‘*I’m actually going to go and change these passwords*’: Analyzing the Usability of Credential Audit Interfaces in Password Managers

Emiram Kablo[*]
ekablo@mail.upb.de
Paderborn University
Paderborn, Germany

Katharina Kader[*]
kkader@mail.upb.de
Paderborn University
Paderborn, Germany

Patricia Arias-Cabarcos
pac@mail.upb.de
Paderborn University
Paderborn, Germany

ABSTRACT

Credential audit interfaces in password managers play a crucial role in enhancing user security by identifying weak, reused, or exposed passwords. However, existing research lacks a comprehensive analysis of the usability and motivations of adopters and non-adopters. To address this gap, we conducted 11 semi-structured interviews with users and non-users of credential audit tools, all of whom use password managers. Our study reveals security as the primary motivator for adoption. Despite a potential challenge in handling overwhelming results by the audit reports, participants showed commitment to security and suggested potential benefits of prioritization techniques for a better overview of important results. Transparency and detailed explanations for password weaknesses were identified as user needs. Our broader discussion on password manager adoption underscores the significance of security and convenience as key adoption factors.

CCS CONCEPTS

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

KEYWORDS

Password Managers, Credential Audit, Usability

ACM Reference Format:
Emiram Kablo, Katharina Kader, and Patricia Arias-Cabarcos. 2024. ‘*I’m actually going to go and change these passwords*’: Analyzing the Usability of Credential Audit Interfaces in Password Managers. In *Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA ’24), May 11–16, 2024, Honolulu, HI, USA.* ACM, New York, NY, USA, 13 pages. https://doi.org/10.1145/3613905.3650889

1 INTRODUCTION

Passwords are the dominant form of digital authentication today and will likely remain prevalent in the foreseeable future [1, 3, 7, 11]. This widespread presence imposes a considerable burden on users, taxing them with the complexity of creating and managing passwords for a multitude of accounts [17, 21, 25]. As a result, people resort to coping strategies that often introduce security risks, such as opting for weak memorable passwords [27]. While password managers (PMs) can help users increase security and usability [4], these apps are not always used efficiently [5, 23, 27]. Beyond generating strong passwords and auto-filling them, a common functionality provided by PMs are *credential audit* (CA) interfaces that inform users about their overall password health. These reports include metrics such as password strength, number of compromised credentials, and reused passwords, usually presented in a dashboard-like fashion pointing at important issues that require action. Some examples are shown in Figure 1. Though potentially highly beneficial to improve password security, the usage of credential audits in the wild is under-explored and some preliminary studies signal that usage might be hampered because people find the reported results to be confusing and overwhelming [18]. To bridge this knowledge gap, we aim to understand whether and how people use credential audits, what usability challenges they encounter, and what can be improved to enhance their adoption and utility. Our guiding research questions are:

- RQ1: What are the main reasons users (do not) use credential audit tools?
- RQ2: How usable and useful are credential audit tools?
- RQ3: How could credential audit tools be improved?

We conducted interviews with 11 PM users, considering a diverse sample in terms of age, gender, used password manager, and usage of credential audits. The main findings indicate that despite feeling overwhelmed by the interface, participants expressed a commitment to enhance their password security by using credential audit tools. CA users attribute potential overwhelm to personal responsibility, not to tool issues. Non-users lack awareness or motivation, while adopters are driven by security concerns and external triggers. From the usability perspective, users would like to have more clarity and transparency in the feedback provided by the CA (e.g., Why is a password weak?). Based on participants’ experiences, we provide recommendations to improve the utility of CA interfaces and increase their adoption, contributing to a more secure authentication ecosystem.

2 BACKGROUND AND RELATED WORK

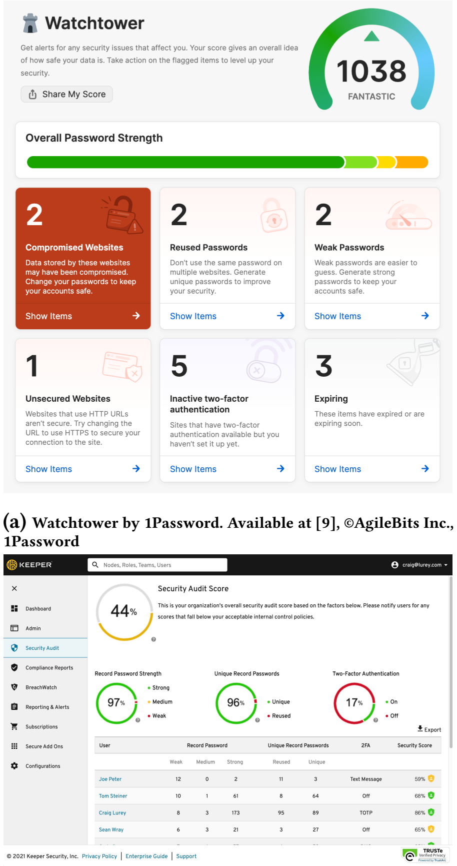
In this section, we discuss CA interfaces in password managers, highlighting relevant findings from past studies on these tools and emphasizing how we extend existing research.

2.1 Credential Audit Interfaces

Currently, credential audit tools are widely integrated across various password managers, each with its different approach to handling them. This includes differences in the tool names, how results

[*]Both authors contributed equally to this research.

CHI EA ’24, May 11–16, 2024, Honolulu, HI, USA
© 2024 Copyright held by the owner/author(s).
ACM ISBN 979-8-4007-0331-7/24/05.
https://doi.org/10.1145/3613905.3650889



Get alerts for any security issues that affect you. Your score gives an overall idea of how safe your data is. Take action on the flagged items to level up your security.

Data stored by these websites may have been compromised. Change your passwords to keep your accounts safe.

Don't use the same password on multiple websites. Generate unique passwords to improve your security.

Weak passwords are easier to guess. Generate strong passwords to keep your accounts safe.

Websites that use HTTP URLs aren't secure. Try changing the URL to use HTTPS to secure your connection to the site.

Sites that have two-factor authentication available but you haven't set it up yet.

These items have expired or are expiring soon.

(a) Watchtower by 1Password. Available at [9], ©AgileBits Inc., 1Password

This is your organization's overall security audit score based on the factors below. Please notify users for any scores that fall below your acceptable internal control policies.

User	Record Password			Unique Record Passwords		2FA	Security Score
	Weak	Medium	Strong	Reused	Unique		
Joe Peter	12	0	2	11	3	Text Message	59%
Tom Steiner	10	1	61	8	64	Off	68%
Craig Lurey	8	3	173	95	89	TOTP	86%
Sean Wray	6	3	21	3	27	Off	65%

© 2021 Keeper Security, Inc.

(b) Security Audit by Keeper. Available at [12],©Keeper Security Inc.

Figure 1: Examples of credential audit interfaces integrated in password managers.

are displayed, and whether these features are accessible for free or require a paid version. As seen in Table 7 in the Appendix, which shows an overview of CA tools in different password managers, nearly all of them provide interfaces that summarize identified weak passwords, reused passwords, and those compromised in data breaches. Additional features may include details on unused multifactor authentication, credentials stored for unsecured websites (http), or credentials that need to be changed due to expiration requirements. We explored the CA features of a set of nine different password managers, namely, 1Password, Bitwarden, Dashlane, KeePass, KeePassXC, Keeper, LastPass, RoboForm, and Zoho Vault. We chose this set based on the work by Simmons et al. [24] and added additional PMs recommended by CNET [6].

A lot of previous works focused on the adoption and usability of password managers [14, 21, 22]. Consistent with our findings (cf. Section 4), these works collectively highlight security, usability, and convenience as principal reasons for password manager utilization.

To our best knowledge, there is no study mainly focusing on the investigation of credential audit interfaces. However, they have been partially covered by a few publications.

Oesch et al. [18] explore user behaviors and motivations related to the use of PMs. Their investigation involves a user study with 32 participants through semi-structured interviews. It is revealed that many users express feeling overwhelmed by the multitude of results presented by credential audits in their password managers. They also stated that users self-audit their credentials instead of using automated processes by tools. Additionally, the study identifies users' positive responses to automatically triggered warnings in Chrome's built-in password manager, specifically those related to compromised passwords.

Our findings indicate that incorporating triggers or warnings could enhance the usability of CA interfaces and potentially boost adoption. While results of Oesch et al. showed that their participants dislike audit features because of overwhelm [18], our participants were less likely to feel overwhelmed, a sentiment that could be influenced by the quantity of results presented. Further, non-users mentioned depending on their own security practices, such as manual audits or active two-factor authentication mechanisms, rather than relying solely on auditing tools.

In 2021, Simmons et al. [24] systematized use cases of password managers and system designs implemented to address those. They describe design paradigms of credential audits including the initiation methods for audits. Furthermore, they suggest prioritizing audit results to prevent users from feeling overwhelmed by the amount of information, allowing them to concentrate on the most crucial recommendations first. Consistent with our research, participants expressed a desire for features that would enable them to prioritize crucial results. This includes functionalities like categorization and the presentation of a subset of issues, aiming to prevent overwhelm and address urgent cases promptly (cf. Section 4.3).

Amft et al. [2] conducted a study on users' strategies and habits when setting up PMs. In the initial phase, an expert review was executed to provide an overview of what setup features PMs offer to users who want to add existing credentials. They examined security centers (referred to as credential audit tools), password strength meters, and breach warning features within PMs. The authors highlighted that security centers or breach reports are often premium features in most password managers and are not accessible to users of free versions. However, no additional usability analyses or user studies on these features were explored.

2.2 Usability and Adoption of Password Managers

In 2019, Pearman et al. [27] investigated factors influencing the adoption of password managers, their effective use, and features related to password generation by conducting 30 semi-structured interviews with PM users and non-users. Their findings indicate that users of built-in password managers primarily adopt them because of convenience, whereas the main adoption reason for users of separately installed password managers is security.

Mayer et al. [14] developed an online survey with 277 participants, recruited from a private university in the US to explore awareness, password strategies, and motivations or barriers for (non-) usage of password managers. They found out that usability is a key factor for adoption, followed by convenience such as the relief of remembering passwords.

In line with our investigation on the reasons for password manager adoption among users, the primary factors cited were security and convenience, each accounting for 64%. Usability was also highlighted, with 27% of interview participants emphasizing its importance (cf. Section 4).

3 METHODOLOGY

We conducted 11 semi-structured interviews to explore participants' usage and understanding of credential audits within password managers, as well as their suggestions for improving usability.

3.1 Recruitment and Ethics

We used Prolific [20] to recruit English-speaking participants in the US. They were asked to take a screening survey to confirm eligibility criteria (being a desktop PM user, available for interview). We also collected answers on usage habits, CA awareness, and demographics, to support the selection of a diverse sample for the interview. We focus on a specific platform (desktop PMs vs browser or mobile-based) because usability can be highly connected to the device. For comprehensive feedback within the desktop context, we do not limit the study to a specific PM but include the experiences of users with a diverse range of managers.

Before running the study, we successfully pretested the screening survey to correct question-wording and estimate the time to complete it. The questionnaire was administered in May 2023 via the LimeSurvey web-based survey tool [8]. The follow-up interviews were conducted by the same researcher in July 2023 using the BigBlueButton (BBB) platform [10]. Both tools are hosted by our university and compliant with GDPR. Participants spent on average 3-5 minutes for the screening and we compensated completed surveys with GBP 1 (GBP 12/hour). The interviews lasted around 35-50 minutes and were compensated at an hourly rate of GBP 15.6, recognizing the additional effort of participants to take part. Participants were >18 years old and provided informed consent. The study was approved by our university's IRB and the survey questions and interview script are available in the Appendix (see Appendix A and B).

3.2 Interviews

Procedure. At the start of each session, we reminded the participants that the interview would be audio-recorded and pointed to the terms of consent. We then started the recording, introduced ourselves, and described the purpose and structure of the interview, offering the possibility to clarify any doubt. To guarantee privacy, participants were instructed not to share any personal details or credentials with us.

The interview script was organized into four blocks. First, we asked about password manager usage and password habits to have a general context. Second, we explained credential audit tools and posed questions about their present or previous adoption, along with the reasons behind their choices. Third, we requested participants to access their PM, inspecting the outcomes presented by the integrated CA tools and probing their understanding and attitudes towards presented data utility and usability. This block included questions about potential behavioral actions in response to the presented information. In the final part of the interview, we explored potential improvements and asked whether the participant would change any aspects of the CA interface.

Data Analysis. Interview audios were transcribed, using the Whisper transcription service by OpenAI [19], and analyzed following an iterative inductive coding approach [16]. To prepare for the qualitative analysis of the interview scripts, we mapped the interview questions to the dedicated research questions (RQ1-reasons, RQ2-usability, RQ3-improvements). For the analysis, we employed QCAmap [15] as our tool. Codebooks were created per research question by one researcher, who coded all the interviews, ensuring a consistent application of the codes. Each phase of the analysis underwent intense discussion with a second researcher, expert on the topic, to validate the coding framework and refine the codebook. This analysis method is similar to approaches used in previous work [26].

4 RESULTS

Of the 350 participants that took part in the screening, 64% (223) use a PM, of which 61% use it primarily on a laptop/desktop PC, while the remaining 39% use PMs on mobile devices. From the PM users, 67% indicated that their manager offers CA tools and half of those (75) currently use CA tools, whereas 51 individuals have previously adopted these tools but have not been using them for several months. Overall, 45 candidates met all the eligibility criteria: (1) PM user, (2) primarily using PM on laptop/PC, (3) only 3rd party PMs (no browser/OS PMs), (4) willing to do an interview. Those were then invited to an interview. In the end, 12 participants agreed to participate in an interview with us. We excluded one interview from the analysis as they did not provide answers to most of the questions. The participant experienced technical issues, distractions, and lack of focus, often leading to difficulty understanding the questions even after repeated explanations.

The interviewee demographics are detailed in Table 2 and their PM usage background in Table 1, covering a wide range of password managers (6) and ranging in years of experience using them (from <1 to >10 years). When asked about why they use a PM, convenience, and security were both highlighted by seven participants each. For security, six participants mentioned using a PM to prevent data breaches, as a response to such incidents or for security best practices, which means avoiding password reuse. Regarding convenience, four participants emphasized the simplicity of not having to remember passwords as the primary benefit. Additionally, five participants mentioned the password generation function as the main reason for adoption.

In the following, we look into participants' perceptions of CA tools.

4.1 Reasons for (Non-) Usage of CA Tools

Security as the main reason. Among the participants who reported currently using credential audit tools, seven out of eleven mentioned security reasons. Specifically, all of them identified **password reuse detection** as a key motivating factor. The second most crucial feature, identified by six participants, was the **detection of weak passwords**. Additionally, five participants mentioned initiating a credential audit with the specific purpose of checking for

Table 1: Interviewee background

Participant	PM	Years of PM Usage	PM provides CA	Using CA
P1	Keeper	7-8	Yes	Yes
P2	LastPass	2	Yes	Yes
P3	1Password	5-6	Yes	Yes
P4	RoboForm	>10	Yes	In the past
P5	LastPass	<1	Yes	No
P6	Zoho Vault	5-6	I don't know	-
P7	Bitwarden	>10	Yes	In the past
P8	LastPass	5-6	No	-
P10	1Password	3	I don't know	-
P11	Bitwarden	>10	Yes	Yes
P12	Bitwarden	5-6	Yes	Yes

Table 2: Interviewee demographics

		n = 11	Perc.
Gender	Male	6	55%
	Female	4	36%
	Non-Binary	1	9%
Age	35-44	5	45%
	25-34	4	36%
	55-64	1	9%
	65-74	1	9%
Education	Master's degree	4	36%
	Some college	3	27%
	Bachelor's degree	2	18%
	Associate degree	2	18%
IT	No	9	82%
	Yes	2	18%

compromised or exposed passwords, intending to take actions based on the audit results.

External and internal triggers support usage. The same seven individuals as above, expressed that their usage is also driven by triggers. All of them indicated that they initiate a credential audit following their **awareness of a data breach** outside the password manager, such as through news sources or notifications from the affected website:

> *"[...] especially if the data breach was on a site that I have a, an ID and password to, absolutely, I will go in as quickly as possible and change things up as much as I can." (P4)*

Other mentioned reasons include **internal prompts** from the password manager, such as pop-ups or email notifications for three individuals, and **personal reminders**, such as calendar notifications or individual decisions to start a credential audit (2). Notably, all participants demonstrated a commitment to improving their security.

Personal reasons for non-usage. Five participants were either **unaware of the feature's existence** in their password manager or had difficulty finding it due to its placement or confusing wording. Three participants expressed a **lack of motivation** to use CA tools. They mentioned needing a specific motivation, such as the urgency or acute danger of compromised credentials, to prompt their use. Otherwise, they indicated not to engage with the feature. Two participants stated that they stay away from using CA tools because they believe their **own practices** in terms of security, such as manually assessing their passwords (without CA support) or having activated two-factor authentication, are sufficient.

4.2 Usability

Nine interview participants expressed positivity regarding the usability of the credential audit interface in their PM, while the other two individuals had more concerns. However, everyone provided mixed feedback regarding their overall user experience with CA tools. In our findings, we only differentiate between password managers when there are notable variations in the interfaces. Otherwise, we present the results collectively.

Overwhelming but manageable. We asked our participants about their perception regarding potentially overwhelming results from credential audits. In a scenario where a substantial number of weak passwords needed changing, most participants admitted **feeling overwhelmed but expressed a commitment** to solving the problem by updating their passwords. Despite the potential stress, they acknowledged the necessity of eliminating security risks. Additionally, two participants mentioned that a high number of vulnerable passwords would be the user's responsibility:

> *"It'd probably be a little bit overwhelming if there was a large number of things that need to be changed, just knowing that each one requires its own customized workflow for changing the password. But [...] if I looked at it and I saw 300 passwords were exposed, that's, that's kind of on me, right? It's my fault. So like, I need to fix it, right?" (P11)*

One participant suggested a practical approach to handling overwhelming situations. They recommended breaking down a large problem into smaller, more manageable tasks and tackling them individually to prevent feeling overwhelmed:

> *"We have a saying here in the United States, 'eating the buffalo one bite at a time'. It's a very, very large meal [...] you can't eat the whole thing, but if you take bites out of it, then it might be more manageable." (P7)*

Lack of transparency. Five individuals expressed dissatisfaction with the **lack of transparency** and the limited information presented which makes this the primary concern for CA tool users. The affected PMs were LastPass, Bitwarden and 1Password. All five raised concerns about the absence of transparency in the password strength check, including a lack of clarity about how a password's strength is determined and why a password is considered weak.

> *"It says 'not strong', and it's not entirely clear to me what 'not strong' means, whether it's like, too short, whether it kind of lacks a certain complexity, which I guess could be kind of incorporated under the same umbrella as too short." (P7)*

> *"So, just give me an explanation of why it thinks it's weak." (P10)*

Additionally, four participants complained about missing or confusing information regarding reused passwords, questioning how often and where they were reused. Other transparency aspects that were perceived as lacking included an unclear security score, with questions arising about how it is computed and understood. This indicates a clear wish for more transparency among the users.

Unhelpful and helpful functionalities. Three participants mentioned hidden information, such as details behind hover-actions or incomplete URLs as issues. Another three liked the direct website navigation feature for password changes. On the other hand, some participants found the password change process in Bitwarden to be manual and involving too many steps, especially when compared to LastPass.

Information rather than presentation leads to actions. When exploring participants' motivations for auditing credentials and making changes, we investigated whether they were influenced more by the information or the visual presentation of statistics. The majority indicated that their primary motivation was the results and the information provided. Two participants mentioned that visual elements played a motivating role, while others stated that it was a combination of both factors. Some participants noted that motivation depended on the number of issues identified and that visual presentation could influence motivation to address problems, e.g., when only a subset of issues was shown (cf. Section 4.3).

> *"It has a lot to do with presentation, [...] if you tell me there's 150 problems, I don't know if I'm gonna be like, oh yeah, let me go tackle those 150 problems right now. But [...] if you package it in a way that makes it feel like quick, doable, [...] you don't have to do it all at once, but that it's moving towards getting you in better sort of like security health, then I think people, including me, would be more willing to do it." (P7)*

4.3 Improvements

Prioritize urgent results. To address the challenge of overwhelming users with extensive information in the results, particularly when dealing with a high number of vulnerable passwords, a desired solution by the majority of participants (8 out of 11) is **prioritization of the results**. This involves the ability to categorize the output by importance, allowing users to focus on urgent cases, and frequently or recently accessed sites. This approach aims to reduce feelings of overwhelm and provide clear guidance on which passwords require immediate attention.

> *"If there was a high number, they're like, OK, this one needs to be changed immediately. This is priority, change it now versus the other ones, I wouldn't feel so overwhelmed and so taken aback." (P5)*

Additionally, five participants wish for extra attention flags or alerts to highlight issues requiring urgent care, aligning with the need for prioritization. Meanwhile, two participants suggested a simplified interface mode, where only essential information is displayed as a summary of results. Another idea is the permanent presentation of a subset of issues which would consistently show a selected portion of the results to enhance motivation for making necessary changes.

> *"It might be nice [...] if there was one extra little flag on there that says [...] immediate attention required." (P4)*

Triggers can support. As mentioned in Section 4.1, reminders serve as a significant motivator for individuals to use these tools. Participants expressed the preference for being reminded to conduct credential audits, either through external personal reminders like emails or calendar alerts (7) or through prompts directly from the password manager, such as popup notifications or emails that allow snoozing, if not already a feature in the PM.

Wish for additional or enhanced functionalities. Seven participants expressed a desire for additional or improved functionalities within the CA tool. Specifically, four participants highlighted a general wish for enhanced methods, such as implementing the process of changing passwords directly within the PM and ensuring automatic synchronization with the respective platforms. However, participants acknowledged the challenges in implementing such a feature and recognizing potential security risks associated with it.

> *"I think really the next step for password managers is going to be some form of automation where it can go in and automatically change your passwords. But that's also asking for a mountain of security issues in and of itself." (P12)*

Another participant expressed a wish for an indicator that notifies when a username has been reused, as this could pose security concerns if leaked. Additionally, one participant desired the PM to display failed login attempts. Furthermore, three participants suggested embedding a sorting list feature, allowing passwords to be sorted based on their weaknesses, enabling users to address the weakest passwords first. Other responses included minor usability improvements such as better wording for credential audit tools and merging similar functionalities to avoid confusion.

5 LIMITATIONS

In recognizing study limitations, several factors deserve consideration. Firstly, when we asked about participants' security habits, social desirability may have shaped their answers to align with perceived expectations rather than expressing their genuine opinions or true experiences [13]. However, the participants are all PM users and already security-minded. Secondly, it is important to note that our survey questions regarding the potential usefulness of prompting and prioritization in password managers were framed with specificity not to prime participants, but to directly probe areas of limited usability as highlighted in previous literature. This deliberate focus aligns with our research objectives to deepen the understanding of these specific usability aspects, which are pivotal yet under-explored in current studies. In contrast, the remainder of our survey comprises neutral questions designed to elicit a broader range of insights, enabling us to capture both targeted and emergent themes in credential audits usability.

The sample size limitations, particularly in age and gender diversity, may impact the broader applicability of our findings. Our participants exclusively come from the United States, and interviews in other countries might yield varied results due to cultural and contextual differences in password management practices.

Excluding participants using mobile or browser-based password managers, based on their common usage according to our screening survey, may impact the study's completeness. Future research should consider including this group, particularly as some password managers provide credential audit tools on mobile devices. Lastly, for the interview participants that were not using CAs or unfamiliar with their features, the researcher demonstrated these functionalities. This may have introduced biased attitudes as participant perceptions are not based in first hand experience. Further, our study involved participants using a variety of password managers, each with different features and visual presentations, potentially influencing the outcomes.

6 CONCLUSION AND FUTURE WORK

In our study, we conducted 11 interviews to explore the reasons behind adopting or not adopting credential audit tools. We also assessed the perceived usability of current tools in various password managers, how they can be improved, and investigated motivations for primary PM adoption. Our findings are in harmony with existing research [14, 18, 21, 22, 24].

Users adopt CAs primarily for security reasons, such as identifying reused, exposed, or weak passwords. Non-adoption often stems from a lack of awareness or motivation. Despite the possibility of being overwhelmed by results, participants expressed commitment to enhancing their security. While individuals stated that results lack detailed explanations, user satisfaction with the overall experience is evident. Suggestions for improvement include implementing prioritization techniques like categorizing results. The study highlights the role of CA tools in supporting user security but underscores the need for enhancements and increased awareness among PM users. For future research, we recommend exploring mobile or browser-based PMs and expanding the study sample, especially including more non-adopters for meaningful comparisons.

REFERENCES

[1] Anne Adams and Martina Angela Sasse. 1999. Users Are Not The Enemy. *Commun. ACM* 42, 12 (1999), 40–46. https://doi.org/10.1145/322796.322806

[2] Sabrina Amft, Sandra Höltervennhoff, Nicolas Huaman, Yasemin Acar, and Sascha Fahl. 2023. "Would You Give the Same Priority to the Bank and a Game? I Do Not!" Exploring Credential Management Strategies and Obstacles during Password Manager Setup. In *Nineteenth Symposium on Usable Privacy and Security (SOUPS 2023)*. USENIX Association, Anaheim, CA, 171–190. https://www.usenix.org/conference/soups2023/presentation/amft

[3] Joseph Bonneau, Cormac Herley, Paul C. van Oorschot, and Frank Stajano. 2012. The Quest to Replace Passwords: A Framework for Comparative Evaluation of Web Authentication Schemes. In *IEEE Symposium on Security and Privacy, SP 2012, 21-23 May 2012, San Francisco, California, USA*. IEEE Computer Society, San Francisco, California, USA, 553–567. https://doi.org/10.1109/SP.2012.44

[4] Patricia Arias Cabarcos, Andrés Marín López, Diego Palacios, Florina Almenárez, and Daniel Díaz Sánchez. 2016. Comparing Password Management Software: Toward Usable and Secure Enterprise Authentication. *IT Prof.* 18, 5 (2016), 34–40. https://doi.org/10.1109/MITP.2016.81

[5] Sonia Chiasson, Paul C. van Oorschot, and Robert Biddle. 2006. A Usability Study and Critique of Two Password Managers. In *Proceedings of the 15th USENIX Security Symposium, Vancouver, BC, Canada, July 31 - August 4, 2006*, Angelos D. Keromytis (Ed.). USENIX Association. https://www.usenix.org/conference/15th-usenix-security-symposium/usability-study-and-critique-two-password-managers

[6] Clifford Colby. 2023. *Best Password Manager to Use for 2023*. https://www.cnet.com/tech/services-and-software/best-password-manager/ Visited on 23/01/2024.

[7] Dinei A. F. Florêncio and Cormac Herley. 2007. A large-scale study of web password habits. In *Proceedings of the 16th International Conference on World Wide Web, WWW 2007, Banff, Alberta, Canada, May 8-12, 2007*, Carey L. Williamson, Mary Ellen Zurko, Peter F. Patel-Schneider, and Prashant J. Shenoy (Eds.). ACM, 657–666. https://doi.org/10.1145/1242572.1242661

[8] LimeSurvey GmbH. 2024. *LimeSurvey*. https://www.limesurvey.org/ Visited on 25/01/2024.

[9] AgileBits Inc. 2024. *Watchtower by 1Password*. https://support.1password.com/watchtower Visited on 25/01/2024.

[10] BigBlueButton Inc. 2024. *BigBlueButton*. https://bigbluebutton.org/ Visited on 25/01/2024.

[11] Blake Ives, Kenneth R. Walsh, and Helmut Schneider. 2004. The domino effect of password reuse. *Commun. ACM* 47, 4 (2004), 75–78. https://doi.org/10.1145/975817.975820

[12] Inc. Keeper Security. 2024. *Security Audit by Keeper*. https://docs.keeper.io/enterprise-guide/security-audit Visited on 25/01/2024.

[13] Jonathan Lazar, Jinjuan Feng, and Harry Hochheiser. 2017. *Research Methods in Human-Computer Interaction, 2nd Edition*. Morgan Kaufmann. https://www.sciencedirect.com/science/book/9780128053904

[14] Peter Mayer, Collins W Munyendo, Michelle L Mazurek, and Adam J Aviv. 2022. Why Users (Don't) Use Password Managers at a Large Educational Institution. In *31st USENIX Security Symposium (USENIX Security 22)*. USENIX Association, 1849–1866. https://www.usenix.org/conference/usenixsecurity22/presentation/mayer

[15] Philipp Mayring and Thomas Fenzl. 2024. *QCAmap*. https://www.qcamap.org/ Visited on 20/03/2024.

[16] Matthew B Miles and A Michael Huberman. 1994. *Qualitative data analysis: An expanded sourcebook*. Sage.

[17] Sean Oesch and Scott Ruoti. 2020. That was then, this is now: A security evaluation of password generation, storage, and autofill in browser-based password managers. In *Proceedings of the 29th USENIX Conference on Security Symposium*. USENIX Association, 2165–2182. https://www.usenix.org/conference/usenixsecurity20/presentation/oesch

[18] Sean Oesch, Scott Ruoti, James Simmons, and Anuj Gautam. 2022. "It Basically Started Using Me:" An Observational Study of Password Manager Usage. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–23.

[19] OpenAI. 2024. *Whisper*. https://openai.com/research/whisper Visited on 20/03/2024.

[20] Stefan Palan and Christian Schitter. 2018. Prolific. ac—A subject pool for online experiments. *Journal of Behavioral and Experimental Finance* 17 (2018), 22–27.

[21] Sarah Pearman, Jeremy Thomas, Pardis Emami Naeini, Hana Habib, Lujo Bauer, Nicolas Christin, Lorrie Faith Cranor, Serge Egelman, and Alain Forget. 2017. Let's Go in for a Closer Look: Observing Passwords in Their Natural Habitat. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security, CCS 2017, Dallas, TX, USA, October 30 - November 03, 2017*, Bhavani Thuraisingham, David Evans, Tal Malkin, and Dongyan Xu (Eds.). ACM, 295–310. https://doi.org/10.1145/3133956.3133973

[22] Hirak Ray, Flynn Wolf, Ravi Kuber, and Adam J. Aviv. 2021. Why Older Adults (Don't) Use Password Managers. In *30th USENIX Security Symposium, USENIX Security 2021, August 11-13, 2021*, Michael D. Bailey and Rachel Greenstadt (Eds.). USENIX Association, 73–90. https://www.usenix.org/conference/usenixsecurity21/presentation/ray

[23] Sunyoung Seiler-Hwang, Patricia Arias-Cabarcos, Andrés Marín, Florina Almenares, Daniel Díaz-Sánchez, and Christian Becker. 2019. " I don't see why I would ever want to use it": Analyzing the Usability of Popular Smartphone Password Managers. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*. ACM, 1937–1953. https://doi.org/10.1145/3319535.3354192

[24] James Simmons, Oumar Diallo, Sean Oesch, and Scott Ruoti. 2021. Systematization of Password ManagerUse Cases and Design Paradigms. In *ACSAC '21: Annual Computer Security Applications Conference, Virtual Event, USA, December 6 - 10, 2021*. ACM, 528–540. https://doi.org/10.1145/3485832.3485889

[25] Elizabeth Stobert and Robert Biddle. 2014. The password life cycle: user behaviour in managing passwords. In *10th symposium on usable privacy and security (SOUPS 2014)*. USENIX Association, 243–255. https://www.usenix.org/conference/soups2014/proceedings/presentation/stobert

[26] Lillian Yang and Carman Neustaedter. 2018. Our House: Living Long Distance with a Telepresence Robot. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW, Article 190 (nov 2018), 18 pages. https://doi.org/10.1145/3274459

[27] Shikun Zhang, Sarah Pearman, Lujo Bauer, and Nicolas Christin. 2019. Why people (don't) use password managers effectively. In *Fifteenth Symposium on Usable Privacy and Security, SOUPS 2019, Santa Clara, CA, USA, August 11-13, 2019*, Heather Richter Lipford (Ed.). USENIX Association. https://www.usenix.org/conference/soups2019/presentation/pearman

A SCREENING SURVEY

A.1 Password Managers

Q1 Which password manager do you use? ○ 1Password ○ Bitwarden ○ Dashlane ○ KeePass / KeePassXC ○ LastPass ○ RoboForm ○ I don't use a password manager ○ Other (please specify)

Q2* On what device do you use your password manager most frequently? ○ Desktop PC ○ Laptop ○ Tablet ○ Phone ○ Other

Q3* How often do you use your password manager? ○ Daily ○ Weekly ○ Monthly ○ Yearly ○ Other (please specify)

Q4* Does your password manager provide you with information regarding password health, password statistics or such? ○ Yes ○ I don't know ○ No

Q5 *If Q4 was answered with Yes.* Do you use this information or the corresponding tools? ○ No, I have never used them ○ I used them in the past, but haven't for several months now ○ Yes, I use them

Q6 How would you estimate your security awareness? ○ Very poor ○ Poor ○ Fair ○ Good ○ Excellent

Q7 How important is security for you? ○ Not important ○ Slightly important ○ Moderately important ○ Important ○ Very important

* These questions are not displayed if the participant answers Q1 with "I don't use a password manager".

A.2 Demographics

Q8 What is your gender? ○ Female ○ Male ○ Non-binary ○ Prefer not to say ○ Prefer to self-describe as: (please specify)

Q9 How old are you? ○ 18-24 years old ○ 25-34 years old ○ 35-44 years old ○ 45-54 years old ○ 55-64 years old ○ 65-74 years old ○ 75 years or older ○ Prefer not to say

Q10 What is the highest level of school you have completed or the highest degree you have received? ○ Less than high school diploma or equivalent ○ High school graduate or GED ○ Completed some college but not a degree ○ Associate degree ○ Bachelor's degree ○ Master's degree ○ Doctorate degree or Professional degree ○ Other ○ Prefer not to say

Q11 Which of the following best describes your educational background or job field? ○ I have an education in, or work in, the field of computer science, engineering, or IT. ○ I do not have an education in, or work in, the field of computer science, engineering, or IT. ○ Prefer not to say

A.3 Closing Questions

Q12 Would you be willing to participate in a one-on-one remote interview with our research group about how you use your password manager? You will receive compensation for your time. The interview will take less than one hour and the audio will be recorded. Your interview data will be anonymized and kept confidential. ○ Yes ○ No

Q13 Do you agree to be contacted for possible future studies by our research group? Your answer to this has no influence on compensation or eligibility for the interview in this study. ○ Yes ○ No

Q14 Please enter your Prolific ID so that we can compensate you for this survey. (Free text)

B INTERVIEW SCRIPT

B.1 General Questions about PM usage and password habits

Q1.1 For how long have you been using a PM?

Q1.2 Do you use a PM for work or privately?

Q1.3 What is your motivation/the reason to use a PM? (E.g., Security / Don't want to remember passwords / requirement at work)

Q1.4 How did you decide which PM to use?

Q1.5 Do you use your PM for all passwords?

Q1.6 If no: How do you decide which passwords to store?

Q1.7 Do you reuse the same passwords between accounts? When/for which type of accounts?

Q1.8 How do you create your passwords?

Q1.9 (If manual): Why and what is your strategy to manually create passwords?

Q1.10 On average, how good do you think is the security of your passwords? ○ Very poor ○ Poor ○ Fair ○ Good ○ Excellent

Q1.11 If relevant for CA functionality: Do you use a paid or free version of your PM?

Q1.12 (If free version): Would you be willing to pay for your PM to get more CA functionality?

Q1.12.1 What features would the paid PM have?

Q1.12.2 How much would you pay for a one-time payment?

Q1.12.3 Many third-party password managers require a monthly fee to use their services. How much would you be willing to pay monthly?

B.2 Questions about CA interface

Q2.1 Do you use (some of) the provided CA tools? If only a subset, which ones?

Q2.1.1 Why (not)?

Q2.1.2 How do you use it (e.g., motivation, purpose)?

Q2.1.3 How often? (When was the last time you used it?)

Q2.1.4 Is your usage somehow triggered (e.g., PM prompts you to do it or prompts you with results of automatic CA, habit (e.g., monthly reminder), specific news (e.g., data breaches))?
If no:

Q2.1.4.1 Do you sometimes change your passwords?

Q2.1.4.2 If yes: what causes you to do so?

Now researcher tells them to unlock their PM to answer the following questions.

B.3 CA results - statistics

Q3.1 How many passwords do you have?

Q3.2 How many of them are compromised? (If you prefer not to answer, that's also fine.)

Q3.3 What is the health score of your password (average, if displayed / roughly how many are considered poor, weak, good, excellent if color-coded)

Q3.4 What is your opinion on these statistics/results?

Q3.5 Were you aware of this information?
Q3.6 If no: After looking at this information, will you take any action?
Q3.7 If you have accessed this functionality in the past, what actions (if any) have you taken based on the information displayed?

B.4 CA results - presentation

Q4.1 What do you think about how the results are presented?
If they only answer very shortly or with something like "I don't know", give examples:
Q4.1.1 Understandable? Easily?
Q4.1.2 Clear layout?
Q4.1.3 Overwhelming?
Q4.1.4 Thoughts on the type of view (e.g., list, dashboard - what would you prefer?)
Q4.1.5 What information do you find (not) useful?
Q4.2 Does the way the results are presented (not the results themselves) lead you to make any changes?
If yes:
Q4.2.1 Which changes?
Q4.2.2 Do you know what you would have to do to improve your password health/to remove the issue?
Q4.2.3 How easy is it for you to make changes starting from the CA results view?
(If no comment, give examples: can you edit directly from result view / do you have to remember the issues, close the view and manually select the affected passwords in another view?)
Q4.2.4 How motivated are you to make changes based on the results or their presentation?
Q4.2.5 Imagine the results show 50% of your passwords to be weak. to be weak.
Q4.2.5.1 Would you change them?
Q4.2.5.2 You said you have a "good/bad/excellent/. . . " score. Would you like to improve it?
Q4.3 Would you be more likely to change any of your passwords if your PWM prioritized the results of a CA (e. g., based on score, password category)?
Q4.4 Would you be more likely to start a CA (or to do it more often) if your PM prompted you to do it?
Q4.5 What is your opinion on such a feature that prompts you (e. g., like/dislike, would you use/disable it)?
Q4.6 Do you have any suggestions for improvement for any of the points in 4.?

B.5 End of interview

Q5.1 Any further comments, ideas etc. that you haven't mentioned yet?
Q5.2 (If not covered yet and if they don't use CA tools: Will you use the CA functionality in the future?)
Q5.3 Please write your Prolific ID into the chat.
Q5.4 Do you have any questions?
Q5.5 Thank you for your participation.

C COMPLETE CODEBOOK

Table 3: Categories and codes used for coding the responses given by the interviewee to answer research question 1 with applied code frequencies, not considering multiple counts per script.

RQ1: What are the main reasons that users (do not) use credential audit tools?		
Category	**Code**	**Description**
Usability (9%)	Easy to use (1)	Interface is easy to use.
Trigger/Prompts (64%)	Internal trigger/prompt (3)	Notification about CA results in form of a mail or popup by the PM is sent.
	External trigger: Data breach (7)	Participant got informed about a data breach, e.g., by the news or website itself.
	External trigger: Reminder (2)	Something or someone reminds or prompts the participant to start a credential audit outside of the PM.
Personal Reasons (82%)	Non-Usage: Lack of motivation (3)	Participant needs a motivation to use it (more frequent), e.g., urgency or acute danger. Will not use it otherwise.
	Clearing up outdated accounts (1)	Participant uses the tool to get knowledge about old accounts and clears them eventually.
	Non-Usage: Laziness (2)	Participant is too lazy to use the tool.
	Non-Usage: Lack of awareness (5)	Participant was not aware that this feature exists in PM or did not find it because of the placement or confusing wording.
	Non-Usage: Own practices are enough (2)	Participant believes that own security practices are enough, e.g., activated 2FA or conducts own credential audits.
	Non-Usage: Lack of interest (1)	Participant is not interested in this particular feature or other features than the main used ones in a PM.
	Interest and curiosity (1)	Participant uses CA simply out of interest and curiosity.
Security (64%)	Password reuse (7)	Participant wants to check for reused/duplicated passwords.
	Security check (4)	Participant is interested in security checks or security knowledge, e.g., general strength of passwords.
	Password changing action (1)	Participant uses tool to know whether passwords need to be changed.
	Weak password (6)	Participant wants to check for weak passwords.
	Improve security (2)	Participant uses tool to improve security in general.
	Passwords at risk (5)	Participant wants to check for exposed passwords/compromised passwords.
	Inactive 2FA notification (2)	Participant wants to check inactive 2FA for websites.

Table 4: Categories and codes used for coding the responses given by the interviewee to answer research question 2 with applied code frequencies, not considering multiple counts per script.

RQ2: How usable and useful are credential audit tools?		
Category	**Code**	**Description**
Layout and Navigation (64%)	Easy to navigate (2)	It is easy to navigate through the interface and it is easy for a participant to find what she is looking for.
	Clear layout (4)	The layout, structure or positioning of elements is clear.
	Logical workflow (5)	Reaching the goal has a logical workflow order or steps to complete tasks are clear.
Functionalities (73%)	Customization (1)	The interface is customizable.
	Sorting (2)	Sorting the results list is a helpful functionality.
	Not helpful: Unsecure websites (2)	Information about unsecure websites in the CA tool is not helpful.

Table 4: Codebook for research question 2 (Continued)

RQ2: How usable and useful are credential audit tools?		
Category	**Code**	**Description**
	Not helpful: 2FA Availability (1)	Feature that indicates if 2FA is available for a specific website is not helpful.
	Not helpful: Age column (1)	Feature that allows sorting a list by age is not helpful.
	Not usable: Hidden functionality/information (3)	A participant is missing a specific functionality here or some information is missable.
	Navigates to affected website directly (3)	Functionality, where PM leads users directly to the website for changing password or does it automatically inside PM.
	Not usable: Changing password action (3)	Functionality for changing the password is confusing or requires too many steps.
Visual Presentation and Wording (45%)	Somewhat helpful coloring (2)	Depending on the use case, colors e.g., for distinguishing different types of passwords, are considered helpful.
	Visual motivation (1)	Presentation of certain parts motivates a participant to make changes or pay attention.
	Not usable: Confusing wording (3)	Confusing names of UI elements or functionalities.
	Not usable: Bad readability (1)	Text has bad readability in terms of visualization.
Information/Content Presentation (91%)	Contains all important information (3)	Interface contains all information that a participant is looking for.
	Information rather than presentation leads to actions (6)	Presented content is the motivator for a participant to make changes, e.g., changing a password, rather than the visual presentation.
	Helpful presentation of vulnerable passwords (5)	The presentation or the information itself of weak, reused, or exposed passwords is helpful.
	Not usable: Overwhelming dashboard (1)	CA dashboard overwhelms or confuses a participant as it contains too much information.
	Good amount of information (3)	The presented amount of information is not too much and not too little.
	Security check statistics (3)	A participant likes the visualization of statistics or thinks it is a helpful functionality with valuable information in general.
	Data density (1)	A participant prefers a higher amount of information, therefore, likes the density of shown information in the interface.
Information/Content Transparency (45%)	Not usable: No transparent password strength check (5)	It is not clear how the PM computes the strength of the password or why a password is weak.
	Not usable: No transparent password reuse check (4)	It is not clear where the password is reused or information is confusing.
	Not usable: Unclear meaning of security score (2)	Either it is not sufficiently explained how the security score was computed or a participant understood it in a wrong way.
	Not usable: Too less explanation (2)	A participant seeks more information or explanation at some point in the interface.
General (91%)	Easy to use (2)	The tool is, in general, easy to use.
	Understandable (9)	Interface is understandable, intuitive, and clear.
	Simple (2)	Interface is simple or clean in a positive way as it does not contain too many details, elements, or functionalities.
	Usable (3)	The interface is usable or user-friendly in general.
	Workable (2)	Interface is easy to work with.

Table 5: Categories and codes used for coding the responses given by the interviewee to answer research question 3 with applied code frequencies, not considering multiple counts per script.

RQ3: How could credential audit tools be improved?		
Category	**Code**	**Description**
Categorization/ Prioritization (73%)	Prioritization/Categorization of results (7)	A functionality that prioritizes or categorizes the results of a CA. Either automatic or manual.
	PM: Early categorization (1)	Participant wishes for a functionality where she can categorize credentials while entering them into PM. Similar to Prioritization-functionality but earlier.
	Prioritization of results of most/recently visited websites (3)	Participant wants to see results of most visited or recently visited websites prioritized in results of CA.
Prompting (91%)	Personal reminder (3)	Personal reminder outside the PM, e.g., scheduled mail or calendar notification, to start a credential audit.
	Prompts by PM (7)	Conditional or scheduled prompts by PM that triggers a participant to use CA tools in the form of popup notifications or emails.
	Extra attention flags/alerts (5)	Extra notification, e.g., popup or mail, for very important/urgent cases where a participant should pay extra attention or should act immediately.
	External prompts about data breach (1)	A participant would like to get an external prompt, e.g., by the website itself when a data breach happened.
Functionalities (64%)	PM: Simpler copy-pasting credentials functionality (1)	A participant wishes for simpler copy-pasting credentials functionality for the PM on the desktop computer.
	Unmask passwords in CA (1)	A participant wishes for the functionality of being able to unmask passwords in CA to easier identify passwords.
	Easier steps to rectify (4)	A participant wishes for easier steps to rectify/reach a solution, e.g., easier steps for changing a password.
	Display failed login attempts (1)	A participant wishes for the functionality where the PM (or the CA) shows failed login attempts for that account.
	Feature to sorting list (3)	A participant wishes for functionality of sorting the result list in CA.
	Indicating reused username (1)	A participant would like to have the feature where the CA indicates whether a username was reused.
	PM: Informing about weak password when entering it (1)	A participant wishes for the feature where PM hints when a password is weak while entering it, e.g., when adding it to the vault or when using it in a login form on a website.
Information Presentation (64%)	Merging similar functionalities (2)	A participant suggests merging similar functionalities or reports.
	Simplified interface (2)	A participant wishes for a simplified interface mode, where not all information is displayed at once but only important ones, e.g., as a summary of results.
	More explanation/information to found vulnerabilities (5)	A participant wishes for more explanation, e.g., why a password is weak.
	Only present a subset of issues (2)	A participant wishes that only a subset of issues is permanently presented.
	Better structure for indicating reused passwords (3)	A participant wishes for an improved visualization of reused passwords, e.g., where it is reused and how many times.
Usability Enhancements (45%)	Clearer visuals for results (3)	Results, e.g., weak/exposed passwords, should be made clearer like adding icons, better layout or colors.
	Use (better) wording (2)	Use (better) wording in the interfaces that leads to more actions.

Table 5: Codebook for research question 2 (Continued)

RQ3: How could credential audit tools be improved?

Category	Code	Description
Potential reasons for future usage (27%)	Potential usage: Prompts by PM (1)	A participant is more likely to start a credential audit when prompted by the password manager.
	Potential usage: Improve security (2)	A participant would use CA tools in the future to improve security.

Table 6: Categories and codes used for coding the responses given by the interviewee for PM Usage reasons with applied code frequencies, not considering multiple counts per script.

What are the main reasons why users use password managers?	
Category	**Code**
Security (64%)	Security (e.g., Fear of breaches, best practices) (6)
	Security reports (1)
	Would pay: More reports (1)
	Password generation (5)
Convenience (64%)	Overview for all passwords (3)
	Exporting passwords (1)
	No need to remember passwords (4)
	Convenience in general (2)
	Storing passwords (2)
	Device-Synchronization (2)
Usability (27%)	Easy to use (3)
Costs and Availability (36%)	Specific PM: Pricing (2)
	Specific PM: Availability (2)
Alternatives (45%)	Specific PM: Changed from other PM
Recommendations (55%)	Personal recommendations (5)
	Search engine results (1)

D CREDENTIAL AUDIT TOOLS OVERVIEW

Table 7: Comparison of tools in different password managers. ❍ = not available, ◗ = available in paid plan, ● = available for free.

	Name(s) of Tools	Weak	Re-used	Compromised	Additional Functions
1Password	Watchtower	◗	◗	◗	expired PWs, unsecured websites (http), inactive 2FA
Bitwarden	Vault Health Reports	◗	◗	●	unsecured websites (http), inactive 2FA
Dashlane	PW Health, Dark Web Monitoring	●	●	●	PW-specific detailed information on breaches
KeePass (v2.x)	PW Quality, Find Similar/ Duplicate PWs	●	●	❍	cluster reused PWs
KeePassXC ($v \geq 2.6.0$)	Database Reports	●	●	●	expired PWs, list of statistics, information where PWs are reused
Keeper	Security Audit, BreachWatch	◗	◗	◗	alert on breach
LastPass	Security Dashboard	●	●	●	expired PWs, missing PWs, inactive 2FA, alert on breach
RoboForm	Security Center	●	●	●	detect complete duplicates
Zoho Vault	Audit	●	●	◗	expired PWs, PWs containing username or dictionary words