TÉCNICO LISBOA

On Usable Security and Verified Password Managers

Carolina Maria da Cunha Carreira

Thesis to obtain the Master of Science Degree in

Computer Science and Engineering

Supervisors: Prof. João Fernando Peixoto Ferreira
Prof. Alexandra Sofia Ferreira Mendes

Examination Committee

Chairperson: Prof. Rui Filipe Fernandes Prada
Supervisor: Prof. João Fernando Peixoto Ferreira
Member of the Committee: Prof. Tiago João Vieira Guerreiro

November 2021

Acknowledgments

I would like to thank my mother and my father for their never-ending support and advice. I am very thankful for their time, their words of support and for giving me the best education I could have had.

I am deeply grateful to Prof. João F. Ferreira and Prof. Alexandra Mendes for for their invaluable advice, support, patience and encouragement without which this Thesis would not be possible. Thank you for believing in me.

I would also like to extend my sincere thanks to all the Professors and colleagues that were a part of PassCert whose inputs and knowledge made my work better.

Finally my thanks and appreciation to all that have assisted me and gave strength during this Thesis. I would like to express my gratitude to my boyfriend, my friends, my parents and my advisors. Without their understanding and encouragement it would be impossible to complete my work.

Thank you.

This work was partially funded by the PassCert project, a CMU Portugal Exploratory Project funded by Fundação para a Ciência e Tecnologia (FCT), with reference CMU/TIC/0006/2019.

Abstract

Password Managers (PMs) are useful tools to manage passwords but they are not widely used. Studies indicate usability problems and distrust from users as the reasons for the low adoption of PMs. As such, we propose extending an existing PM by implementing relevant usability best practices and increasing transparency by educating users about how PMs work. This project is part of the PassCert research project, which aims to build a formally verified PM. Therefore, another goal is to explore ways that effectively convey to users the formally verified properties. We performed user studies that suggest that our solution improves the usability of the PM and that we were able to convey relevant information about its formally verified features. We contribute with the first study on users' perceptions of formal verification on PMs and hope that our findings can help the formal verification security community better communicate with end-users.

Keywords

Usable Security; Password Manager; Formal Verification; Password Security

Resumo

Os gestores de passwords (PMs) são ferramentas úteis para gerir passwords, no entanto não são amplamente utilizados. Estudos indicam problemas de usabilidade e falta de confiança como barreiras à sua adopção. Como tal, propomos a extensão de um PM existente, implementando as práticas de usabilidade relevantes e aumentando a transparência educando os utilizadores sobre a forma como os PMs funcionam. Este projecto faz parte do projecto de investigação PassCert, que visa a construção de um PM formalmente verificado. Outro objectivo é explorar formas de transmitir informação aos utilizadores sobre verificação formal. Realizámos estudos de utilizadores que sugerem que a nossa solução melhora a usabilidade do PM e conseguimos transmitir informações relevantes sobre as suas funcionalidade formalmente verificadas. Contribuímos com o primeiro estudo sobre a percepção dos utilizadores da verificação formal em PMs e esperamos que as nossas conclusões possam ajudar a comunidade de verificação formal a comunicar melhor com os utilizadores.

Palavras Chave

Segurança Usável; Gestor de Passwords; Verificação Formal; Segurança de Passwords

Contents

List of Figures

List of Tables

Acronyms

OS	Operating System
PM	Password Manager
FAQ	Frequently Asked Questions
SUS	System Usability Scale
UI	User Interface

1

Introduction

Contents

Context. Text passwords are one of the most used security mechanisms [1]. However, they are not always used correctly. As Whitten and Tygar pointed out, security mechanisms are only effective when used correctly [2]. To be used correctly, text passwords should not be reused, simple, or easy to guess – this presents a challenge for users. In a study by Stobert et al. [3], only one of the 26 participants reported not reusing passwords between accounts and 73% reported reusing passwords either "always" or "frequently". Furthermore, a 2018 Eurobarometer Survey [4] concluded that only 28% of citizens stated that they used "different passwords for different sites". Not only is password reuse a problem but users also struggle with generating random passwords. Gaw et al.'s [5] study about password usage found that 51.79% (of 56 responses) believed that a friend had a higher chance of guessing their password, suggesting that they used non-random passwords with personal information.

It is in this context of insecure password usage that Password Managers (Password Manager (PM)) become an essential solution. When using a PM the user only needs to remember one password (the master password or primary password) and all others are saved in secure storage. Regarding password reuse and uniqueness, PMs allow secure password generation preventing users from having to generate the passwords themselves. Several governmental institutions such as the National Cybersecurity Center [6] and the European Union Agency for Cybersecurity [7] recommend the usage of PMs.

PMs' problems. As stated by Li et al. [8], although there have been some security problems in PMs [9], they are important tools that enable users to use stronger passwords, freeing them from the cognitive burden of remembering them. Despite this, there are still many users who do not fully trust password managers. Although PMs are recommended and seem to help users manage their passwords, they are not widely used. A study from Alkadi et al. [10] about smartphone PMs examined online reviews of password managers and elicited opinions from 352 respondents. This study reported that only 6.8% of respondents used a dedicated PM. Several studies tried to find the reasons for this phenomenon and have reached different conclusions: some stating unawareness of the existence of PMs [3, 10, 11], lack of trust [12, 13] or lack of motivation [10, 13]. One factor that was mentioned by all studies in some degree was usability problems [10, 11, 14–16].

Proposed solution. Formal verification is valuable when considering password security [17, 18] and so this project is part of the PassCert research project that aims to build a formally verified PM, guaranteeing properties on data storage and password generation [19]. These features will be implemented by other members of the PassCert project. A challenge that is addressed in this project is how to effectively convey the relevance of formal verification in a way that increases the confidence of users in verified PMs.

To improve trust and relay the importance of formal verification in a PM we propose the implementation of several usability best practices [2, 16, 20]. It is important to ensure that users understand what the system will formally assure; for this, we propose to explore the implementation of status symbols

that indicate that a certain action is formally verified. But even if users are aware that the PM has formally verified features they may still not know what formal verification is, so, it is also relevant to provide concise explanations about what it is. These explanations should not have technical jargon [21, 22] as to be understandable by beginner users in order to build trust in the tool. As misinformation can be a source of mistrust in PMs [10, 11, 13, 16], educating users about the benefits and usage of PMs is also something we want to implement in our solution with a Frequently Asked Questions (FAQ). Most users are familiar with FAQs and know how they work [21, 23].

In regards to improving usability features of the PM, we aim to implement tutorials about how the user interface works (for beginner users) [21]. Another important addition is to provide information about what each security setting means [24]: we implemented this with the use of tooltips. Tooltips are a good way to take context into account by monitoring the cursor location and providing helpful information about the icon under the cursor [23] we intended to use tooltips for this.

We implemented a proof-of-concept PM interface, taking as the starting point the interface of Bitwarden, an established and widely used PM. When choosing the PM to be extended, we considered several different ones and concluded that the Bitwarden browser extension PM would be the most impactful as it is easy to install, open-source, and easily extended. In our prototype we: expanded and added missing tooltips; added the formal verification icon; implemented explanations on formal verification; designed a tutorial, a FAQ and finally solved inconsistencies in Bitwarden's interface.

Evaluation. To evaluate the success of our ideas and proof-of-concept prototype, we performed qualitative usability tests and non-structured interviews with 15 users (see Section 5.2.1). We used the System Usability Scale (SUS) [25] and questionnaires to learn about the impact of our proposals and about what users think about formal verification in the context of PMs. An effective technique during usability testing that we used is to invite users to "think aloud" (sometimes referred to as "concurrent think-aloud") about what they are doing as they are performing the task [21].

1.1 Work Objectives

The goals of this project are to:

- survey usable security techniques that can be applied to improve password managers;
- ensure that the password managers developed in the context of the PassCert project integrate best practice guidelines developed by the usable security community;
- explore ways that effectively convey the formally verified properties of the password managers;
- learn more about users perception of PMs and formal verification.

1.1.1 Research Questions

We aim to answer the following research questions:

RQ1. What are the usable security techniques that can be applied to PMs?

RQ2. How can we effectively convey formally verified properties of a PM to its users?

RQ3. What are users' perceptions of formal verification and PMs?

1.2 Contributions

This project is part of the PassCert research project,[1] a CMU-Portugal exploratory project that aims to build an open-source, proof-of-concept password manager that through the use of formal verification, is guaranteed to satisfy properties on data storage and password generation. These features will be implemented by other members of the PassCert project.

The usable security techniques studied in this thesis were implemented by extending Bitwarden. Our results suggest that we have succeeded in developing an usable PM in the context of PassCert. Moreover, our findings seem to hint that our solution is more usable than the base Bitwarden PM. As such, we recommend that Bitwarden should implement some of our extensions such as more tooltips and a tutorial. These are the user support features that users interacted with more and that were most frequently mentioned in the user studies.

Another goal was to communicate with users about formal verification. The results suggest that our solution was successful in transmitting to users that certain features of the PM were formally verified. Likewise, we were also able to study users' perceptions of formal verification before and after using the PM. We found that most users did not know the concept before the study and by the end of the study most of them associated formal verification with security and some were able to give an accurate (but non-technical) explanation about what formal verification is. We hope that our interface extensions can be used by other projects and that the insights we gathered can help the formal methods security community communicate with end-users about its assurances.

1.2.1 Research Papers

Parts of the work presented in this thesis were used in the following research papers:

- ***Carolina Carreira**, João F. Ferreira, and Alexandra Mendes. Towards Improving the Usability of Password Managers. Presented at INFORUM 2021 (Comunicação). 2021 [26]*

[1] PassCert is a CMU Portugal Exploratory Project funded by Fundação para a Ciência e Tecnologia (FCT), with reference CMU/TIC/0006/2019

- ***Carolina Carreira**, João F. Ferreira, Alexandra Mendes, and Nicolas Christin. Exploring Usable Security to Improve the Impact of Formal Verification: A Research Agenda. In 1st International Workshop on Applicable Formal Methods (co-located with Formal Methods 2021). Beijing, China. 2021 [27]*

1.3 Organization of the Document

This thesis is organized in 6 chapters as follows: we begin in Chapter 1 with the introduction providing a brief overall description of our work, identifying the problem, solution, testing methodology and results. In the next section, Chapter 2 we provide more in-depth background on PMs, their advantages, problems and relevant usable security techniques that can be applied to them. In Chapter 3 we go over our goals and explain the design process behind our solution going through the choice of base PM and describing how we plan to ensure the PM's usability and convey the formally verified features of PassCert to users. In the next section, Chapter 4 we describe our iterative implementation process of the features described in the previous section. In Chapter 4 we also address the technical details of the extensions done to Bitwarden. After extending the PM we then evaluate our solution in Chapter 5 by performing user tests whose results are analyzed. Finally, we end the thesis with Chapter 6 presenting our conclusion and future work.

2

Background and Related Work

Contents

In this section, we provide a summary of relevant background and related work. We begin this section by addressing directly the concept of what is a password manager. We then talk about the usability challenges they face, followed by a section describing user retention and adoption of PMs. The last part of this chapter is about usability best practices and techniques. We focus mainly on the usability of PMs and the effect this had on their usage.

2.1 Password Managers

Nowadays most websites require a minimum password length and complexity and request periodical password changes. Inglesant et al. [28] investigated password use in two major organizations and they concluded that generating new passwords that conform with strict security policies is not trivial for users. Moreover, when faced with these password requirements, users would resort to unsafe practices (e.g., writing passwords down on paper). Another study, by Das et al. [29] and published in 2013, estimated that 43-51% of users reuse the same password across multiple sites. Furthermore, Das et al. demonstrated that many users introduce small modifications to their passwords across sites, and many also share the same procedures for introducing these modifications. However, these modifications are simple enough that an attacker who is aware of typical user behavior can significantly improve guessing efficiency.

Password Managers (PMs) present a solution. PMs are applications that help users manage and generate their passwords. Most PMs offer password security features like secure password generation, multi-factor authentication, and secure storage with a master password. Other common features contemplate productivity, such as automatically filling password fields (usually referred to as auto-fill) and cloud synchronization across devices. Not all PMs offer all these features but most of the ones mentioned in this document do.

PMs normally require a master password that the users have to remember — this is the only password they are required to memorize. All others are kept in a secure "vault" protected by this master password.

Advantages of password managers, when used correctly, include:

- Reduction of password reuse — by generating and saving different passwords for each site.
- Improvement of password complexity — by generating passwords that are more complex than the users would generate for themselves (e.g., random passwords with special characters).
- Secure storage — by saving the password in a secure "vault" instead of less secure storage (e.g., some users who do not use PMs report using spreadsheets to save their passwords [11]).

- Increase in productivity — in PMs that provide the auto-fill feature by automatically filling the users' credentials in websites.

Although the more popular PM applications have millions of downloads [10] and despite their benefits, PMs are not widely used. One study, from Alkadi et al. [10] about smartphone PMs, examined online reviews of password managers and elicited opinions from 352 respondents. This study reported that only 6.8% of respondents used a dedicated PM.[1]

PMs can be classified according to various criteria. For example, we can classify them according to the platform they run on, in three main categories: *browser built-in*, *mobile*, and *desktop*. PMs can also be classified according to where they store their data: locally, in a portable storage (e.g., pen drive), remote self-hosting, and cloud-based.

PMs can be very useful, however, there are some usability challenges that make some users unwilling to use them. These challenges will be explored in the next sections.

2.2 Usability Challenges of Password Managers

The usability of PMs is an important aspect that can increase the adoption of these tools and that has been studied by the research community. For instance, Stobert et al. [3], in a study about password usage, were surprised to find that none of their participants used a dedicated PM and that most of them appeared unaware of prominent PMs. A few participants expressed distrust in this software. The authors suggested that a good integration of PMs into operating systems and browsers would help with visibility and trust.

2.2.1 Password Manager Usage

Pearman et al. [11] studied the usage of password managers and other password management methods. A 30-participant interview study was conducted with people who do not use password managers at all (7 people), people who use password managers built into their browsers, or operating systems (12 people), and people who employ separately installed password-manager applications (4 people).[2]

The study also found that **people who do not use PMs** rely mostly on memory or unsafe methods (e.g., saving passwords on Excel sheets). The reason for not using PMs was mostly **unawareness** of their existence. Despite some users being aware of the existence of PMs, they were still reluctant in using them due to a **lack of understanding of their security properties**. These findings were also

[1] A dedicated PM is a PM that users can name. In contrast with, for example, a browser built-in PM, that the users may use unknowingly.

[2] There were two participants that "were difficult to place in the aforementioned categories" [11]

backed by the work of Ion et al. [12] where non-expert users expressed a lack of trust in password managers.

Some participants in the Pearman et al. [11] study used browser built-in PMs. The main reason for this was the auto-fill feature, which is enabled by default in most modern browsers. Having the PM enabled by default, in a tool that users are already familiar with, was considered a convenient and fast way to save and manage passwords. Some of the barriers to effective use of PMs were similar to those of the participants that did not use password managers at all. One major complaint was related to **lack of awareness of how the tool and its security worked**. By not understanding the features offered, some users could not, for example, synchronize passwords between devices. This lack of information also made the users wary of password managers' security. So, this raises the question of how to communicate information to users. However, even with the usability problems reported, the adopters of browser built-in password managers found them convenient.

The motivation for users of separately installed password managers was primarily security and even though some reported poor usability (e.g., having difficulty in navigating the interface), they were satisfied with the security provided by the password managers they used. It is important to mention that these participants had a better understanding of how password managers worked and trusted them more.

A consistent theme that emerged is the **trade-off between security and convenience** [11]. Perceived security is also relevant as it can motivate users to overcome the initial overhead of using password managers. About half of the participants replied that they would continue to use the PMs after the study. In the follow-up question which asked why the participant decided to continue using them, the reasons for positive answers were mainly convenience and enhanced security.

Convenience, usability, and security were the main concerns of the users in this study and a problem identified was the users' lack of information regarding how PMs work. The study also calls for better usability testing and focus on non-expert users [11].

2.2.2 Password Manager Usage with Older Users

The participants in Pearman et al.'s study were skewed towards young people, with a disproportionately high percentage of participants with technical backgrounds. As such, Ray et al. [13] expanded Pearman et al.'s findings by replicating their protocol and interview instrument but applied to a sample of strictly older adults. A 26-participant interview study was then conducted with older adults (aged above 60) who do not use password managers at all (10 people), who use password managers built into their browsers, or operating systems (9 people), and with people who employ separately installed password-manager applications (7 people). Across all, secure access to financial accounts was valued above other types of online accounts. Control over one's private information was also valued by older and younger adults (from Pearman et al. [11]).

Regarding users that do not use PMs, according to Pearman et al.'s study, both older adults and younger adults were concerned about **a single point of failure** when using PMs (e.g., losing access to all passwords stored in one place). Concerning the participants that used browser built-in password managers, both older adults and younger adults were worried about others having access to their passwords and about where they were stored. Specifically older adults did not trust separately-installed PMs to be invulnerable. Similar to the findings of Pearman et al., Ray et al. [13] found that users who adopted separately installed PMs were motivated by their desire for better security. What they found is that although older adults, in general, expressed more favorable experiences using PMs, they also had a higher **mistrust of cloud storage** of passwords and cross-device synchronization [13].

2.2.3 Lack of self-efficacy

Lack of self-efficacy when dealing with software was one of the main barriers to the adoption of PMs. A higher level of transparency (e.g., showing users how secure their passwords are) could also help towards alleviating concerns and increasing levels of trust [13]. Lastly, concerns about being dependent on technology and having no motivation to learn how to use a new tool were mentioned.

The suggestion given by Ray et al. was to encourage advocacy, particularly from family and close friends, but also by trusted organizations (e.g., AARP).[3] Encouraging the adoption of PMs by younger adults may in turn increase adoption among older adults, as these users become advocates to their close friends and, particularly, their older family members [13]. Another suggestion was education and helping older adults better understand the urgency of secure practices (e.g., classes at senior centers and libraries). Erroneous and **incomplete mental models** of how PMs work (e.g., encryption, cloud storage, etc.) also surfaced in this study [13].

2.2.4 Password Managers in Smartphones.

Usability in smartphones presents different challenges from conventional desktop interfaces. For example, in a study focused on PMs for mobile devices by Seiler-Hwang et al. [24], users' unawareness of the existence of PMs was not a rejection factor, as most of the participants knew about them. Seiler-Hwang et al. conducted a usability study comparing 4 popular smartphone PMs (Dashlane, Keeper, Lastpass, and 1Password) with 60 participants. They used the System Usability Scale [25] to compare the PMs' usability. Dashlane got the highest score, but, overall, looking at the small sample of analyzed applications, PMs appear as software tools that can be subjectively considered "ok" or "good", but far from being "excellent" [24].

Participants often complained about **lack of guidance**, instructions, tutorials, or help pages. This

[3] American Association of Retired Persons, www.aarp.org

meant that sometimes they were unable to achieve their goals within the PM. Also, for participants that were unfamiliar with PMs, this lack of guidance is translated into a lack of understanding about how PMs work.

One of the most problematic areas in mobile PM usability, for users, was **poor integration** with other applications and browsers. The recommendation given by Seiler-Hwang et al. [24] is for better performance, as in, the password managers should be less prone to errors and the features should work as intended. Functionalities like password generation, auto-fill, and device synchronization are core and need to be well implemented [24].

Alkaldi et al. [10] investigated the factors impacting the adoption or rejection of smartphone PMs based on Play Store and App Store reviews. They found factors such as **awareness, no perceived usefulness, security, and privacy concerns** to be detrimental to the adoption of PMs. These factors can be linked with a lack of knowledge about how PMs work and their advantages. New or inexperienced users also reported uncertainty in how to use PMs and they lacked the motivation to learn.

They state that even if people become aware of the apps, they might still not embark on a search process to consider installing one. Many people mistakenly think their current password practice is secure. It seems that, in addition to making people aware of PMs, users should also be educated to understand that their current behavior might be making them vulnerable to attacks [10].

A failure to reassure potential users about the trustworthiness of these applications was identified as a main factor behind the rejection of PMs.

2.2.5 Usability studies

A comparative analysis of PMs usability and security was conducted by Arias-Cabarcos et al. [14] on five different mainstream PM applications. For the usability study they used a set of evaluation criteria known as the 5 E's (Efficient, Effective, Engaging, Easy to learn and Error tolerant). The study reveals that users rate all the PMs high on the scale in regard to efficiency, effectiveness, and error tolerance. Although the PMs studied did not have negative ratings of usability, important differences arose when users rated PMs according to the engaging and easy-to-learn features. An interface is engaging if it is pleasant and satisfying to use and it is easy to learn if it allows users to build on their knowledge without deliberate effort. KeePass was the worst evaluated manager in both these categories [14]. The best rated PM, in all categories, was Dashlane.

Karole et al. [15] did a comparative analysis of types of password managers: desktop managers (that store passwords locally), online managers (that store passwords in the cloud), and portable device managers (that store passwords on a portable device, e.g., smartphone). Desktop managers are less portable but provide more control for the users as all information is stored locally. Online PMs, although portable, require the user to trust a third-party service provider, relinquishing control. Lastly, the portable

device managers are divided into two categories: smartphone PMs and USB-based PMs. Smartphone PMs are portable but less usable as when the users are on another device they have to copy the password manually. USB-based PMs are easier to use on a desktop but when the users are on a smartphone they also have to copy the password manually.

The authors conducted a study with 20 participants to assess the usability of an online manager, a phone manager, and a USB manager. They found that non-technical users preferred to manage passwords on their mobile devices rather than relinquish control to a web-based password manager [15]. They also found that users overall preferred the two portable managers over the online manager, despite the better usability of the latter. Also, technical people were more inclined towards the USB manager in comparison to the online manager. Overall the users preferred to retain control over their passwords and felt that the two portable PMs offered more control. Karole et al. [15] conclude the study by stating that portable managers represent a more promising password management approach than online managers.

Chiasson et al. [16] did a usability test with 26 participants evaluating two PMs. Multiple issues arose because users' **mental models did not match the reality** of the system, e.g., a number of participants felt that they had successfully completed tasks when in reality they had not. A suggestion given by Chiasson was to give more feedback to users about whether their actions were successful or, if unsuccessful, what they could do to use the software correctly. This would help situate the users and improve their mental map of the PMs.

A usability issue related to the users' **mental maps** was about the tools' activation. Users believed that the PMs would, after initial activation, stay working for the rest of their computer session. As this was not the case, the users felt a false sense of security. Inconsistency in the interface of the PM also hinders the mental model of the users. This happened, for example, when one of the PMs studied (PwdHash) had a specific command that was irrelevant and, whether used or not, would give the same output.

This challenge is also confirmed by Ion et al. [12], who suggest that users' reluctance to adopt PMs may also be due to an ingrained mental model that passwords should not be stored or written down—advice users have been given for decades.

Not all usability problems encountered by Chiasson et al. were a direct result of the PMs' interfaces. Some problems were due to bad website design. The sites used on the usability tests were not, for example, consistent with the login page location. These are valid usability issues that provide context and insight into the circumstances and environments where people will be using PMs [16].

Control was also an important issue for users. When the PMs on the study did not show the passwords that they were generating, users felt frustration as they felt as though they had no control over their passwords [16].

Long security messages also posed a usability issue. PwdHash would alert users when they tried to enter a password into a non-password field. Nonetheless, this message was long, and as such most

users would dismiss it.

A major problem arises from the developers' assumption that users are going to use the tool correctly. Users sometimes thought they were secure but were using the application in an insecure way. This is problematic as new users make frequent mistakes and may be deceived into thinking they are safe when they are not. For example, when the password generated by the PMs was not "strong enough" (by the website's standards) users felt that if they chose their passwords they would be able to produce a strong password [16].

If the systems are very secure but do not have good usability, users may opt to use a different, less secure system that lets them do what they want [20].

2.3 User Retention in Password Managers

User retention can be seem as a consumer's intention to keep using an application, as a favorable attitude towards maintaining a long-term relationship with an application [30].

Retention metric, as defined in the Google HEART Framework [31], tracks how many of the users from a given time period are still present in some later time period (for example, the percentage of seven-day active users in a given week who are still seven-day active three months later). What counts as "using" a product can vary depending on its nature and goals. In some cases just visiting its site might count. In others, you might want to count a visitor as having adopted a product only if they have successfully completed a key task, like creating an account. In more detail, retention metrics can be used to provide stronger insight into the number of unique users in a given time period (e.g. seven-day active users), addressing the problem of distinguishing new users from existing users [31].

As we have seen in the previous section, user acquisition and retention have proved to be problematic for PMs. In the Pearman et al. [11] study, 3 participants reported trying PMs but then stopping. Their problems were mostly related with:

- usability problems (e.g. problems navigating the interface, not understanding how to use the PM);
- bugs in the PM (e.g. passwords not saving properly).

In another paper, previously mentioned, Seiler-Hwang et al. [24] conducted a usability study comparing 4 popular smartphone PMs. They gathered insights about users' continued intention to use the PM. About half of the participants replied that they would not continue to use the PMs after the study. The reason for not wanting to use the PMs in the study was grouped into these three categories:

- Usability issues were reported by 40% of the participants (e.g. poor integration with browsers and applications and not understanding how the PM worked);
- "No perceived need" was given as a reason for not using a PM by 40% of the participants;
- "Already use another PM" was another reason provided by 26% of the participants that chose not

to continue using the PM under evaluation.

Alkaldi et al. [10] investigated the factors impacting the adoption or rejection of smartphone PMs based on Play Store and App Store reviews. Besides the usability problems addressed in Section 2.2, Alkaldi et al. found that some users would install smartphone PMs, and try them, but would ultimately not use them due to problems. Those were:

- Device Speed - Smartphone users complained about the efficiency of their devices after installing and using PM applications;
- Device Memory and Battery - Participants expressed concerns about apps consuming the battery and RAM;
- Connectivity - Some users indicated that having a poor Internet connection was a usage barrier because some PMs would not let them access their password without the internet. This is something that can be addressed during the development process of a PM and that should have been detected as a problem and rectified;
- Linkage with Other 3rd Party Services - Some PMs required an additional account with a specific service provider such as DropBox in order to be able to synchronize their passwords. This presented a challenge to those users that did not have any previous account in these services;
- They are not supported by other web accounts - This happened when sites did not support the PMs. An example given in the study was that British Gas that did not support PMs on their login page;
- Uncertainty in how to use the PM - This problem was due to usability issues. An example would be when a user was required to log in with a PM and was not sure how to do that;
- Lack of control - Specifically due to the fact that some of these applications do not provide a recovery plan for the master key so if the user loses this password they lose access to the whole vault.

These problems can be categorized into three categories: i) usability; ii) lack of knowledge about PMs; iii) and badly implemented PMs (with integration/performance problems).

2.4 Best Practices and Techniques

As described in the previous sections, in general, PMs have usability issues that have to be addressed. Moreover, these problems have affected user retention and usage of PMs [24]. These best practices address some problems identified, for example, users' lack of control is directly addressed in the rule "Keep users in control". A good guideline to follow when designing an interface is to follow "*The Eight Golden Rules of Interface Design*" according to Shneiderman [23]. These are intended to be used during design but can also be applied to the evaluation of systems [32]. They are:

1. *Strive for consistency*: consistency in action sequences, interface design, terminology, and so on.
2. *Seek universal usability*: by adding features for novices, such as explanations, and features for experts, such as shortcuts and faster pacing, enriches the interface design and improves perceived quality.
3. *Offer informative feedback*: for every user action, there should be an interface feedback.
4. *Design dialogs to yield closure*: informative feedback at the completion of a group of actions gives users satisfaction and prepares them for the next group of actions.
5. *Prevent errors*: ideally, prevent users from making mistakes. If users make an error, then the interface should offer simple, constructive, and specific instructions for recovery.
6. *Permit easy reversal of actions*: as much as possible, make actions reversible. This relieves anxiety since users know that errors can be undone, and encourages exploration of unfamiliar options.
7. *Keep users in control*: for example, experienced users strongly desire the sense that they are in charge of the interface and that the interface responds to their actions.
8. *Reduce short-term memory load*: by keeping the interface consistent and memorizing sequences of repeated actions.

Regarding usability, a good practice for PMs is to make default settings the most secure choice [24]. This helps all users but especially inexperienced users that may not know the meaning of every setting in the interface.

The auto-fill was found to be the most useful feature of the PMs by the users in Seiler et al.'s [24] study, but it was also problematic as its integration with applications and browsers is not always well implemented or possible. The recommendation was to improve the interfacing between PMs and 3rd parties to improve the auto-fill integration. Better guidance could also lead users to understand how to configure and use the auto-fill feature to get the most out of it [24].

A summary of the challenges presented here can be found in Table 2.1.

2.4.1 Usability beyond the interface

An important point to address is that usability issues can go beyond the interface of the system. Although traditionally one can increase usability by providing a better interface, this is not always the case. One of the most well-known cases is presented in the paper "Why Johnny Can't Encrypt" by Whitten et al. [2]. The researchers tested a security program (PGP 5.0) that had a good user interface by general standards. However, despite PGP 5.0 being attractive and despite the fact that it is simple to use for those who already understand the basic models of public-key cryptography and digital signature-based

trust, Whitten et al. concluded that it was not sufficient to make computer security usable for people who are not already knowledgeable in that area. This shows that usability in security must go beyond the interface: applications must also be easy to learn and must communicate an accurate conceptual model of the security to the user as quickly as possible.

2.4.2 Abstraction

Abstraction can be a solution to simply a complicated interface but, even then, using a well-designed security mechanism is still more effort than not using it at all, and users will always be tempted to cut corners [20]. As such, they need to be motivated to learn how to use the systems and navigate the interface. This can be achieved by educating users on how their current actions are insecure or by simply educating them on the benefits of using a certain mechanism. To motivate users, it is important to inform users that PMs can simplify their life. As Ion et al. [12] state, the low adoption rate of PMs among non-experts might stem from a lack of understanding of its security benefits.

2.4.3 Educate users

Pearman et al. [11] reported that participants that had a better understanding of how PMs worked, trusted them more. Nonetheless, currently, most users are uncomfortable with using PMs and do not trust them because they do not understand them [16]. As an intermediary tool, the PM needs to appear reliable, consistent, and predictable. Chiasson et al. [16] recommend that users be educated about passwords and PMs. Educating users should encompass information about the inner working of PMs: explaining how they work, why they are secure, and providing basic cryptography explanations. Learning about how PMs work would increase transparency, improve users' mental models and their sense of control (especially regarding where the passwords are stored). A higher level of **transparency** (e.g., showing users how secure their passwords are) could help towards alleviating concerns and increasing levels of trust [13]. Once users understand that PMs can potentially simplify their tasks (e.g., by helping them remember fewer passwords) user acceptance would also increase [16].

Also regarding educating users, Balfanz et al. [33] state that adding explanatory dialog boxes to a confusing system is not the solution. Developers must think about usability, security, and their interplay during the very first stages of system design. This can include features like instructions, tutorials or help pages [24].

A suggestion by Seiler et al. is to provide **tutorials** adapted to beginners and advanced users. The tutorials should be naturally integrated with the interface so they could be promptly accessible when required, but they should not interfere negatively with the user experience [24].

The problem here is how to pass information to users in an efficient way. In relation to teaching users

about the inner workings of a PM, there are alternatives like a wiki section of the PM [21], informative videos or even a quick "learn more" button on the screen of the application. Although it is important to educate users, long messages also pose a usability issue [16, 22]. As such, a good idea would be to give concise warnings that users would be more alert to. Another suggestion is to try to give these warnings fewer times reducing false positives. This would improve the users' alertness as they would not see the warnings as much.

2.4.4 Mental Models

Mental models are formulated when users interact with a system and are naturally evolving. That is, through interaction with a target system, people formulate mental models of that system. These models need not be technically accurate (and usually are not). A person, through interaction, will continue to modify the mental model [21, 34].

Regarding mental models, Chiasson et al. [16] gave some suggestions. They state that it is important to ensure that users have the cues needed to form an accurate and complete mental model of a system. Users do not need to fully understand the details of complex security programs, but rather need a mental model that is consistent, and that will allow them to predict program behaviour and the results of their own actions [16]. They suggested providing strong feedback about the success status of an action. Also, if something goes wrong, the feedback should be short and help the user address the issue. Education could also help towards correcting mental models [13].

2.4.5 Improving user acquisition and retention

When acquiring users it is important to assure them of the PMs' trustworthiness and when trying to retain them is crucial to provide a product with minimal usability issues and trustworthiness.

In a study from 2008, Cyr [35] studied user retention in the form of e-loyalty and defined it as the intention to revisit a Website or to consider purchasing from it in the future. The author concluded that in the Western countries studied (Canada and Germany) trust and satisfaction have a positive impact on e-loyalty. Furthermore, she also concluded they are more likely to revisit the website if they are satisfied with the design of a website. Alternative ways to improve user retention have also been studied. An example is a study by Vaibhav et al. [36] about the Gamification of Massive Open Online Courses (MOOCs) to improve user engagement and retention. They performed the study in two environments: a *non-gamified* environment and a *gamified* environment. The non-gamified environment was based on the conventional method of learning, i.e. "reading" a document. The Gamified environment was based on a web-based gamified learning platform with quizzes and a less conventional way of teaching. What they found was a 28% increase in users successfully completing the tasks in the gamified platform and

a 14% decrease from candidates who denied taking up the final test. All in all, there is a noticeable increase in user retention in the gamified platform.

Another alternative method proposed to increase user retention was the incorporation of socialization tools in the product. Ciampaglia et al. [37] studied this method by trying to increase new user retention in Wikipedia through lightweight socialization. To achieve this they used an extension that allows newly registered users to send their feedback (or share their "mood") about their first edit experience on Wikipedia. They concluded that this lightweight socialization tool improves users' engagement and retention in Wikipedia.

User retention in PMs can also be improved by solving the problems described in Section 2.3. The key problems behind users' low retention rates were divided into three categories: i) usability; ii) lack of knowledge about PMs; iii) and badly implemented PMs (with integration/performance problems). These categories have been addressed with possible solutions from the literature in Table 2.1. Users value a strong, usable, and reliable software experience, especially in security software.

2.5 Factors that influence usage

In this subsection, we hope to summarize some factors found in the literature and in previous sections that influence user retention, adoption, and usage of PMs. They include:

- **Mental Models** — People commonly construct implicit mental maps to understand complex systems when the systems' functionality goes beyond their technical knowledge [38]. Such tacit knowledge influences human behavior, even unknowingly. As such, it is important to engage with users to identify their tacit knowledge, mental models, and misconceptions.

- **Usability** — A products' usability can impact *adoption* and *retention* rates. For example, despite experts recommending password managers, the adoption of these tools is still low, partly because users complain about usability problems (e.g. problems navigating the interface, not understanding how to use the software, bugs) [11].

- **Perceptions** — Perceptions, being in regards to perceived privacy, usability, satisfaction, etc, affect end-users. Pearman et al. [11] cited perceptions of increased security as a reason for the adoption of password managers and perceptions of lack of usability as a reason for dissatisfaction.

- **Knowledge** — Lack of knowledge about a product can also discourage users from using it. It can, in some cases, prevent users from trusting it [11].

- **Motivation** — Aspects such as the security guarantees of formally verified software have the potential to impact users' when choosing what software to use. Motivation is a crucial factor in

encouraging users to overcome the initial overhead of using new software [23]. As some users value some online accounts more than others (e.g. financial accounts over others) [11] the impact of users' perception of safety and privacy can be higher or lower in security products over other types of software.

Table 2.1: PMs' challenges and proposed solutions

Challenge	Proposed Solution	Description of Solution
Lack of trust and understanding [12, 13, 24]	Provide a higher level of transparency (e.g., showing users how secure their passwords are)	• Educate users about how PMs work [24] • Advocacy from trusted organizations about the use of PMs (e.g., schools) [13]
Lack of motivation to use PMs [10, 12, 13]	Educate users about the benefits of using a PM	• Provide information related to the dangers of unsafe password habits [11, 12, 16], about the increased productivity and security of using PMs [12]
Bad performance, poor integration with other applications and browsers [3, 24]	Solid implementation of all PM's features	• Functionalities like password generation, auto-fill, and device synchronization are core and need to be well implemented [24] • Usability testing of the PMs and their integration with other applications and browsers [3, 24]
Difficulty of use (lack of usability) [10, 11, 14–16]	Simplify the interface and provide support for users	• Tutorials about how the interface works (for beginner and expert users). They should be naturally integrated with the interface so they could be promptly accessible when required, but they should not interfere negatively with the user experience [24] • Explain what different options in the security settings mean [24] • If the user is unsuccessful, feedback should be short and help them address the issue [16, 39] • The PM should be error tolerant, this is especially important for new users. The PM must be permissive and allow the users to recover and learn from mistakes [14, 39]
Inadequate Mental Models [12, 13, 16]	Provide a precise interface	• Give feedback to users about the status of their action (if they were successful or not) [16] • Navigation should be as clear as possible [20, 21, 39]

3

Improving Password Managers

Contents

The main goal of this project is to explore ways to effectively convey the formally verified properties of the password manager developed in the PassCert project and to ensure it is usable. In this chapter we explore ways to answer the following research questions:

RQ1. What are the main usable security techniques that can be applied to PMs?

RQ2. How can we effectively convey formally verified properties of a PM to its users?

To achieve this, in this chapter we begin in by studying ways to convey information on formal verification to users and to ensure the PM is usable. We then explore several PM options to extend. Finally we end the chapter with an analysis of the chosen PM and a brief description of some preliminary problems identified in the PM. The solutions reached in this chapter were implemented and are described in the next chapter, Chapter 4.

3.1 Information on Formal Verification

PassCert's PM differs from other PMs by using formal verification to guarantee safety properties on data storage and password generation. Therefore, a primary concern is educating the users about formal verification. To achieve this we plan to:

- Provide a way to clearly convey what the system is formally assuring. To do this we plan to implement status symbols (e.g. icons in Figure 3.1) to indicate that a certain action is formally assured [21].

- Concise explanations about formal verification and how it can guarantee certain security properties. A good way to take context into account is to monitor the cursor location and provide helpful information about the icon under the cursor [21, 23] (often called a tooltip, pop-up box, ScreenTip, or balloon help). By having the users control this action it becomes less disruptive to them [23] (see Figure 3.2). It is important to use the correct language in order to prevent alienating users, as such, we will avoid the use of jargon and more technical language that non-technical users may not understand [21, 22].

- Further information may be required by the more inquisitive users and it should also be provided. So it is also relevant to recommend resources where they can learn more about formal verification beyond its use in the PM. This can be done by providing links to introductory books and resources about this topic.

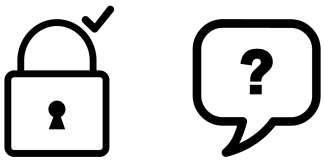


Figure 3.1: Examples of icons: formal verification (on the left) and help (on the right)

3.2 Usability of the Password Managers

Several suggestions have been made concerning a possible solution for the usability problems of PMs. We plan to implement a solution based on the findings presented in Section 2.4, where we gathered some usability best practices from the existing literature. We plan to implement the following (based on Table 2.1):

- Tutorials about how the PM interface works (for beginner and expert users). None of the PMs compared in Table 3.1 has this feature implemented. These should be naturally integrated with the interface as to be promptly accessible when required, but they should not interfere negatively with the user experience [24] and should be optional. The tutorials would depend on the PM chosen to implement the solution (see Section 3.3); furthermore, tutorials are useful and offer step-by-step teaching modules. When implementing tutorials it is important to keep in mind the heuristic "speak the users' language" [39] and to avoid the use of technical jargon [21, 22].

- Explain what the different options in the PM represent, for example in the security settings [24]. This can be achieved through the use of tooltips [21], (as seen in Figure 3.1 and Figure 3.2). Tooltips offer contextual interactive help, this is, they are triggered by user actions (in this case hovering over an item in the interface) and the help they offer is directly related to what the user is doing [32, 40].

We also planned to implement a way for users to learn more about PMs in general. This can be achieved by providing a direct link to the frequently asked questions (FAQs) regarding PMs. Most user interfaces have FAQs [23] and most users are accustomed to them [21]. FAQs consist of a simple series of questions and answers that provide ready access to commonly needed information. The FAQ should contain useful information about the security and benefits of using the PM.

3.3 PM Choice

Our implementation consists of a proof-of-concept prototype. We extended the interface of an established PM and then tested it for usability and user acceptance.

[1] LibreOffice is an open-source office suite project of The Document Foundation, www.documentfoundation.org

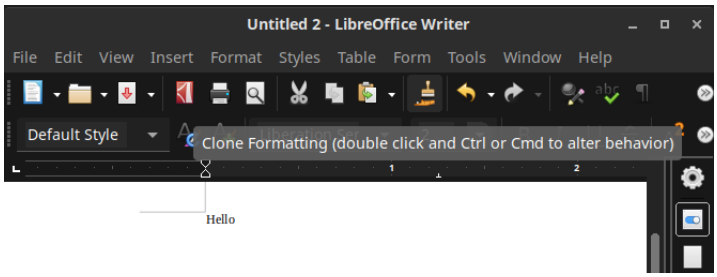


Figure 3.2: Example of tooltip in LibreOffice.[1]

When planning the implementation we considered three popular open-source PMs: Google Chrome Password Manager, Bitwarden, and KeePass (see Table 3.1). Each one is a different type of PM: one is Browser Built-in, one is a Browser Extension, and the other is a Standalone Application (see Section 2.1).

Table 3.1: Feature and programming language comparison among popular PMs

PM	Type	Password generation has fine configuration [2]	Auto--fill	Tutorials	Internal[3] documen-tation for users	Source Code
Google Chrome PM	Browser Built-in	✗	✓	✗	✗	C++/ Javascript
Bitwarden	Browser Extension	✓	✓	✗	✗	JavaScript TypeScript
KeePass	Standalone application	✓	✗	✗	✗	C#

Regarding the **fine password generator** mentioned in Table 3.1, we believe users should be made aware of what each setting means (see Section 3.2). Concerning the **auto-fill** feature, which most PMs in Table 3.1 implement, it is important to inform users how it works to improve transparency and to explain the risks associated with it (e.g. some vulnerabilities were found in previous studies [9]). Furthermore, we can see that none of the PMs in Table 3.1 have **documentation** in the PM itself. This is important, for example, if a user is using the PM offline and therefore cannot access information about it. Regarding **tutorials**, the compared PMs also do not perform well. Although other PMs offer such tools (e.g. LastPass[4]), these do not and, as such, could be improved by adding them. These topics will be further expanded in the next sections.

We analyzed the three PMs as follows.

[4]A popular closed-source PM, `www.lastpass.com`

3.3.1 Keepass

KeePass, according to their website [41], is "a free open source password manager, which helps users to manage their passwords in a secure way." Officially Keepass is a standalone application for Windows, but it can be used on other Operating System (OS)'s like Linux and macOS with Mono (from the OpenSUSE Mono repository). They currently support two versions of the software 1.x and 2.x written, respectively, in C++ and C# (as can be seen in Table 3.1).

KeePass could be a good choice because it has been around for a long time, since 2003, and as such has acquired a dedicated following. Several contributors have developed extensions such as an android application, browser extensions, and other non-official products. This PM has usability issues [14].

As stated before, KeePass is a standalone application and these types of PMs are mainly used by expert users [11]. As beginner users experience more barriers to effective PM usage we decided to not extend KeePass.

3.3.2 Google Chrome Password Manager

We then focused on studying Google Chrome PM. Google Chrome PM is a browser built-in, i.e., it is bundled with each Google Chrome browser. The browser market share leader is Google Chrome according to several sources [42, 43] and any improvement on this PM would impact a large number of users. Google Chrome code is based on the code of the Chromium browser, an open-source Google project. Chromium would be the browser we extended and is written in C, C++ and, JavaScript (as can be seen in Table 3.1). Nonetheless, this type of deployment could be troublesome because if a non-chrome user wants to use the PM he will not be able to do so in his browser. Additionally, from a usability standpoint, if this user really wants to use Google Chrome's PM he would need to install a whole new browser and migrate all his data and bookmarks. For these reasons, we choose not to extend Google Chrome's PM.

3.3.3 Bitwarden

Lastly, we analyzed the Bitwarden browser extension. Bitwarden is a free and open-source PM. Bitwarden is available in a variety of platforms, such as Android, iOS, Linux, Windows, and browser extension. We focused our analysis on the Bitwarden browser extension that can be used on Google Chrome (and Chromium), Firefox as well as several others like Tor, Edge, and Brave [44].

Bitwarden was released in 2006 and according to their website it has a very active community [44]. The browser extension is written in JavaScript and TypeScript. From an implementation standpoint, of the three PMs studied, Bitwarden seems to be the easier to extend. Because it is a browser extension,

the installation process for the user is also straightforward through, for example, the Google Web Store, or Firefox Addons.

If our extensions are not accepted by the Bitwarden official team, we will be able to provide our own separate PM, with local storage and it will be easier to install than if we extended Google Chrome's PM (where the user would have to install a whole browser). Additionally, the Bitwarden license is GNU General Public License v3.0 and allows modification and distribution of the product. For these reasons, we choose to extend the Bitwarden browser extension.

The PM's formally verified features will be implemented by other members of the PassCert project.

3.4 Analysis of Bitwarden's browser extension

As stated in the previous section we extended Bitwarden's browser extension. Before starting the implementation it was important to analyze the features this tool provides in more detail. The main features of Bitwarden's PM are:

- Secure vault: where the user can save login information (i.e. a username-password pair), notes (i.e. secure notes that can be shared with others), credit card information, and identities (i.e. information used to fill online forms, such as first and last name, email address, etc).
- Password generator: with fine configuration, this is, the user can choose in detail how they want the password to be generated. As we can see in Figure 3.4.
- Autofill: This feature is disabled by default. It is disabled by default because they claim it is "currently an experimental feature. Use at your own risk."
- Bitwarden Send: a way for users to share encrypted text messages with each other.
- Cross-device synchronization: Bitwarden supports most OS's and platforms and provides synchronization across these devices.
- Self-host or cloud host (in Bitwarden's servers).

A thorough analysis of the Bitwarden interface was conducted and some problems were found such as:

- Lack of user support: Bitwarden does not provide any tutorial or initial walkthrough for new users and it is recommended to have these [24].
- Lack of consistent tooltip information: some technical settings have no user-support or tooltips (e.g. the input box "Authenticator Key (TOTP)" which is clearly jargon and where further explanation was owed to the user).
- Lack of consistent behavior: we found, for example, buttons that looked the same but behaved differently. Inconsistencies in navigation and interfaces can hinder the user experience and users' mental models [16, 23].

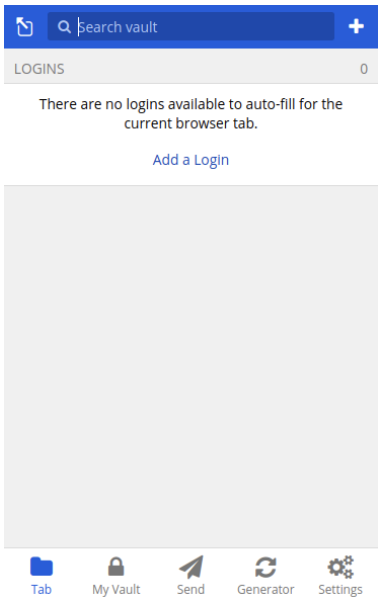


Figure 3.3: The home screen of Bitwarden provides information about the credentials saved in the website the user is currently on

In the next subsections, we will go through a description and analysis of Bitwarden's browser extension.

3.4.1 Current Tab

After entering the extension and logging in the user is greeted with the Tab Screen (as can be seen in Figure 3.3). On the lower part of the interface, we can see a tab bar that lets users switch between the sections: *Tab, My Vault, Send, Generator and Settings*.

On the main body of this page, there are the login credentials for the website the user is currently on. Because the autofill feature is experimental it is disabled by default and instead, to autofill the credentials, the user can click on the ones that appear on this page for the current website. Disabling the automatic autofill by default is something we recommend as this feature has had some known vulnerabilities [9] and the default settings should be the most secure [24]

Most of the tooltips of the extension present themselves on this screen, however, most have redundant text and are non-descriptive. An example is a tooltip for the settings section, which says the same as the icon label "Settings" and, as such, is redundant and non-descriptive.

3.4.2 Secure Vault

In regards to the secure vault, Bitwarden presents the saved item in categories such as "Log in" about login credentials, "Card" about credit card information, "Identity" about form information, and "Secure

Note" about Sends and notes saved in the vault (for more information about Sends see Section 3.4.4).

These categories are predefined, cannot be changed, their names can be misleading (e.g. "Identity" means form information) and no tooltips are provided to contextualize users about them.

When analyzing the credentials themselves, we found some tooltips present, such as the ones labeling the icons to "Copy a password", but others were missing (a more thorough analysis of the tooltips present in this version of Bitwarden can be seen in Section 4.1).

3.4.3 Password Generator

Bitwarden's generator has two modes: password and passphrase. The user can choose between generating a fully random password or a passphrase composed of words (passphrases can be easier to memorize [45]). The password generator provides a fine configuration for the user (see Figure 3.4). This fine configuration includes:

- Set password length (from 5 to 126 characters).
- Enable the use of upper and lowercase letters.
- Enable the use of numeric characters. If enabled the user can also choose the number of numbers the generated password will have.
- Enable the use of special characters (e.g. !@#$%&̂*). If enabled the user can also choose the number of special characters the generated password will have.
- Disable "ambiguous" characters (i.e. characters that may be mistaken for another, for example, uppercase i and lowercase l).

A password generator is an integral feature of a PM and, after analyzing the current interface of the generator, we realized that **Bitwarden's Password Generator does not have any tooltip** or support for users. While in some settings this was not crippling because the interface was self-explanatory (for example when choosing the "Length" of the password) in other settings tooltips were needed (see Section 4.1).

3.4.4 Send

Regarding Bitwarden Send, this feature is not available in most PMs and is outside the scope of this project. It consists in a secure way to send files or texts messages between individuals or within organizations.

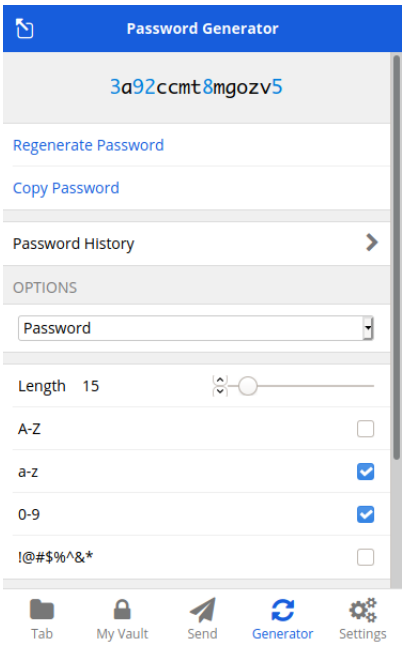


Figure 3.4: Bitwarden password generator

4

Extending Bitwarden

Contents

In this chapter, we describe the implementation of the solutions described in the previous chapter. We previously chose Bitwarden browser extension as the PM we want to extend (see Section 3.3) and in this chapter, each section consists of one type of extension done to the interface of Bitwarden. We will go through the design and implementation processes of the: (a) the tooltips; (b) formal verification icon; (c) descriptions of the formally verified features; (d) FAQ; (e) Tutorial; (f) and some other improvements not directly related to formal verification or user support.

The extensions implemented in this chapter will be evaluated in the following sections.

4.1 Tooltips

As stated in Section 3.4 we analyzed the tooltip usage in Bitwarden's application and found that improvements and additions were needed.

Tooltips are a type of interactive help and are a good way to provide help to users [21]. This type of help should be: (a) Available without interfering, only when users need help [40]; (b) Succinct yet descriptive, and in plain language [32, 40]; (c) Unobtrusive [40].

4.1.1 Tooltip Analysis and Development

We categorized existing tooltips as **Well implemented**, **Non-descriptive**, or **Missing**. In this subsection, we'll go through each part of the PM and describe its tooltips.

Toolbar. In the bottom toolbar, through which the user can navigate through the PM (i.e. vault tab, password generator tab, settings tab) the tooltips were **non-descriptive**. For example, the tab "My Vault" has icons and a label. This is a good practice according to Wiedenbeck [46] but its tooltip is "My Vault", exactly the same text as the icon label. This does not help the user, is redundant, and goes against tooltips guidelines [47].

Current tab. The first tab/section of the PM is the "Current Tab" section where the user can see the credentials and information saved about the website it is currently visiting (i.e. if the user is on Facebook this tab displays the saved login information related to Facebook).

This was the section where more tooltips could be seen, and some were useful and well implemented (for example, the tooltip to copy the username to the clipboard).

Other tooltips were implemented but were **non-descriptive** and could be improved. For example, the tooltip to trigger the auto-fill behavior was just "Auto-fill - Name" (where the name was the name of the information). This could be more helpful and we replaced it with "Click to auto-fill Name". This new tooltip directly prompts users to click on the credential and gives users information about what will happen if they click on it. This may prevent errors in the interface and, as such, is advisable under Shneiderman's rules of interface design [23].

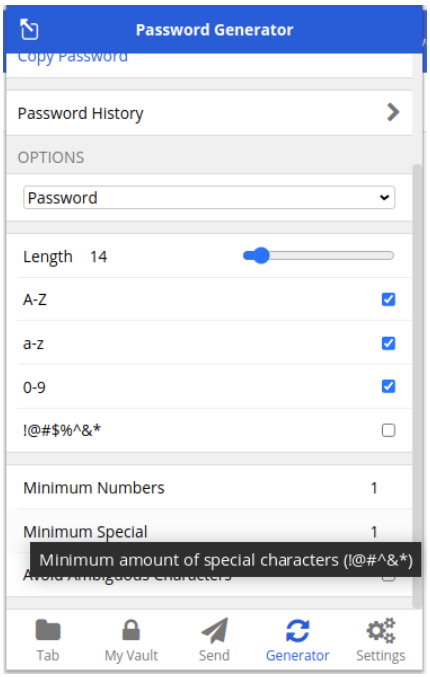


Figure 4.1: Tooltip implemented for the “Special Characters” button

Lastly, some tooltips were **missing**. This was especially jarring when juxtaposed with difficult-to-understand input boxes. An example is the input box “Authenticator Key (TOTP)” which is jargon and where further explanation was owed to the user. For this, we suggest the tooltip: “This key is used to generate a single-use password and verify your identity for websites that use Two-step Login”

Password Generator tab. A password generator is an integral feature of a PM and Bitwarden’s Password Generator did not have any tooltip, they were **missing**. While in some settings this was not crippling because the interface is self-explanatory (for example when choosing the “Length” of the password), in other settings tooltips are needed. The most egregious example was the “Avoid Ambiguous Characters” checkbox where no information about what are ambiguous characters was supplied to the user. We suggest the implementation of a tooltip with the text “Letters that may be mistaken for others, e.g. i and l”. An example of the tooltip implemented about the Special Characters can be seen in Figure 4.1. As this is a main feature of the PM we want to implement its missing tooltips.

Setting tab. Tooltips were also missing from the settings tab. As these are settings related to the security of the password, they should have tooltips or other types of user support. An example is the setting regarding vault timeout. Vault timeout behavior determines how the vault will behave after a customizable period of inactivity. A user can set its vault to never timeout, thus making him vulnerable, as the next person using the computer may be able to access his passwords. This is an example of a setting that should have a tooltip and, accordingly, we implemented one with the following text “How long Bitwarden can be inactive before timing out”.

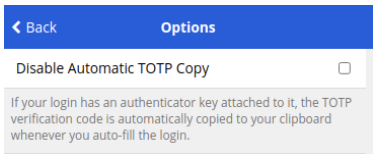


Figure 4.2: Disable TOTP text description

Other settings did not have tooltips but had a description below them, an example is the "Disable Automatic TOTP Copy" that did not have a tooltip but instead a description (as can be seen in Figure 4.2). We found this solution acceptable but that if overdone could overwhelm the user and clog the interface.

To summarise, we added tooltips where they were missing and extended them where they were non-descriptive. We followed guidelines on writing support messages (such as avoiding jargon and keeping the language simple).

4.2 Icon

As stated before, as a way of transmitting to the user what was formally verified we implemented status symbols or icons. To make sure users were aware of the formally verified features of the PM we decided to design a new set of icons to represent formal verification.

We decided to use icons, instead of labels or descriptions because:

- Bitwarden already uses icons throughout the interface and as such using more icons would keep our extensions consistent with the current interface, and this is a good practice [40]
- and because Wiedenbeck [46] suggested that users have less favorable perceptions of the usability and usefulness of text-only interfaces. While text labels play an important role briefly in the early stage of learning, they soon lose their value [46]. As such, and to match the already established interface of Bitwarden, we choose to implement just the icons without any label (besides the tooltips).

4.2.1 Icon Font and Size

The current icons used in Bitwarden are designed with Font Awesome 4[1], so to have a unified design we decided to use the same font for the formal verification icon as it is important to have a unified design throughout the User Interface (UI). As an example, Bitwarden's logo (see Figure 4.3) also uses Font Awesome.

The size of the icon was designed to be the same as the other icons in Bitwarden. Similarly, the color of the icon is an important choice as to not clash with the existing Bitwarden color pallet.

[1] Font Awesome is fully open source and GPL font friendly. https://fontawesome.com

Figure 4.3: Bitwarden Logo

4.2.2 Icon Color

One of the color's roles is to contribute to the aesthetic value of the interface, but it can also be used to transmit information [21] and influence the user emotionally [21, 23].

Some colors rely on learned conventions [48], like the color red that is commonly considered to indicate stop or danger, yellow a warning, and green an all-clear or go sign [23]. This is not universal, while red can mean danger in some cultures it can also represent life (India), happiness (China), and royalty (France) [32]. As such, the meaning of the color will depend on the context within the interface and also on the cultural context of the user himself [21].

To ensure the interface is consistent one needs to respect the established color pallet. A useful tool to choose color is the color wheel [49].

Bitwarden color scheme has mostly cool colors, i.e., colors like blue and gray [49]. Their color scheme is also minimalist and the interface of the browser extension has Bitwarden's blue and shades of gray, without any other color. This is a good practice as color should be used conservatively in the interface as not to overwhelm the user [23].

To differentiate and draw attention to the formal verification logo we used a new color, specifically green. We also choose this color because it is important to take into consideration cultural and social conventions [21]. Although it is difficult to assume any universal interpretation of color, green is associated with safeness in the Anglo-American cultural color convention [32].

Bitwarden's icons use two main colors, Bitwarden's blue (`#175DDC`) (see Figure 4.3) and grey (`#7C7C7C`). Using a quadratic schema [21] from Bitwarden's blue and with the help of the Adobe Color[2], we derived a shade of green (`#0BDB0B`).

4.2.3 Icon Accessibility

Tidwell et al. [49] recommend putting the design into an image tool such as Photoshop, desaturating it (make it grayscale), and checking if it is legible. We then tested the derived shade of green with the method described by Tidwell et al. [49] where we desaturated it and compared its legibility with Bitwarden's original logo (see Figure 4.3). The end result was a darker green (`#009605`) that still matches the original Bitwarden's blue.

[2] A pallet tool referenced by Tidwell et al. [49]. https://color.adobe.com/create

Figure 4.4: Icon variations

It is also important to consider the needs of color-deficient users [23]. Color impairment is a very common condition. Approximately 8% of males and less than 1% of females in North America and Europe have some permanent color deficiency in their vision [23]. These individuals may confuse some shades of orange or red with green or not see a red dot on a black background [23]. Currently, in the Bitwarden interface, we do not have any colors that could be mistaken by color-blind users. Although we are using green for the bitwarden symbol it remains legible (it is not a green text in a red background, for example).

Additionally, there is no other green icon in the interface and this icons' functionality is unique. Also, by having an unique color we hope to prevent users from associating this icon with others in Bitwarden and help them understand that this icon is singular.

4.2.4 Design and Implementation

We began the design of the icon by brainstorming several icons ideas and asking for feedback from the rest of the PassCert team.

To design the formally verified icon we employed other icons, like a lock and a check icon as can be seen in Figure 4.4. Icons like a lock and a check can be associated with a physical vault or security metaphor already used by some PMs. Bitwarden's own symbol is a blue shield (see Figure 4.3). Interface metaphors are important to convey information [23,49]. To generate a vector version of the icon we used the Inkscape[3] software.

The process of choosing and designing the icon went through the following iterations and phases:

[3]Inkscape is a vector graphics open-source software. www.inkscape.org

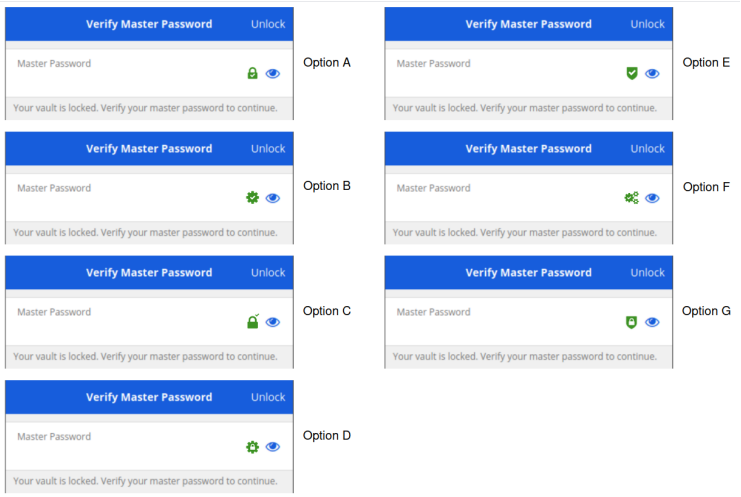


Figure 4.5: Master/Primary Password icon variations

A – First phase - brainstorming sessions We began by brainstorming icons designs and ideas. After reaching about 15 different variations, we had a meeting with the PassCert team and gathered feedback. With this feedback we iterated over the 15 icons and we reached the 7 that can be seen in Figure 4.4.

B – Second phase - team feedback session After reaching the 7 icon variations, we created variations of the interface with different icons that can be seen in Figure 4.5. With the icons in their context, we met again with the PassCert team and asked for their feedback. The team gave specific feedback about icon size, subjective aesthetic value, icon position in the tool, places where the icon should be, and lastly about what icons represented more accurately formal verification. The icons the team reacted to more positively were icons D and E from Figure 4.4.

C – Third phase - feedback from outside the team The previous design iterations of the icon were all with members of the team, but to gather more unbiased feedback we asked 20 users outside the team to give their opinion.

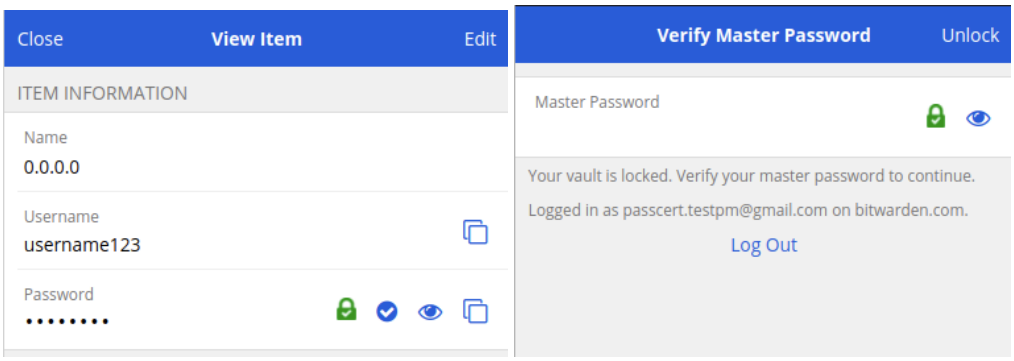
Because the icon means that a certain feature is formally verified it is very important to the underlying message we want to transmit. As such, we performed user studies to determine the best variation of the icon. The form where we asked for feedback about the icons was divided into 3 parts:

1. An *attractiveness test* without context, where we asked users to choose the icons they liked more without knowing what they were meant to represent;

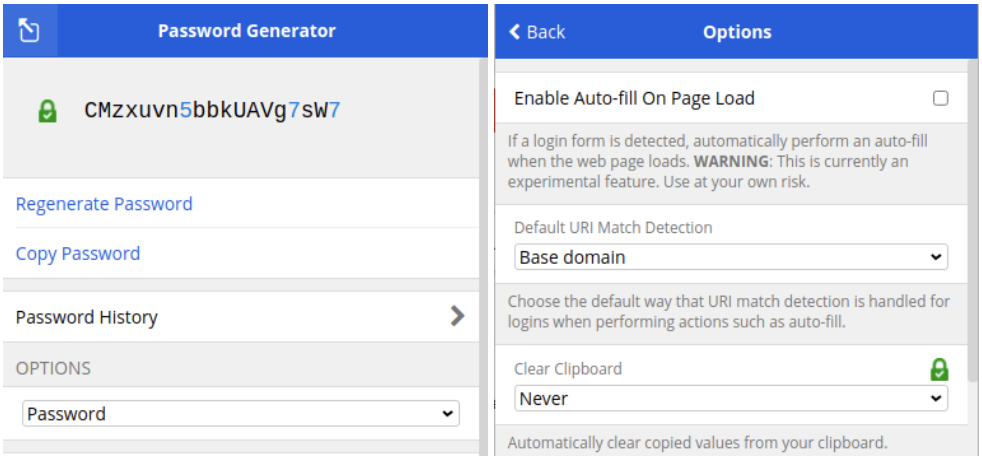
2. A *preference test*, where we explain what the icon is trying to convey and ask users to again rate the icons according to the ones they prefer [50];

3. A *preference test*, where we show the icon in the context it will be in the final interface and ask users to rate it (as can be seen in Figure 4.5 for the master/primary password). We repeated this procedure for 3 icons locations in the interface: for the master/primary password, in the password vault, and in the password generator.

The icons users preferred more were B and D (from Figure 4.4). The icon we ended up choosing was icon D because it was also a favorite of the team. This icon was then distributed through the interface in all formally verified features: (a) Password vault by the password field ; (b) Primary password input box; (c) Password generator by the generated password; (d) Clipboard by the setting to toggle it on. The placement of each icon can be seen in Figure 4.6.

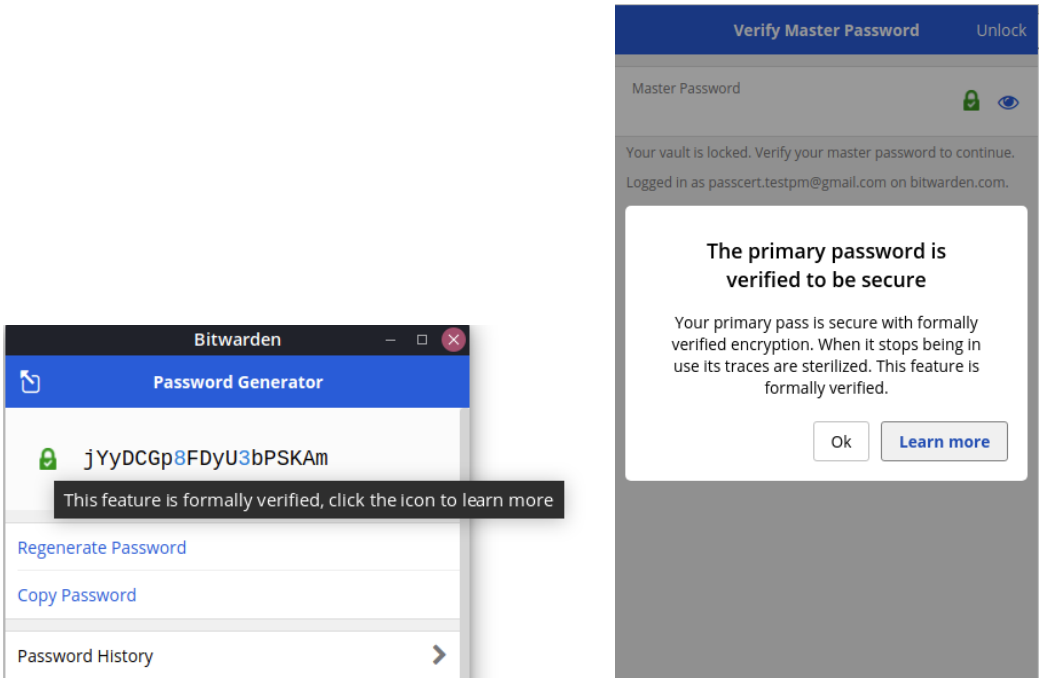


(a) Icon in vault **(b)** Icon in master password



(c) Icon in generator **(d)** Icon in clipboard

Figure 4.6: Formal verification icon in the interface



(a) Password Generator formal verification icon and respective tooltip

(b) Icon in master password with pop-up that appear after clicking on it

Figure 4.7: Icon and formal verification pop-up

4.3 Description of formally verified features

PassCert's PM has the added challenge of explaining to the user what security guarantees formal verification provides. As stated before in order for users to understand these concepts, it is important to keep the language simple and without jargon [21, 22].

Having distributed the formal verification icon throughout the interface in all places where a feature is formally verified, we now wanted to add a contextual explanation about the formal verification of that specific feature (see Figure 4.8). We used the icon described in the previous section (see Section 4.2) and when a user clicks on the formal verification icon it shows a pop-up with contextual help. Furthermore, we also address the formal verification icon and what it symbolizes a the tutorial (see Section 4.5).

4.3.1 Design and Implementation

To design this pop-up we used the pop-ups that were already used in some parts of Bitwarden (for more detail about these pop-ups and the improvements done upon them see Section 4.6).

We began the design of the explanations by going through every feature where the formal verification

icon is. The icon is in: (a) Password vault by the password field ; (b) Master password input box; (c) Password generator by the generated password; (d) Clipboard by the setting to toggle it on; (and can be seen in Figure 4.6).

For each of these places, we added a tooltip asking users to click on the icon to learn more – this tooltip can be seen in Figure 4.7(a).

The development process went through the following phases:

A – First phase - initial description We began this process by having a one-on-one discussion with the members of the team that were developing the formally verified features and asked them for an explanation about what they were effectively doing. The answers were varied and had jargon, so we "simplified" them and reached the first set of explanations.

B – Second phase - first feedback In the next meeting with the team we presented our explanation and gathered and applied relevant feedback (but somewhat general) to shorten some of the messages, change some terminology (e.g. not using the word random on the password generator and instead use unpredictable), and add nuances that were missing (e.g. in regards to the generator to state that the generated password was verified to be generated accordingly to the settings given).

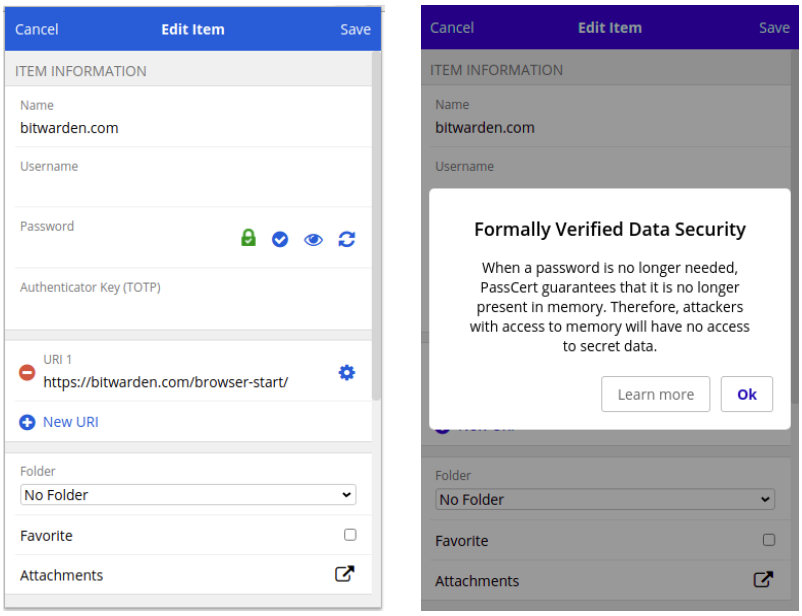
C – Third phase - second feedback After applying the feedback from the previous meeting we redid the messages, reduced the amount of jargon, and added the explanation to the interface. Next, we took screenshots of each explanation in the interface and created a shared document. In this document we presented both short and long descriptions of each feature and again, asked the team to read through the description of the features they were implementing and to give their honest feedback. After reaching a consensus regarding the explanations we updated the ones on PassCert's PM.

It is important to mention that we had both short and long explanations for each of the descriptions. We use the short explanations in the pop-up that opens after clicking the formal verification icon (see Figure 4.8) and the long explanations are used in the FAQ description of the formally verified features (see Section 4.4). We also address the formal verification icon, its role, and explanation in the tutorial (see Section 4.5).

The final description of the formally verified features can be seen in Table 4.1.

4.4 FAQ

Right now, users that have questions or need help while using Bitwarden are redirected to Bitwarden's website. While the website provides a large quantity of useful information, in order for the users to get to it they need to have an internet connection.



(a) Formal verification icon in the password vault (green icon)

(b) Pop-up after clicking formal verification icon

Figure 4.8: Formal verification icon and subsequent pop-up

Table 4.1: Description of formally verified features

Feature	Short description
Password Generator	The generator is truly unpredictable. It makes it more difficult for attackers to discover a password since they have to try all possible passwords. This feature is formally verified.
Password Vault	Your passwords are verified to be secure. Your passwords are vaulted with formally verified encryption. When the password is returned to the vault, its traces are sterilized. This feature is formally verified
Primary Password	The primary password is verified to be secure. Your primary password is secure with formally verified encryption. When no longer in use, its traces are sterilized. This feature is formally verified.
Clipboard	PassCert's PM guarantees that the clipboard is cleared. When a password is copied to the clipboard, PassCert guarantees through formal verification that it will be cleared from memory (within the time frame you choose), reducing the chance of it being leaked.

Most browser users are connected to the internet when using the browser but some may be offline and still want to use Bitwarden's vault. The web vault cannot be accessed offline and said user may only have Bitwarden's browser extension installed. Supposing that the user requires help navigating Bitwarden, he will not be able to receive any as the help provided is exclusively online on Bitwarden's website. Furthermore, if the user is searching for information regarding formal verification, he will not find any on Bitwarden's website.

Having confirmed the need for offline help we suggest the implementation of a FAQ page on the extension. This help would:

- Minimize the effort required of users to get help or information about Bitwarden;
- Provide help for offline users;
- Provide information about formal verification not available on Bitwarden's site.

4.4.1 Design and implementation

Users can access this FAQ page from the "Settings" tab or by clicking on the formal verification icon and selecting "Learn More" (see Figure 4.8).

We implemented a FAQ following Redish's [22] recommendation of going through every topic of interest in Bitwarden's browser extension, thinking about what users may want to know about that topic and about how to give them that information as clearly and concisely as possible. We also used pre-existing FAQs from Bitwarden's website[4] and focused on their Security FAQs.

Following these guidelines, we devised 6 new frequently asked questions and respective answers. After designing the questions and answers, we asked feedback from the team about the explanations on the formally verified features and iterated over them – as described in Section 4.3.1.

The new questions and answers can be read in Appendix A and the screen where the FAQ is in PassCert can be seen in Figure 4.9.

4.5 Tutorial

The tutorial is focused on beginners [24] and intermediate users as those are the ones that struggle the most with PMs. As such we decided to turn our attention to first-time users and the support they need to begin using the PM. As PMs are security applications we hope that by supplying users with tools to learn how to use them we improve long-term user retention.

We designed a tutorial for users in the form of a walkthrough. This is a type of guided tour that is common in other products. Walkthroughs are usually in the form of a lightbox or other layer on top of

[4] https://bitwarden.com/help/article/security-faqs/

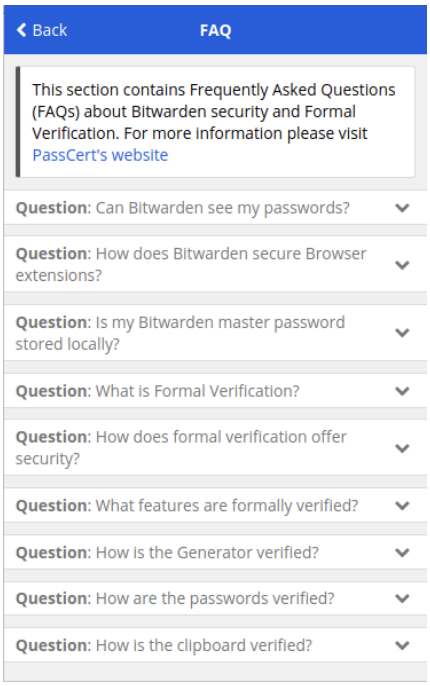


Figure 4.9: Frequently Asked Questions

the application itself. These display as a series of pop-ups or pointers that take the user through a tour or help them complete a process in a stepwise fashion [49].

To implement the tutorial we used the ngx-guided-tour angular package,[5] as can be seen in Figure 4.10.

For the tutorial, we followed some of the same guidelines used in the previous section such as *keeping the language simple and without jargon* [21, 22].

We began the implementation by identifying the features we wanted to go through in it. We ended up going through 3 main sections of the PM: (a) current tab; (b) password generator; (c) and settings. Each of these has some sort of formally verified feature and is crucial while using the PM. However, as not to overload the user, we decided to go through the task and only mention formal verification, and the icon, on one of the sections.

For the **current tab**, we describe how to login into a website with the autofill feature and how to add credential information. In the **password generator**, we address the formal verification icon, mention that it is clickable, and explain its role. Within the generator, we also explain how to copy the password and change the length. In regards to the **settings tab**, here we explain how the syncing process can be changed, how the user can log out, and how the vault timeout works (after a set time the vault times out and logs out automatically).

The user goes through the section in order and can skip the tutorial at any time. After implementing

[5] https://github.com/lsqlabs/ngx-guided-tour

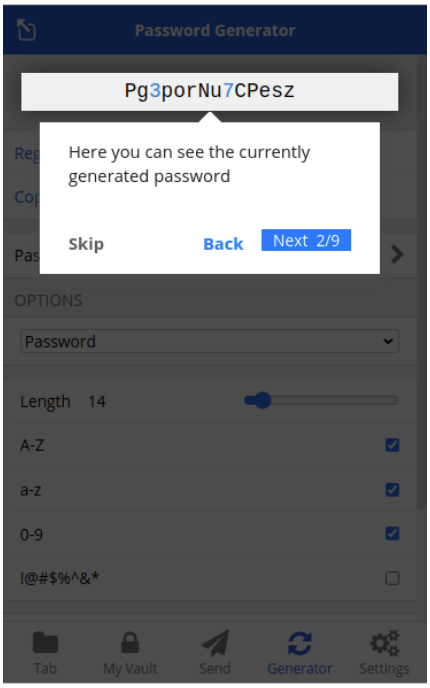


Figure 4.10: Interface extensions: tooltips and tutorial walkthrough

the tutorial we asked for feedback through a shared google docs. From this, we gathered some minor feedback and applied it.

4.6 Other Improvements

Besides the previous problems identified other inconsistencies were found in Bitwarden's interface. Inconsistencies in navigation and interfaces can hinder the user experience and users' mental models [16].

4.6.1 Buttons in Settings

The first set of inconsistencies are in the behavior of certain buttons, for instance, in buttons that redirect the user to Bitwarden's webpage. These buttons are in the settings tab and can be separated into two groups:

- Purchase Premium, Two-step Login, Change Master Password
- Import Items, Bitwarden Web Vault, Help and Feedback and Rate the Extension.

For the first group, Bitwarden warns users that it is going to redirect them to its web page and users have to agree. In the second group, Bitwarden does not warn users that it is going to redirect them to its web page.

The buttons in these two groups look the same. They are all followed by the same symbol, this can be

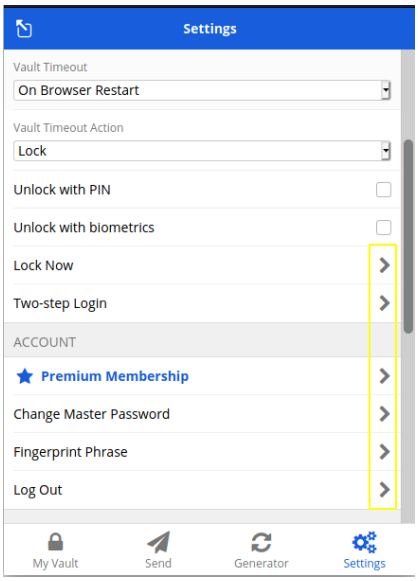


Figure 4.11: Buttons in Bitwarden's setting highlighting the same functionality

seen in Figure 4.11 where the buttons have the same "plus" icon. The buttons also do the same, this is, they all open a new tab in the users' browser and redirect to Bitwarden's website. Although they look and do similar actions they behave differently. When clicking a button from the first group a popup appears warning the user that a new tab will be open and asking for the users' permission. In Figure 4.12 we can see the "Share Vault" button and the popup that appears after clicking on it. This behavior contrasts with the second group's buttons actions. Even though these look the same and have the same underlying functionality, they do not open a popup and instead redirect the user directly to Bitwarden's web page without warning or asking the user.

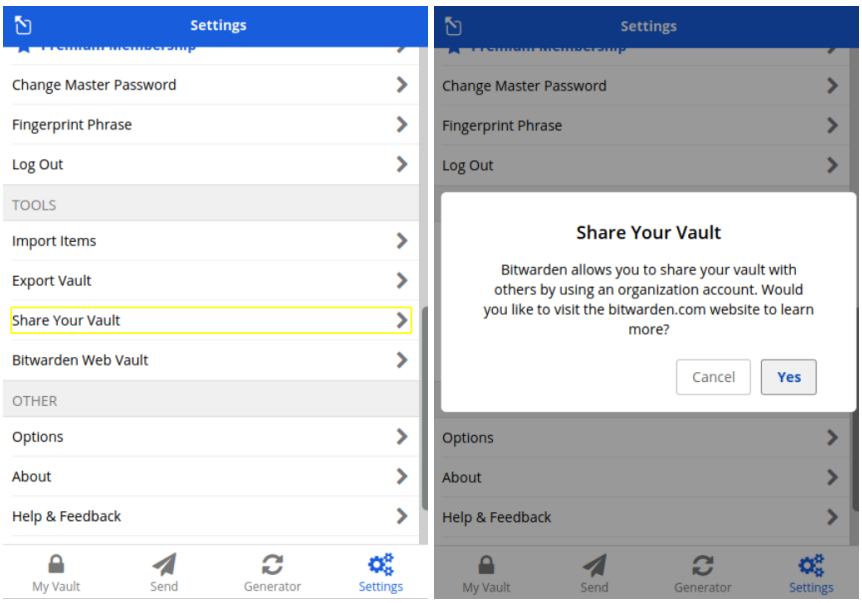
This inconsistency goes against two of the "Golden Rules" of Interface Design as stated by Shneiderman et al. [23] and referenced in Section 2.4: **Strive for consistency**, which states that consistent sequences of actions should be required, which does not happen because the options behave differently; and **Keep users in control**, because experienced users strongly desire the sense that they are in charge of the interface and that the interface responds to their actions, and as such a prompt should be available before redirecting the user to Bitwarden's website.

To rectify these problems all the buttons from the second group were expanded. To these buttons, we added a prompt like the one in the first group (see Figure 4.12(b)).

4.6.2 Pop-up button

Another problem in Bitwarden's interface was its own pop-up style. It is important to mention that this pop-up was the same that we used in Section 4.3.1 to display the explanations on formal verification.

As a follow-up to one of the feedback sessions one of the team members raised our attention to this

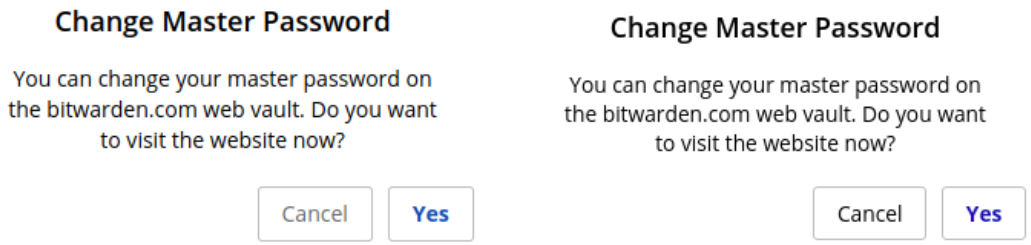


(a) Screen before clicking "Share Vault"

(b) Redirection popup

Figure 4.12: Behaviour of "Share Vault" button

button. The button to "Cancel" is greyed out and some team members stated that they were not sure if it was clickable. To rectify this issue we changed the style of the buttons to a darker color in order for it to appear clickable. You can see the before and after in Figure 4.13.



(a) Cancel button before

(b) Cancel button after

Figure 4.13: Cancel button before and after

5

Evaluation

Contents

With our work, we hope to, on the one hand, ensure that the final PM is usable – ideally by providing an improvement in usability – and on the other hand, make sure users understand what is formally verified and how formal verification impacts the PM. With these goals in mind, we implemented several extensions, such as extended tooltips, new icons, and more user-support features.

In this section, we describe the evaluation process designed to test if the extensions were successful. The user studies described aim to:

- ensure the usability of the final PM;
- understand if users are able to understand that a feature is formally verified and what formal verification is;
- understand users perceptions of formal verification and PMs;
- understand if formal verification can help improve adoption of password managers.

As well, aims to answer the research question: **RQ3.** What are users' perceptions of formal verification and PMs?

The chapter is divided into three parts. We begin the chapter by going through the design of the user studies and testing methodology, we then present the results, and finally discuss them.

5.1 Design of the User Studies

The user tests were divided into 4 parts. First, we provided users with a brief introduction on what the study is about. A script of this introduction is needed to guide how users are greeted, and to tell them the goals of the study, how long it will last, and to explain their rights [48]. The script used was based on Macintosh Human Interface Guidelines [51]. After this introduction, we asked users to fill the "Pre-Task Questionnaire", so that we could learn information about their past experience with PMs and demographics (see Section 5.1.1.A). Then, we began the tasks, which were presented in a random order (see Section 5.1.3) and in between each one we asked users to fill a quick "Task Questionnaire" (see Section 5.1.1.B) relating to it. When the participants finished all the tasks they were asked to fill a "Final Questionnaire" (see Section 5.1.1.C). Before ending the testing session we asked for further feedback from users. Each of the previously mentioned questionnaires was used as a base for an informal interview where we discussed with users the reasons for their answers.

The evaluation framework is based on the one designed by Chiasson et al [16]. Data was collected in two ways: through questionnaires and observation. We chose this methodology because it is applied directly to PMs' usability studies and can be easily replicated. We conducted our user testing remotely due to Covid-19 restrictions.

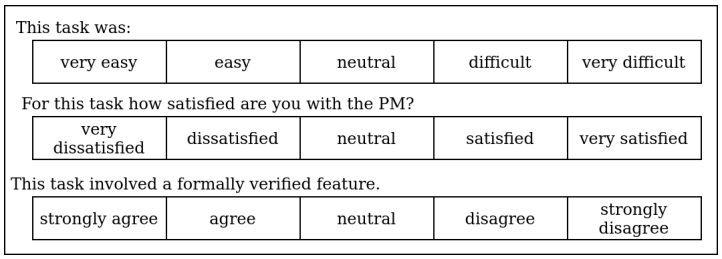


Figure 5.1: Example of task questionnaire

The PM's formally verified features were implemented by other members of the PassCert project. By the time we performed the evaluations, the formally verified features were not integrated into our prototype so, we included a disclaimer to warn the users that: "This is a product in development and some formally verified features are not fully implemented.". We did this because users should be made aware that the features that appear to be formally verified in the prototype, may not be in reality.

5.1.1 Questionnaires

The questionnaires used during the testing sessions were also used as a script for a non-structured informal interview where we asked users why they answered the way they did certain questions and about their feedback on the PM. Special emphasis was given to questionnaires questions related to perceived security and usability.

5.1.1.A Pre-Task Questionnaire

To begin the session we asked users to fill a **"Pre-Task Questionnaire"** to learn information about the user's past experience with PMs, if they had ever used one, and what is their perception of them. For this type of question, we used Likert Scales [52] while keeping the conversation informal and prompting more information from users. Within these questionnaires we also wanted to learn if users know what the concept of PM means. One of the findings was that some users stated not using PMs but simultaneously admitted to saving passwords in the browser, and thus using a browser PM (for more information see Section 5.2.2).

We also asked questions related to formal verification. If users stated knowing formal verification we asked them to explain what it is. The Dunning-Kruger Effect happens frequently and consists of users overestimating their skills and knowledge in self-assessments [53].

We deemed important demographic questions such as level of expertise with computers, education, and gender. The questionnaire is in Appendix B.

5.1.1.B Task Questionnaire

For the **"Task Questionnaire"**, we also used a Likert [52] scale (identical to the "Final Questionaire") to ask about the perceived difficulty and satisfaction when performing the task. We presented this questionnaire after each task. It is also important to learn about the users' perception of formal verification during the task and we did so with the last question of this questionnaire – see example in Fig. 5.1.

After the first three questions, we questioned users about further feedback concerning the task with an informal interview.

5.1.1.C Final Questionnaire

For the **"Final Questionnaire"**, we used the System Usability Scale (SUS). The SUS was developed by Brooke [25] as a "quick and dirty" survey scale that would allow the usability practitioner to quickly and easily assess the usability of a given product or service [21, 54]. Although there are alternatives (e.g., QUIS, CSUQ, SUMI) the SUS was used in more than 2,300 individual surveys collected while conducting more than 200 studies and, it has proven itself a valuable and robust tool in helping assess the quality of a broad spectrum of user interfaces [55]. The SUS is composed of 10 questions and results in a usability score between 0 and 100.

The SUS was originally designed in English but we use the European Portuguese translation by Martins et al. [54] verbally in parallel with the English version for our Portuguese participants. The translation in this study proved to have a good percentage of agreement (76.67%) and can be used to distinguish between usable and non-usable applications [54]. In the "Final questionnaire", we also asked questions related to formal verification and the users' perception of it in a PM. The participants answer according to the commonly used Likert scale [23, 52] (Strongly Agree, Agree, Neutral, Disagree, Strongly disagree).

Finally, in this questionnaire, we also asked some questions from the pre-task questionnaire. With this we hoped to understand if users' opinions and perceptions of PMs changed after using PassCert. The full final questionnaire is in Appendix C.

5.1.2 Observation

In the observation part of the evaluation, our approach is similar to the one designed by Chiasson et al. [16]. An experimenter was with each participant throughout the session, taking notes of users' observations, noting any difficulties, any obvious misconceptions in the participant's mental model, any comments made by the participant and, whether they successfully completed the task.

An effective technique during usability testing is to invite users to "think aloud" (sometimes referred to as "concurrent think-aloud") about what they are thinking as they are performing the task [21]. The

informal atmosphere of a “think-aloud” session is pleasant and often leads to many spontaneous suggestions for improvements [23]. As such, participants were also asked at the beginning of the session to “think aloud”.

Using the methodology designed by Chiasson et al. [16], the outcome of each task is recorded by the observer according to the following possibilities:

- **Successful**: The participant completes the task without difficulty.
- **Difficult success**: The participant eventually completes the task after several attempts (i.e., had difficulty).
- **Failed**: The participant gives up on the task without completing it.
- **False completion**: The participant fails to complete the task but erroneously believes that they are in fact successful.

A key difference between Chiasson et al.’s experiment and ours is that our tasks are all independent of each other, i.e., even if a user fails in one task he can reach success in all the others. As such, within the outcomes of our tasks we do not have the option “Failed due to previous” present in Chiasson et al.’s experiment. This conscience design choice was made so that the task could be presented randomly to the users (see Section 5.1.3).

5.1.3 Tasks

In each session, the participants perform the following set of tasks:

- **Login in the PM**: use the primary password and login in the PassCert. In this task, the user goes through the formal verification icon by the primary password field (implemented in Chapter 4).
- **Register in a website**: register a new user on a website and save the credentials in the PM’s vault. This task is one of the most commonly done tasks in a PM, with Bitwarden a pop-up appears prompting users to save the password to the PM, nonetheless, they can dismiss the pop-up and choose to manually add the credential. In this task, users may be exposed to the secure formally verified vault.
- **Generate a random password**: use the PM’s generator to generate a new random password. As the password generator is a formally verified feature this task allows us to understand how they integrate with this key feature of the PM.
- **Log in to a website**: login to a website that has a previous password saved in the PM. In this task, users explore the autofill feature of the PM.

- **Update password**: update a password saved in the PM's vault to a new one. Here they explore Bitwarden's vault and can also see the formal verification by the password field.

These tasks are based on the ones implemented by Chiasson et al. [16] in their usability study of PMs. They were evaluated according to the possibilities described in Section 5.1.2. As mentioned before, a key difference between Chiasson et al.'s experiment and ours is that our tasks are all independent of each other. This allowed us to randomize the order of the tasks and, as such, in the user test, the tasks were presented to the users in a random order to prevent bias. If they were not randomized, users' performance in the last tasks could be affected by their previous experience with the tool in the first tasks.

5.1.4 Metrics

While in this study we were more focused on the qualitative data from the informal interview, we measured metrics related to the performance of the participants in the tasks. These metrics included clicks and time to complete the task. Furthermore, because we hoped to take notice of to which degree the user interacts with the new Bitwarden features (e.g. if the user clicks on the formal verification icon and spends time on that screen), we constructed websites to test the PM. These needed to be well integrated with the PM [24] because, as Chiasson et al. [16] concluded, some PM's usability problems are due to bad website design.

5.1.4.A Formal Verification

To evaluate the users' perception of formal verification in the PM, we included questions about it in all three questionnaires ("Pre-Task Questionnaire", "Task Questionnaire", and "Final Questionnaire"). These questions used a Likert scale (see Section 5.1.1). The Likert scales were then transformed into numeric values (1 = most negative, 3 = neutral, 5 = most positive). This type of statistical analysis is the most common and accepted way of reporting Likert-scale data as the difference in results between parametric and non-parametric analysis is usually minimal [16].

5.1.4.B Usability

The answers to the SUS on the "Final questionnaire" were aggregated to reach a usability score from 0 to 100. An acceptable SUS score, according to Bangor et al. is above 70, with better products scoring in the high 70s to upper 80s. Products with scores of less than 70 should be considered candidates for increased scrutiny and continued improvement and should be judged to be marginal at best [55].

According to the outcomes of each task, the mean was also calculated. The outcome was measured according to the possibilities described in 5.1.2.

Feedback from the “think-aloud” and from users’ feedback at the end of the testing sessions that may be relevant to usability problems are also are to be put into categories and taken into consideration.

5.1.4.C Baseline

We also compared the base PM (i.e. without any extensions or formal verification) and the extended PM (i.e. Bitwarden with formal verification and the interface extensions). With this, we wanted to understand if the use of formal verification has an impact on users’ perceptions of security and privacy and also if we had been able to improve the final usability of the PM.

The user testing protocol for the baseline was similar to the extended interface but excluded all mentions of formal verification from the questionnaires.

5.1.5 Pilot Studies

We performed two Pilot User Tests with the aim of refining the testing protocol and script. An example of an improvement suggested was to reduce the number of tasks, which we did, by removing a task related to deleting a credential in the PM.

When users performed the tasks we also noticed that they did not explore the interface or click on the formally verified icon; after asking them the reason for this, one replied “*I was focused on the tasks*”. As a result, in the next round of tests we began the session by going through the tutorial for all users (see Section 4.5). Another preliminary result we found was that, even though one of the users stated knowing what formal verification is, they were not able to identify how it was used in PassCert. This may be due to the lack of user interaction with the new help tools. When asked to explain what they understood by formal verification, the user stated “*Something that guarantees security*”, so even though the user did not understand the concept fully, they associated the concept with security. A positive result we found was the use of tooltips as both users used them.

It is important to keep in mind that, according to Alkaldi et al. [10], even if people are aware that PMs exist, they might still not embark on a search process to consider installing one. Further results regarding usability were gathered in the next round of tests (see next section).

5.2 Results

We implemented the previously described protocol (see Section 5.1) in two phases the: i) extended interface, and; ii) base interface. All the user tests were done remotely through Zoom by giving the participants remote access to the researcher’s screen and thus being able to perform the tasks. One limitation of this method is that problems in network speed affect user performance in terms of time.

5.2.1 Participants

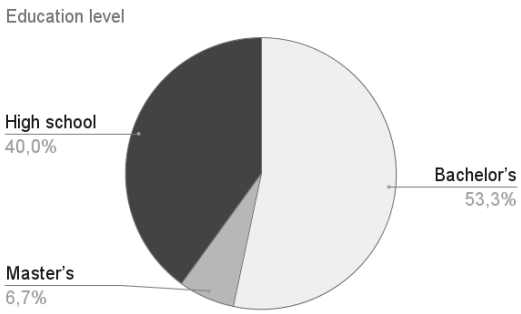
The participants should accurately represent the users who would use the actual system, have similar experience, and knowledge [16]. Although in the past 5 participants have been an acceptable number to identify usability problems [56], recent studies have found that increasing the number from 5 to 10 can result in a dramatic improvement in data confidence [57]. This evaluation is heavily focused on qualitative data and not quantitative, as such we used **15 users in total, 10 for the extended interface and 5 for the baseline interface**. Because we used a small number of users, and because in previous studies on usability of PMs the criticism was on the sample (e.g. the participants in Pearman et al.'s [11] study were skewed towards young people, with a disproportionately high percentage of participants with technical backgrounds; as such, Ray et al. [13] repeated their experiment with older adults), it was important that the chosen sample of users was varied.

The participants were all Portuguese and because the minimum education in Portuguese in Portugal is high school, all had that level, or more, of education (see Figure 5.2(a)). Nonetheless, 60% of the participants had higher education, 53.3% had a Bachelor's degree, and 6.7% had a Master's. Of the 15 participants, only 2 had a technical background related to IT. Regarding age, the most frequent age group was 25-34, 40%. Overall we had 60% of users with less than 34 years and 40% with more or equal to 35 years of age (see Figure 5.2(b)). Gender in our sample was mostly evenly divided because we had 15 participants the percentage was 46.7% women vs. 53.3% men (see Figure 5.2(c)).

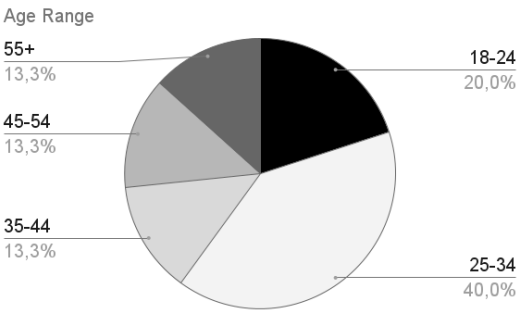
5.2.2 Knowledge and perceptions about Password Managers

We asked participants if they knew what a PM was before the study, and in total 6 of the 15 participants gave correct explanations. Of these 6, 3 were somewhat vague using terms like "a password manager helps to manage passwords". This explanation is right but maybe simply due to logic and not intrinsic knowledge. Other users stated more accurately that a "PM helps secure passwords and unlocks with a single password" and mentioned the auto-fill feature. A piece of important information is about the previous use of PMs, which can be seen in Figure 5.3(a). In this case, only 20% of users (3 users) stated using a PM currently. One participant mentioned using a PM in the past but giving up. When asked why, they stated "I started to use it because a friend recommended their use, and my Facebook was hacked but I couldn't understand how it worked and uninstalled it after a few days". In the case of this participant, difficulty of use was a major barrier to the adoption of the PM. We also found that some users (11 participants, 53%) stated not using PMs in the past but when asked whether they saved passwords in the browser (i.e. used the browser PM) the majority (8 participants of 11) said yes. This suggests that they do not understand what a PM is.

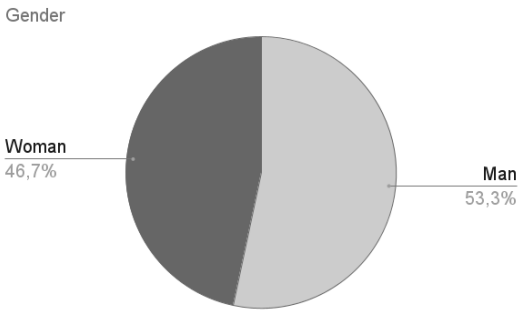
If we group the users that reported using a PM with the users that were unaware they were using a



(a) Participants' education level



(b) Participants' age range



(c) Participants' gender distribution

Figure 5.2: Participants' demographic information

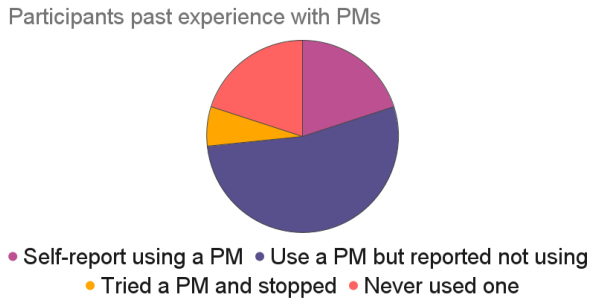
browser PM, then 11 of the 15 users (73%) used a PM.

Furthermore, most participants reported repeating passwords in different accounts (90%), and even 2 of the 3 participants that self-reported using PMs still stated that they had some password repeated across devices. When asked why, one stated "It's easier this way and (...) just for websites I know I'll never return to" and the other, a user of a browser PM, stated "I didn't even know the PM had a generator". This suggests that using a PM does not necessarily lead to more unique passwords. It is important to mention that most users stated that for high-value accounts, such as bank accounts or main social media accounts, they tried to use unique passwords.

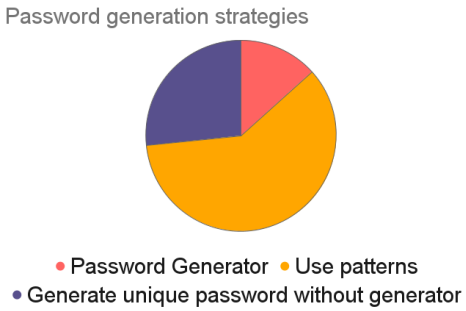
Regarding password generation, only 2 users reported using a password generator for their password (see Figure 5.3(b)). Four participants self-reported generating themselves passwords that are unique, and 9 reported using a pattern for generating a password (e.g. a word and a number), and of these 9, 4 stated they used personal information (e.g. birth dates).

Concerning storage, all users that self-reported using a PM also reported saving their password in it, and of the other participant's common methods included memory, writing down on paper, and 2 even reported using locked files (one with an iPhone and another with a word document). We found that some

users that used a browser PM were not aware of how to access their password vault and would write passwords down to make sure they were not “lost”.



(a) Participants’ past experience with PMs



(b) Participants’ password generation strategies

Figure 5.3: Participants’ PM usage information

The majority of users preferred PMs that were not open-source. When asked why some stated “I don’t want my passwords to be seen” and “Open source means that other people may be able to attack my passwords easier”. This suggests that users do not understand the concept of open-source. Again this may be due to bias on the part of the script, the phrasing of the question, users’ lack of knowledge or even previous misconceptions they had about open source software. We also asked participants what factors were important for them when choosing a PM to use. We provided some factors like “Ease of use” and “Price” and used a Likert transformed into numeric values (1 = most negative, 3 = neutral, 5 = most positive). With a one-way ANOVA analysis we found that not all factors were equal ($p < 0.01$). With a Tukey HSD test, one factor of the set stood out – the PM being open-source: 70% of participants did not want an open-source PM and would rather their next PM was closed-source. When questioned about why, one of the participants said “if everyone can see the code then it is not private”. This presents a misconception about open-source security software. This suggests that a lack of understanding about what is open-source software caused users to reject it.

Other factors that were important for users, but not statistically significant, were the PM being easy to use, feeling that their data is secure, and having formally verified features.

5.2.3 Task Analysis

As stated in Section 5.1.4 we measured key metrics for the evaluation of PassCert, like time to complete the task, number of clicks, interaction with support tools, and comments made by the users while thinking aloud.

After the users performed each task we asked them to fill a task questionnaire, with a one-way

ANOVA analysis ($F(4, 70) = 5,724, p < 0.01$) we found that not all tasks were equal regarding users' difficulty perception: task 5 was considered more difficult than 1 and 2 (with a Tukey's HSD test, ($p < 0.05$). It is important to mention that our work focused on qualitative data and not quantitative as we have a small sample of users (5 for the baseline and 10 for the extended PM) and, as such, our goal is to gather insights on topics that could be used for posterior large scale quantitative studies.

In this questionnaire, we also asked users if they were satisfied with the PM and if the feature they were using had formal verification. With these two questions, we found no statistically relevant difference between the means and, overall, users though formal verification was present in most tasks and were satisfied with the PM.

Next, we will go through each task and provide a brief analysis.

5.2.3.A Task 1

Task 1 consisted in asking the users to login into the PM by entering the primary password. During the think-aloud process, users stated that this task was easy and no users failed this task. This task was considered the easiest by users in the task questionnaire and was overall the shortest task. Although this task is introductory, it was nonetheless useful to gauge users' interaction with the formal verification icon in the login screen. We found that 30% of users clicked on the formal verification icon in this task.

The background of this task was mostly the same between the baseline and the extended interface as the only difference was the use of the formal verification icon by the password field.

5.2.3.B Task 2

Task 2 consisted in asking participants to register a new account on a website and save the credentials to the PM. The expected way of solving this task was creating the new account and clicking on the pop-up that appears on the page asking if they want to save the password in the PM.

The users that had problems with this task were the ones that did not notice this pop-up and instead had to save the password manually in the PM. In this task, two users were unable to complete it (both did not notice the pop-up).

One of the users who noticed the pop-up actually stated that "When I'm in google a box appears asking me to save the password and I was expecting this to happen here too". Other users that did not notice the pop-up did not look for it, and this may be related to less contact with other PMs.

This may suggest that users unfamiliar with PMs will not look for the popup box or that perhaps the pop-up should be more noticeable to make users click on it more.

The background of this task was mostly the same between the extended interface and the baseline. Nonetheless, if users were to manually save the credential, the users from the extended interface would see the formal verification icon in the PM, while users of the baseline would not.

5.2.3.C Task 3

Task 3 consisted in asking participants to generate a random password in the password generator. This password should follow some guidelines like having numeric characters, a minimum length, and not using special characters.

The participants that had gone through the tutorial mentioned remembering the tutorial during the think-aloud procedure. Passing through the tutorial before the task may have helped them find the generator easier. Two users from the baseline, on the other hand, complained about not finding the generator initially (but found it after a moment). All users succeeded in this task. Participants used the tooltips that were absent from the base PM. The tooltip for the special character was especially useful as it does not have a label and most users (70%) used this tooltip.

The background of this task between the baseline and the extended interface was fairly different. Users of the extended interface had a different experience: 1) these users had gone through the tutorial explaining how the password generator worked; 2) a formal verification icon was present in the password generator; 3) there were several tooltips in the extended interface missing from the baseline (see Section 4.1).

5.2.3.D Task 4

Task 4 consisted in asking participants to login into a website with a username and password pair that were already saved in the PM.

This task could be done in two ways: 1) by selecting the credential on the PM, this way it auto-fills immediately; 2) by manually copying the username name and the password and then login into the PM. The autofill solution is much quicker and requires fewer clicks – this solution was addressed in the tutorial where participants were given an explanation about the autofill. When asked why they clicked on the credential one said “I was exploring the interface and saw the text” (with text the user was referring to the tooltip). The tooltips on the credential were also updated on the extended interface telling users to “Click to auto-fill”, something that does not happen in the base PM.

The background of this task was fairly different between the baseline and the extended interface by two factors: 1) the autofill feature being addressed in the tutorial; 2) the autofill tooltip was extended in the extended interface.

5.2.3.E Task 5

Task 5 consisted in asking participants to update a password previously saved in the PM.

To succeed in this task, participants should enter the vault, click on the credential, click on edit and then edit the password. The users struggled with this task due to not finding the edit button on the task.

The background of this task was mostly the same but users of the extended interface were exposed to the formal verification icon in the password vault when they were changing the password.

5.2.3.F Discussion

Overall users of both interfaces performed well in the proposed tasks and the differences are not statistically relevant. We suggest that further studies should be done with a larger number of users.

Nonetheless, our results suggest that the users from the extended interface benefited from having gone through the tutorial in the beginning. Tutorials are usually directed at beginner users and most of our users had no previous contact with a browser-extension PM. Moreover, during the think-aloud, users of the extended PM mentioned remembering where a certain feature was because of the tutorial (e.g. where the generator was) and users of the base PM during the informal interview stated that they would have liked to have more support during their first-time use of the PM. One user even clearly mentioned the need for a tutorial if they were to use this PM in the future.

5.2.4 Usability

As stated before, to make sure our PM is usable we used the SUS, which is composed of 10 questions and outputs in a usability score between 0 and 100 [55].

Overall users' experience with the extended interface was more positive than with the base PM. The average usability score of the extended interface was 75 points (this constitutes an acceptable SUS [55]). The results for the baseline were lower with only 48 points. According to Bangor et al. [55] with this score, the base product should be considered a candidate for increased scrutiny and continued improvement and should be judged to be marginal at best. The difference between the average SUS of the baseline and the extended interface is about 28 points and can be seen in Figure 5.4. Because we have two samples with different sizes (10 users for the extended PM and 5 for the baseline) we used a t-test with unequal sample variance $p = 0.0011$ to compare the average SUS results. More detail can be found in Figure 5.4 and Table 5.1. Due to the think-aloud and the informal interview, our results suggest that the satisfaction and SUS score may be related to the tutorial, as users of the base PM asked for more guidance and users of the extended PM mentioned remembering the tutorial when doing tasks and seemed satisfied when that happened. Nonetheless, for future work, we suggest repeating our testing procedure with more users to reach stronger statistical conclusions.

Although our focus was on the qualitative data and not quantitative, we found that, of the 5 users from the baseline, 4 suggested more guidance on the interface before the tasks, and 2 users stated they would have liked to see some sort of introduction to the tool or tutorial.

The task where users, of the baseline and extended interface, seemed to struggle the most was task 5 where they were asked to generate a password. One user failed this task and 2 others had a difficult

success. On the opposite end, the task where users seemed to succeed more quickly and considered easier was the first one. This, though, was expected seeing as the first task is logging in to the PM.

So, recurring themes emerged from the user testing session such as whether users use the browser or not to save passwords they do not associate saving in the browser with saving in a password manager. Moreover, we also noticed that some users were not willing to save their password in the browser but stated that would be willing to save the password in a PM.

The support features users used more frequently were the tooltips. Regarding these one user stated "I'm using this text to know what will happen when I click the button this way I don't have to risk going to the wrong screen". Frequent uses of the tooltips were associated with the auto-fill feature from tasks 3 and 4. For task 3 the tooltips on the settings of the generator were used frequently and by all participants but they were absent from the base PM. On task 4 users mostly used the tooltip to click on the auto-fill button and login into the website. On the extended interface, the tooltip clearly prompted users to "Click to autofill" while in the base PM the tooltip just said "Autofill".

Regarding the tutorial, all users of the extended PM used it and of the 5 users from the baseline, 4 suggested more guidance on the interface, and 2 users stated they would have liked to see some sort of introduction to the tool or tutorial.

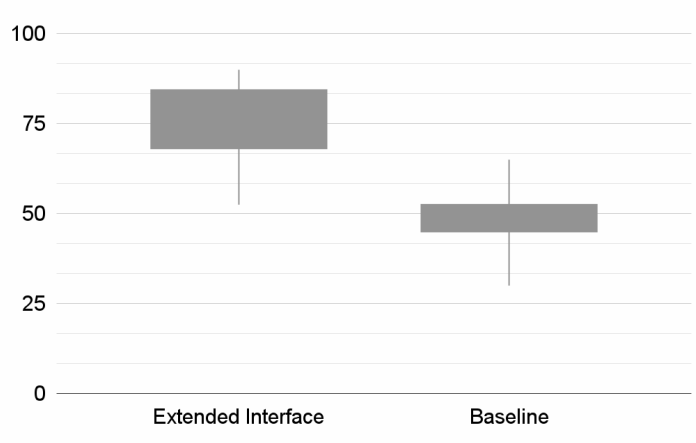


Figure 5.4: SUS Comparison between Baseline interface and Extended interface

Table 5.1: Comparison between Extended Interface and Baseline SUS Scores

	Minimum	1st Quartile	3rd Quartile	Maximum	Standart Deviation	Mean	Difference	p-value
Extended Interface	52,5	68,125	84,375	90	11,93	75,25	27,7	0,0011
Baseline	30	45	52,5	65	12,75	47,5		

Figure 5.5: Formal Verification icon

5.2.5 Formal Verification

Regarding formal verification, we identified some themes from the non-structured interview and from the questionnaires.

A – Users associated formal verification with security While this can be true in some formal verification usages (such as guaranteeing the correctness and security of the cryptographic algorithms on the password vault), formal verification does not necessarily guarantee security. Formal verification itself mathematically guarantees that the program behaves in a certain way, thus ensuring that the formal verified features work as intended.

The reason why users associated the concepts may be bias from our script, the icon (it looks like a green lock, see Figure 5.5), because of the descriptions of formal verification (see Table 4.1) or because of the name formal verification in plain English (users may associate it with some concept resembling the definition of formal verification without fully understanding it). While this is not always the case, in PassCert, formal verification sometimes ensures security. For example by ensuring that the vault's cryptographic algorithms are well-implemented. We recommend that further work should analyze this trend to understand why is this so.

One of our key goals is to learn about users' perception of formal verification and if its use in the PM has an impact on trust. The questions related to formal verification were only directed at the users of the extended interface (i.e. the version of the PM that has formal verification). With this in mind, it was critical to learn about users' previous knowledge of formal verification, and because user self-assessment of technical concepts is often biased (users commonly overestimate their skills and knowledge in self-assessments [53]). After asking users if they knew what formal verification is, we also asked them to explain the concept. Only 2 users had preconceived concepts of it, but only one effectively knew what it is (i.e. gave an explanation matching the concept). So 90% of participants were unfamiliar with the concept.

After using the PM and thus experiencing the tutorial, the icon, and respective explanations on formal verification, we asked users if they now knew what formal verification is. It is important to keep in mind that all explanations of the concept were given in the context of the PM and most users lacked a technical background. After using the PM 80% of participants stated knowing what formal verification is and the remaining 20% stated "not sure". Similar to the previous question on the pre-study questionnaire,

we asked users to explain what they thought formal verification is. Most explanations were non-technical and without jargon, this was expected, and only 3 explanations mentioned that formal verification "guarantees a certain behavior". However, even though most descriptions lacked a general concept of formal verification, 80% of users mentioned the word "security" in their explanations. The remaining 20% were users that successfully gave a general (i.e. non-technical) explanation of formal verification.

This corroborated the results of the pilot tests where users lacked a formal understanding of formal verification but still associated the concept with security (see Section 5.1.5). It is also important to mention that our goal was not for users to develop a deep technical understanding of formal verification and this can be seen in our explanations of the concept in Section 4.3.1. We aimed to transmit to users the concept of formal verification without jargon or technical language.

B – Some users correctly identified the formally verified features Another factor we would like to understand is related to the formally verified features, so, we asked users about what was specifically verified in the PM: 60% identified the generator and the password storage as formally verified; 30% stated that the whole PM was formally verified, when asked why one user stated "I saw the icon in several places". And one user (10%) was not able to give a explanation, stating they did not know.

C – Some users thought the whole PM was formally verified Some users (30%) stated that the whole PM was formally verified which is a misconception. This may have happened due to bias in the script or description of formal verification (see Table 4.1). This misconception can be damaging as it can potentially give users a false sense of security and as such should be studied further in future studies.

D – Formal verification may have affected some users trust After using the PM, and when asked if they felt safe using the PM and if they felt their passwords were secure, 90% agreed or strongly agreed. When asked why, 5 participants mentioned the formal verification icon, 2 users mentioned formal verification, and one user mentioned that the layout of the PM seemed trustworthy. One participant stated that they did not feel their passwords were secure in the PM, when asked why they stated "I'm skeptical of this kind of software (...) it's difficult for me to trust third parties with my passwords".

When comparing perceptions of formal verification users stated trusting the formally verified PM more than the non-formally verified (using paired t-test with $p = 0,0469$, see Figure 5.6 and Table 5.2).

One important user support element related to formal verification was its icon Figure 5.5. To understand if participants knew its meaning we asked them if they knew what the icon was supposed to represent and for them to explain what it was. Of the 10 participants, 5 mentioned formal verification and the other 5 mentioned concepts related to the security of the PM like "the icon means that the passwords were safe".

Overall 70% of users interacted with the formal verification during the tasks using the tooltip. Of this 70%, 43% clicked on the icon. When we asked why users did not click on the icon the answers fell into two categories: i) users stated that they were focused on the tasks, “I wanted to finish the task”; ii) users stated that they knew what the icon meant from the tooltip “I thought I understood (...) the meaning of the symbol, the label said it was verified” (with label the participant meant the tooltip).

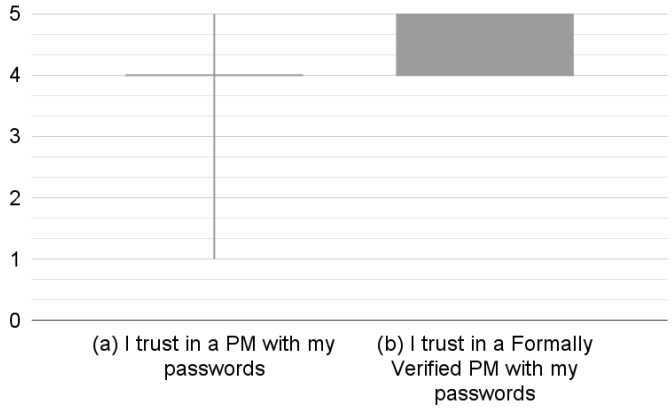


Figure 5.6: Comparison between Likert scale results trust in PM with formal verification

Table 5.2: Comparison between Likert scale results trust in PM with formal verification (higher is better) (a) I trust in a PM with my passwords (b) I trust in a Formally Verified PM with my passwords

	Minimum	1st Quartile	Median	3rd Quartile	Maximum	Standart Deviation	Difference	p-value
(a)	1	4	4	4	5	1,03	0,8	0,0469
(b)	4	4	5	5	5	0,53		

5.3 Discussion

We propose several extensions to the base PM with two clear goals: (a) ensure the usability of the final PM; (b) provide information on formal verification. In our evaluation, we wanted to assess our solution, and learn more about users’ perception of formal verification and PMs.

As mentioned before, it is important to mention that our work focused on qualitative data and not quantitative as we have a small sample of users (5 for the baseline and 10 for the extended PM). As such our goal is to gather insights on topics that could be used for posterior large-scale quantitative studies.

Regarding the usability of the PM we found that PassCert scored higher than the baseline in SUS (see Section 5.1.1.C). Although these results suggest an improvement, this study should be repeated with more users to get a more significant statistical result. Nonetheless, the expanded PM’s score

suggested that our end product is usable. We strongly recommend that Bitwarden implements some of our extensions such as the tooltips and the tutorial.

5.3.1 Perceptions of formal verification and Password Managers

In regards to users perception of formal verification, we found that most (90%) were unfamiliar with the concept after the study and when asked to explain what formal verification was most (80%) mentioned the word security in their explanations of formal verification – this may suggest that they associate the two concepts. Formal verification does not always imply security so future research should study in more depth users' perceptions of formal verification.

We wanted to make sure users understood what was formally verified in the PM and this was the case as some participants correctly identified the generator and the storage as formally verified (see Section 5.2.5). No participant identified the master password as a formally verified feature, maybe due to low expose to that feature (only in one task).

A different perspective was found when asking users to state what would be important for them in a PM: the majority of users stated that they did not want to use a PM that was open-source, and the reason why revealed a lack of understanding of what open-source software is. We suggest future research should also look into users' perceptions of open-source security software.

Like in literature [3, 11] we found that some participants (60%) reported that they were not aware of the existence of PMs before the study. One user stated "I never saw any publicity (...) like I see for anti-virus". The majority of users also stated that they were open to trying to use PMs in the future. This suggests that unawareness is a barrier to the use of PMs. One of the participants had used a PM in the past but stopped using it because they were not able to understand how it worked stating "I felt everything was very complicated to do". This is corroborated by the literature where lack of usability was identified as a barrier to effective use of PMs [11, 16].

5.3.2 Adoption of Password Managers

Our results demonstrate that most users correctly identified the formally verified features. The results also suggest that we have improved the usability of the PM. By improving the usability of the tool and providing formal verification, user adoption of PMs may improve, but further long-term studies must be done in order to gather insights on the impact of formal verification in the adoption of PMs.

6

Conclusion

Contents

The advantages of using PMs are undeniable. As such it is important to make users trust and want to use them. This project is a part of the PassCert research project that will build an open-source, proof-of-concept formally verified PM. Previous work from members of the PassCert project identified the need for better user-experience design and thorough usability test of password managers [11].

We have surveyed usable security techniques that can be applied to improve password managers; aimed to ensure that the password managers developed in the context of the PassCert project integrate best practice guidelines developed by the usable security community; explored ways to effectively convey the formally verified properties of the password managers; performed user studies to determine the impact of formal verification on the adoption of password managers; and finally gathered insights on users perception of PMs and formal verification.

In our solution, we proposed ways to educate users about formal verification and increase their trust in the software. Regarding usability, we wanted to implement the relevant usability best practices in the solution such as informing them about what each security feature does.

We implemented several support tools for users such as more tooltips, a tutorial, FAQ page, and solved inconsistencies in the base PM's interface. To ensure users understand the formal verification concepts present in our solution, we designed and implemented a new icon in PassCert meant to represent them. With this icon, we provide explanations on the various formally verified features of the PM.

After implementing the proposed solution we performed user tests where we learned about the usability of the solution, users' perception of PMs, and of formal verification. Our results suggested that our solution has better usability than the base PM. Additionally, some of the insights gathered suggest that there is a general unawareness of both PMs and formal verification. Moreover, our results suggest that we were able to effectively convey the formally verified features as most users were able to successfully identify them. While users did not present a formal understanding of formal verification in general most associated the concept with security.

PassCert's PM is composed of our interface extensions and the work of other project members to ensure a full PM that aims to provide a usable and secure experience for users. We contribute with first user study on perceptions of formal verification on PMs. We hope our insights can help the formal methods security community better communicate with end-users about its assurances.

6.1 Research Questions

In this section, we summarize the answers to the research questions answered in previous chapters and identified in Chapter 1.

6.1.1 RQ1. What are the usable security techniques that can be applied to PMs?

Several suggestions have been made concerning a possible solution to ensure the usability of PMs such as: providing tutorials about how the PM interface works [24]; ensuring a precise interface that is consistent and well-designed [20,21,39]; and providing a solid implementation of all PM's features [3,24].

User tests can be used to identify problems in PMs and to test solutions. All extensions done to a PM with the goal of improving usability should be done while following usability guidelines such as "*The Eight Golden Rules of Interface Design*" [23], the usability recommendations of Whitten and Tygar [2] or even Chiasson et al [16]'s extensions to the previous. Additionally using a well-designed security mechanism is still more effort than not using it at all, and users will always be tempted to cut corners [20]. As such, to ensure users are motivated to use PMs, they should be transparent and provide users with information about how they work. Techniques such as status icons, tooltips, explanations, and tutorials should be included in PMs. Moreover, these additions should be implemented while following the usability guidelines mentioned previously.

6.1.2 RQ2. How can we effectively convey formally verified properties of a PM to its users?

In PassCert, a primary concern is educating the users about formal verification. To convey that a certain feature is formally verified one can use an interface symbol. These include status icons (e.g. the formal verification icon in PassCert's PM) but also colors or pop-ups. Concise explanations about formal verification and how it can guarantee certain security properties are also important and can be provided through tutorials, explanations in the interface, FAQs, and support pages. User tests should be done to ensure that the PM is effectively conveying its formally verified properties to users. These tests should include questions about users' understanding of formal verification before and after using the PM.

6.1.3 RQ3. What are users' perceptions of formal verification and PMs?

Our results suggest that most users are unfamiliar with the concept of formal verification. Moreover after using our PM most associated formal verification with security. Our results also suggest that users identify better the formally verified features they were more exposed to (e.g. the password vault was more often identified as formally verified than the master password – users were less exposed to the master password). In regards to PMs, some users knew of their existence before the study but few stated using PMs. The results also suggest that users had a lack of understating about what PMs are (we found that most users who used a browser-built in PM did not identify it as a PM).

6.2 Threats to validity

User studies such as these may suffer from bias. Bias can arise from the questions asked, the questionnaires, the description of formal verification, the formal verification icon (it looks like a green lock, see Figure 5.5) or even because the name formal verification in plain English (users may associate it with some concept resembling the definition of formal verification without fulling understanding it).

To mitigate these risks we went through a long design process from the descriptions, the icon, and having reviewed the testing protocol with members of the PassCert team. We also made sure to randomize the order of the tasks and provide approximately the same experience for all participants. Nonetheless, problems related to bias can still occur, such as the Dunning-Kruger Effect where users overestimate their skills and knowledge in self-assessments [53]. To mitigate this problem, after asking users to self-assess their knowledge, we always ask them to explain what is their understanding of the topic (e.g. if they state they know what is a PM we ask them to explain what it is to assert their knowledge).

Other biases we must take into consideration when analyzing the results include the Hawthorne effect where users may be inclined to agree with the researchers [58].

Additionally, the sample of participants can also induce bias in the results, and we have a small sample of users in the users' studies. To mitigate this problem we aim for a diverse sample of users. Nonetheless, due to the sample size, we are not able to gather strong statistically significant findings. As stated before this evaluation is heavily focused on qualitative data as we are trying to gather insights on relevant research paths for the future.

6.3 Future Work

Future work could include a translation into Portuguese and this was something mentioned by members of the PassCert team and from users in the user's tests as a barrier to the adoption of PMs.

Another major problem in PMs is the transition into a PM as it can be a cumbersome task for some users. There are two main ways:

1. The user gathers all his previously saved passwords (from all his password management tools such as memory/excel sheets/notebooks) and then adds them one by one to the PM. This approach is labor-intensive and longer and requires a high motivation on the part of the user.

2. The other method is by passively adding passwords to the PM. One can do this by enabling a PM and using the browser/smartphone as normal and waiting for the PMs' prompts to save passwords. A user then chooses to save the password to the PM for future use and gradually and through the next days/weeks he will keep saving his passwords. This method although slower is more passive

than the first one and thus less labor-intensive. The problem is that some accounts may be used rarely and will take a long time to be saved in it.

A problem remains, the passwords that were saved in the PM would still be the same that the user-generated, only new passwords would be generated by the password manager. As user-generated passwords are more vulnerable to guessing attacks and likely to be reused, several of the user's accounts would still be vulnerable. So we suggest that further work should be done to study user transition into PMs and if this has an impact on the adoption of PMs.

It is important to mention that our work focused on qualitative data and quantitative (but subjective) data, as we have a small sample of users. As such we have gathered insights on topics that could be used for posterior large-scale quantitative studies. Future work could include topics such as users' unawareness of formal verification, misconceptions about PMs, and users' perceptions of formal verification in PM.

Our results seem to indicate that most users are unaware of formal verification as such we suggest that future research should study users' pre-conceptions on formal verification in general. This study could be done in a qualitative way with semi-structured interviews and by focusing on formal verification in different domains.

Formal verification in this study was strictly applied to PMs and the method of transmitting information was also specific to PMs (in this case the formal verification icon). Future research on this topic should study different approaches of conveying formal verification in different contexts (i.e. other formally verified software).

Finally, due to the limited time frame of a master thesis, we were not able to perform longitudinal user studies to measure PM adoption. To understand the impact that formal verification has on adoption and user retention in PMs, future work on this topic should include long-term user studies.

Bibliography

[1] C. Herley and P. C. van Oorschot, "A research agenda acknowledging the persistence of passwords," in *Published in IEEE Security and Privacy Magazine, Volume 10 Issue 1, Jan.-Feb.* IEEE, 2012, pp. 28–36.

[2] A. Whitten and J. D. Tygar, "Why johnny can't encrypt: A usability evaluation of pgp 5.0." in *USENIX Security Symposium*, vol. 348, 1999, pp. 169–184.

[3] E. Stobert and R. Biddle, "The password life cycle: user behaviour in managing passwords," in *10th Symposium On Usable Privacy and Security (SOUPS 2014)*, 2014, pp. 243–255.

[4] K. P. Brussel, "Special eurobarometer 480 - europeans' attitudes towards cybersecurity," European Commission, Directorate-General for Migration and Home Affairs, Report, March 2019.

[5] S. Gaw and E. W. Felten, "Password management strategies for online accounts," in *Proceedings of the second symposium on Usable privacy and security*, 2006, pp. 44–55.

[6] C. N. D. Cibersegurança, "Boas práticas no uso de palavras-passe," 2019, [Online; accessed 12-December-2020]. [Online]. Available: https://www.cncs.gov.pt/content/files/bp_pp_nov19.pdf

[7] E. U. A. for Cibersecurity, "Authentication methods," [Online; accessed 12-December-2020]. [Online]. Available: https://www.enisa.europa.eu/topics/csirts-in-europe/glossary/authentication-methods

[8] Z. Li, W. He, D. Akhawe, and D. Song, "The emperor's new password manager: Security analysis of web-based password managers," in *23rd USENIX Security Symposium (USENIX Security 14)*, 2014, pp. 465–479.

[9] D. Silver, S. Jana, D. Boneh, E. Chen, and C. Jackson, "Password managers: Attacks and defenses," in *23rd USENIX Security Symposium (USENIX Security 14)*, 2014, pp. 449–464.

[10] N. Alkaldi and K. Renaud, "Why do people adopt, or reject, smartphone password managers?" in *The 1st European Workshop on Usable Security (EuroUSEC 2016)*, 2016.

[11] S. Pearman, S. A. Zhang, L. Bauer, N. Christin, and L. F. Cranor, "Why people (don't) use password managers effectively," in *Fifteenth Symposium On Usable Privacy and Security (SOUPS 2019). USENIX Association, Santa Clara, CA*, 2019, pp. 319–338.

[12] I. Ion, R. Reeder, and S. Consolvo, ""... no one can hack my mind": Comparing expert and non-expert security practices," in *Eleventh Symposium On Usable Privacy and Security (SOUPS 2015)*, 2015, pp. 327–346.

[13] H. Ray, F. Wolf, R. Kuber, and A. J. Aviv, "Why older adults (don't) use password managers," in *30th USENIX Security Symposium)*, 2021.

[14] P. Arias-Cabarcos, A. Marín, D. Palacios, F. Almenárez, and D. Díaz-Sánchez, "Comparing password management software: toward usable and secure enterprise authentication," *IT Professional*, vol. 18, no. 5, pp. 34–40, 2016.

[15] A. Karole, N. Saxena, and N. Christin, "A comparative usability evaluation of traditional password managers," in *International Conference on Information Security and Cryptology*. Springer, 2010, pp. 233–251.

[16] S. Chiasson, P. C. van Oorschot, and R. Biddle, "A usability study and critique of two password managers." in *USENIX Security Symposium*, vol. 15, 2006, pp. 1–16.

[17] J. F. Ferreira, S. Johnson, A. Mendes, and P. Brooke, "Certified password quality: A case study using coq and linux pluggable authentication modules," in *13th International Conference on Integrated Formal Methods*, 2017.

[18] S. Johnson, J. F. Ferreira, A. Mendes, and J. Cordry, "Skeptic: Automatic, justified and privacy-preserving password composition policy selection," in *Proceedings of the 15th ACM Asia Conference on Computer and Communications Security*, 2020, pp. 101–115.

[19] M. Grilo, J. F. Ferreira, and J. B. Almeida, "Towards formal verification of password generation algorithms used in password managers," *arXiv preprint arXiv:2106.03626*, 2021.

[20] M. A. Sasse and I. Flechais, "Usable security: Why do we need it? how do we get it?" O'Reilly, 2005.

[21] M. J. Fonseca, P. Campos, and D. Gonçalves, "Introdução ao design de interfaces," *FCA-Editora de Informática*, 2017.

[22] J. G. Redish, *Letting go of the words: Writing web content that works*. Morgan Kaufmann, 2012.

[23] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist, and N. Diakopoulos, *Designing the user interface: strategies for effective human-computer interaction*. Pearson, 2016.

[24] S. Seiler-Hwang, P. Arias-Cabarcos, A. Marín, F. Almenares, D. Díaz-Sánchez, and C. Becker, ""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*, 2019, pp. 1937–1953.

[25] J. Brooke, "Sus: a "quick and dirty'usability," *Usability evaluation in industry*, p. 189, 1996.

[26] C. Carreira, J. F. Ferreira, and A. Mendes, "Towards improving the usability of password managers," *INFORUM*, 2021.

[27] C. Carreira, J. F. Ferreira, A. Mendes, and N. Christin, "Exploring usable security to improve the impact of formal verification: A research agenda," *First Workshop on Applicable Formal Methods (co-located with Formal Methods 2021)*., 2021.

[28] P. G. Inglesant and M. A. Sasse, "The true cost of unusable password policies: password use in the wild," in *Proceedings of the sigchi conference on human factors in computing systems*, 2010, pp. 383–392.

[29] A. Das, J. Bonneau, M. Caesar, N. Borisov, and X. Wang, "The tangled web of password reuse." in *NDSS*, vol. 14, no. 2014, 2014, pp. 23–26.

[30] J. Kim, "The effect of design characteristics of mobile applications on user retention: an environmental psychology perspective," 2012.

[31] K. Rodden, H. Hutchinson, and X. Fu, "Measuring the user experience on a large scale: user-centered metrics for web applications," in *Proceedings of the SIGCHI conference on human factors in computing systems*, 2010, pp. 2395–2398.

[32] A. Dix, A. J. Dix, J. Finlay, G. D. Abowd, and R. Beale, *Human-computer interaction*. Pearson Education, 2004.

[33] D. Balfanz, G. Durfee, D. K. Smetters, and R. E. Grinter, "In search of usable security: Five lessons from the field," *IEEE Security & Privacy*, vol. 2, no. 5, pp. 19–24, 2004.

[34] D. Gentner and A. L. Stevens, *Mental models*. Psychology Press, 2014.

[35] D. Cyr, "Modeling web site design across cultures: relationships to trust, satisfaction, and e-loyalty," *Journal of management information systems*, vol. 24, no. 4, pp. 47–72, 2008.

[36] A. Vaibhav and P. Gupta, "Gamification of moocs for increasing user engagement," in *2014 IEEE International Conference on MOOC, Innovation and Technology in Education (MITE)*. IEEE, 2014, pp. 290–295.

[37] G. L. Ciampaglia and D. Taraborelli, “Moodbar: Increasing new user retention in wikipedia through lightweight socialization,” in *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 2015, pp. 734–742.

[38] A. R. Kearney and S. Kaplan, “Toward a methodology for the measurement of knowledge structures of ordinary people: the conceptual content cognitive map (3cm),” *Environment and behavior*, vol. 29, no. 5, pp. 579–617, 1997.

[39] R. Molich and J. Nielsen, “Improving a human-computer dialogue,” *Communications of the ACM*, vol. 33, no. 3, pp. 338–348, 1990.

[40] K. Sherwin, “Pop-ups and adaptive help get a refresh,” Mar 2015. [Online]. Available: https://www.nngroup.com/articles/pop-up-adaptive-help/

[41] D. Reichl. [Online]. Available: https://keepass.info/

[42] “Browser market share,” [Online; accessed 21-December-2020]. [Online]. Available: https://www.netmarketshare.com

[43] “User agent breakdowns,” [Online; accessed 31-December-2020]. [Online]. Available: https://analytics.wikimedia.org/dashboards/browsers/#all-sites-by-browser

[44] “Bitwarden open source password manager.” [Online]. Available: https://bitwarden.com/

[45] J. Yan, A. Blackwell, R. Anderson, and A. Grant, “The memorability and security of passwords–some empirical results,” University of Cambridge, Computer Laboratory, Tech. Rep., 2000.

[46] S. Wiedenbeck, “The use of icons and labels in an end user application program: An empirical study of learning and retention,” *Behaviour & Information Technology*, vol. 18, no. 2, p. 68–82, Jan 1999.

[47] A. Joyce, “Tooltip guidelines,” Jan 2019. [Online]. Available: https://www.nngroup.com/articles/tooltip-guidelines/

[48] Y. Rogers, H. Sharp, and J. Preece, *Interaction design: beyond human-computer interaction*. John Wiley & Sons, 2011.

[49] A. V. Jenifer Tidwell, Charles Brewer, *Designing interfaces: Patterns for effective interaction design*. ” O’Reilly Media, Inc.”, 2020.

[50] A. Harley, “Usability testing of icons,” 1 2016. [Online]. Available: www.nngroup.com/articles/icon-testing/

[51] I. Apple Computer, *Macintosh Human Interface Guidelines*, 10 1992.

[52] R. Likert, “A technique for the measurement of attitudes.” *Archives of psychology*, 1932.

[53] K. Mahmood, “Do people overestimate their information literacy skills? a systematic review of empirical evidence on the dunning-kruger effect,” *Communications in Information Literacy*, vol. 10, no. 2, p. 3, 2016.

[54] A. I. Martins, A. F. Rosa, A. Queirós, A. Silva, and N. P. Rocha, “European portuguese validation of the system usability scale (sus),” *Procedia Computer Science*, vol. 67, pp. 293–300, 2015.

[55] A. Bangor, P. T. Kortum, and J. T. Miller, “An empirical evaluation of the system usability scale,” *Intl. Journal of Human–Computer Interaction*, vol. 24, no. 6, pp. 574–594, 2008.

[56] R. A. Virzi, “Refining the test phase of usability evaluation: How many subjects is enough?” *Human factors*, vol. 34, no. 4, pp. 457–468, 1992.

[57] L. Faulkner, “Beyond the five-user assumption: Benefits of increased sample sizes in usability testing,” *Behavior Research Methods, Instruments, & Computers*, vol. 35, no. 3, pp. 379–383, 2003.

[58] F. Merrett, “Reflections on the hawthorne effect,” *Educational Psychology*, vol. 26, no. 1, pp. 143–146, 2006.

A

Frequently asked questions - Formal verification

A – What is Formal Verification? Formal verification is a process in which developers prove (mathematically) that a certain part of a program behaves as intended.

To say that a feature is formally verified means that it is guaranteed to behave a certain way.

A concrete example is guaranteeing that the length of generated passwords satisfies the minimum requirements.

B – How does formal verification offer security? Formal verification is a process in which developers prove (mathematically) that a certain part of a program behaves as intended.

To say that a feature is formally verified means that it is guaranteed to behave a certain way.

A concrete example is guaranteeing that the length of generated passwords satisfies the minimum requirements.

C – What features are formally verified? The password generator, the password vault, the copy to clipboard feature, and the security of the master password.

D – How is the Generator verified? PassCert's password generator guarantees unpredictability: if an attacker wants to discover a password, they can only do so by trying all possible passwords. Moreover, the passwords generated always respect your specification. For example, if you select the option to include numbers, the generated password will necessarily contain numbers.

E – How are the passwords verified? When a password is no longer needed, PassCert guarantees that it is no longer present in memory. Therefore, attackers with access to memory will have no access to secret data.

To use a password we need to access it and then submit it to the site where we want to login. PassCert will guarantee that when a password is no longer needed it stays exposed the minimum time possible, reducing the chance of it being leaked.

The cryptographic tools used to keep your password secure are also formally verified. We verified the correctness and security of the algorithms used.

F – How is the clipboard verified? When a password is copied to the clipboard, PassCert guarantees through formal verification that it will be cleared from memory (within the time frame you choose), reducing the chance of it being leaked.

B

Pre-study Questionnaire

Pre-study Questionnaire

The goal of this survey is to information about you, your experience with Password Managers and specifically with PassCert's Password Manager.

This project is part of the PassCert research project, a CMU Portugal exploratory project that is building an open-source, proof-of-concept password manager that through the use of formal verification, is guaranteed to satisfy properties on data storage and password generation.

All the data collected is anonymous and will be used solely by the researchers of PassCert. The data may be used to present insights at conferences, academic events, or similar events and for scientific publications.

Your participation is voluntary, and you may always quit at any time, without any kind of penalty. By selecting "Yes" below, you are consenting for your data to be processed, stored, and used as described above. You also are confirming that you have read this consent form.

*Obrigatório

1. Do you consent with your data to be used as described in the consent form and confirm that you have read it? You must click Yes in order to take the survey. *

   Marcar apenas uma oval.

   - ◯ Yes
   - ◯ No

PM

2. Did you know what a password manager (PM) was before this study? *

   Marcar apenas uma oval.

   - ◯ Yes
   - ◯ No
   - ◯ I'm not sure

3. If you answered "yes" to the previous question please explain what a Password Manager is below: *

_____

4. Do you use a password manager? *

Marcar apenas uma oval.

- ◯ Yes, I currently use a PM *Avançar para a pergunta 11*
- ◯ No, but I've used one in the past *Avançar para a pergunta 8*
- ◯ No, I've never used one *Avançar para a pergunta 15*
- ◯ I'm not sure *Avançar para a pergunta 15*

5. How do you manage/remember passwords?
(open ended question, answer verbally)

_____

Never used a PM

6. Before this study have you ever been interested in using a PM?

Marcar apenas uma oval.

- ◯ Yes
- ◯ No, and I don't know what they are
- ◯ No, but I know what PMs are

7. How much do you agree with the following statements

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I want to use a PM in the future					

Avançar para a pergunta 15

Has stopped using a PM

8. What PM did you use?

9. Why did you stop using it?

10. How much do you agree with the following statements

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I want to use a PM in the future	◯	◯	◯	◯	◯

Avançar para a pergunta 15

Uses PM

11. What PM do you use?

_____

12. How often have you used your PM? (e.g. using the credential in PM to log in websites?

Marcar apenas uma oval.

- ◯ More than once a day
- ◯ Once a day
- ◯ A few times a week
- ◯ A few time a month
- ◯ Fewer than once a month

13. In which context do you use a PM?

Marcar tudo o que for aplicável.

- ☐ Work context
- ☐ Personal context

Outra: ☐ _____

14. What types of account do you save in the PM?

Marcar tudo o que for aplicável.

- [ ] Social Media
- [ ] Homebanking
- [ ] Credit Cards
- [ ] Online shopping
- [ ] Mail
- [ ] News/Entertainment

Outra: [ ] _____

Demographics

15. How much do you agree with the following statements

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I trust password managers (with my passwords).					
I think PMs are, overall, difficult to use					
I feel my password are safer in a PM					
I feel PMs are, overall anoying to use					

16. When using a browser do you allow the browser to remember your password?

Marcar apenas uma oval.

- Yes
- No

17. Do you know what formal verification is?

18. How much do you agree with the following statements

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I am willing to pay for software products					
I understand how PMs work					

19. Age?

Marcar apenas uma oval.

- 18-24 years old
- 25-34 years old
- 35-44 years old
- 45-54 years old
- above 55

20. Gender?

Marcar apenas uma oval.

- Man
- Woman
- Non-binary
- Prefer not to say
- Outra: _____

21. How many different internet accounts do you have?

Marcar apenas uma oval.

- 1-5 accounts
- 6-20 accounts
- 21+

22. What devices do you use daily?

Marcar tudo o que for aplicável.

- [ ] Laptop
- [ ] Desktop
- [ ] Smartphone
- [ ] Tablet
- [ ] Outra: _____

23. What is the highest degree you have completed? (If currently enrolled, highest degree received.)

Marcar apenas uma oval.

- ◯ Less than a high school diploma
- ◯ High school degree or equivalent
- ◯ Bachelor's degree
- ◯ Master's degree
- ◯ Doctorate degree

C

Final Questionnaire

1. How much do you agree with the following statements:

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I think that I would like to use this system frequently.					
I found the system unnecessarily complex.					
I thought the system was easy to use.					
I think that I would need the support of a technical person to be able to use this system.					
I found the various functions in this system were well integrated.					
I thought there was too much inconsistency in this system.					
I would imagine that most people would learn to use this					

system very quickly.					
I found the system very cumbersome to use.	◯	◯	◯	◯	◯
I felt very confident using the system.	◯	◯	◯	◯	◯
I needed to learn a lot of things before I could get going with this system.	◯	◯	◯	◯	◯

Section 2 - Knowledge about PMs

2. What does this icon symbolize (in the context of PassCert)?
(open ended question, answer verbally)

3. How much do you agree with the following statements:

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I feel my password are less safe in a PM					
I feel I could trust in a PM to save my passwords for me					
I know how a PM works					
It is more convenient to use a PM than to memorize passwords					
I think PMs are, overall, difficult to use					
If I were to use a PM it would not be important for me to understand how the PM works					

4. When using a PM, how important is:

Marcar apenas uma oval por linha.

	Not at all important	Slightly Important	Important	Fairly Important	Very Important
Ease of use					
Support material (tutorials, help pages)					
Security and Privacy/Feeling that my data is secure					
Having formally verified features					
Price					
Understanding how the PM works					
Being open source (i.e. the code is open for everyone to see)					
The company that made and manages the PM					

Section 3 - Impact of Formal Verification on PMs

5. How much do you agree with the following statements:

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I feel my password are less safe in a formally verified PM					
I feel I could trust in a formally verified PM to save my passwords for me					
I know how a formally verified PM works					
I prefer PMs that are not formally verified					

Section 4 - Impact of Formal Verification in general

6. How much do you agree with the following statements:

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I would be willing to pay more for a software product if I knew it was formally verified					
I would not trust a product just because it was formally verified					
I do not value the use of Formal Verification in software products					

Section 5 - Questions about PassCert's PM

7. After using PassCert's PM how much do you agree with the following statements:

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I know how this PM works					

8. Do you know what formal verification is?

_____

9. In your understanding, what is formal verification?
(open ended question, answer verbally)

_____

10. What specific features were formally verified?
(open ended question, answer verbally)

_____

11. The formal verification in * is used for:
(open ended question, answer verbally)

_____

Section 6 - Perception of Safety

12. After using PassCert's PM how much do you agree with the following statements:

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I felt safe using the PM					
I felt my password were secure in the PM					

13. Why?

Section 6 - End Questions

14. Will you use/continue to use a PM in the future?

15. After using PassCert's PM how much do you agree with the following statements:

Marcar apenas uma oval por linha.

	Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
I would be more willing to use a PM in the future if I knew it was formally verified					

16. If you have any further feedback feel free to share