption that their funds were tied to mobile apps.

The Foundation for Interwallet Operability surveyed 200 crypto users on wallet usability [30]. 55% of them had more or less concern about their transactions, with public address accuracy (35%) as the most popular one. Emotionally, many users were nervous about transactions. In another survey of 395 crypto-asset users of different levels of experience in crypto [12], people's security behaviors were probed, which were often related to usability issues of the crypto infrastructure. For example, rookie users appeared to refrain from managing their own private keys and often rely on third parties, which could be attributed to the unusable keys management process.

Fröhlich et al. conducted semi-structured interviews to understand how the onboarding process worked in mobile crypto wallets, and how it could be improved for novice users [33]. Users' expectations about the onboarding process turned out to be short, skippable, focused on the most relevant features, well integrated into the app, and lightweight with concise information. Some users even desired no onboarding, expecting wallet apps to be intuitive and self-explanatory.

In a user experiment exploring challenges first-time cryptocurrency users faced when using crypto wallets [34], participants were asked to conduct three tasks, i.e., account registration, the first acquisition of Bitcoin, and spending them in an online shop. The experiments revealed that user interfaces of popular wallets were not optimized for novice users, with primary actions difficult to access. One crypto-specific challenge was that of dealing with cryptocurrency itself, which required mental effort from users. For instance, sub-comma amounts were regarded as hard to deal with. Similarly, Moniruzzaman et al. evaluated the usability of five popular crypto wallets with a controlled experiment [49]. The failure rate on mobile wallets was lower than their desktop counterparts, but the rate of usage of mobile-based wallets was lower than that of desktop-based wallets.

2.2.3 Pain Points. Across all of the studies exploring non-custodial cryptocurrency wallet usability, a few key pain points are evident that span nearly all surveyed implementations:

- *Portability.* Moving a cryptocurrency wallet from one device to another is a consistent difficulty. This process typically requires either correctly storing and typing a long pseudo-random keyphrase or copying a key file from one device to another without inadvertently exposing it.
- *Recovery.* Recovering from a lost factor, such as a forgotten password, is an additional pain point across many implementations. In most cases, there is no way to recover from forgotten core factors, and notorious instances of lost funds due to forgotten passwords are widespread [54, 56].
- *Resilience.* Finally, many users are concerned with the resilience of their cryptocurrency assets to the failure of a single hardware or software component. Unlike custodial wallets, which have built-in redundancy, a single broken device or lost file could result in a total loss of funds.

2.2.4 Consequences. By taking advantage of trusted centralized infrastructure, custodial wallets do not present many of the usability challenges of decentralized wallets. They are architected around resilient cloud storage systems, can be accessed over the internet from any device in the world, and can use centralized infrastructure, such as emails and SMS, for usable account recovery. It is no wonder that many cryptocurrency wallet users, particularly novice users, have chosen to leave their assets in the care of a centralized custodian rather than utilizing non-custodial solutions.

Unfortunately, this centralization of what are otherwise supposedly decentralized assets has also led custodial wallets to be overrepresented in their share of major security incidents and fraud. From the infamous collapse of Mt. Gox in 2014 [7] to the recent downfall of FTX [6], custodial services have proven notoriously prone to catastrophic failure, diminishing the public reputation of blockchain technologies as a whole.

This problem can be rectified by improving the usability of non-custodial solutions to achieve parity with custodial wallets and

promote their use by novice users. Thus, we are motivated to explore solutions from the broader field of key management that may be applicable to decentralized applications.

2.3 Authentication and Key Derivation

2.3.1 Password-Based Key Derivation. In conventional centralized systems, password-based key derivation has been a vital tool for addressing the issue of usable key management. Fundamentally, a password-based key derivation function serves as a deterministic one-way function, converting a password, salt, and optional configuration parameters into a fixed-length key. Thus, users are no longer burdened with safely storing and managing a cryptographic key, and instead only need to remember a password, a task that most users are more familiar with. Most modern PBKDFs also feature a degree of intentional computational inefficiency to increase the difficulty of performing brute-force attacks.

Today, password-based key derivation functions like PBKDF and PBKDF2 [45] are used in a wide variety of centralized systems, including use in the Windows [20] and iOS [15] operating systems, LastPass [2] and Dashlane [25] applications, and WPA [43] and WPA2 [44] wireless protocols. The adoption and usability of these systems has greatly benefited from the ability of users to interact with them using passwords rather than managing cryptographic keys.

Despite the widespread success of password-based key derivation in centralized systems, PBKDFs have seen limited adoption in decentralized applications, such as cryptocurrency wallets. Indeed, there are several reasons why using PBKDFs in decentralized applications could have negative consequences, which we explore in the following sections.

2.3.2 Multi-Factor Authentication. While PBKDFs are effective at binding a user's secrets to their password, they are generally insufficient to protect a user's account. Due to the well-known insecurity of passwords as a sole authentication factor [29, 39] and their susceptibility to attacks such as credential stuffing [3], multi-factor authentication (MFA) is typically used in conjunction with password-based key derivation to strengthen centralized applications.

Popular authentication factors for MFA include "soft tokens" like HMAC-based One-Time Password (HOTP) [64] and Time-based One-Time Password (TOTP) [65], "hard tokens" like YubiKeys [71], and Out-of-Band Authentication (OOBA) factors like email and SMS [42]. While these factors have no effect on password-derived keys, centralized applications benefit from the ability to validate these factors before granting access to an account.

By contrast, decentralized applications typically lack the infrastructure to protect secrets using multi-factor authentication. Thus, a PBKDF-based decentralized application would be entirely reliant on passwords as a sole factor, and would be susceptible to credential stuffing and offline brute-force attacks. In applications like cryptocurrency wallets, where a compromised key could entail a significant loss of funds, this risk has correctly been seen as untenable, and has prohibited the use of PBKDFs in most decentralized contexts.

2.3.3 Account Recovery. An additional consideration made by centralized applications using password-based key derivation is account recovery in the event of a forgotten password. Without additional provisions for account recovery, systems using PBKDFs may experience a complete loss of user data in the event of a lost password. Therefore, key management standards have been developed to facilitate account recovery in systems using derived keys. In particular, the commonly-used NIST SP 800-57 [17] standard suggests the use of a master key stored in a central hardware security module (HSM) to recover account data in the event of a lost password.

Once again, we find that the solutions used by centralized applications to resolve issues with PBKDFs are not applicable to decentralized applications, which cannot use a central master key to recover a lost password. Because passwords are, in fact, often forgotten by end users [1], this risk is again considered untenable, as using PBKDFs in cryptocurrency wallets could result in a complete loss of access to stored funds in the event of a forgotten password.

2.3.4 Multi-Factor Key Derivation. Thus far, we have established that password-based key derivation has had a significant impact on the usability of centralized applications, but is unsuitable for decentralized applications, such as cryptocurrency wallets, due to the difficulty of supporting multi-factor authentication and account recovery. The Multi-Factor Key Derivation Function (MFKDF) [51] is a recent improvement over PBKDFs that incorporates multiple authentication factors into the key derivation process. MFKDF aims to address the lack of MFA support and account recovery in PBKDFs while being fully compatible with decentralized applications and supporting popular MFA factors.

The MFKDF specification contains two major architectural components. The first is a set of "factor constructions," which convert *factor witnesses*[1] and public parameters into static key material. The public parameters require no security assumptions and can safely be stored in the open, such as on a public blockchain. Constructions are given for many popular authentication factors, including TOTP, HOTP, OOBA, and YubiKey. These factor-specific implementations either do not involve central servers at all (e.g., TOTP, HOTP, and YubiKey), and thus do not require a trusted third party, or use end-to-end encryption to avoid implicating additional trust assumptions (e.g., OOBA). For some factors, the construction requires the trustless public parameters to be updated upon each key derivation.

The second major architectural component is the key derivation function itself, which adds a secret sharing layer to allow for key recovery in the event that a factor is lost.

By incorporating support for multi-factor authentication and account recovery, MFKDF makes it feasible, for the first time, to realize the usability advantages of PBKDFs in decentralized applications without implicating their associated security pitfalls. Using the non-custodial cryptocurrency wallet as a prototypical use case, we are motivated to explore the question of whether MFKDF presents a practical solution to usable key management in decentralized applications.

[1] *Witness* refers to the message used to authenticate, such as a 6-digit OTP.

3 METHOD

The goal of this study is to evaluate the usability of MFKDF-based decentralized key management in the context of a non-custodial cryptocurrency wallet use case. Specifically, we aim to present a fair comparison with existing custodial and non-custodial wallets that isolate the key management mechanism from other aspects influencing usability, such as the user interface, branding, feature set, and platform.

To this end, we implemented three open-source Ethereum web wallets using the Sepolia testnet, shown in Fig. 1. These wallets represent a typical custodial (1a) and non-custodial wallet (1b), to serve as controls, and an experimental MFKDF-based custodial wallet (1c) that we wish to evaluate.

All three wallets use a React.js frontend and Cloudflare Workers backend with identical UI components, colors, fonts, and can be accessed via a standard web browser, with the only key difference being the key management mechanism.

3.1 Control Wallet Designs

3.1.1 Custodial Control Wallet. Our custodial reference wallet uses server-side key storage and management in a distributed key-value database provided by Cloudflare. Users authenticate using an email address, password, and one-time pin, for which we chose to use TOTP (via Google Authenticator, due to its popularity). Users can recover a lost factor via their email. In these respects, the user experience is largely modeled after that of Coinbase [22], with the only difference being the lack of a Know Your Customer (KYC) identity verification process, as is the case in wallets like KuCoin [47]. During our presentation of the findings, we refer to this as *"Wallet A."*

3.1.2 Non-Custodial Control Wallet. Our non-custodial reference wallet uses client-side key storage, in the form of a password-protected JSON file, with support for recovery via a BIP39 key phrase. This key management mechanism is consistent with MetaMask [5] versions v8.0.0 through v10.11.3. One distinction of our non-custodial control wallet is the inclusion of the "key file" or private key on the same UI page as the seed phrase during account creation, which differs from the actual MetaMask implementation [5]. In MetaMask, users have to navigate through multiple layers in the security settings to retrieve the key file. However, because our streamlined implementation does not have a settings page, this feature is moved directly into the authentication process. This is particularly important in our experiment for when users want to import their wallet from a different device or recover their wallet in the event of a lost authentication factor. It is worth noting that this streamlined implementation may give this control wallet an artificial advantage in the task completion metrics by potentially reducing the overall time required. During our presentation of the findings, we refer to this as *"Wallet B."*

3.2 MFKDF Wallet Design

The MFKDF-based non-custodial wallet is implemented based on the proposed architecture of Nair and Song [50]. Specifically, rather than storing keys anywhere, users derive their keys on the client side as needed using a password and TOTP code. Only trustless, public material is stored in any location. Because no trust assumptions are associated with these public parameters, they can be stored openly on a blockchain or in IPFS, as shown in Fig. 2.

In our implementation, passwords and TOTP are used as primary authentication factors, with email OOBA allowing for key recovery in the event of a forgotten password or lost TOTP device. Thus, users can simply "log in" to the wallet with an email address, password, and multi-factor authentication, as if it were a custodial wallet. However, the wallet is in fact non-custodial and does not require trust in any centralized entity. Instead, MFKDF is used to derive their wallet key directly from their authentication factors. During our presentation of the findings, we refer to this as *"Wallet C."*

3.3 Wallet Functionality

All three of our wallet implementations support basic features such as sending and receiving cryptocurrency and viewing the history of previous transactions, using an identical interface shown in Fig. 3. None of the wallets support advanced functions, such as ERC20 tokens, NFTs, and smart contract integrations, as these features are orthogonal to the key management aspects of the wallet that we wish to evaluate.

In summary, by developing three cryptocurrency wallets from scratch, corresponding to three separate key management paradigms, we have been able to isolate the key management mechanism from all other aspects that may impact the user experience in a way that would not have been possible using existing cryptocurrency wallets. We have done so while ensuring our control wallets remain faithful, in the key management domain, to widely-used wallets like MetaMask and Coinbase.

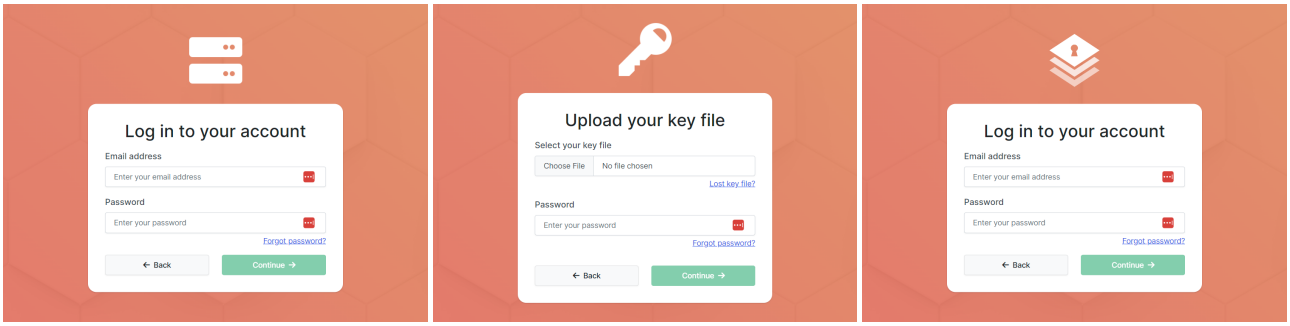
We offer our three wallet implementations as open-source research artifacts[2] for any other researchers who wish to perform a similarly well-controlled user study in the future.

3.4 User Study Protocol

We aimed to evaluate the MFKDF-based non-custodial wallet (Wallet C) for usability (in addition to security) with two control wallets (Wallets A & B). The study was reviewed and approved by our organization's Institutional Review Board (IRB), and each participant was compensated with a $30 Amazon gift card. We conducted the study online over Zoom. Below, we provide a detailed description of the recruitment process, experiment setup, and data analysis methods. The full protocol is available in the Supplementary Materials and anonymous open-source link.

3.4.1 Participant Recruitment. For this study, we recruited both novice and experienced crypto users. We defined novice users as those who may have heard of cryptocurrencies but have not traded or used them yet. We defined experienced users as those who had used crypto wallets at least once before. The screening survey included questions asking about prospective participants' experience with cryptocurrency, exchanges, and wallets, as well as demographic information such as gender identity, age, educational level, and country currently living in. The participants were recruited from various cryptocurrency channels and forums, including Discord, Twitter, and educational institution mailing lists. Of the 27

[2] https://anonymous.4open.science/r/research-wallets



(a) Custodial control wallet. **(b) Non-custodial control wallet.** **(c) MFKDF-based non-custodial wallet.**

Figure 1: Three competing representative cryptocurrency wallet architectures (a) custodial control wallet; (b) Non-custodial control wallet, (c) MFKDF based non-custodial wallet.

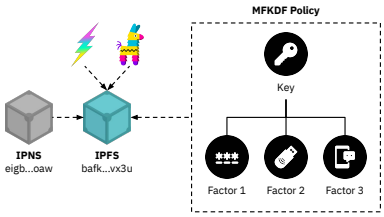


Figure 2: Network architecture of the MFKDF-based wallet.

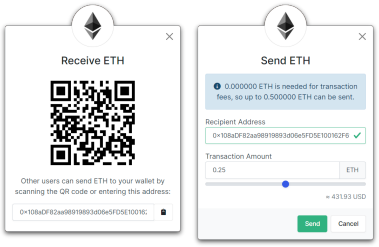


Figure 3: All three wallets support basic features such as sending and receiving ETH, and viewing transaction histories.

participants, seven were recruited through Twitter, seven through a cryptocurrency-related forum, and four through word of mouth. The remaining participants did not specify their sources of reference. While our original aim was to have 30 participants, three were unable to complete the experiment due to scheduling conflicts.

3.4.2 Pilot Study. We conducted a pilot study with four participants in order to gather feedback on our study protocol. None of them were among the 27 participants who completed the main study. During the pilot study, participants tested the MFKDF wallet, while for the control wallets, they utilized the ones available online. However, there were notable differences in the wallet installation process. Specifically, the non-custodial control (Metamask) only has a browser extension, which required additional time and cognitive load, resulting in an onboarding process that was incomparable to the other two wallets (MFKDF and Coinbase). Overall, our pilot results showed that the different platforms and installation methods had an untenable confounding effect on perceived usability.

Since our study is focused on key management, encompassing wallet account creation, account importing, and account recovery workflows, we decided to develop controlled versions of all three wallets based on the pilot observations. Accordingly, we examined the onboarding workflow of the control wallets (Coinbase and Metamask) before proceeding with development. We acknowledge potential challenges concerning external validity, as the user interfaces may not be precisely identical to the latest versions of Coinbase and MetaMask. Nevertheless, for our experiment, we aimed to isolate specific features of the wallets. Consequently, we developed the custodial and noncustodial wallets with the aim of being representative of the user flow of MetaMask and Coinbase. The three wallets shared the same branding, user interface, and features, thereby eliminating extraneous variables. Specifically, the following measures were taken:

- Homogeneous Platform: All three wallets were web-based, removing the platform and installation variables.
- Homogeneous Branding: Perceptions of the wallets were not influenced by recognition of their respective brands.
- Homogeneous Interfaces: Because the three wallets used the same fonts and UI component library, the usability results were solely based on key management aspects.
- Homogeneous Features: All wallets presented an identical set of features, equilibrating cognitive load.
- Detailed Analytics: Click-by-click event analytics were collected for all three wallets being tested.

While the possibility of introducing biases in the study design cannot entirely be eliminated, the control measures listed above ensure that to the extent possible, observed differences in wallet usability correspond only to differences in the underlying key management mechanisms, as all other implementation aspects are completely homogenous.

We employed Google Authenticator (TOTP) as one of the factors for the custodial wallet and MFKDF wallet. However, three out of four pilot participants did not have Google Authenticator installed, resulting in additional time during the study and creating a potential distraction for participants. Consequently, we updated the

study invitation instructions to include the requirement of having Google Authenticator installed beforehand. Further, pilot participants expressed the need for clarifications on certain technical terms throughout the wallet interface while testing. They expected to have quick text descriptions for terms such as "TOTP" and "seed phrase." Additionally, we included a pre-survey question to asses participants' current usage of TOTP, in order to better interpret the task results. We also updated the instructions within the wallet interfaces to specify the use of only the Google Authenticator app for scanning the QR code for TOTP. Finally, we used the pilot process as an opportunity to identify and eliminate any unknown bugs in the wallet implementations. At the conclusion of the pilot study, we observed that users were consistently able to utilize all three of our controlled wallets without encountering implementation flaws or glitches.

3.4.3 *Main Study Setup.*

Pre-Study Survey. We asked our participants to complete a pre-study survey and install Google Authenticator. We designed the pre-survey to collect information on participants' technical backgrounds, self-efficacy in managing crypto wallets, and experience with crypto wallets. We also asked participants about their understanding of custodial and non-custodial crypto wallets, and whether they recognized the differences between them. Participants were directed to watch an explanatory video on wallet types and then we assessed their knowledge afterward with multiple-choice questions. For instance, we posed questions, *"Do you know what types of crypto wallets you use currently?"* with options of custodial, non-custodial, and other, and *"Do you know know the difference between Custodial and Non-custodial crypto wallets?"* Following the video, we verified their understanding by asking *"Which following are correct to the best of your knowledge? [Please select all that apply]."* The detailed protocol description can be found in our main wallet repository.[3] We then asked participants about their past negative encounters with crypto wallets, their perceptions of security, trustworthiness, and the associated risks related to custodial and non-custodial wallets, using Likert scale ratings, multiple-choice questions, and open-ended questions.

Tasks & Scenarios. Our study followed a within-subject design, whereby all 27 participants tested all three wallets, with identical instructions given for each wallet. We randomized the order of wallets for each participant, mitigating any ordering effect. Our expectation was that wallets A and C would exhibit similar usability results due to their nearly identical UX; nevertheless, both wallets were included to validate this hypothesis. Wallet B was included as a point of comparison to wallets A and C as an additional contribution. We did not tell users that Wallet C (the MFKDF wallet) was the main focus of the study to avoid any social desirability effect. The instructions for each wallet were as follows:

First, each participant was asked to watch a 60-second explainer video detailing the features of the wallet. The video for each wallet contained a bland recitation of facts about the wallet, following a similar set of talking points for each wallet to avoid biasing participants. Participants were then asked knowledge questions to assess if they understood the basic principles of the wallet; namely, if the wallet was centralized or decentralized, and custodial or non-custodial.

Participants were then given a series of tasks to complete, which were the same for each wallet (listed below). They were asked to think aloud and answer questions during the experiment. With their permission, we captured audio and video recordings of the participants performing the tasks for later analysis. We also asked several questions about their experience after each task.

- Task T1: Configuration/Account Creation - Creating a new "account" (or address) within the wallet.
- Task T2: Import/Login - Participants were asked to assume they are traveling and they don't have their primary device and need to access their existing wallet from their mobile phone or another personal device.
- Task T3: Account Recovery - Participants were asked to assume they forgot a primary authentication factor, such as a password, and need to recover access to their wallet.

For Task T3, in wallets A and C, participants were in equal proportions instructed to assume either forgetting their password or no longer having access to their device used for TOTP. Similarly, for Wallet B, participants were in equal proportions asked to assume either forgetting their password or losing their key file. We note that for security reasons, participant credentials were not retained by the researchers in plaintext for any of the three wallets. After each wallet test, all 27 participants completed the System Usability Scoring (SUS) survey to reflect on their experience.

Data Analysis. We performed both qualitative and quantitative analyses on our collected data from the pre-study survey, transcripts of the audio recording of study sessions, task performance, and observations, and the SUS survey. We coded the transcripts of the conversations during the experiments as well as observational notes. We (two) researchers independently read through the transcripts of 20% of the interviews, developed codes, and compared them until we developed a consistent codebook. We met regularly to discuss the coding and agreed on a shared codebook before coding the remaining data. After completing the coding for all interviews, both researchers spot-checked the other's coded transcripts and did not find any inconsistencies. We grouped lower-level codes into sub-themes and further extracted main themes. Finally, they organized codes into higher-level categories. We also used XMind [69], a mind-mapping tool, to arrange and organize codes and corresponding quotes into a hierarchy of themes. After several iterations of analyses, we arrived at the current themes of the findings. We use observational notes and participant quotes to illustrate our points. All quotes have been anonymized to protect the privacy of the participants. For quantitative analyses, recordings of the participants performing study tasks were analyzed to measure task success rates and task completion time.

We computed descriptive statistics of the pre-study survey data to gain insights into participants' pre-existing perceptions of usability, security, trustworthiness, and risks associated with crypto wallets and their current use of different crypto wallets, providing a baseline. For task time analytics, we performed descriptive statistics as well as inferential tests (two-tailed paired t-tests) to measure the significance of differences. For System Usability Scoring (SUS), we

[3] https://github.com/multifactor/research-wallets

calculated the result by converting the original scores to a scale of 0-100 [19].

4 RESULTS

4.1 Participant Background and Demographics

We had a total of 27 participants (21 male, 5 female, 1 non-binary). Table 1 summarizes our participants' demographics. Seven of them were students, 5 of them were operations or product or logistic managers, 4 of them were teachers or educators, 4 of them were software engineers, 2 of them were in web3 product development, and the rest were investors, management consultants, or research & development professionals. The majority (14 out of 27) had a Master's degree, 11 of them had a Bachelor's degree, remaining 2 had a Doctoral degree and a Diploma in Accounting, respectively. 14 out of 17 had college degrees or work experience in computer science, software development, web development, or similar technical fields. The majority (16 of them) were in the 25-34 age range; 6 of them were 35-44 years old, 5 of them were 18-25 years old, and the rest of them were 35-44 years old. 17 out of 27 had a technology background. 23% of participants did not own cryptocurrency while 77% did. 74% of the total participants had used crypto exchanges, such as Coinbase, Binance, Uniswap, Sushiswap, BitPanda, Bittrex, Crypto.com, Curve, Huobi, OKEX, Kucoin, Changelly, Gate.io, Purecoin, etc. 66% mentioned using web or browser-based crypto wallets, including Coinbase, Binance, MetaMask, and Trust Wallet, Unisat, Sui Wallet, Yoroi and 59% of the total participants indicated that they had used mobile crypto wallets. 13 participants used a crypto wallet for more than 2 years, five for 1-2 years; three had used wallets for about one month, and two had used wallets for less than one week, four chose others.

4.2 Participants' Pre-Task Security, Trust & Risk Perception

Among the participants, 53% reported using non-custodial wallets, while 47% reported using custodial wallets. Of those, 37% stated that they used both types of wallets interchangeably. 74% of participants who mentioned using custodial or non-custodial wallets also demonstrated knowledge about the difference between the two, both in multiple-choice and in-depth responses to open-ended queries. For instance, P21 explained *"I delegate handling of my private key in a custodial wallet mechanism, I take it on my side when non-custodial."* Furthermore, 24 (out of 27) participants were able to correctly answer multiple choice knowledge questions after watching the video[4] on the differences between custodial and non-custodial. Notably, 5 of them responded to not having known the differences prior to watching the video.

When it comes to securing their crypto wallets, the majority of participants expressed a lack of confidence, while 40% claimed to be competent in security but still had concerns. 25% (7 out of 27) mentioned being most skilled in protecting keys and crypto wallets on their own, having crypto wallet experience ranging from one year to over two years. Additionally, around 48% of the participants shared past negative experiences related to crypto wallets. Common issues included forgetting passwords, losing devices with logged-in wallets, and experiencing financial losses due to fraud or attacks, with subsequent difficulties in wallet recovery.

In terms of prior user experience, participants provided neutral ratings on a 5-point Likert scale for both custodial (mean: 3.3) and non-custodial (mean: 3.0) wallets. There was no statistically significant difference between the users' levels of prior experience with each type of wallet according to a two-tailed paired t-test ($p = 0.51$).

In terms of security perception based on prior and current experiences, participants rated custodial wallets lower in security (mean: 2.7) compared to non-custodial wallets, which were perceived as relatively more secure (mean: 3.3). This difference was highly statistically significant ($p = 0.01$). Similarly, non-custodial wallets were deemed more trustworthy (mean: 3.2) compared to custodial wallets (mean: 2.6), which was moderately statistically significant ($p = 0.04$).

With respect to risk, the general narrative of participants throughout their responses is that the risk within custodial wallets (mean: 3.4) resides in the 3rd party, such as the potential of an exchange being hacked, while the risk of the non-custodial wallets (mean: 3.1) originates in the user.

Thus, while the presence of different risks was acknowledged, there was no statistically significant difference in the overall level of perceived risk ($p = 0.57$).

4.3 Reported vs. In-Task Key Management Practice

We observed variations in the reported practices of managing different credentials for crypto wallets compared to the strategies employed during the task. Table 2 in the Appendix reflects the reported practices (current or intended) and behaviors related to managing credentials, such as the seed phrase and key file for Wallet B, and the password for all three wallets. During the task, participants saved the seed phrase in text files, notepads, iCloud notes, and documents. However, during the exit interviews, they mentioned employing various practices. The predominant methods for storing seed phrases were Google and Apple cloud drives, with only a few participants mentioning physical paper or pen drives. In contrast, participants appeared to be more cautious when handling key files, mentioning the use of pen drives, encrypted cloud storage, micro-SD cards, and similar methods. Regarding passwords, participants reported either memorizing them or using password managers, with two participants specifically mentioning LastPass.

Participants further shared their distinctive methods for saving seed phrases. P17 illustrated *" I draw its actual representation, such as a banana or a kitten. Then I take a photograph for backup. Even if someone sees it, deciphering it is challenging."* P21 mentioned another unique way to save seed phrase *"I assign number to each word of seed phrase, converting 24 words into numbers. The challenge is remembering number corresponds to word, but I've memorized it due to frequent use."* P1 employs a unique tactic for safeguarding credentials by nesting it within multiple subfolders in a drive. P1 explained-*"if my Google account compromised, it's a long tree actually to find that folder first of all, and I actually use different types of names, for example, education, homework or something unrelated."*

[4] https://www.youtube.com/watch?v=0eq59zTGZFc

4.4 Answering RQ1: MFKDF (Wallet C) in Addressing Usability Concerns of Non-Custodial Wallet Key Management

4.4.1 Is it Really Non-Custodial? Participants (P0, P7, P10, P11, P12, P14, P24, P25), including experienced crypto wallet users, expressed confusion regarding the decentralization of Wallet C and whether it was actually noncustodial, particularly due to its reliance on email and TOTP, factors that have never previously been seen in non-custodial wallets. P12 highlighted this confusion by stating, *"I guess conceptually [Wallet C] is not relying on some centralized server. I have a subjective perception of it as being more secure ... operations like Google Authenticator or Gmail OTP are happening on the client side, which gives me a perception of better security, but I guess in the practice of using it day-to-day, I could imagine myself kind of forgetting that or not really having that presently on the mind because experiences are so very similar to a custodial wallet."* Similarly, P10 expressed a lack of understanding about the non-custodial nature of Wallet C by saying, *"Your email provider, like Google, somehow stores your password. I'm not sure about this part of how this is non-custodial."* P11, P14, P24, P25 expressed the curiosity regarding the technology behind how Wallet C operates in terms of generating new keys and seed phrases. P14 said *"I'm not very sure in terms of how those keys are generated. I'm curious about what happens behind the scene, and what happens with my seed phrase or keys."* P25, looking from a company's viewpoint, commented, *"Given the diverse user base of crypto, many without cryptographic knowledge, it'll be interesting to see how the company convey this information on Wallet C's non-custodial nature to its users."*

4.4.2 Is Recovery a Concern? Participants (P2, P4, P5, P11, P13, P15, P16, P21, P23) expressed a tendency to forget their passwords frequently and relied on various recovery methods, such as text messages, emails, and time-based one-time passwords (TOTP), in their current practices. Interestingly, a few participants (P2, P3, P13) even forgot their passwords while performing Task 2 of Wallet C, which involved logging in to wallets from a different device (mobile phone or personal device) than the one they had used before (laptop). P2 described his approach as follows: *"I don't usually remember passwords at all. Recovering my password for C was quite simple — I just needed to verify my email and then complete the authentication process using an authenticator."* Similarly, P13 stated that he didn't know any of the passwords he used for various sites and preferred to rely on frequent password recovery as his normal behavior. In his own words *"I've done it many times before. I don't know most of my passwords; only a few are stored in password apps or saved in my Google account, which I have access to. I believe this is a healthy course of action. I prefer changing passwords regularly."* This behavior aligns with previous research on usability and memorability, as forgetting passwords is common and can lead to significant financial losses [36, 61]. When passwords are difficult to remember, users may resort to compromising security measures, such as password reuse or insecure storage practices [24], in a similar fashion to our participants who use non-custodial crypto wallets due to the absence of usable recovery options (see §4.3).

4.4.3 Usability Regarding UI Workflow Familiarity. Participants (P2, P5, P7, P8, P11, P13, P19, P20), particularly novice users, felt their experience of Wallet C was similar to their day-to-day application usage. P11 said *"[Wallet C's onboarding] was the same process I use regularly. On my phone, I sometimes use a different factor like my face recognition and verification over text message."* By contrast, P7 highlighted the difficulty of creating an account in Wallet B compared to Wallet C, emphasizing the significant attention required just to understand the new factors (e.g., seed phrase, private keys) and workflow. She mentioned, *"There was the key download and requirement to keep it, along with a secret phrase. At first, I thought I could create my own phrase, but it turned out that I had to use the provided phrase. It took me some time to comprehend the process. I had to manually copy and input all the numbers and words. Overall, I was not familiar with the process at all."*

Some experienced participants also found Wallet C's usability to be advantageous. They expressed the view that it is crucial for new users to easily manage assets by lowering the cognitive load. For example, P13 expressed this viewpoint, stating, *"While security is important to me, I believe that usability holds even greater significance for new users. As more assets accumulate, I think there will be a certain point where in the end comfort of being in the space, they might switch to more security versus usability as I did throughout my experience."* This lack of familiarity with seed phrases has been identified as a barrier to onboarding new users in previous research, as it can lead to frustration and disengagement and even poses risks, related to loss of the device and seed phrase [67]. Additionally, this can result in security vulnerabilities, particularly for blind users who often use public library computers and screen readers to access web application components [72].

4.4.4 Activity Patterns: Frequency & Purpose of Using Crypto Wallets. Some participants (P1, P4, P10, P11, P22, P23, P26) also evaluated the usability of crypto wallets based on their frequency and specific use cases. P10, an experienced crypto wallet user, mentioned using both custodial and noncustodial wallets, but primarily relies on custodial wallets for conversion and exchanges, and indicated that he would not use non-custodial wallets for regular use. He mentioned keeping only a small amount in noncustodial wallets, using them occasionally for transactions with certain decentralized applications (dApps), and then eventually burning the wallet. In his words, *"If I am checking my wallet every single day, then C, but if I'm only doing it once a month, less activity, then I'm probably choosing B. I'm not using A if I have the option of C."* Similarly, P4, shared his experience of wallet usage for asset management, saying that *"I've used 100% cold storage. when I've bought from exchanges, there's a wallet that exists that I've only used to transfer to cold storage."* He mentioned doing a little bit of staking and used MetaMask, not for its usability, but for its transparent documentation for users. Furthermore, P26 expressed that *"I would use Wallet C for some regular activities such as making transaction, buying NFT, staking, auctioning, I would say small to medium size asset, not for storage."*

4.4.5 Impact of Accessibility on Usability. A few participants (P0, P5, P14) associated accessibility with the level of usability. P5, who identified themselves as dyslexic and having other disabilities, expressed difficulty with the seed phrase confirmation process during account creation, particularly in terms of technology interaction, design, and word order. At some point during account creation, he

asked *"Can I give up [on account creation]? I'm pretty sure I confirmed the seed phrase, but it did not show the continue button."* He further stated, *"Seed phrase confirming can be difficult for people with disability. I'm a keyboard-heavy person, more than mouse ... for the older generation, which I certainly fall into ... with respect to the words generated for seed phrases, I found similar words like a tree that are confusing since I'm a dyslexic."* It is important to note that P5 ultimately failed to complete the account creation process for Wallet B despite investing a significant amount of time. Thus, he highlighted his preference for the usability of Wallet C, which met his need for accessibility, on par with other web applications. This preference might be the result of his lack of prior experience with non-custodial crypto wallets. We also noticed subjective preferences among participants, such as P0 suggesting that seed phrases be arranged in alphabetical order for improved searchability (in Wallet B).

4.4.6 Impact of Perceived Risk on Usability. Participants who had past negative experiences (e.g., fraud, asset loss) with crypto wallets, particularly with non-custodial wallets (P0, P4, P6, P10, P11, and P14, P18, P22, P25 as indicated in the pre-task survey), associated usability with a mechanism to mitigate financial risks commonly encountered in the crypto space. They highlighted the improved usability of Wallet C for financial asset management and associated risks. They appreciated the easier workflow of the recovery process, which provided similar security to Wallet B. P14 expressed his satisfaction, stating, *"Less risk in losing assets than decentralized wallets like MetaMask, made me more comfortable. I would say I like the fact that it had that backup factor, and then I was able to access it again, while if I ever lose my private key and mnemonic, I will basically lose that wallet and my funds."*

4.5 Answering RQ2: Security and Trust Perceptions of MFKDF

4.5.1 Should I Trust Myself More Than I Trust a Custodian? Many participants chose Wallet C when asked to consider the trade-off of who to trust with keeping their keys safe. To illustrate this, P10, who had previously used a non-custodial wallet, said, *"I found the process pretty stressful saving seed phrases by myself. I think I don't trust myself to save it properly or [avoid] being stolen. Noncustodial [wallets], like MetaMask, are pretty inconvenient. You have to manually set your transaction and you have to think about it a lot before you do it, versus a custodial [wallet], where you don't have to do all that."* P0 expressed distrust towards central authorities when it comes to safeguarding data, saying, *"I trust myself more to keep my seed phrase and password safe."* He acknowledged the challenges of avoiding the use of centralized applications in practice, mentioning, *"It's difficult though, especially because my school e-mail is through Gmail and it always has some of [my data], but I try to find alternatives."*

Participants' trust perceptions were also influenced by notable incidents, such as FTX's mishandling of funds, as mentioned by P10, and the disappearance of funds associated with the Quadriga crypto exchange (the *"crypto king"* case), as mentioned by P3. Additionally, P3 noted that community endorsements play a significant role in his trust towards centralized platforms: *"I definitely go on Reddit and Twitter, looking at what the community uses and says, and if there's anything about the leadership which is negative. I would trust people who are well-versed in the space."*

4.5.2 Preferences Influenced by Preconceived Notions of Trust. Participants' wallet usage preferences are shaped by their level of trust in centralized institutions, which can vary greatly. Some participants expressed complete trust in most custodians, while others generally only trust well-known centralized providers, and some participants don't trust any centralized providers. P3 expressed a preference for the security of Wallet B by saying, *"I would prefer the security of Wallet B because it seems I need to rely on my email and the Google Authenticator app for Wallet C... I generally trust Google since they are the company I have my email with. However, there is still a sense of involvement from third parties."* P3 also questioned if Google has access to their credentials, but if not, said that they would prefer Wallet C. In contrast, P8 preferred a centralized wallet saying, *"I prefer centralized wallets because if I lose my device or credentials, I can contact the company for assistance, similar to a bank, and prove my identity. With decentralized wallets, if I lose everything, there is no way for the company to help me retrieve my credentials; same if my credentials are stored on my personal device and I lose them."* In contrast, P24, who currently uses decentralized wallets, preferred Wallet C due to their trust in the underlying technology, stating, *"Though I cannot see what's under the hood for all the decentralized mechanisms, we take those at face value. I trust and value the technology more and feel that Wallet C is truly decentralized."* This sentiment appears to resonate with many crypto enthusiasts [11, 59].

4.5.3 Dilemma of Advising Others on Crypto Wallet Choice. When recommending crypto wallets to others, most participants considered security and trust important factors, but recommendations varied. P2, a novice crypto wallet user, mentioned that he would suggest using Wallet B, while he would himself use Wallet C, stating, *"If you're recommending someone else, you are taking that burden of responsibility that if something goes wrong, they can curse at you. I would probably go for the most secure option when suggesting."* In contrast, P12 indicated she would recommend Wallet C if she chose to onboard a friend, as there would be a steep learning curve for Wallet B, and there is a chance for newcomers to face difficulties in keeping the key and seed phrase secure: *"C is super easy to use and more secure than Coinbase because every time you log in, new keys are created on your device, no need to write down any seed or anything."*

4.5.4 Prior Preferences for Known Authentication Factors. Some participants, predominantly those who are novices in the crypto wallet space, appreciated Wallet C for its authentication mechanisms, which were similar to those of Wallet A and other familiar apps they had used in the past. They expressed confidence in Wallet C due to the added layer of security provided by decentralization. For example, P1 indicated, *"I mentioned earlier in B that I didn't need to put any e-mail, create username, receive any text or no biometrics for authentication, what I usually experience in most other authentication processes. I believe traditional 2FA or multi-verification is more secure. What came to my mind was that B is adding risk to the account if I only had to provide the passphrase to recover; one can take control of my account."*

4.6 Direct Comparisons of Wallet C to Wallets A and B

4.6.1 "Safety of Decentralized with the Ease of Centralized". Participants, both novice and experienced, frequently expected a balance between usability and security. P7 explained her preference for Wallet C, stating, *"I would choose Wallet C. It's easier to use and more secure compared to Wallet B. Additionally, I prefer having my information stored on my own device for added security. Adding a centralized server introduces risk. So I could choose C which has a combination of non-custodial and user-friendly features."* P1 also made a similar comparison, appreciating the safety of a decentralized wallet with the ease of the centralized wallet, such as having custody of keys without the hassle of writing down a seed phrase. P1 further noted that, *"You should find a way to make people understand Wallet C more to draw attention."*

4.6.2 Comparison of Risk and Threat Landscape. Although participants expressed a clear preference for the usability of Wallet C, stating that it is comparable to Wallet A and superior to Wallet B, they also considered the risks associated with Wallet C compared to Wallet B on a case-by-case basis. P13 discussed the perception of risk based on the specific threat landscape and various factors such as the number of credentials to target and the number of failure points involved in the scenario. In P13's words, *"[Wallet C] is little less than B security-wise in some cases. It falls somewhere in the middle comparing A and B. It still has a weak link that is a mobile phone, while on the other hand the key file is obfuscated. But for hacking, one only needs to target the key file for Wallet B, while for C there are few things to target. So in one scenario, C is safer than B. Additionally, I have a passcode for the email app and authenticator, which enhances safety. However, in another scenario, B is safer than C."* Similarly, P2 and P14 mentioned the concept of a single point of failure for Wallet C, even though the keys are not stored on a server. He noted the human factor of users demonstrating less secure practices when using mobile devices: *"[Wallet C] appears to be safer, because it's not stored on the server. It's not through a centralized server like Wallet A, but there's still a single point of defeat in case of a lost device. You're only as strong as your weakest link."* P4 echoed this sentiment, but added the condition that the lost device would pose a risk only if unauthorized actors were aware of the phone's passcode.

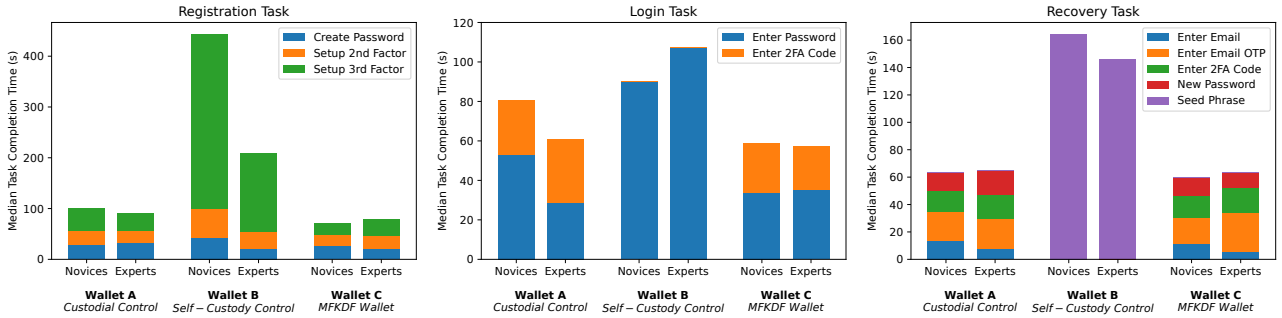
P16 also highlighted the risk of Wallet C compared to Wallet B. With Wallet B, there is a recovery process that involves exposing the seed phrase on the web browser whenever the password or key file is forgotten, which creates a perceived threat. In contrast, he emphasized the two layers of security provided by Wallet C, which include email and TOTP: *"Seed phrase is very dangerous, because it is basically your private key, and I don't want to expose my seed phrase every time I forget my password. Seed phrase idea is safe for offline/cold storage. I believe Wallet C, the key is sharded, and one of the shards is encrypted to my password."* On the same note, P4 explained the prevalence of targeted attacks on Wallet B, mentioning *"People gain access to machines all the time, somebody sends you Trojan or you accidentally upload the wrong file somewhere. For C, even if someone gets access to your laptop, they would need TOTP which is on your phone; it's less likely to get hacked in all devices at the same time."*

4.6.3 Usability Impact of Number of Devices Required. Two participants out of 17 focused on the number of devices involved in creating an account and the recovery process to evaluate the usability of Wallet C. P3 compared Wallet C with Wallet B and expressed frustration, stating, *"Recovery is easy with Wallet C, but I'm a bit frustrated that I always need to have a mobile device with me."* In contrast, some participants (P1, P2, P4, P7, P11, P8) questioned the notion of security (in the case of the number of devices used) of Wallet B in conjunction with usability. They highlighted the assumption of keeping the seed phrase offline for better security during recovery, as well as storing the key file on a pen drive or hard drive for added security, which essentially requires access to external means (devices). They further pointed out that they, and many others, keep their key file in the cloud (e.g., Google Drive or iCloud) for convenience while traveling, or for when their primary device is unavailable, which challenges the concept of keeping credentials offline. They also mentioned keeping the seed phrase in a notepad or text file. To further illustrate this, P7 said, *"Normally in our generation, we end up using online platforms, like using email, and so many other media. Even if the philosophy is of protecting accounts by saving the seed phrase offline, it's just not convenient for me. Let's say, for the worst case, I'm sending the seed phrase to my email, considering I don't have my personal device with me. So it's compromising my security, because it defies the whole idea of saving it securely offline. In your day-to-day life, you don't always do that; it's not convenient. So, I also don't see what's the point of downloading the whole thing manually."*

To maintain the security of wallet B, P4, who uses a cold wallet, shared their experience indicating that the security assumption of using fewer devices is not entirely valid: *"having an authenticator factor off the device is like cold storage, whereas there's plenty of malware for browsers and computers; any targeted attacks, keystroke attacks, can happen while many people store backups of seed phrases in normal drive or cloud and even platforms like emails or notepad."* He further suggested that he wouldn't recommend Wallet B to non-technical users newcomers as it could potentially have negative consequences. Further, he indicated a preference for using Wallet C for short-term staking while he uses a cold wallet for long term storage. P1 further wished to contextualize the number of the devices used in Wallet C as being more convenient, saying, *"For security purposes, it's better to carry the seed from [Wallet B] on a pen drive and it's better to have multiple copies on a pen drive, or micro-SD, or other devices because the physical drive can crash any time. But for C, it's a flexible choice, either use one or two devices as per my setup."*

4.7 Quantification of Usability (SUS)

After participants completed tasks such as account creation, account recovery, and logging in from a different device, we administered the System Usability Scale (SUS) survey to gauge the overall usability of each wallet. The survey responses were evaluated on a quantitative scale ranging from 0 to 100. The scores for each question were converted to numerical values, summed up, and then multiplied by 2.5 to convert the original scores of 0-40 to a scale of 0-100 [19]. Based on the results, Wallet C received the highest average usability score of 87.94, followed by Wallet A with a score



(a) Registration task completion time. **(b) Login task completion time.** **(c) Recovery task completion time.**

Figure 4: Task completion time for novice and expert users across three wallets for three different tasks.

of 84.85. Wallet B received a usability score of 60.88, which is considered substandard. When comparing the SUS scores of wallets B and C with a paired two-tailed t-test,[5] we found that t=6.59, p<0.0001. Thus, the subjective usability assessment of the MFKDF-based wallet is higher than that of the non-custodial control wallet with high statistical significance.

In addition to the standard 10 items in the SUS survey, we included wallet-specific questions for each task to assess participants' experience and preference when using the different wallets. The results consistently showed that participants preferred using and recommending Wallet C over Wallets A and B. They perceived Wallet C as more secure (mean: 4.17) compared to Wallet A (mean: 3.00) and Wallet B (mean: 3.41), with highly statistically significant p-values of p = 0.001 and p = 0.009 respectively according to two-tailed paired t-tests.

4.8 Task Completion Time Analytics

To collect additional objective metrics of usability for each of the three prototype wallets, we embedded a detailed analytics script in each wallet that measured the precise step-by-step completion time for each of the assigned tasks. The measured task completion times for the registration, login, and recovery tasks are shown in Fig. 4. We split the results by novice and expert users; "expert" is defined here as any level of prior cryptocurrency wallet experience.

The results show that the MFKDF wallet (Wallet C) consistently outperforms the non-custodial control wallet (Wallet B) and performs on par with the custodial control wallet (Wallet A) in all three scenarios. This trend appears to hold for both novice and expert users. These results align with our intuitive expectations, as the MFKDF wallet provides a nearly identical user experience to the custodial control wallet.

We performed statistical tests to measure the significance of the observed results. Specifically, we wanted to compare the task completion times for Wallet C, the MFKDF-based wallet, and Wallet B, the equivalent non-custodial wallet. We again chose to use two-tailed paired t-tests due to the within-subject experimental design. The results were as follows:

- Registration was, on average, 67% faster when using the MFKDF wallet (N=27, t=4.02, p=0.0004).
- Logging in was, on average, 32% faster when using the MFKDF wallet (N=27, t=2.50, p=0.0187).
- Password recovery in was, on average, 86% faster when using the MFKDF wallet (N=27, t=3.46, p=0.0019).
- Overall task completion time was, on average, 71% faster when using the MFKDF wallet (N=27, t=6.52, p<0.0001).

[5]We chose to use a paired test in this case due to the within-subject design.

5 DISCUSSION

Our study provides a nuanced understanding of how MFKDF addresses usability concerns associated with non-custodial wallet key management as well as how users' perceptions of risk, security, and trust are shaped by technical and human-centric components (e.g., users' understanding of decentralization). In this section, we expand on the broader applications of MFKDF beyond just cryptocurrency wallets, informed by our study results.

5.1 Addressing open challenges of human-computer interaction in decentralized applications

The storied history of key management usability research has unfolded across several chapters. First came notoriously cumbersome early cryptographic tools, such as PGP, that largely relied on manual key storage and management by the end-user [68]. Next, password-based key derivation functions like PBKDF2 [45] emerged as a reasonably effective key management approach in centralized systems, such as password managers [2, 10, 25] and network protocols [43, 44]. However, with the advent of cryptocurrencies and decentralized finance, we have entered a third era of key management in which centralized systems usually cannot be relied upon. As it stands, this era largely resembles the first, with most of the advancements made in centralized systems failing to transfer to non-custodial applications, forcing non-custodial applications to revert to more cumbersome key management approaches.

The resulting disparity between the usability of custodial and non-custodial wallets has had dramatic negative consequences for the reputation of blockchain and cryptocurrency technologies as a whole. Just as an over-reliance on centralized password managers has proven dangerous in light of recent major security incidents [53, 63], an over-reliance on custodial cryptocurrency wallets

breeds distrust in the entire ecosystem when these custodians experience catastrophic failures [6]. HCI researchers have explored and focused on some of the design challenges of blockchain applications, striving to develop inclusive systems that accommodate varying expertise levels [31, 32], languages, and accessibility needs [21, 67, 72]. Much of the prior HCI research has delved into the perception and usability of different cryptocurrency wallets through both qualitative and quantitative analyses. These studies have identified challenges and suggested design implications, but many remain unexplored, leaving a gap for future HCI research.

Our study focused on testing a crypto wallet based on a novel cryptographic scheme, MFKDF [51], which offers a design space to incorporate various user-centric components in key management, including intuitive interfaces, error tolerance, and flexible recovery mechanisms. Our user study results demonstrate that the current usability benefits of custodial wallets, including portability, resilience, and recoverability, are not intrinsic advantages of custodial key management but rather constitute gaps that can be filled with adequate cryptography and engineering. Our findings indicate that not only do most users subjectively prefer MFKDF-based key management to recovery phrases and similar competing solutions according to several standardized metrics, but also that they can perform a range of basic tasks, such as registration, login, and recovery, far more efficiently when using MFKDF-based applications. In many cases, users choose to use custodial solutions over non-custodial options due to their perceived usability advantages, despite knowing the security trade-off that they make in doing so. Thus, we have conducted this study with the hope that true parity in usability between custodial and non-custodial applications will help convince a large number of users to consider adopting non-custodial options. Our results suggest that the use of MFKDF [51] has allowed us, for the first time, to achieve this level of true parity in the usability of non-custodial wallets compared to that of custodial solutions without compromising the security of the resulting system. Beyond cryptocurrency, we believe that this result could apply to other applications requiring strong trustless key management. While this is the first user study examining the usability of multi-factor key derivation, we believe future work may apply MFKDF as a general solution to the key management problem.

5.2 Addressing the Discrepancy Between Security Intentions & Behavior

Our findings also reveal that participants predominantly opted for cloud-based solutions like Google Drive, iCloud, and Google Cloud, when managing the secrets (seed phrases and private keys) associated with non-custodial wallets. Although non-custodial wallet stakeholders have made concerted efforts to promote secure key management practices (e.g., using a secure password manager or a safe deposit box, or writing down and storing keys in multiple secret locations), via instructional videos and text descriptions, participants still exhibited a preference for convenience over safety risks. This discrepancy between what participants said they would keep their keys in private and secure places, but prioritize convenience in practice, is not a new phenomenon, which has been studied in the prior literature (e.g., [52, 58].

These behaviors resonate the long-standing tension between security and usability. Users often prioritize convenience over security, leading to behaviors such as reusing passwords across multiple platforms [37], writing down passwords or using simplistic ones [26], believing that security breaches happen to others but not to them [40], an implicit trust in digital platforms' security infrastructure, thus avoiding personal responsibility to act securely [28]. Consequently, the majority of non-custodial wallets primarily rely on users alone for security, leading to increased complexity in usability and posing challenges for onboarding and maintaining security, especially when users bear the exclusive responsibility for the protection and storage of keys and associated credentials. In our study, the discrepancy is tackled by the underlying mechanism, MFKDF [51].

5.3 Should I Trust Myself More Than I Trust a Custodian

Our study has also identified participants' perceptions of security, and trust regarding existing custodial and non-custodial wallets, highlighting a distinct preference for our MFKDF wallet. Qualitative insights further suggest that factors such as pre-existing trust levels (e.g., confidence in centralized entities, reliance on renowned centralized parties, or complete distrust in centralized systems), historical authentication inclinations, and trust in self-efficacy for key management play a pivotal role in participants' crypto wallet preferences.

Findings from our study resonate with existing literature on trust dynamics in the realm of the increasingly networked era. A central theme that emerges from our participants' responses is the tension between trust in oneself versus trust in a custodian or centralized authority. This tension is not unique to our study but has been documented in prior research on digital trust, playing a key role in users' security behavior [23, 28, 40]. The apprehension about managing seed phrases on their own and the perceived inconvenience of non-custodial wallets like MetaMask mirrors the broader sentiment in the user community. This sentiment is rooted in the inherent complexity and responsibility associated with the self-management of cryptographic keys. Existing literature has highlighted the cognitive load and stress users experience when tasked with the sole responsibility of their digital assets, especially in the absence of a centralized authority or intermediary [31, 72]. Conversely, distrust towards central authorities and preference for self-management echoes the ethos of the cryptocurrency movement, which champions decentralization and autonomy. Notable incidents, such as the FTX mishap [6] and the Quadriga crypto exchange debacle [55], further amplify these concerns and highlight the ripple effects of such events on user trust.

Our study, when juxtaposed with established literature [28], offers a multifaceted exploration of trust, usability, and security, and can offer a lens to view the broader shifts in digital trust in our increasingly networked era. This can provide a roadmap for stakeholders in the digital finance ecosystem, aiming to craft solutions that resonate with users.

While many participants appreciated, there were preferences expressed for specific authentication factors. For instance, some participants favored using Yubikey or text OTP over email as a factor.

There is potential for future iterations to offer users a selection of authentication factors during account creation. This could enhance both the perceived security and trustworthiness of the system. Another challenge highlighted by participants was the requirement to use a minimum of two devices during the account creation and recovery phases. Introducing flexibility in factor choice might reduce this workload, especially for those who find methods like QR code scanning or using multiple devices cumbersome.

5.4 Limitations

There are some notable limitations implicated by our approach. First, although we had 27 participants for the main study, we cannot generalize our findings to the broader population, including marginalized populations that require further exploration. Additionally, our study only focused on participants from the United States, which restricts the generalizability of our results to users in other countries. Moreover, our participant pool was not gender-balanced, aligning with the existing literature that suggests cryptocurrency users are mostly male [32].

Secondly, in terms of study design, we developed three wallets: custodial, non-custodial, and MFKDF-based. We aimed for these wallets to represent the user flow of MetaMask and Coinbase/KuCoin accurately, ensuring consistent branding, UI, and experiment-focused features while eliminating extraneous variables. However, we acknowledge that there may be limitations in precisely replicating the real wallets, which could impact task completion time and user experience. Nevertheless, our carefully designed and developed wallets aimed to mitigate interface effects and brand recognition biases, while providing detailed click-by-click analytics. This approach would have been challenging if we had used real MetaMask and Coinbase or KuCoin wallets, with MFKDF only as a prototype.

Finally, we used prior crypto wallet usage as a proxy for distinguishing novice and experienced users, following prior literature [72]. However, alternative metrics such as years of usage could also be considered. It is worth noting that our research did not aim to identify latent clusters of wallet users, which remains a valuable topic for future investigations.

5.5 Future Work

The focus of this study has been on the usability of decentralized key management mechanisms, with cryptocurrency wallets being an archetypal example of a domain in which usable key management is vital. However, there are many other dimensions of cryptocurrency wallet usability that we did not explore in this paper due to our focus on key management, and we encourage researchers to continue producing usability studies on these other aspects.

Beyond cryptocurrency wallets and decentralized finance, MFKDF has the potential to improve the usability of private key management across a wide variety of centralized and decentralized applications alike. Unlike public key management, which often requires highly application-specific solutions, private key management can generally be reduced to securely storing and accessing a series of bits constituting a cryptographic key. MFKDF solves this problem for a generic bit string by eliminating any key storage, and instead deriving the correct series of bits if and only if valid authentication factors are provided. Importantly, it does so in a way that is agnostic to any particular cryptographic scheme and is thus theoretically compatible with a variety of existing applications.

Given our focus on isolating key management from all other variables, our results strongly suggest that differences in the key management mechanism alone account for observed differences in usability across the three evaluated wallets. Future research is however needed to further examine the generalizability of this finding to other application domains. In particular, we hope to see future research that evaluates MFKDF as a general-purpose key management solution in applications such as disk encryption, password management, cloud storage, wireless network protocols, and more.

Further work can also be done on analyzing the usability of MFKDF for members of marginalized populations. For instance, although we focused on studying a general population in this study as an initial proof of concept, it is equally important to ensure that the use of MFKDF does not have negative consequences for users of accessibility technologies.

Finally, given that MFKDF is a relatively new key management approach, it is important to place a focus on awareness and education to increase understanding of this technology. Along with usability, increased consumer awareness and trust in MFKDF is critical to enabling the wider implementation and adoption of usable key management.

6 CONCLUSION

Key management has for decades now been a particularly thorny problem for cryptographers and HCI researchers alike and is likely to continue to evolve as a field over time. Just as PBKDFs served as an important step in this evolution for centralized applications, MFKDF may also constitute an important step for decentralized applications.

In light of its previously seen security advantages and newly-demonstrated usability advantages, we hope to see MFKDF continue to receive academic interest and eventually be implemented as a major key management solution in various applications. This paper demonstrates that users both perform better and subjectively prefer MFKDF-based experiences, reducing key management as a point of friction in non-custodial applications.

7 AVAILABILITY

The source code and documentation for our MFKDF-based Ethereum wallet, and custodial and non-custodial control wallets, are available for review at the following URL:

https://github.com/multifactor/research-wallets

ACKNOWLEDGMENTS

We appreciate the advice of Björn Hartmann. This work was supported in part by the National Science Foundation, the National Physical Science Consortium, the Fannie and John Hertz Foundation, and the Berkeley Center for Responsible, Decentralized Intelligence. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors

and do not necessarily reflect the views of their employers or the supporting entities.

REFERENCES

[1] [n. d.]. 78% of People Reset a Password They Forgot in Past 90 Days. https://blog.hypr.com/hypr-password-study-findings
[2] 2016. Lastpass. https://www.lastpass.com/security/zero-knowledge-security
[3] 2022. 2020 State of the Internet. https://www.akamai.com/site/en/documents/state-of-the-internet/soti-security-credential-stuffing-in-the-media-industry-report-2020.pdf
[4] 2022. Top Cryptocurrency Decentralized Exchanges Ranked. arXiv:CoinMarketCap
[5] 2022-12-02. The crypto wallet for Defi, Web3 Dapps and NFTs MetaMask. https://metamask.io/.
[6] 2022-12-02. How Sam Bankman Fried's FTX Crypto Empire. https://www.nytimes.com/2022/11/14/technology/ftx-sam-bankman-fried-crypto-bankruptcy.html.
[7] 2022-12-02. The Inside Story of Mt. Gox. https://www.wired.com/2014/03/bitcoin-exchange/.
[8] 2022-12-02. Top Cryptocurrency Exchanges Ranked By Volume. https://coinmarketcap.com/rankings/exchanges/.
[9] 2022-12-02. Trezor Hardware Wallet. https://trezor.io/.
[10] 1Password. 2021. 1Password Security Design. , 100 pages. https://1passwordstatic.com/files/security/1password-white-paper.pdf
[11] Svetlana Abramova, Artemij Voskobojnikov, Konstantin Beznosov, and Rainer Böhme. 2021. Bits under the mattress: Understanding different risk perceptions and security behaviors of crypto-asset users. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–19.
[12] Svetlana Abramova, Artemij Voskobojnikov, Konstantin Beznosov, and Rainer Böhme. 2021. Bits Under the Mattress: Understanding Different Risk Perceptions and Security Behaviors of Crypto-Asset Users. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. https://informationsecurity.uibk.ac.at/pdfs/CHI2021_Bits_Under_the_Mattress.pdf.
[13] Noelle Acheson. 2022. After FTX: Rebuilding Trust in Crypto's Founding Mission. https://www.coindesk.com/layer2/2022/11/14/after-ftx-rebuilding-trust-in-cryptos-founding-mission/.
[14] Carlisle Adams and Steve Lloyd. 1999. *Understanding public-key infrastructure: concepts, standards, and deployment considerations*. Sams Publishing.
[15] Apple. 2012. iOS Security. https://web.archive.org/web/20121021133728/http://images.apple.com/ipad/business/docs/iOS_Security_May12.pdf
[16] Jean-Philippe Aumasson. [n. d.]. Murphy's laws of cryptography. https://www.aumasson.jp/murphy.html
[17] Elaine Barker. 2016. *Recommendation for Key Management Part 1: General*. Technical Report NIST SP 800-57pt1r4. National Institute of Standards and Technology. NIST SP 800–57pt1r4 pages. https://doi.org/10.6028/NIST.SP.800-57pt1r4
[18] Binance 2022-12-02. Binance US. https://www.binance.us/.
[19] John Brooke et al. 1996. SUS-A quick and dirty usability scale. *Usability evaluation in industry* 189, 194 (1996), 4–7.
[20] Elie Burzstein and Jean Michel Picod. 2010. Recovering Windows Secrets and EFS Certificates Offline. In *Proceedings of the 4th USENIX Conference on Offensive Technologies* (Washington, DC) *(WOOT'10)*. USENIX, USA, 1–8.
[21] You-Ping Chen and Ju-Chun Ko. 2019. CryptoAR wallet: A blockchain cryptocurrency wallet application that uses augmented reality for on-chain user data display. In *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services*. 1–5.
[22] Coinbase 2022-12-02. Coinbase - Buy and Sell Bitcoin. https://www.coinbase.com/.
[23] Cynthia L Corritore, Beverly Kracher, and Susan Wiedenbeck. 2003. On-line trust: concepts, evolving themes, a model. *International journal of human-computer studies* 58, 6 (2003), 737–758.
[24] Anupam Das, Joseph Bonneau, Matthew Caesar, Nikita Borisov, and XiaoFeng Wang. 2014. The tangled web of password reuse.. In *NDSS*, Vol. 14. 23–26.
[25] Dashlane. 2021. Security White Paper. https://www.dashlane.com/download/whitepaper-en.pdf
[26] Rachna Dhamija and Adrian Perrig. 2000. Deja {Vu–A} User Study: Using Images for Authentication. In *9th USENIX Security Symposium (USENIX Security 00)*.
[27] Shayan Eskandari, David Barrera, Elizabeth Stobert, and Jeremy Clark. 2018. A First Look at the Usability of Bitcoin Key Management. In *arXiv*. https://arxiv.org/pdf/1802.04351.pdf.
[28] Dinei Florencio and Cormac Herley. 2007. A large-scale study of web password habits. In *Proceedings of the 16th international conference on World Wide Web*. 657–666.
[29] Dinei Florencio and Cormac Herley. 2007. *A large-scale study of web password habits*. Technical Report. 657–666 pages.
[30] Foundation for Interwallet Operability. 2019. Blockchain Usability Report. https://fioprotocol.io/wp-content/themes/fio/build/files/blockchain-usability-report-2019.pdf.
[31] Michael Fröhlich, Maurizio Raphael Wagenhaus, Albrecht Schmidt, and Florian Alt. 2021. Don't stop me now! exploring challenges of first-time cryptocurrency users. In *Designing Interactive Systems Conference 2021*. 138–148.
[32] Michael Fröhlich, Franz Waltenberger, Ludwig Trotter, Florian Alt, and Albrecht Schmidt. 2022. Blockchain and Cryptocurrency in Human Computer Interaction: A Systematic Literature Review and Research Agenda. *arXiv preprint arXiv:2204.10857* (2022).
[33] M. Fröhlich, C. Kobiella, A. Schmidt, and F. Alt. 2021. Is it better with onboarding? improving First-Time cryptocurrency app experiences. In *Designing Interactive Systems Conference 2021*. 78–89.
[34] M. Fröhlich, M. R. Wagenhaus, A. Schmidt, and F. Alt. 2021. Don't stop me now! exploring challenges of First-Time cryptocurrency users. In *Designing Interactive Systems Conference 2021*. 138–148.
[35] X. Gao, G. D. Clark, and J. Lindqvist. 2016. Of two minds, multiple addresses, and one ledger: characterizing opinions, knowledge, and perceptions of Bitcoin across users and non-users. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 1656–1668.
[36] Xianyi Gao, Yulong Yang, Can Liu, Christos Mitropoulos, Janne Lindqvist, and Antti Oulasvirta. 2018. Forgetting of passwords: Ecological theory and data. In *27th {USENIX} Security Symposium ({USENIX} Security 18)*. 221–238.
[37] Maximilian Golla, Miranda Wei, Juliette Hainline, Lydia Filipe, Markus Dürmuth, Elissa Redmiles, and Blase Ur. 2018. " What was that site doing with my Facebook password?" Designing Password-Reuse Notifications. In *Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security*. 1549–1566.
[38] Ameya Hanamsagar, S Woo, Christopher Kanich, and Jelena Mirkovic. 2016. How Users Choose and Reuse Passwords.
[39] Ameya Hanamsagar, S Woo, Christopher Kanich, and Jelena Mirkovic. 2016. How Users Choose and Reuse Passwords. *Information Sciences Institute* (2016).
[40] Marian Harbach, Markus Hettig, Susanne Weber, and Matthew Smith. 2014. Using personal examples to improve risk communication for security & privacy decisions. In *Proceedings of the SIGCHI conference on human factors in computing systems*. 2647–2656.
[41] Shuangyu He, Qianhong Wu, Xizhao Luo, Zhi Liang, Dawei Li, Hanwen Feng, Haibin Zheng, and Yanan Li. 2018. A Social-Network-Based Cryptocurrency Wallet-Management Scheme. *IEEE Access* 6 (2018), 7654–7663. https://doi.org/10.1109/ACCESS.2018.2799385
[42] Ping Identity. [n. d.]. What is Out-of-Band Authentication (OOBA)? https://www.pingidentity.com/en/resources/blog/post/what-is-out-of-band-authentication-ooba.html
[43] IEEE. 1997. IEEE Standard for Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications. *IEEE Std 802.11-1997* (1997). https://doi.org/10.1109/IEEESTD.1997.85951
[44] IEEE. 2004. IEEE Standard for information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN MAC and PHY specifications: Amendment 6: MAC Security Enhancements. *IEEE Std 802.11i-2004* (2004), 1–190. https://doi.org/10.1109/IEEESTD.2004.94585
[45] Burt Kaliski. 2000. *PKCS #5: Password-Based Cryptography Specification Version 2.0*. Request for Comments RFC 2898. Internet Engineering Task Force. https://doi.org/10.17487/RFC2898 Num Pages: 34.
[46] Kraken 2022-12-02. Kraken Cryptocurrency. https://www.kraken.com/.
[47] KuCoin 2022. Crypto Exchange Kucoin. https://kucoin.com/.
[48] Ledger 2022-12-02. Hardware Wallet State of the art security for crypto assets. https://www.ledger.com.
[49] Md Moniruzzaman, Farida Chowdhury, and Md S. Ferdous. 2020. Examining Usability Issues in Blockchain-Based Cryptocurrency Wallets. In *International Conference on Cyber Security and Computer Science (ICONCS)*. 631–643. https://link.springer.com/chapter/10.1007/978-3-030-52856-0_50.
[50] Vivek Nair and Dawn Song. 2023. Decentralizing Custodial Wallets with MFKDF. In *2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*. 1–9. https://doi.org/10.1109/ICBC56567.2023.10174998
[51] Vivek Nair and Dawn Song. 2023. Multi-Factor Key Derivation Function (MFKDF) for Fast, Flexible, Secure, & Practical Key Management. In *32nd USENIX Security Symposium (USENIX Security 23)*. USENIX Association, Anaheim, CA, 2097–2114. https://www.usenix.org/conference/usenixsecurity23/presentation/nair-mfkdf
[52] Patricia A Norberg, Daniel R Horne, and David A Horne. 2007. The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs* 41, 1 (2007), 100–126.
[53] Kaiti Norton. 2021. LastPass: Is it a Safe Password Manager? https://www.esecurityplanet.com/products/lastpass-review/
[54] Nathaniel Popper. 2021. Lost passwords lock millionaires out of their bitcoin fortunes. *The New York Times* 12 (2021).
[55] quar Accessed on 2022. quadrigaCX. https://www.cbc.ca/news/canada/nova-scotia/quadrigacx-cryptocurrency-bankruptcy-ernst-and-young-1.4364467.
[56] C. B. C. Radio ·. 2021. This man owns $321M in bitcoin — but he can't access it because he lost his password | CBC Radio. https://www.cbc.ca/radio/asithappens/as-it-happens-friday-edition-1.5875363/this-man-owns-321m-in-bitcoin-but-he-

ID	*Gender*	*Age Range*	*Occupation*	*Educational*	*Crypto Wallet Experience*
P0	Male	25-34	PhD student	Master's degree	1 - 2 Years
P1	Male	25-34	Teaching Assistant	Master's degree	Novice
P2	Male	18-25	Student	Master's degree	<1 Week
P3	Non-Binary	18-25	Graduate Student	Bachelor's degree	Novice
P4	Male	25-34	Engineer	Master's degree	>2 Years
P5	Male	35-44	Product Manager	Master's degree	1 Week - 1 Month
P6	Female	18-25	Work education in web3	Bachelor's degree	>2 Years
P7	Female	25-34	Teaching Assistant	Master's degree	Novice
P8	Male	18-25	Software Engineer	Master's degree	<1 Week
P9	Male	25-34	Graduate Student	Bachelor's degree	Novice
P10	Male	25-34	Graduate Student	Master's degree	>2 Years
P11	Male	25-34	Student	Bachelor's degree	1 - 2 Years
P12	Female	35-44	Product Manager	Master's degree	>2 Years
P13	Male	35-44	Operations Manager	Master's degree	1 - 2 Years
P14	Male	25-34	Software Engineer	Master's degree	>2 Years
P15	Male	25-34	Teacher	Bachelor's degree	>2 Years
P16	Male	35-44	Product Manager	Master's degree	>2 Years
P17	Male	35-44	Logistics Manager	Bachelor's degree	>2 Years
P18	Male	25-34	Web3 Product	Bachelor's degree	1 - 2 Years
P19	Female	25-34	Gradute Student	Bachelor's degree	1 Week - 1 Month
P20	Female	25-34	US Forestry	Master's degree	1 Week - 1 Month
P21	Male	25-34	Product in a crypto firm	Master's degree	>2 Years
P22	Male	25-34	Investor	Bachelor's degree	>2 Years
P23	Male	25-34	Engineer	Bachelor's degree	>2 Years
P24	Male	35-44	Management Consultant	Diploma in Accounting	>2 Years
P25	Male	25-34	Research & Development	Doctorate degree	>2 Years
P26	Male	18-25	Student	Bachelor's degree	1 - 2 Years

Table 1: Participant demographics and background.

ID	Observed: Seed Phrase	Reported: Seed Phrase	Reported: Key File	Reported: Password
P0	txt	physical paper	pendrive	memorize
P1	notepad	notepad, doc	multiple backups	password manager
P2	notepad	cloud	cloud, zip file	memorize
P3	icloud	drive, airdrop	cloud	password manager
P4	txt	encrypted cloud	encrypted cloud	lastpass
P5	notepad	notepad	encrypted	password manager
P6	notepad	notepad	drive	document
P7	doc	send to email	pendrive	notepad
P8	txt	cloud	cloud	password manager
P9	txt	memorize	pendrive	notepad
P10	notepad	pen paper	cloud	document
P11	icloud	drive locked	drive locked	paper
P12	txt	google cloud	google cloud	lastpass
P13	icloud	icloud	icloud	notepad
P14	icloud	siloed device	siloed device	paper
P15	notepad	cloud	cloud	document
P16	notepad	google cloud	google cloud	paper
P17	txt	words in sketch	encrypted drive	memorize
P18	notepad	pen drive	google cloud	password manager
P19	doc	N/A	N/A	password manager
P20	notepad	paper	google cloud	password manager
P21	screenshot	Words to numbers	pen drive	memorize
P22	txt	google cloud	google cloud	paper
P23	paper	physical paper	encrypted cloud	memorize
P24	paste ubuntu	physical paper	pendrive	memorize
P25	txt	pendrive	pendrive	password manager
P26	doc	physical paper	pendrive	memorize

Table 2: Reported/actual credential management methods.

can-t-access-it-because-he-lost-his-password-1.5875366.

[57] Scott Ruoti, Jeff Andersen, Daniel Zappala, and Kent Seamons. 2015. Why Johnny still, still can't encrypt: Evaluating the usability of a modern PGP client. *arXiv preprint arXiv:1510.08555* (2015).

[58] Ben D Sawyer and Peter A Hancock. 2018. Hacking the human: The prevalence paradox in cybersecurity. *Human factors* 60, 5 (2018), 597–609.

[59] Tanusree Sharma, Zhixuan Zhou, Yun Huang, and Yang Wang. 2022. " It's A Blessing and A Curse": Unpacking Creators' Practices with Non-Fungible Tokens (NFTs) and Their Communities. *arXiv preprint arXiv:2201.13233* (2022).

[60] Steve Sheng, Levi Broderick, Colleen Alison Koranda, and Jeremy J Hyland. 2006. Why johnny still can't encrypt: evaluating the usability of email encryption software. In *Symposium on usable privacy and security*. ACM, 3–4.

[61] Elizabeth Stobert and Robert Biddle. 2014. The password life cycle: user behaviour in managing passwords. In *Proc. SOUPS*.

[62] Saurabh Suratkar, Mahesh Shirole, and Sunil Bhirud. 2020. Cryptocurrency Wallet: A Review. In *2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP)*. 1–7. https://doi.org/10.1109/ICCCSP49186.2020.9315193

[63] Karim Toubba. 2022. Notice of Recent Security Incident. https://blog.lastpass.com/2022/12/notice-of-recent-security-incident/

[64] Mountain View, David M'Raihi, Frank Hoornaert, David Naccache, Mihir Bellare, and Ohad Ranen. 2005. *HOTP: An HMAC-Based One-Time Password Algorithm*. Request for Comments RFC 4226. Internet Engineering Task Force. https://doi.org/10.17487/RFC4226 Num Pages: 37.

[65] Mountain View, Johan Rydell, Mingliang Pei, and Salah Machani. 2011. *TOTP: Time-Based One-Time Password Algorithm*. Request for Comments RFC 6238. Internet Engineering Task Force. https://doi.org/10.17487/RFC6238

[66] Artemij Voskobojnikov, Oliver Wiese, Masoud Mehrabi Koushki, Volker Roth, and Konstantin Beznosov. 2021. The U in Crypto Stands for Usable: An Empirical Study of User Experience with Mobile Cryptocurrency Wallets. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. https://informationsecurity.uibk.ac.at/pdfs/CHI2021_Bits_Under_the_Mattress.pdf.

[67] Artemij Voskobojnikov, Oliver Wiese, Masoud Mehrabi Koushki, Volker Roth, and Konstantin Beznosov. 2021. The U in crypto stands for usable: An empirical study of user experience with mobile cryptocurrency wallets. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. 1–14.

[68] Alma Whitten and J Doug Tygar. 1999. Why Johnny Can't Encrypt: A Usability Evaluation of PGP 5.0.. In *USENIX security symposium*, Vol. 348. 169–184.

[69] xmind Accessed on 2022. xmind. https://www.xmind.net.

[70] Martin Young. [n. d.]. Coinbase custodies 11 of entire crypto capitalization. https://cointelegraph.com/news/coinbase-custodies-11-of-entire-crypto-year={2022-12-01},.

[71] Yubico. [n. d.]. YubiKey: Strong Two-Factor Authentication. https://www.yubico.com/

[72] Zhixuan Zhou, Tanusree Sharma, Luke Emano, Sauvik Das, and Yang Wang. 2023. Iterative Design of An Accessible Crypto Wallet for Blind Users. *In SOUPS at USENIX Security Symposium* (2023).

[73] Fangdong Zhu, Wen Chen, Yunpeng Wang, Ping Lin, Tao Li, Xiaochun Cao, and Long Yuan. 2017. Trust your wallet: A new online wallet architecture for Bitcoin. In *2017 International Conference on Progress in Informatics and Computing (PIC)*. IEEE, 307–311.

A FIGURES & TABLES