Nov 14th, 10:30 AM - 10:50 AM

Towards the Development and Assessment of a Method for Educating Users into Choosing Complex, Memorable Passphrases

Juan M. Madrid
Universidad Icesi, jmadrid@icesi.edu.co

Yair Levy
Nova Southeastern University, levyy@nova.edu

Laurie Dringus
Nova Southeastern University, laurie@nova.edu

Ling Wang
Nova Southeastern University, lingwang@nova.edu

Madrid, Juan M.; Levy, Yair; Dringus, Laurie; and Wang, Ling, "Towards the Development and Assessment of a Method for Educating Users into Choosing Complex, Memorable Passphrases" (2022). *KSU Proceedings on Cybersecurity Education, Research and Practice*. 5.
https://digitalcommons.kennesaw.edu/ccerp/2022/Research/5

Abstract

The currently most used method for authentication is the password because it is simple to implement, and computer users are very familiarized with it. However, passwords are vulnerable to attacks that can be mitigated by increasing the complexity of the chosen password, particularly in terms of length. One possible approach to accomplish this is through the usage of passphrases, which can be easier to remember than a standard password, thus reducing the loss of work time and productivity related to forgotten passwords. To achieve the required balance between complexity and memorability, the concept of passphrase categories can be used, i.e. more sensitive accounts or services should have more complex passphrases, and vice versa. This work-in-progress study proposes to develop and assess a method for educating users into creating complex, yet easy to remember passphrases, according to the category of account or service they want to protect. The work-in-progress study will be developed in three phases, including validation of the method by a panel of subject matter experts, a pilot test, and a main data collection and analysis phase.

Disciplines

Information Security | Other Education

Towards the Development and Assessment of a Method for Educating Users into Choosing Complex, Memorable Passphrases

Juan M. Madrid
Information and Communication Technology Department
Universidad Icesi
Cali, Colombia
jmadrid@icesi.edu.co
ORCID: 0000-0001-5094-4868

Yair Levy
College of Computing and Engineering
Nova Southeastern University
Davie, FL, USA
levyy@nova.edu
ORCID: 0000-0002-8994-6497

Laurie Dringus
College of Computing and Engineering
Nova Southeastern University
Davie, FL, USA
laurie@nova.edu
ORCID: 0000-0002-9834-6098

Ling Wang
College of Computing and Engineering
Nova Southeastern University
Davie, FL, USA
lingwang@nova.edu
ORCID: 0000-0002-9202-6501

***Abstract*— The currently most used method for authentication is the password because it is simple to implement, and computer users are very familiarized with it. However, passwords are vulnerable to attacks that can be mitigated by increasing the complexity of the chosen password, particularly in terms of length. One possible approach to accomplish this is through the usage of passphrases, which can be easier to remember than a standard password, thus reducing the loss of work time and productivity related to forgotten passwords. To achieve the required balance between complexity and memorability, the concept of passphrase categories can be used, i.e. more sensitive accounts or services should have more complex passphrases, and vice versa. This work-in-progress study proposes to develop and assess a method for educating users into creating complex, yet easy to remember passphrases, according to the category of account or service they want to protect. The work-in-progress study will be developed in three phases, including validation of the method by a panel of subject matter experts, a pilot test, and a main data collection and analysis phase.**

Keywords*— passphrases, passphrase complexity, passphrase memorability, cognitive load theory, passphrase levels, account categories, user

I. INTRODUCTION

Many authentication methods have been devised for verifying a user's identity [2]. Such methods include knowledge factors (something the user knows), ownership factors (something the user has), and biometrics (something the user is) [20]. Despite all this variety, the most currently used authentication method is the password, because of its simplicity of implementation and the familiarity of the average user with the method [26]. According to [24], many multiple factor authentication systems use passwords as one of their factors. However, passwords are vulnerable to brute-force and dictionary attacks [2], thus leading to account compromise. Along with malware, account take-over (ATO) and credential abuse attacks continue to be of the most concern for organizations [6].

The research problem this work-in-progress study will address is the issue of password memorability and complexity, compounded by the high number of password-protected accounts and services an average person manages. This can be mitigated by increasing the complexity of the chosen password, in terms of choice of characters and length [5]. Nevertheless, very complex passwords are easy to forget, leading to lost time and productivity [18, 25].

Two important attributes of passwords are complexity and memorability. According to [5], complexity or password entropy is defined in terms of the password length and choice of characters used in it (alphabetic, numbers, symbols), with length as the most important factor. Memorability is defined in [16] as how easy is for the user to correctly recall the password. Overly complex passwords can be easily forgotten; this can be explained by the cognitive load theory, which researches human limitations for learning new information and suggests improved methods to present information for enhanced memorization and learning [13, 23].

One possible approach to achieve good complexity and memorability is the use of *passphrases*, i.e. a sequence of related words and/or a phrase in natural language, which can be easier to remember than a standard password [19]. In addition, [9] introduced the concept of *password categories*, which groups passwords according to the account or service they protect. This idea could be further developed, by suggesting that not all password-protected services have the same requirements in terms of password complexity, i.e. less sensitive services require less complex passwords, and highly sensitive services require more complex passwords. A study which applies this concept to passphrases could lead to an effective method for educating users into creating complex enough, yet easy to remember passphrases for the different types of services they are trying to protect.

This work-in-progress study will lead to the development and assessment of a method for educating users into creating adequate passphrases, according to the type of information they want to protect. The method will define a series of account categories and passphrase levels that will be validated by subject matter experts (SMEs). The main research question this work-in-progress study will address is: What is the effect of the proposed method in a person's choices of memorable, yet complex enough passphrases, adequate for the type of information s/he is trying to protect?

In addition, this proposed work-in-progress study will address the following specific research questions:

RQ1: What are the elements identified by the SMEs composing the passphrase complexity and memorability constructs?

RQ2: What are the account categories and passphrase levels identified by the SMEs a user should consider when choosing a passphrase for a certain service, and the complexity thresholds associated to the proposed passphrase levels?

RQ3: What is the method validated by the SMEs for educating users into choosing an adequate passphrase for a certain service, according to the level of protection the user wants to achieve?

RQ4: What will be the effect of users experiencing the proposed method on the complexity of the passphrases they choose, when compared to a control group?

RQ5: What will be the effect of users experiencing the proposed method on the memorability of the passphrases they choose, when compared to a control group?

RQ6: What will be the effect of users experiencing the proposed method on choosing a passphrase level which is adequate to the account category they wish to protect?

RQ7: Will there be statistically significant mean differences for: (a) complexity of the passphrases, (b) memorability of the passphrases, and (c) matching of the passphrase levels chosen by users to the account categories, after controlling for their age, gender, or computer experience level?

II. Background

Password authentication is currently among the most popular authentication methods, because it is one of the easier methods to deploy [4] and users perceive it as familiar and easy to use [26]. However, password authentication is vulnerable to attacks which have more chance of being successful if the system user makes a weak password choice, such as only numbers, words that are popular as passwords and/or can be found in a dictionary, or the user's login name [2, 22]. To minimize such attacks, two approaches have been devised: one based upon coercion, and other based upon education. The coercion-based approach uses policies to force users into creating passwords with a certain level of complexity, by using mechanisms such as password meters and complexity validation [5]. However, this approach can be counterproductive, because it does not explain the added value or the security reasons for enforcing such policy to the end users; Furnell et al. [10] stated that when users know the reasons for implementing a security policy and receive help to comply with it, they exhibit a more secure behavior. Thus, the education-based approach has a better chance of closing the gap between the demands of security policies and the user's behavior [16].

Another factor leading to poor password choices is memorability. Users tend to choose passwords that are easy to remember [25]. However, when faced with a policy asking for a too long or complex password, the user is more prone to forget it, losing productive work time while the password is reset [18]. According to Miller [17], immediate memory impacts the amount of information a person can remember. Cognitive load theory [13] delves into this issue, asserting that humans first store information in their short-term memory. Learning occurs when the brain processes such information and commits it to long-term memory through practice and repetition; cognitive load measures the brain's effort to learn something. There are three different types of cognitive load: *Intrinsic*, related to processing information in the short-term memory; *extraneous*, caused by useless information that overloads the short-term memory; and *germane*, related to committing the information to long-term memory [7, 13, 23]. Learning can be more challenging if the complexity of the material increases the intrinsic load, and/or the extraneous load is increased by non-relevant or redundant material [23].

Several research studies have documented the influence of cognitive load in the information technology field. For instance, Ahmed et al. [1] concluded that user interfaces with lower cognitive loads led to faster data input with less errors. Jenkins et al. [14] concluded that lower extraneous cognitive loads in security training led to improved policy compliance by users. Mujeye et al. [18] found out that high password complexity imposes a high cognitive load upon user, thus increasing the possibility of forgetting their own password.

One way of improving memorability is to use passphrases instead of passwords. Nielsen et al. [19] stated that passphrases may contain several words that make sense to the user, or they may be a phrase in natural language. Several approaches for adopting passphrases have been researched, such as loose matching mechanisms that account for typing errors when logging into the system [19], or guiding users to choose words from a random set to create a highly memorable passphrase [3].

III. Methodology

Since this work-in-progress study is seeking to build a method to address the research problem, the most adequate research design for this case is a developmental methodology [8]. The research will include a combination of qualitative and quantitative methods [21], and will be executed in three phases, as shown in Fig 1.

The first phase will be qualitative, aiming to validate the passphrase complexity and memorability constructs and their elements (e.g. passphrase length and choice of characters for passphrase complexity), the account categories and passphrase levels, as well as the method for educating users, i.e. to answer research questions RQ1, RQ2 and RQ3. The participants for this first phase of the work-in-progress study will be a group of 15-20 SMEs in the field of cybersecurity. The Delphi technique [11, 21] will be used to elicit the experts' opinions, by means of a series of questionnaires that will be refined through several iterations, until an agreement is reached.

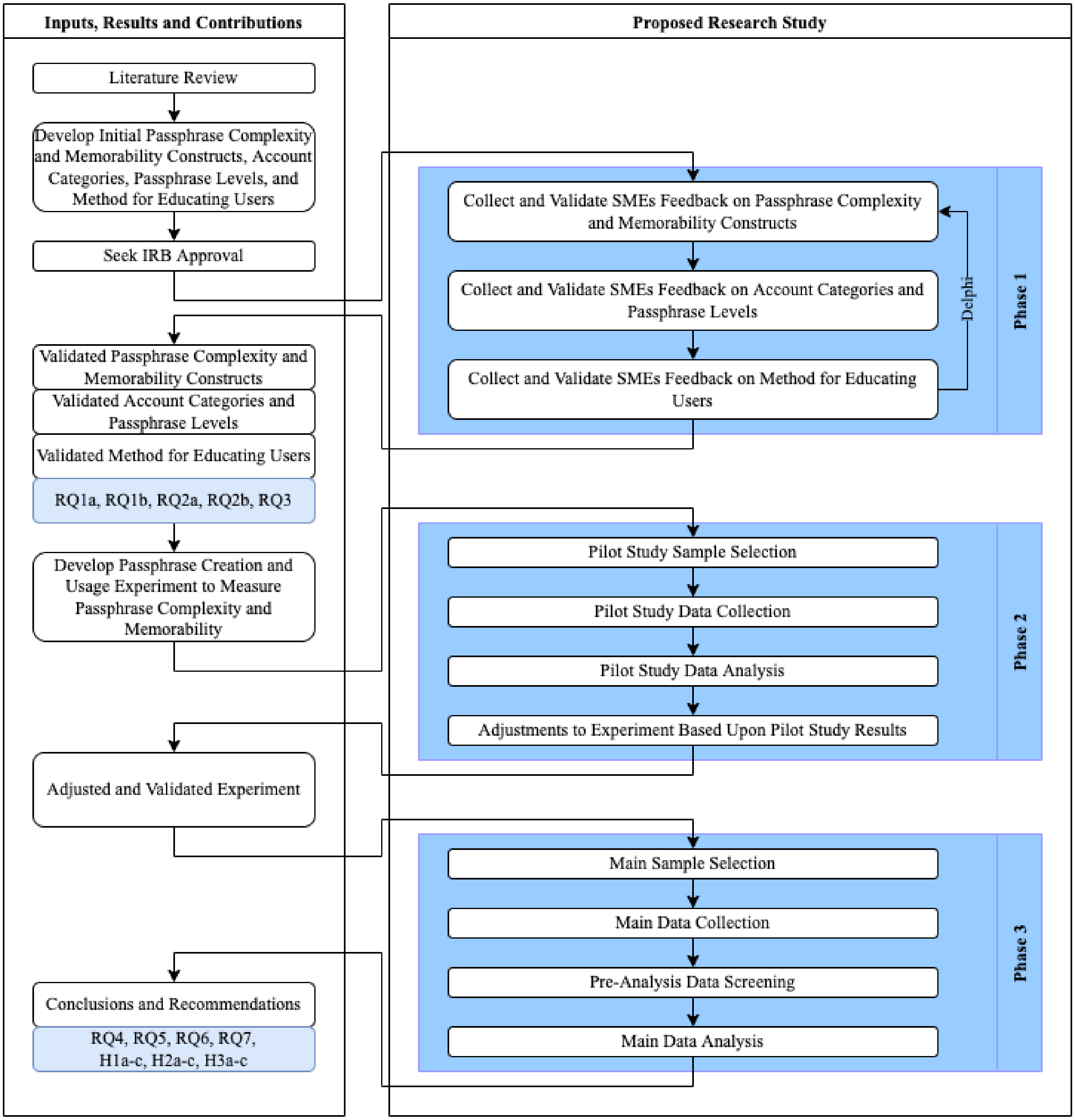
Inputs, Results and Contributions
Literature Review
Develop Initial Passphrase Complexity and Memorability Constructs, Account Categories, Passphrase Levels, and Method for Educating Users
Seek IRB Approval
Validated Passphrase Complexity and Memorability Constructs
Validated Account Categories and Passphrase Levels
Validated Method for Educating Users
RQ1a, RQ1b, RQ2a, RQ2b, RQ3
Develop Passphrase Creation and Usage Experiment to Measure Passphrase Complexity and Memorability
Adjusted and Validated Experiment
Conclusions and Recommendations
RQ4, RQ5, RQ6, RQ7, H1a-c, H2a-c, H3a-c
Proposed Research Study
Collect and Validate SMEs Feedback on Passphrase Complexity and Memorability Constructs
Collect and Validate SMEs Feedback on Account Categories and Passphrase Levels
Collect and Validate SMEs Feedback on Method for Educating Users
Delphi
Phase 1
Pilot Study Sample Selection
Pilot Study Data Collection
Pilot Study Data Analysis
Adjustments to Experiment Based Upon Pilot Study Results
Phase 2
Main Sample Selection
Main Data Collection
Pre-Analysis Data Screening
Main Data Analysis
Phase 3

Fig. 1. Research design process

The second phase of the work-in-progress study will be quantitative, experimental pilot study. It will serve to test the passphrase creation and usage experiment on a smaller group of participants, and will allow to perform adjustments before running the experiment with a larger group of participants.

The third phase of the work-in-progress study will be experimental and quantitative, aiming to answer research questions RQ4, RQ5, RQ6 and RQ7. This phase will feature a quasi-experimental design [15], which will allow to assess the effect of the proposed method on the complexity, memorability and level/category matching of passphrases participants choose, when compared to a control group. A sample of 100 users will be taken from the students, faculty and staff from a university in Cali (Colombia).

IV. Anticipated Results

The first phase of the work-in-progress study, which will be done with the help of SMEs, will yield three main results: (a) the validated passphrase complexity and memorability constructs, which will allow to design a first proposal of the account categories and the passphrase levels; (b) the validated account categories and passphrase levels which will be used in the method for educating users; and (c) the validated method for educating users, which will be tested in the third phase of the study.

Once the experiment to measure password complexity and memorability is designed, the second phase of the work-in-progress study will allow to adjust such experiment, according to the results of the pilot study.

The third phase of the work-in-progress study will allow to know the goodness of the method for educating users into choosing adequate passphrases for the account or service they want to protect; once this phase is finished, the method will be tested, validated and ready to implement.

V. Discussion and Conclusions

The issue of password memorability and complexity is still an open research problem. By considering the advantages of passphrases over passwords [19] and the concept of password categories [9], this work-in-progress study proposes to increase both complexity and memorability of passphrases chosen by users, by developing a method to educate users on how to choose a passphrase which is adequate to the category of the account or service they are aiming to protect. SMEs opinions will be gathered to validate the definition of the passphrase complexity and memorability constructs, as well as the account categories and passphrase levels to be used, and the method to educate users into choosing suitable passphrases for the categories of services being protected. The method will be further validated through a pilot study with a reduced number of participants, and a main study featuring at least 100 participants from a university. Data collected through these studies will be used to determine if the participants using the method are able to produce more complex and memorable passphrases matched to the category of the account or service being protected, when compared to a control group.

Future work includes producing the method's materials in a format that could be easily distributed to organizations interested in improving their employees' cybersecurity practices, and implementing the method in an organization, so that the effects of educating the employees could be measured in a longer time window.

References

[1] A. Ahmed, S. Chandra, V. Herasevich, O. Gajic and B. Pickering, "The effect of two different electronic health record user interfaces on intensive care provider task load, errors of cognition, and performance", Critical Care Medicine, vol. 39, no. 7, pp. 1626-1634, 2011. https://doi.org/10.1097/CCM.0b013e31821858a0

[2] M. Barkadehi, M. Nilashi, O. Ibrahim, A. Zakeri Fardi and S. Samad, "Authentication systems: A literature review and classification", *Telematics and Informatics*, vol. 35, no. 5, pp. 1491-1511, 2018. https://doi.org/10.1016/j.tele.2018.03.018

[3] N. Blanchard, C. Malaingre and T. Selker, "Improving security and usability of passphrases with guided word choice", *Proceedings of the 34th Annual Computer Security Applications Conference*, 2018. https://doi.org/10.1145/3274694.3274734

[4] J. Bonneau, C. Herley, P. Oorschot and F. Stajano, "The quest to replace passwords: A framework for comparative evaluation of web authentication schemes", *2012 IEEE Symposium on Security and Privacy*, 2012. https://doi.org/10.1109/SP.2012.44

[5] K. Chanda, "Password security: An analysis of password strengths and vulnerabilities", *International Journal of Computer Network and Information Security*, vol. 8, no. 7, pp. 23-30, 2016. https://doi.org/10.5815/ijcnis.2016.07.04

[6] Cyberedge Group 2022 Cyberthreat defense report. https://www.netwrix.com/2022_cyberthreat_defense_report.html

[7] K. DeLeeuw and R. Mayer, "A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load.", *Journal of Educational Psychology*, vol. 100, no. 1, pp. 223-234, 2008. https://doi.org/10.1037/0022-0663.100.1.223

[8] T.J. Ellis and Y. Levy, Y. (2009). "Towards a guide for novice researchers on research methodology: Review and proposed methods". *Issues in Informing Science & Information Technology*, vol. 6, 2009.

[9] S. Farrell, "Password policy purgatory", *IEEE Internet Computing*, vol. 12, no. 5, pp. 84-87, 2008. https://doi.org/10.1109/MIC.2008.108

[10] S. Furnell, W. Khern-am-nuai, R. Esmael, W. Yang and N. Li, "Enhancing security behaviour by supporting the user", *Computers & Security*, vol. 75, pp. 1-9, 2018. https://doi.org/10.1016/j.cose.2018.01.016

[11] T. Grisham, "The Delphi technique: a method for testing complex and multifaceted topics", *International Journal of Managing Projects in Business*, vol. 2, no. 1, pp. 112-130, 2009. https://doi.org/10.1108/17538370910930545

[12] C. Herley, P. van Oorschot and A. Patrick, "Passwords: If we're so smart, why are we still using them?", *Financial Cryptography and Data Security*, pp. 230-237, 2009. https://doi.org/10.1007/978-3-642-03549-4_14

[13] N. Hogg, "Measuring cognitive load", *Handbook of Research on Electronic Surveys and Measurements*, pp. 188-194, 2007. https://doi.org/10.4018/978-1-59140-792-8.ch020

[14] J. Jenkins, A. Durcikova and M. Burns, "Simplicity is bliss", *Journal of Organizational and End User Computing*, vol. 25, no. 3, pp. 52-66, 2013. https://doi.org/10.4018/joeuc.2013070104

[15] Y. Levy and T. J. Ellis. "A guide for novice researchers on experimental and quasi-experimental studies in information systems research." *Interdisciplinary Journal of Information, Knowledge, and Management*, vol. 6, 151 pp, 2011.

[16] A. A. Kaur and K. K. Mustafa, "A Critical appraisal on password based authentication", *International Journal of Computer Network and Information Security*, vol. 11, no. 1, pp. 47-61, 2019. https://doi.org/10.5815/ijcnis.2019.01.05

[17] G. Miller, "The magical number seven, plus or minus two: Some limits on our capacity for processing information.", *Psychological Review*, vol. 63, no. 2, pp. 81-97, 1956. https://doi.org/10.1037/h0043158

[18] S. Mujeye, Y. Levy, H. Mattord and W. Li, "Empirical results of an experimental study on the role of password strength and cognitive load on employee productivity", *Online Journal of Applied Knowledge Management*, vol. 4, no. 1, pp. 99-116, 2016. https://doi.org/10.36965/OJAKM.2016.4(1)99-116

[19] G. Nielsen, M. Vedel and C. Jensen, "Improving usability of passphrase authentication", *2014 Twelfth Annual International Conference on Privacy, Security and Trust*, 2014. https://doi.org/10.1109/PST.2014.6890939

[20] A. Ometov, S. Bezzateev, N. Mäkitalo, S. Andreev, T. Mikkonen and Y. Koucheryavy, "Multi-factor authentication: A survey", *Cryptography*, vol. 2, no. 1, p. 1, 2018. https://doi.org/10.3390/cryptography2010001

[21] U. Sekaran and R. Bougie, R. "Research methods for business: A skill-building approach" (7th ed.). Wiley, 2016.

[22] C. Shen, T. Yu, H. Xu, G. Yang and X. Guan, "User practice in password security: An empirical study of real-life passwords in the wild", *Computers &amp; Security*, vol. 61, pp. 130-141, 2016. https://doi.org/10.1016/j.cose.2016.05.007

[23] J. Sweller (1994). "Cognitive load theory, learning difficulty, and instructional design". *Learning and instruction*, vol. 4, no. 4, pp 295-312, 1994.

[24] I. Velásquez, A. Caro and A. Rodríguez, "Authentication schemes and methods: A systematic literature review", *Information and Software Technology*, vol. 94, pp. 30-37, 2018. https://doi.org/10.1016/j.infsof.2017.09.012

[25] N. Woods and M. Siponen, "Too many passwords? How understanding our memory can increase password memorability", *International Journal of Human-Computer Studies*, vol. 111, pp. 36-48, 2018. https://doi.org/10.1016/j.ijhcs.2017.11.002

[26] V. Zimmermann and N. Gerber, "The password is dead, long live the password: A laboratory study on user perceptions of authentication schemes", *International Journal of Human-Computer Studies*, vol. 133, pp. 26-44, 2020. https://doi.org/10.1016/j.ijhcs.2019.08.006