



**HAL**  
open science

## Know their Customers: An Empirical Study of Online Account Enumeration Attacks

Maël Maceiras, Kavous Salehzadeh Niksirat, Gaël Bernard, Benoît Garbinato, Mauro Cherubini, Mathias Humbert, Kévin Huguenin

► **To cite this version:**

Maël Maceiras, Kavous Salehzadeh Niksirat, Gaël Bernard, Benoît Garbinato, Mauro Cherubini, et al.. Know their Customers: An Empirical Study of Online Account Enumeration Attacks. ACM Transactions on the Web, 2024, 18 (3), pp.37:1-37:36. 10.1145/3664201 . hal-04562671

**HAL Id: hal-04562671**

**<https://hal.science/hal-04562671v1>**

Submitted on 12 May 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Know *their* Customers: An Empirical Study of Online Account Enumeration Attacks

MAËL MACEIRAS, University of Lausanne

KAVOUS SALEHZADEH NIKSIRAT, University of Lausanne & EPFL

GAËL BERNARD, EPFL

BENOÎT GARBINATO, University of Lausanne

MAURO CHERUBINI, University of Lausanne

MATHIAS HUMBERT, University of Lausanne

KÉVIN HUGUENIN, University of Lausanne

---

Internet users possess accounts on dozens of online services where they are often identified by one of their e-mail addresses. They often use the same address on multiple services and for communicating with their contacts. In this paper, we investigate attacks that enable an adversary (e.g., company, friend) to determine (stealthily or not) whether an individual, identified by their e-mail address, has an account on certain services (i.e., an *account enumeration attack*). Such attacks on *account privacy* have serious implications as information about one's accounts can be used to (1) profile them and (2) improve the effectiveness of phishing. We take a multifaceted approach and study these attacks through a combination of experiments (63 services), surveys (318 respondents), and focus groups (13 participants). We demonstrate the high vulnerability of popular services (93.7%) and the concerns of users about their account privacy, as well as their increased susceptibility to phishing e-mails that impersonate services on which they have an account. We also provide findings on the challenges in implementing countermeasures for service providers and on users' ideas for enhancing their account privacy. Finally, our interaction with national data protection authorities led to the inclusion of recommendations in their developers' guide.

CCS Concepts: • **Security and privacy** → *Social aspects of security and privacy*; **Web application security**; *Human and societal aspects of security and privacy*; • **Information systems** → *World Wide Web*;

Additional Key Words and Phrases: usable security and privacy, web privacy, account enumeration attacks, online accounts, phishing

## ACM Reference format:

Maël Maceiras, Kavous Salehzadeh Niksirat, Gaël Bernard, Benoît Garbinato, Mauro Cherubini, Mathias Humbert, and Kévin Huguenin. 2024. Know *their* Customers: An Empirical Study of Online Account Enumeration Attacks. *ACM Trans. Web* 1, 1, Article 1 (May 2024), 37 pages.

<https://doi.org/10.1145/3664201>

---

## 1 INTRODUCTION

Online accounts have become mainstream in the Web ecosystem. Such accounts facilitate users' recurring interactions with online services and enable services to link such interactions. For instance, users are required to input their personal data (name, address) only once—when creating an account—and can use it repeatedly (e.g., whenever they place an order). It also offers users a portal for tracking their history of interactions with services (e.g., previous orders). Such accounts are mandatory on social networks, which are based on the very notion of user profiles and where generated content (e.g., posts) is linked to a user. Overall, most services enable users to create accounts; these services cover a wide variety of domains, including dating, e-commerce, social networking, and streaming.

Some websites offer standard services (e.g., social networks) but specialize in certain categories of users or certain categories of content. For instance, a dating service could specialize in users with a specific sexual orientation or religion, and a video streaming service could specialize in adult content. The fact that a user has an account for a given service is personal information by nature, but the aforementioned specialization of services makes such information even more sensitive. For most services, online accounts are linked to an e-mail address—required to create the account. This e-mail address is often used as a username for logging in. In 2015, a survey revealed that the same e-mail address was linked to more than 90 online accounts on average and that this number was rapidly growing.<sup>1</sup> E-mail addresses are also used for communicating with contacts (this is their original purpose in fact). This means that one’s e-mail address is usually known to many.

This situation paves the way for privacy attacks, where curious entities try to determine whether there exists an account associated with a given e-mail address on a list of online services or, equivalently, where they try to determine from a list of e-mail addresses, which ones have an associated account on a given service. Such attacks are commonly referred to as *account (or username) enumeration attacks* [5, 16, 21, 34, 38].<sup>2</sup> In such attacks, the account owners may receive an e-mail notification indicating that someone might have tried to infer the existence of their account. Not sending such e-mails by service providers can make the adversaries keep their attacks *stealthy*.

Authoritarian governments may profile and repress their citizens based on their political, religious, or sexual orientation. In the context of surveillance capitalism, private companies may try to derive more accurate profiles of their customers by determining which online services they have accounts with. Moreover, such companies can try to determine which of their customers also have an account with their competitors to target those customers more aggressively or identify other potential customers to approach them. In the context of corporate espionage, companies may try to acquire information about other companies’ users and/or customers by enumerating their user account lists. Phishing may also motivate such profiling: by discovering which online services the owner of a given e-mail address has an account with, an attacker could impersonate one of those services and significantly augment its chances of having the owner click on a link contained in the e-mail (i.e., spear phishing) [11].

The incidents related to account enumeration attacks often occur discreetly; these attacks are more commonly individual-to-individual rather than publicly reported breaches, and attackers employ the attack stealthy where victims may not even be aware of the initial enumeration attempts. Therefore, there are few real-world examples of such attacks on mainstream media. One notable case comes from Azure Advanced Threat Protection (ATP), which observed enumeration and brute force attacks over 12 months.<sup>3</sup> Attackers leveraged NTLM (NT LAN Manager) or Kerberos authentication protocols to access servers to identify valid user accounts within an organization. This example highlights the severity and potential consequences of account enumeration attacks.

While account enumeration attacks have been known for more than a decade and documented in various online resources (e.g., [5, 16, 34, 38]), to the best of our knowledge, very few empirical studies on their effectiveness have been published so far (essentially [17]). We believe that such studies are needed, especially as new data protection laws (e.g., GDPR) were recently enacted in many countries and that they often imply that service providers should protect their websites against these attacks. Besides, no studies have been published on the users’ perceptions of such

---

<sup>1</sup>See [blog.dashlane.com](https://blog.dashlane.com), last visited: Feb. 2024.

<sup>2</sup>See also [owasp.org](https://owasp.org), last visited: Feb. 2024.

<sup>3</sup>See [techcommunity.microsoft.com](https://techcommunity.microsoft.com), last visited: Feb. 2024.

attacks and on what ideas users would recommend to help better protect their accounts. This paper investigates this urgent—yet overlooked—issue by addressing four central research questions:

- **RQ1.** To what extent can an adversary determine, stealthily or not, whether an account is associated with a given e-mail address on a given online service?
- **RQ2.** How would users perceive the sensitivity of their list of online accounts and how they react when receiving unsolicited account creation or password reset e-mails?
- **RQ3.** What design tweaks would users suggest to improve the effectiveness of existing counter-measures?

To answer these questions, we followed three complementary approaches to assess the responses to such privacy attacks both on the service side and on the user side. First, on the service side, we manually tested the vulnerability of 63 websites using three different attack vectors related to the (poor) design of the service’s website, namely *login*, *password reset*, and *account creation*, to assess the success and stealthiness of the attack (see [Section 4](#)). Second, on the user side, we conducted an online survey with  $N = 318$  respondents to understand users’ perceptions of account privacy and their reactions toward unsolicited e-mails<sup>4</sup> (see [Section 5](#)). Third, on the user side, we conducted a focus group session with  $N = 13$  Internet users to shed light on new ideas for privacy-enhancing technologies (see [Section 6](#)).

We made the following findings. On the service side, our experiments show that a tremendous fraction of popular online services (93.7% in total in our experiments) are still vulnerable to the most basic account enumeration attacks and more than a third of the tested services leaked information when undergoing login attacks. Furthermore, the login attack was stealthy for all online services, but one. As for password reset attacks, they were even more successful, as they resulted in more than half of the services leaking information. Almost all the services leaked information with the account creation attack.

Regarding user responses to such attacks, we first learned that the general population regularly receives password-reset and account-creation e-mails, some of which might be the symptom of the attack we describe here. Yet, the respondents would generally not suspect the attack under study should they receive such an e-mail. Besides, the respondents found the list of their accounts to be quite sensitive, but only a small fraction of them took (effective) measures to conceal it. Finally, they reported being twice more likely to click on a link in an e-mail impersonating a service for which they have an account, demonstrating that a hacker could leverage our attack as a first step in a (spear) phishing attempt. Lastly, during the focus group session, participants highlighted the sensitive contexts for attacks, considering cultural and situational factors. They also provided insights into their mitigation strategies and offered design suggestions to enhance account privacy while noting the importance of balancing security enhancements with user experience considerations.

Based on these findings, we contacted the data protection officers of the (vulnerable) tested services to responsibly inform them about the vulnerability and offer our help to fix it. Out of the 59 contacted services, five already provided answers (beyond automatic replies), and so far, one fixed the vulnerability. To raise awareness and maximize the impact of our research, we further contacted the data protection authorities of several countries where data protection laws make service providers responsible for protecting their websites against account enumeration attacks. Our interactions with one of them led to the inclusion of recommendations (about protection against account enumeration attacks) in their developers’ guide aimed at helping service providers make their websites compliant with current data protection laws.

---

<sup>4</sup>As studied in a recent work on user interaction with login notifications [27].

## 2 RELATED WORK

In this section, we first cover the closely related literature on account enumeration attacks (including attacks based on e-mail and phone number) and then the related work on single sign-on (SSO) services, which can be used as protective measures against such attacks.

### 2.1 Account Enumeration Attacks Using E-mails

To the best of our knowledge, the problem of account (or username) enumeration attacks was first mentioned back in 2007 [5, 34]. It is notably mentioned that such attacks can be easily performed against web applications by querying their login, password reset, or account creation forms and analyzing the returned message. Other blog posts and webpages provide recommendations on how to avoid such enumeration attacks [16, 38]. For instance, to prevent the login attack, Hacksplaining [16] recommends returning a generic message when a login failure occurs and to make sure the HTTP response and the time taken to respond are not significantly different whether an account exists. They provide similar simple recommendations to prevent password reset and account creation attacks. Stuttard [38] mentions that, besides fixing obvious leakage like returned messages, it is crucial to check every aspect of a service's behavior, such as timing differences when existing and non-existing usernames are entered. He further provides recommendations, such as defining application-specific usernames, on how to prevent enumeration attacks through the account creation form, which is the most difficult to counter.

In terms of scientific publications, Bortz and Boneh [5] show that response times of websites can reveal private information, such as the validity of a *username* on a web login page or the number of private photos on photo-sharing websites. They demonstrate that such a user account validity timing attack is effective by experimenting with it against two popular, high-traffic websites. The authors discuss countermeasures such as controlling the time taken to respond to any request, either through careful server-side coding or through a web server module that automatically regulates the time at which responses are sent. Balduzzi et al. [3] show that an attacker can query popular social networks for registered e-mail addresses to automatically gather private information about their users. Starting with a list of about 10.4M e-mail addresses, they are able to automatically identify more than 1.2M user profiles associated with these addresses.

A major consequence of account enumeration attacks is the rise of spear-phishing e-mails, a prominent cybersecurity concern. Several recent studies have addressed this issue. Distler [9] examine how contextual factors influence responses to spear-phishing attempts, showing that task alignment, time pressure, and social context can alter users' susceptibility to phishing attacks. Marin et al. [26] examine how people's attitudes affect whether they report phishing e-mails. They find that self-efficacy, subjective norms, and altruism positively influence reporting intentions, while sportsmanship inhibits reporting. Tally et al. [40] focus on anti-phishing awareness among mid-career office workers, highlighting the importance of informal sources such as news, podcasts, and social media in increasing workers' knowledge and perceptions of phishing threats. All of these studies highlight the need to consider the multifaceted aspects of user behavior, contextual factors, and psychological factors when addressing the consequences of account enumeration attacks.

Finally, closest to our work, Hasegawa et al. [1, 17] studied login, password reset, and account creation attacks on a various set of online services. These services were selected based on their sensitivity (including both sensitive and popular platforms) and identified from the data collected through a user survey. They also collected and analyzed data about users' perceptions of the threat. The main differences between our work and Hasegawa et al.'s work are the stealthiness of the attack, the use of single-sign-on, the perceptions studied through the survey, and the legal

aspects. Additionally, we conducted a focus group session to assess users' ideas about future privacy-enhancing technologies against account enumeration attacks.

## 2.2 Account Enumeration Attacks Using Phone Numbers

Phone numbers are becoming an alternative vector for account enumeration attacks. Mobile applications and various online services often use phone numbers as primary identifiers, which can be just as vulnerable to enumeration attacks as e-mail addresses.

Kim et al. [21] propose a new phone number enumeration attack to automatically identify Facebook users' private data. In particular, they show that one can leverage Facebook's search option and enumerate the entire range of phone numbers to collect users' information such as location, birthdate, or phone numbers. They notably manage to collect more than 25K Facebook profiles from 214K Californian phone numbers. This attack is tailored to Facebook's search directory (with phone numbers) that is not applicable to most online services. McDonald et al. [30] study the security and privacy risks of phone recycling and phone numbers being used as identifiers by online services and mobile applications. The authors conducted a qualitative user study ( $N = 195$ ) to better understand the negative consequences faced by users due to phone number use, e.g., as identifiers. The participants elicited problems caused by phone number recycling, unwanted exposure, and temporary or permanent loss of access to their phone numbers. The authors argue that online services should stop requiring users to connect a phone number to their account whenever possible and use an e-mail address or username only.

The aforementioned body of research sheds light on the dynamic nature of online security threats and emphasizes the importance of understanding different attack vectors and their respective countermeasures. In this paper, our investigation will focus only on e-mail identifiers, enabling us to maintain a targeted and manageable scope within our experimental design. E-mail-based attacks have unique dynamics, and the phone number domain introduces unique technical considerations<sup>5</sup> requiring different protective strategies for each vector. Future research must explore the challenges and potential strategies associated with phone number attacks. Despite this focus, [Section 7.2](#) will address the applicability of our discussed countermeasures to phone-number-based attacks, evaluating their practicality and effectiveness in this context.

## 2.3 Single Sign-On (SSO) Services

Wang et al. [42] were the first to study the security of major SSO services. By capturing the flow of web traffic going through the browser, they discovered eight security-critical logic flaws in several identity providers. All discovered flaws enabled an adversary to sign in as the original user. Wang et al. [42] responsibly disclosed these vulnerabilities to the affected services that subsequently fixed them. Ghasemisharif et al. [15] investigated account management flaws in SSO deployments and proposed a tool to automate their detection. In particular, they focused on the flaws that stem

---

<sup>5</sup>E-mail- and phone-number-based account enumeration attacks differ significantly in their use and the effort required to manage them. E-mail addresses are relatively private and easily managed; users can effortlessly create a new e-mail or delete an old one for free, allowing them to frequently change their online identifiers. On the other hand, phone numbers are typically linked to a SIM card rather than a specific mobile device. Additionally, phone numbers are often reassigned or reused when users cancel their subscriptions. Despite these factors, phone numbers still provide a more consistent identifier across various services, making them more fixed and traceable compared to e-mail addresses. While e-mail addresses offer relatively easy management, phone numbers require more effort to change, often involving the purchase of a new SIM card. Lastly, phone numbers are crucial for security features like SMS-based two-factor authentication, making them targets for specific attacks such as SIM swapping, where attackers convince mobile providers to transfer a victim's phone number to a new SIM card under their control. This type of attack is specific to phone numbers and highlights the unique risks they pose compared to e-mail addresses.

from the concurrent use of “regular” credentials (e-mail/password) and identities provided by SSO, for instance, GMail e-mail address and Google SSO. Recently, Dimova et al. [8] investigated the privacy implications of OAuth authentication (i.e., a standard for federated SSO) on the Web. By evaluating a significant number of OAuth-based logins, they show that identity providers (IdPs) provide websites with various data resources (scopes), many of which are unnecessary for user identification. Interestingly, they find that certain websites offer alternative login methods that require less user information and that revoking access to non-essential information often does not affect the website’s functionality. Lastly, Morkonda et al. [33] developed the SSOPrivateEye extension to improve privacy in SSO services, offering users insights into SSO providers’ permission requests.

Several studies have investigated the acceptance of SSO and how users perceive it. For instance, Sun et al. [39] explored user perceptions and concerns regarding Web SSO through various user studies. Their research revealed that many participants had misconceptions about SSO, believing that it shared their identity provider login credentials with relying parties. Furthermore, concerns about personal data exposure and the functions relying parties could perform with their identity provider accounts led to hesitancy in adopting SSO. Similarly, Bauer et al. [4] studied the perceptions and willingness to use SSO services through an online survey ( $N = 424$ ). Their study notably shows that users do not understand what personal information is shared by the identity providers (such as Google or Facebook) with the service providers. Moreover, both self-reported data and users’ actions indicate a need for better insight and control over the shared data, which would lead to greater adoption of SSO. Egelman [12] studied the trade-off between privacy and convenience, specifically in Facebook Connect, through a controlled experiment ( $N = 65$ ). The study shows that 15% of the participants refuse to use Facebook Connect to authenticate to online services. It further reveals that the vast majority of participants (88%) understand the type of Facebook profile data that is shared with online services.

Through an online user survey ( $N = 364$ ), Cho et al. [7] studied the willingness to rely on SSO services for privacy-sensitive applications. They tested four applications, from a low-sensitivity “class reunion” app to a high-sensitivity “affair” app. The study shows that users tended to choose the SSO service (Facebook in that case) when they sign up for a low-sensitivity app due to ease of sharing and lower fear of their security being compromised. Moreover, users prefer not to use Facebook to log into the “affair” app. Interestingly, users with strong security concerns tend to prefer using e-mail for logging into highly sensitive apps. Lastly, Morkonda et al. [32] investigated the impact of displaying permission-related information on Web SSO login decisions to understand the factors influencing users’ choices among different login options. After a user study ( $N = 200$ ), they found that usability preferences and familiarity are the primary drivers of login decisions. Still, when participants were given permission-related information, many shifted to more privacy-conscious login options.

### 3 THREAT MODEL

We consider an adversary whose objective is to determine whether a target user has an account on a specific online service.<sup>6</sup> Several adversaries would be interested—for various reasons—in knowing this information. For instance, authoritarian governments could profile their citizens and repress a subset based on their political and sexual orientation (revealed by the fact that they possess accounts on specific services associated with opposition parties or with the LGBTQ+ community). Even in democratic countries, law enforcement could determine the list of online accounts of potential suspects and further request access (e.g., through a subpoena) to the data of these accounts. In fact,

<sup>6</sup>Note that a user could have an account on a service but not use it.

our contact in law enforcement confirmed that this is a standard technique in the open-source intelligence (OSINT) toolbox. Intimate partners could target online dating services to determine if their partner might be looking for other partners.<sup>7</sup> Such surveillance could in fact come from any individuals (i.e., stalkers) who obsessively track their targets. Companies could profile their customers or identify other potential customers, including those of their competitors, in order to approach them. They could also identify the customers of their competitors for economic espionage purposes. Overall, the list of online services for which a user has an account is quite telling and thus sensitive from a privacy perspective. Finally, hackers could use this information to improve the effectiveness of their (spear) phishing campaigns by impersonating services for which the targeted user has an account.<sup>8</sup>

The adversary is assumed to know only an e-mail address of the target user and we therefore focus on online services where users are identified by their e-mail addresses.<sup>9</sup> This *a priori* knowledge of e-mail addresses might be due to the fact that the adversary knows the target (e.g., colleague, partner, customer) and possibly interacts with them via e-mail, or because it obtained a list of e-mail addresses from a private directory (e.g., that of its contacts/customers), from a public directory (e.g., that of an organization), or from a leaked dataset (e.g., Microsoft’s 2020 leak).<sup>10</sup> We consider that the adversary conducts the attack by relying on the different features offered on the service’s website. When doing so, the adversary might want its attack to go undetected (i.e., stealth) with respect to the target user, following the well-established covert adversary model [2]. For instance, a covert adversary would be reluctant to take action that would lead to the sending of a notification (e.g., e-mail) to the targeted users, thus causing those users to suspect that they are under surveillance.

## 4 SERVICE BEHAVIOR

We evaluate the extent to which an adversary can infer the existence of an account:

**RQ1.** *To what extent can an adversary determine, stealthily or not, whether an account is associated with a given e-mail address on a given online service?*

We designed and conducted an experiment targeted at online services. In a nutshell, we proceeded as follows. We considered three different basic attack vectors that exploit common (poor) design elements of the websites, built a dataset of 63 online services, and tested the attacks on the selected services (1) for an e-mail address for which an account existed and (2) for an address for which no account existed. Because the procedures for making the different steps of our experiment were not homogeneous across the different websites (e.g., account creation forms differ from one website to another), we had to perform them manually.<sup>11</sup> Even though the process was manual, it followed a strict and systematic procedure making it rigorous and reproducible.

### 4.1 Attack Vectors

We considered the following basic attack vectors (see Figure 1 for illustrations of the attack). Note that ad-hoc attacks exist for specific websites, such as the recent attack against Duolingo. Here, we focus on attacks that exploit common features of websites.<sup>12</sup>

<sup>7</sup>Intimate partner online surveillance has recently received attention [6, 14, 18, 24, 41].

<sup>8</sup>It was shown that susceptibility to phishing is higher when the sender is known to the recipient [31]. Our survey (Section 5) also shows that users are more likely to click on a link in an e-mail seemingly sent by service on which they have an account.

<sup>9</sup>Note that previous works investigated the case where users are identified by their phone numbers (e.g., Facebook) [21, 30].

<sup>10</sup>See [informationisbeautiful.net](https://informationisbeautiful.net), last visited: Feb. 2024.

<sup>11</sup>Automating the attack for large numbers of accounts on a given service, however, is doable, as shown by our proof-of-concept implementation based on Selenium ([selenium.dev](https://selenium.dev)). We discuss attack automation in Section 7.1.

<sup>12</sup>The personal data of 2.6M Duolingo users—their e-mail addresses, usernames, and actual names—were compromised and became available for scraping on BreachForums. See <https://cybernews.com/security>, last visited: Feb. 2024.





- (a) Failed login attack. The attack is stealthy as no e-mail is sent.
- (b) Successful login attack for an existing account. The attack is not stealthy as an e-mail is sent.
- (c) Failed password reset attack. The attack is not stealthy if an account exists as an e-mail is sent.
- (d) Successful account creation attack for an existing account. The attack succeeds early (before the form is submitted) and thus stealthy as no e-mail is sent.

Fig. 1. Synthetic illustrations of the considered attack in four sample designs/scenarios (not exhaustive).

- **Login (L)** The adversary makes a single attempt to log into the online service with the targeted e-mail address and a random password. As the login fails, the adversary tries to infer the existence of an account from the (error) message returned by the service (e.g., “invalid e-mail address” vs. “invalid password”; see Figure 1b). Should the returned message be the same, whether an account associated with the e-mail address exists or not (e.g., “invalid e-mail address and/or password” in both cases), the attack fails (see Figure 1a).
- **Password reset (P)** The adversary makes a password-reset request (a special case of fallback authentication [23]) to the online service for the targeted e-mail address. Note that such a feature exists for most websites; in fact, it existed on all the online services tested in the experiment. Again, the adversary tries to infer the existence of an account from the message returned by the service (e.g., “no account associated with this e-mail address” vs. “a password-reset link has been sent to the e-mail address”). Should the returned message be the same, whether an account associated with the e-mail address exists or not (e.g., “if an account associated with the e-mail address exists, a password-reset link has been sent to the e-mail address” in both cases), the attack fails (see Figure 1c).
- **Account creation (C)** The adversary makes an attempt to create a new account on the online service with the targeted e-mail address. Again, it tries to infer the existence of an account from the returned message (“an account has been successfully created” vs. “an account associated with this e-mail already exists”; see Figure 1d). If the same message is returned in both cases (e.g., “if no account associated with the e-mail address already exists, the account has been created successfully”), the attack fails.

Note that these attacks could trigger notifications (e.g., e-mails) to the owner of the e-mail address. For instance, in the password reset attack, the owner of the e-mail address—should an account associated with the address exist—is likely to receive an e-mail with a password-reset link, thus compromising the *stealthiness* of the attack. Similarly, in the account creation attack, the owner of the e-mail address—should *no* account associated with the address exist—is likely to receive an e-mail confirming the creation of the account or asking to confirm ownership of the e-mail address.

However, it should also be noted that messages leaking information about the existence of an account might be displayed *before* the adversary completes the attack procedure. For instance, a message indicating the existence of an account might be displayed as soon as the adversary inputs

the targeted e-mail address in the account creation Web form (see Figure 1d), before the adversary submits the form. In this case, the attack can thus be conducted stealthily.

We decided to focus on these basic and relatively well-known attacks in order to obtain a first (under)estimation of the extent of the threat. As part of future work, more advanced attacks such as those based on timings [5] (e.g., when the response time of a website upon submission of the login form is faster when there is no account associated with the submitted e-mail address)<sup>13</sup> and those based on user lookup features (e.g., when sharing a group on Zotero or a project on Overleaf) could be considered. Finally, we focus only on the message displayed on the websites; in particular, we do not take into account the raw HTTP responses to asynchronous fetch requests as mentioned in previous work [5, 16]. Such advanced attacks are discussed in Section 2. As described in Section 4.6, a tremendous proportion of popular online services are already vulnerable to at least one of those basic attacks.

## 4.2 Service Selection

In order to build a sample of online services on which to test the aforementioned attack, we relied on the similarweb,<sup>14</sup> a website analytics service.<sup>15</sup> Similarweb offers various statistics for websites, per country and per category (24 categories in total, including E-commerce and Shopping, Finance, and Adults). In each category, we focused on the most popular websites—with respect to their numbers of unique monthly visitors in Switzerland for the period Apr.-Jun. 2021. Even though popular online services are used by a large fraction of Internet users, the fact that one has an account on such a service is still informative and potentially privacy-sensitive (e.g., streaming service for adult content). Also, it is particularly relevant for phishing attacks, as these popular services affect a large fraction of Internet users. As the analysis of the selected websites is manual, we limited the selection to 63 services.

We selected approximately three services per category. We excluded websites from similarweb’s ranking based on the following criteria: (1) the website does not offer the possibility to create an account or the creation is restricted or not linked to an e-mail address (e.g., tax services), (2) accounts on the website are tied to an e-mail provider (e.g., YouTube: users can only connect with a Google account and any user with a GMail e-mail address has, *de facto*, a YouTube account), (3) the website is redundant with a previously selected website (e.g., Amazon.com and Amazon.ca). In each category, whenever a website was excluded, we considered the next one in the ranking as a replacement. Our final sample was diverse, including services related to audio and video streaming (incl. adult content), classified ads (misc, real estate, lodging), dating, e-commerce (misc, clothes, DIY, food, furniture, high-tech), finance, gambling, gaming, health, jobs, delivery, news, social network, and transportation. The complete list of services is available in Table 1 in Appendix A.

## 4.3 Procedure

Because forms for creating accounts, resetting passwords, and logging in differ substantially across services, we conducted the analysis manually. We created a total of six (corresponding to the 3×2 conditions) *fresh* e-mail addresses for the sake of the experiment: two addresses per attack vector, one for which an account exists (e.g., login\_account@mail.com, pwdreset\_account@mail.com, etc.) and one for which no account exists (e.g., login\_noaccount@mail.com, etc.). For three of the six

---

<sup>13</sup>In fact, we tested Bortz and Boneh [5]’s attack against one of the services for which the other attacks did not succeed. We report on this additional experiment at the end of Section 4.6.

<sup>14</sup>See [similarweb.com](https://www.similarweb.com), last visited: Feb. 2024.

<sup>15</sup>Because of this choice, we might have missed popular services that are accessed through channels other than their websites, such as mobile apps.

e-mail addresses, we created an account on each of the selected online services. We reused the same six e-mail addresses for the different services. We used the same set-up (operating system, browser, language, *etc.*) for all e-mail addresses and attack vectors, but we used different (private) browser sessions and different IP addresses when creating the three accounts on the same online service, so as to circumvent the limitations imposed by some of the services (e.g., one account creation per day from a given IP).<sup>16</sup> In order to assess the stealthiness of the different attacks, we monitored the mailbox of the e-mail addresses used (in the case where the service sends notifications by e-mail) for several days after conducting the attack.

For each service, we conducted the attacks iteratively: In case of success, we interrupted the attack immediately. For instance, for the account creation attack, if a message confirming the (non-)existence of an account was displayed *before* submitting the form, we did not pursue and marked the attack as successful. If no message was sent after a couple of days, the attack was marked as stealth. We also marked whether the attack required human intervention (e.g., CAPTCHA). The experiment took place in 2021-2022. We recorded the computer screen while conducting the attacks. The videos will be made available with the final version of the paper.

#### 4.4 Ethics

We strove to ensure high ethical standards [10] for our research. This study was approved by our institution’s ethics committee. To minimize the harm to individuals, we conducted our attacks on e-mail addresses created for the sake of the experiment.

The raw results of our study are presented in [Table 1](#) (see [Appendix A](#)), with a detailed explanation provided in [Section 4.6](#). We emphasize our awareness of the ethical implications associated with disclosing the names of service providers in light of the vulnerabilities identified. To ensure ethical integrity, we strictly adhered to established standards and guidelines for responsible disclosure,<sup>17</sup> which included notifying the affected service providers well in advance and engaging in transparent dialogue about the vulnerabilities (for details, see [Section 4.7](#)). We initially provided these service providers with 30 days to respond to our findings before proceeding with further actions, such as contacting national data protection agencies. Our decision to publish these findings more than a year after the initial notification aligns with guidelines that advocate for public disclosure when vulnerabilities still need to be addressed. This practice of naming entities is commonly adopted in the security field, and it enhances accountability and improves security practices. Considering the widespread nature of these vulnerabilities and their frequent neglect by companies, we deemed the full disclosure of service provider names in [Table 1](#) ethically justifiable and necessary.

#### 4.5 Experiment Limitations

One limitation of our study is that we tested “only” 63 online services. The first reason for that is the complexity of automating the attack vectors *across* online services (as discussed above), since each service has different UI elements of login, account creation and password reset forms. For instance, while tools such as PVACreator<sup>18</sup> can automate the creation of hundreds of online accounts, they can do so for only a few dozens of online services for which specific plug-ins have been manually developed (including Amazon, Facebook, Netflix, and PornHub). The second reason, even more challenging, lies in the creation of accounts on all the online services under study. Indeed, going

---

<sup>16</sup>The fact that we reused the same IP addresses, for different services, might have affected the presence of CAPTCHAs as the underlying mechanisms sometimes rely on IP reputation, which is updated based on network activity (including browsing activity).

<sup>17</sup>See [medium.com/@ptcrews](https://medium.com/@ptcrews), [kaspersky.com/blog](https://kaspersky.com/blog), or [enisa.europa.eu/publications](https://enisa.europa.eu/publications), last visited: Feb. 2024

<sup>18</sup>See [pvacreator.com](https://pvacreator.com), last visited: Feb. 2024.

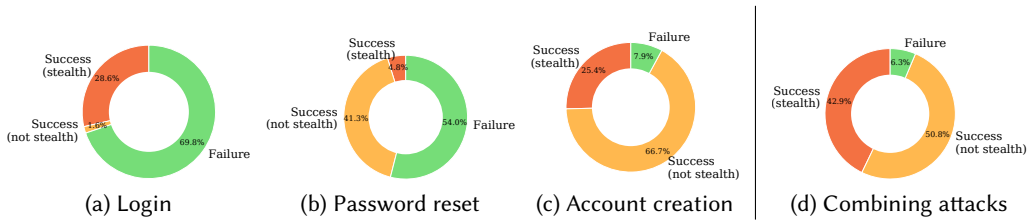


Fig. 2. Experimental results of the considered attacks. We do not specify whether the cases of failure are stealth or not as the adversary would not conduct the attack in such cases (the adversary would test the attack on e-mail addresses they control beforehand, as we did in the experiment).

through the whole process of account creation often requires to fill different fields in multiple webpages and to click on a link received by e-mail. This makes automating the whole process quite challenging (note that this is required for the experimenter, but not for the attacker). Another limitation of our study includes the fact that the adversary has to have an *a priori* knowledge of e-mail addresses to be tested and that the existence of an account does not mean that the latter is active.

#### 4.6 Results

We first look at the success rate of the considered attack vectors *taken individually* and then discuss the success rate of combining those three attack vectors. The results are summarized in Figure 2. Finally, we discuss features offered by the online services (i.e., single-sign-on and account deletion) that help counter these attacks.

**The login attack** revealed successful for 19 services (30.2%), including an adult video streaming service (i.e., Xhamster) and a religion-based dating service (i.e., InshAllah). None of the tested services for which the attack succeeded included a CAPTCHA<sup>19</sup> in the login procedure (this would have limited the automation of the attack by requiring human intervention). In fact, only 2 included a CAPTCHA for logging in. This could be explained by the fact that logging in is a frequent task and including a CAPTCHA would severely hurt its usability. None of the services displayed a message leaking the (non-)existence of an account before submitting the login form. Yet, the attack was stealth for *all* services but one. For this one service, when an account existed, an e-mail with a one-time link for logging in was sent. This was indicated on the webpage: “Your password is wrong. You have received an e-mail to access your account in a click. You can also try to login again”. For the other services, no e-mail was sent after the (failed) login attempt, regardless of whether an account existed.<sup>20</sup>

**The password reset attack** revealed more successful than the login attack, i.e., it was successful for 29 services (46.0%), including an adult video streaming service (i.e., Xhamster) and a dating service (i.e., Badoo). Only 4 of the services for which the attack succeeded included a CAPTCHA in the procedure (13.8%). Although the attack was rather successful, it was rarely stealth. It was successful and stealth for only 3 services (4.8%), including Amazon and Facebook. This was expected as the traditional action for a password reset request is to send an e-mail (with a link for resetting

<sup>19</sup>Note that even though the service does use CAPTCHA, it might not systematically include a puzzle upon registration. Indeed, services like reCAPTCHA do not systematically include puzzles: they rely on several factors (e.g., IP, activity) to make this decision.

<sup>20</sup>Note that some services send a notification upon a successful, yet unusual (e.g., from an unusual location), log in.

the password) to the requested e-mail address *if (and only if) an account exists*.<sup>21</sup> However, as the adversary does not know whether there exists an account associated with the targeted e-mail address (it is actually precisely what the adversary is trying to infer), mounting this attack could reveal to the e-mail owner that some curious or malicious entity is targeting their account. For the 3 services for which the attack was successful and completely stealth (i.e., stealth whether an account existed or not), the information was leaked before the form was submitted: typically, the procedure was in two steps (submitting the e-mail address first and then confirming the request) and the (non-)existence of the account was confirmed in the first step. This was the case for a major online social network.

None of the targeted services sent an e-mail when no account existed. Yet, we could personally witness this behavior outside of this experiment. For instance, Zoom notifies non-users of password reset attempts targeted at their e-mail address without leaking the non-existence of the account on the website: “We sent a reset password e-mail to XXX. Please click the reset password link to set your new password.” (message on website); “Hello XXX, You tried to reset your password but there is no account associated with this e-mail address. If you want to sign up a Zoom Account, please click the button below to sign up.” (e-mail). Such notifications warn users about potential attacks against them.

**The account creation attack** was the most successful, i.e., it was successful for 58 services (92.1%). Naturally, account creation failed (and the adversary was notified) in the vast majority of the cases if an account associated with the targeted e-mail address already existed. The few services that did not behave this way used usernames instead of e-mail addresses as identifiers. And the creation of an account with an address already associated with an account simply resulted in the creation of a new account associated with the same e-mail address but with a different username. This was the case for three services, one related to a religious community (i.e., JW, a website targeted at Jehovahs’ witnesses), one related to a lifestyle magazine, and one related to an encyclopedia (i.e., Wikipedia). Note that it could be argued that, on such services, users are not *really identified* with their e-mail address and that, as such, these services are outside of the scope of our system model (see [Section 3](#)). Interestingly, for the lifestyle magazine, the e-mail sent in response to a password-reset request contains one reset link *for each* of the usernames associated with the submitted e-mail address.

While more successful, the account creation attack was less prone to automation than the other techniques as CAPTCHAs were relatively often included in the procedure (for 28.6% of the services). Taking this parameter into account, the attack was successful for 65.1% of the services, with no need for human intervention. Similar to the password reset attack, the account creation attack was not stealth in most cases, as the traditional action for an account creation is to send a confirmation e-mail, either welcoming the user to the service or asking them to confirm their e-mail address and the account creation. Note that, in certain cases, an e-mail indicating the creation of an account (confirmation or newsletter) was sent only a few days after the creation of the account. Unlike with password reset, however, with account creation the attack was exposed *if (and only if)*<sup>22</sup> the targeted user *does not* already have an account on the considered service. The attack was successful and stealth for 16 services (25.4%), including an adult video streaming service, a gambling service, and a health-related forum. For most of these services, a message was displayed *before the submission of the account creation form* if the e-mail indicated by the adversary on the form was associated with an existing account. This was usually implemented through an asynchronous fetch request to

---

<sup>21</sup>In our experiment, no e-mail was sent for a password reset attempt targeted at an e-mail address for which no account existed.

<sup>22</sup>In our experiment, no e-mail was sent for an account creation attempt targeted at an e-mail address for which an account already existed.

a specific endpoint of the service's backend (e.g., "https://login.service.com/validate", which yields: {"field": "email", "valid": false, "code": "error.register.email.exists", "message": "This e-mail address has already been used. Would you like to <a href="/login">log in</a> or <a href="/passwordreset">reset your password</a>?"}). Such a method opens the door to large-scale automated attacks, as we show in [Section 7.1](#). The other services for which the attack was stealth simply did not send any e-mail confirming the creation of the account.

As with the password reset attack, none of the targeted services sent an e-mail when an account already existed, but we could personally witness this behavior outside of this experiment: "Dear LEGO® user, This e-mail is just to inform you that someone tried to register a LEGO® Account using your e-mail. ". Yet, unlike for Zoom, the non-existence of the account was, unfortunately, leaked on the website.

**Combining attacks** could enable the adversary to increase their chances of success as well as their chances of succeeding *stealthily*. In our sample, out of the 5 services for which the account creation attack failed, only one was not also protected against the two other attacks. Therefore, combining the attacks modestly increased the chances of success. When considering stealthiness requirements, however, the chances of success were much higher when combining attacks. Indeed, the number of services for which *at least* one of the three attacks succeeded stealthily was much higher than for each individual attack, reaching a total of 27 services (42.9%). Although the modest number of considered services does not allow us to conduct a proper statistical analysis of the link between security practices and the category/sensitivity of the online services, it is worth noting that among the 4 services that are vulnerable to none of the three attacks, one is a religion-related website targeted at Jehovah's witnesses (i.e., <https://www.jw.org>), which can be considered sensitive, and one is Wikipedia's website, which can also be considered sensitive in certain regions of the world.

**The timing attack (advanced)** [5] can be used for the websites that are not vulnerable to the attacks mentioned above. As it provides a probabilistic output (a probability that an account exists), deterministic techniques such as the aforementioned attacks should be favored. We selected one such website (a lifestyle magazine) and ran a timing attack on the webpage that enables users to retrieve their (forgotten) username from their e-mail address. We used a total of 6 different e-mail addresses (3 for which an account existed). We submitted 500 requests (with an average delay of 0.5 seconds between requests) with an automatic script<sup>23</sup> based on Selenium<sup>24</sup> and measured the response time (i.e., the delay between requestStart and responseStart). We did not observe any significant difference between the mean response times of requests for existing accounts and for non-existing ones: In short, the attack did not succeed.

**Protective measures** against the considered attack can be taken by users. We present two of them and assess their feasibility on the considered online services.

We first looked at the possibility to register/log in to the considered service with a third-party service (a.k.a. SSO) such as Facebook and Google. Out of the 63 services, 21 services (33.3%) offered this feature. This feature is certainly convenient and potentially defeats the attacks presented in this paper, but it raises other security and privacy issues [12, 42].<sup>25</sup> Note also that SSO does not always defeat the considered attack. For instance, when SSO is implemented by an e-mail provider (e.g., Gmail), for some services, if an account is created through SSO with the e-mail provider, it is no longer possible to create an account with the associated e-mail address and the account creation attack, at least, succeeds. In fact, it was the case for all the services we tested (and that offered the option to use Google's SSO, 18 in total): We registered an account using Google's SSO and tried

---

<sup>23</sup>The source code is available on an OSF repository: <https://doi.org/10.17605/osf.io/5equw>.

<sup>24</sup>See [selenium.dev](https://selenium.dev), last visited: Feb. 2024.

<sup>25</sup>Users' willingness to adopt this feature has been studied in previous work (e.g., [4, 7]).

to create an account with the corresponding (GMail) address: In all cases, the account creation failed. Note that the effectiveness of SSO at protecting individuals and services against account enumeration attacks depends on the implementation of the account management by the service provider, as investigated by Ghasemisharif et al. [15]. When SSO identity providers do not use e-mail addresses as user identifier or when services handle accounts created with e-mail/password and SSO separately, the use of SSO should protect against the considered attacks.

We then looked at the possibility to delete one's account on the considered services. Such a feature is useful to conceal the existence of one's account that is no longer used. Note that this is a right of individuals enshrined in several data protection laws, including (implicitly) under the General Data Protection Regulation of the European Union (EU GDPR, Article 17, "Right to erasure ('right to be forgotten')").<sup>26</sup> Only one service (Wikipedia) explicitly stated that account deletion was impossible. Yet, it was possible to change all account information, including the e-mail address, and to delete all the articles contributed by the user. Among the remaining services, 42 (66.7%) offered the possibility to delete the account through an online form. For the services that did not (20 in total), we sent a message to the e-mail address indicated in the privacy policy asking them to delete the account.<sup>27</sup> So far, we have received 14 responses, all of them positive and effective: The service providers simply deleted our account without any complication.

#### 4.7 Responsible Disclosure and Outreach

We contacted by e-mail (typically `dpo@service.com` or `privacy@service.com`; we usually obtained the contact e-mail from the privacy policies found on the online services' websites) the data protection officers of all the vulnerable services, detailed the vulnerability, and issued recommendations (see [Section 7.2](#)) for fixing it.<sup>28</sup>

Out of the 59 services we contacted, five answered (beyond automatic replies) so far. Specifically, a gambling website mentioned that they started investigating the issue. A DIY e-commerce website fixed the issue, but it was not completely satisfactory, so we contacted them again with further recommendations. A food e-commerce website asked us to submit the vulnerability to Bugcrowd, which we did, but we did not receive any news since then. A video streaming service responded that there was no record of the e-mail address we used to contact them in their system and simply stated that "We take information security seriously and use reasonable administrative, technical, and managerial measures to protect your personal information against unauthorized access. Accordingly, we follow many industry best practices to monitor, protect, detect, block, and react to inappropriate access to personal information." Finally, a grocery shop asked for more details, which we provided, but we did not receive any news.

We further contacted (more than 30 days after contacting the services' DPOs) three national data protection authorities (DPAs) in countries where the current data protection laws make service providers responsible for protecting their websites against account enumeration attacks. Two DPAs have answered so far. Both were unaware of the reported vulnerability and agreed to communicate about it to raise awareness and thus increase compliance among online services. One decided to include a description of the vulnerability and the associated recommendations in an updated version of their guide for developers. They further reported considering the possibility of auditing the vulnerable websites. Yet, because there are many vulnerable websites (in fact, the vast majority

---

<sup>26</sup>Note that Apple recently forced developers who allow account creation in their (iOS) mobile apps to also allow account deletion from within the app. See [developer.apple.com/news](https://developer.apple.com/news), last visited: Feb. 2024.

<sup>27</sup>The message was sent from the e-mail address associated with the account.

<sup>28</sup>A copy of the e-mail sent to service providers is also available in the OSF repository.

of websites are vulnerable, as shown by our results), selecting the ones to audit remains a tricky question.

## 5 USER PERCEPTION AND BEHAVIOR

To consider the users' perspective of online services concerning the considered attack, we complemented our experiment with a user survey.<sup>29</sup> Indeed, as the considered attack vectors may produce feedback that is visible to the targeted users (i.e., e-mail), it is important to understand users' perceptions of these e-mails as well as their reactions. Additionally, it is important to understand users' privacy perceptions with respect to the list of their online accounts. Therefore, we set the following research questions:

**RQ2a.** *How frequently do Internet users receive unsolicited e-mails concerning account creation or password reset?*

**RQ2b.** *What is their mental model regarding such e-mails? Specifically, what do they think might be the cause and purpose of these e-mails?*

**RQ2c.** *How do they react when they receive these unsolicited e-mails?*

**RQ2d.** *Are users more likely to click on a link in an e-mail from a service for which they have an account?*

**RQ2e.** *Do they take actions that might defeat the considered attack (e.g., using different e-mail addresses for different services)?*

**RQ2f.** *How sensitive do users perceive the fact that they have an account on an online service?*

### 5.1 Methodology

We conducted an online user survey<sup>30</sup> through a specialized vendor<sup>31</sup> we contracted. The vendor built a representative sample in terms of demographics (i.e., age, gender, region of residence, education, and income) of Internet users in the French-speaking region of Switzerland and distributed our questionnaire to the selected users. The respondents were recruited via partnerships and went through quality controls. Respondents received points in exchange for their participation. Every month, the vendor organizes a raffle for prizes of up to CHF 1,000 (~USD 1'150) among respondents of all studies and the total number of points they collected in the past month influences their probability of winning.

The survey was designed to collect quantitative data regarding the considered attacks: perception, occurrences, typical responses, and self-reported appreciation of the connected risks. Additionally, the survey aimed to collect qualitative data on the mental models of the respondents. The survey transcript is available in [Appendix B](#). To prime respondents to think about privacy across various domains, including sensitive ones (e.g., adult, dating, LGBTQ+, politics, religion), we first presented them with a list of services that are extremely popular in Switzerland and asked them to select services for which they owned an account. Responses to this question were also used to drive the logic of the subsequent questions (and personalize them), where we asked respondents to compare occurrences of the attacks described in [Section 4.1](#) and responses to these attacks. In the questionnaire, we presented respondents with two real examples of e-mails that could result from

<sup>29</sup>We also interviewed a security expert working on the login system of one of the most popular online retailers in our region. The main objective was to investigate the service provider's perspective, particularly to understand the corporate security expert's awareness of account enumeration attacks, their perception of the threat's impact on services and users, and the countermeasures they recommend, along with potential implementation challenges. However, further interviews could not be conducted. Given the limited sample size, we have included this interview as supplementary material on OSF.

<sup>30</sup>Surveys are often used to study Internet users' security and privacy-related attitude/behaviors and mental models [20, 22, 35].

<sup>31</sup>DemoSCOPE. See [demoscope.ch](https://demoscope.ch), last visited: Feb. 2024.



the attacks considered in this paper (i.e., from the password reset and account creation flows of popular services). Furthermore, the text of some of the questions adapted automatically to whether respondents owned (or did not own) accounts with the services we used in the examples (e.g., “Imagine that you do have an account on this service and that you received this e-mail today...” vs. “Imagine that you received this e-mail today...”). Also, the order of the sections containing the examples was randomized to control for presentation bias. Similarly, the order of options for multiple-choice questions was randomized.

In our survey, we consciously refrained from utilizing traditional attention checks. This choice was influenced by the work of Matsuura et al. [28], which suggests that such checks may introduce bias by excluding respondents less attentive to security threats—a key demographic for security research. To preserve the quality of our survey while maintaining a representative sample, we assessed respondent honesty through their answers to open-ended questions, examining both the sincerity and relevance of their responses. Additionally, we evaluated response times and patterns to recognize speeders (i.e., respondents who completed the survey too quickly) and straightliners (i.e., those who consistently selected the same response option), indicative of insufficient engagement with the survey content.

Before deployment, we ran five cognitive pretests with regular Internet users from the French-speaking region of Switzerland. These tests were useful to finalize the wording of some questions. The questionnaire was deployed in two stages: we initially deployed it to a short sample of 20 respondents (i.e., soft launch). These answers were checked for consistency with expected outcomes. The survey took an average time of 5 minutes and 56 seconds to complete.

Qualitative answers provided for the open-ended questions were categorized by two coders. The codebook was developed through inductive analysis [37]. We measured inter-coder agreement through Cohen’s kappa at 0.87 and judged it sufficient. The few cases of conflict were resolved via discussion. The survey was approved by our institution’s ethics committee and the questionnaire was deployed in the Fall of 2021.

## 5.2 Survey Limitations

Survey methodology is susceptible to response bias [43]. As for most user surveys on security/privacy-related behavior, the information reported by the respondents might not reflect their actual behaviors; this fact is well documented (e.g., [13]). Therefore, we crafted the survey with neutral and non-leading language. This strategy aimed to minimize the influence of participants’ awareness of the survey’s focus on their responses.

## 5.3 Demographics and General Statistics

The final dataset comprises 318 complete answers. The proportion of female respondents was 53.8% and the mean age of the respondents was 45.2 ( $SD = 15.6$ ). The respondents had diverse education and income levels as well as diverse privacy concern scores (IUIPC [25]:  $M = 3.6$ ,  $SD = 0.7$ , 5 agreement levels).

The large majority of respondents had accounts with one or multiple services that we listed at the beginning of the survey, namely 278 respondents (87.4%). The list also contained services that are typically considered more sensitive from a privacy perspective (e.g., adult content, dating, politics, religion). A total of 27 respondents (8.5%) declared having an account for at least one of these services.

## 5.4 Results

After presenting respondents with screenshots of a typical password reset e-mail (for Badoo, a dating service) and account creation e-mail (for PornHub, an adult video streaming service), we instructed the respondents to imagine that they had received this e-mail even though they had not made any request to the service that could have triggered the e-mail. We then asked them a few questions.

**RQ2a.** More than half of the respondents stated they had received one or multiple of either a password reset (168, or 52.8%) or an account creation e-mail (97, or 30.5%) that they had not triggered in the 6 months preceding the survey. About a fourth of the respondents stated they had received both types of e-mails in the past 6 months (85, or 26.7%). **This indicates that the general population might regularly receive these types of unsolicited e-mails, which might be the symptom of the attacks considered in this paper or general phishing attacks.**

**RQ2b.** Concerning the password reset e-mail, almost half the respondents (152, or 47.8%) thought that they received these e-mails because of an *hacking attempt* from a third party trying to access their accounts. We coded as “hacking attempt” the responses that explicitly mentioned hacking or referred to the situation where someone tries to unduly *gain access* to their account; as such, these attempts differ from the attack considered in this paper, which simply aims at determining the *existence* of an account. The second most cited explanation was *phishing* (48, or 15.1%), while the third was an (aggressive) marketing attempt (or *spam*) from companies to nudge recipients to visit their website (25, or 7.9%). Other respondents thought they were receiving these e-mails because they had *lost their password* (18, or 5.7%). This explanation reveals a misconception: respondents did not realize a legitimate password reset is typically triggered by their action on the website of the concerned service. A minority of respondents reported a *service error* as the root cause of these messages (14, or 4.4%) or the fact they had been *inactive* with the service for some time (7, or 2.2%). Finally, the remaining respondents either reported *not being sure* about the cause of these messages (29, or 9.1%), or their reply was unclear and could not be classified (25, or 7.9%).

Regarding the account creation attack, a large number of respondents reported explanations similar to the password reset, namely *hacking attempt* (77, or 24.2%), *spam* (52, or 16.4%), *phishing* (47, or 14.8%), or some sort of *error* (6, or 1.9%) due to either the service malfunctioning, inactivity of their account, or a typing mistake by someone else. In addition to these similarities, the account creation scenario primed respondents to think about the *unauthorized use* of their e-mail addresses (54, or 17.0%). In addition, some respondents simply described the reason for receiving that kind of e-mail as one of the intermediate steps of an *account creation* process, missing to consider that someone else had created the account using their e-mail address (17, or 5.3%). Interestingly, one respondent (0.3%) guessed that one possible reason for receiving this type of e-mail could be connected with someone testing whether that e-mail *address existed*. Concerning the remaining respondents, they either reported *not being sure* about the cause of these messages (42, or 13.2%), or their answer was unclear and could not be classified (22, or 6.9%). **These results show that almost none of the respondents thought that the most likely reason for receiving these e-mails could be the attacks considered in this paper.**

**RQ2c.** We then asked respondents whether they took any action when they received one of these e-mails (or *if* they were to receive one). Note that, unfortunately, at this stage there is nothing they can do to protect themselves against the attack (i.e., to prevent the adversary from inferring the existence of their account): it is too late. Yet, they could delete their account in order to prevent further investigations or to lower the accuracy of the profile built by the attacker. Many respondents stated they would not take any action (143, or 45.0% for the password reset, and 154, or 48.4% for

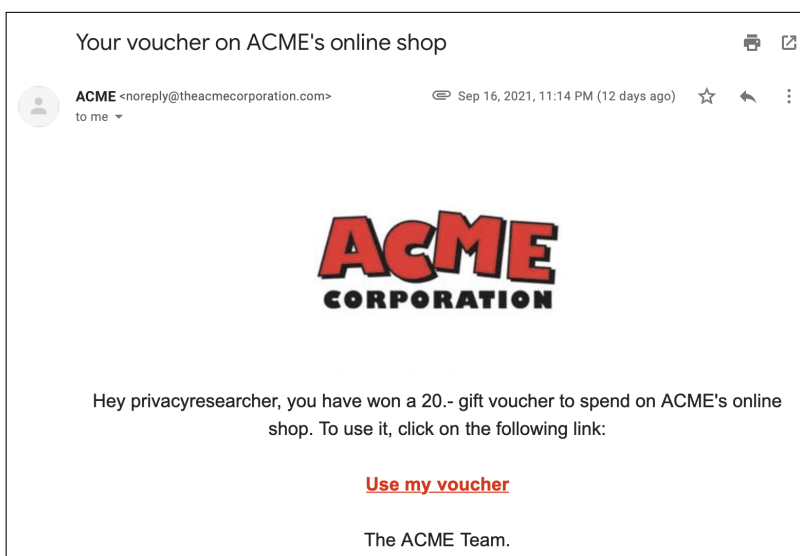


Fig. 3. Screenshot of the synthetic e-mail shown in the survey questionnaire.

the account creation e-mail). The remaining respondents described taking actions that were quite similar across the two scenarios (password resets and account creations) that we tested. Therefore, we merged the codes to these answers. We analyzed a total of 339 responses across the two scenarios. The most frequent action reported by respondents was to simply *delete* the received e-mail (117, or 34.5%), or to *mark it as spam* (55, or 16.2%). In addition, several respondents reported being concerned when receiving this kind of e-mail and taking protective actions: *changing the password* of either their e-mail account or of the service for which they received the e-mail<sup>32</sup> (39, or 11.5%), *contacting the service* to report the reception of an unsolicited e-mail (37, or 10.9%), *removing the account* from the service for which they received the unsolicited e-mail (12, or 3.5%), or *going to the website* of the service to check whether their account was compromised, and eventually, change the password (12, or 3.5%). The fact that a substantial fraction of the respondents would contact the service providers indicates that **a large-scale attack (in terms of the number of targeted e-mail addresses on a given service) would probably be detected by the service providers**. Further to these, fewer respondents reported taking other actions with smaller frequencies (cumulatively 24, or 7.1%): unsubscribing from the mailing list the message might come from, asking a peer for help, or trying to identify the sender of the message. Concerning the remaining respondents, they either reported *not being sure* about the cause of these messages (6, or 1.8%), or their answer was unclear and could not be classified (37, or 10.9%).

**RQ2d.** In the questionnaire, we presented a screenshot of a synthetic e-mail containing a link for a ~USD 20 voucher, sent by a generic e-commerce online service (see Figure 3). Then, we asked respondents to rate how likely it would be that they would click on the link, on a 7-point Likert scale, from 1 (*very unlikely*) to 7 (*very likely*). We asked respondents to rate this twice: imagining they had an account with the service and not having an account with the same service

<sup>32</sup>Responses coded in this category did not clarify whether this action was taken by clicking on the link received in the suspicious e-mail or not.

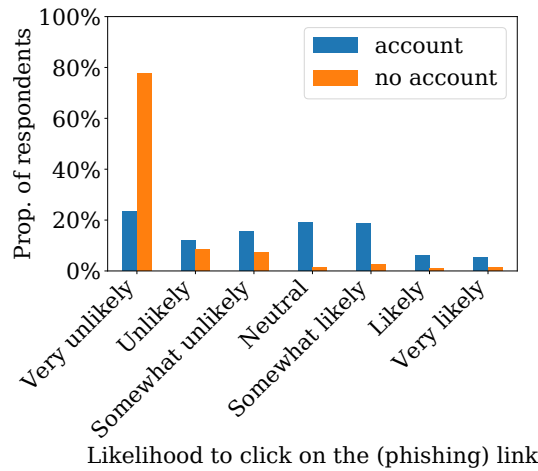


Fig. 4. Respondents’ likelihoods to click on the link contained in an e-mail depending on whether they have an account with the service that (supposedly) sent the e-mail.

(the order of presentation of these two scenarios was randomized).<sup>33</sup> The mean ratings across the two scenarios were: having an account 3.4 ( $SD = 1.8$ ); not having an account 1.5 ( $SD = 1.2$ ). A paired, one-tailed Wilcoxon Signed-rank test shows that the distributions differed significantly ( $W = 386.5$ ,  $Z = -12.7$ ,  $p < .001$ ,  $r = 0.85$ ).<sup>34</sup> **In other words, respondents claimed to be more likely to click on an e-mail from a service for which they had an account than on an e-mail from an unknown service, confirming the usefulness of the attack for (spear) phishing.** The complete distributions are depicted in Figure 4. Furthermore, we asked respondents to rate the likelihood they would check whether the e-mail was sent to the address they had used to register to that service. Respondents claimed to be slightly unlikely to check ( $M = 3.8$ ,  $SD = 2.2$ ). This indicates that they would still be more likely to click on the link even if the message was sent to an address different from the one they used to register their account.

**RQ2e.** To assess the susceptibility of the respondents to the considered attack, we asked them about their usage of e-mail addresses, in particular when creating online accounts. The majority of the respondents reported that they had more than one e-mail address: one 62 (19.5%); two 134 (42.1%); three or more 122 (38.4%). This is likely due to respondents differentiating between e-mails for certain contacts. **However, when registering on services from various categories, the majority of the respondents stated that they use mainly one e-mail address:** one 133 (41.8%) (combined with the results of the previous question, it means that a total 19.5% + 41.8% of the respondents reported using a single e-mail address for registering all their accounts); two or more 123 (38.7%). **This indicates that knowing the e-mail address an individual uses for one online service is enough to carry out the attack on most of the other services they use.**

<sup>33</sup>At first, we considered objective measurements through phishing campaigns with AB-testing—which would have a higher ecological validity than self-reported attitudes collected in our survey; however, this would have involved deception, which is not allowed by our institution’s ethics committee.

<sup>34</sup>An R-workbook with the details of the analysis is available in the OSF repository for reproducibility.

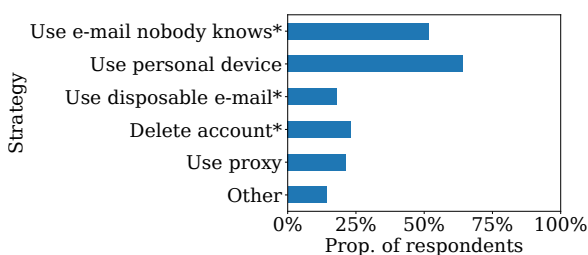


Fig. 5. Strategies used by the respondents for preventing others from knowing about the existence of some of their online accounts ( $N = 56$ ). Note that only strategies marked with a ‘\*’ are effective against the considered attack.

We also asked the respondents whether they had used SSO to login to services in the past 6 months: **196 respondents (61.6%) reported using SSO at least once, thus protecting themselves against the considered attack.** Furthermore, we asked them whether they used specific strategies to prevent others from knowing they have an account on a given service. The majority declared not using any strategy (196, or 61.6%). Out of those who did use a strategy (56, or 17.6%), the majority declared connecting to these services exclusively from a *personal device* (36, or 64.3%) or using an e-mail *address that nobody knew\** (29, or 51.8%). Other strategies included *deleting the account* after its use\* (13, or 23.2%), using a *disposable e-mail address\** (10, or 17.9%),<sup>35</sup> and using a *proxy* (12, or 21.4%). Of the 8 respondents who selected ‘Other’ (14.3%), 6 declared to create e-mails with pseudonyms\* and keep them secret, while the remaining 2 respondents declared to browse the service using a private browsing mode. These results are depicted in Figure 5. Note that even if all these strategies make sense, only those with a ‘\*’ are effective against the considered attack. **These results show that the attacks reported in this paper could potentially be used to profile the majority of the respondents. Unfortunately, only a minority of them use effective strategies to prevent such profiling.**

**RQ2f.** We then asked respondents to rate how sensitive they considered the disclosure of the services they have an account for with regard to distinct adversaries. Figure 6 depicts these results. Respondents reported being mostly concerned with *companies* ( $M = 3.4, SD = 1.4$ ), their *government* ( $M = 3.2, SD = 1.4$ ), and their *employer* ( $M = 3.2, SD = 1.5$ ). They were slightly less concerned with their *colleagues* gaining access to this information ( $M = 3.0, SD = 1.4$ ). Finally, respondents were less concerned with their *friends* ( $M = 2.4, SD = 1.3$ ), *relatives* ( $M = 2.2, SD = 1.3$ ) and *partner* ( $M = 1.8, SD = 1.2$ ). **These results indicate that the respondents’ level of concern increases with social distance.** Note that the opposite was observed in other cultures, e.g., research conducted in Africa and India showed that people tend to keep their online activity hidden from partners and relatives to maintain social reputation [36]. Also, it was surprising to see that the intimate partner raised so few concerns, considering the prevalence of intimate partner surveillance [24, 29, 41].

## 6 USER IDEAS

The online survey results helped us understand users’ perceptions and behaviors about receiving unsolicited e-mails, as well as the frequency and sensitivity of such incidents. In addition to the survey, we conducted a focus group to gain deeper insight into users’ understanding and concerns about account enumeration attacks and to explore the process of creating privacy-enhancing

<sup>35</sup>Note that the (careless) use of disposable e-mail addresses raises privacy issues [19].

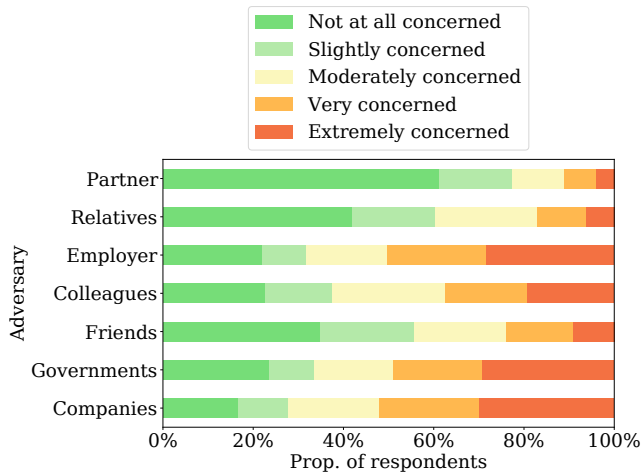


Fig. 6. Respondents’ concerns, with respect to different adversaries, regarding the list of the online services for which they have an account ( $N = 318$ ).

countermeasures against such attacks. A focus group is a group interview gathering a few people to discuss and gain insights. Therefore, we pose the following research questions:

**RQ3a.** *What do users think about such attacks? In particular, what are the contexts in which disclosure of account information is perceived as sensitive? Can users identify the typical attack vectors?*

**RQ3b.** *How would users protect themselves? More importantly, what design tweaks would they suggest to improve the effectiveness of existing countermeasures?*

## 6.1 Methodology

We conducted a focus group session to gather a group of participants to answer specific questions about account enumeration attacks. We recruited participants through LABEX, a dedicated recruiting organization at our institution with a pool of approximately 8,000 university students. LABEX advertised the study to their pool. We used a screener survey to select participants. The screener survey included questions about the respondent’s age, gender, institution, nationality, and English proficiency. Because account enumeration attacks may be more sensitive in countries with authoritarian governments, we asked about respondents’ nationality. For this, we used the ranking of countries based on the Democracy Index published by the Economist’s sister company, the EIU.<sup>36</sup> We included a healthy mix of participants from countries classified as full democracies, flawed democracies, hybrid regimes, and authoritarian regimes. In terms of language proficiency, we only recruited participants who were fluent in English. We also provided participants with a list of common online services and asked them which of these services they had an account for. We collected responses from 76 students. Based on their responses to the screener survey, we contacted 27 students, and based on their availability, we recruited  $N = 13$  participants.

Two researchers conducted the focus group session. Some activities were done individually, and some collectively, either at the group level or with all participants. The participants were divided into four groups of three or four people, and each group gathered around a round table. We welcomed the participants and asked them to read and sign the consent form. We began the session by explaining the benefits of having accounts with various online services. To get participants

<sup>36</sup>See [eu.com/n/campaigns/democracy-index-2022](https://eu.com/n/campaigns/democracy-index-2022), last visited: Feb. 2024.

involved in the discussion, we asked them to think about how many accounts they had. We then asked them to think about each online service for which they had an account and to think about those for which they would be uncomfortable if others knew they had an account. We also asked them to think about sensitive accounts and to discuss a “hypothetical” example where they would be unhappy if they had an account there and others knew about it. We also asked them to discuss how important this issue is to them.

Next, we evaluated whether the participants comprehended the various attack methods that exist. We initiated this evaluation by asking them: “*imagine your friend wants to know if you have an account on a certain website,*” and then requested that they identify the ways in which others can determine if they have an account on that particular website. Following that, we asked them to discuss their attack strategies with other participants in the group and share the different attack methods they had identified. Lastly, we inquired about the measures they would take to protect themselves from such attacks and whether they had already taken any protective steps.

In the last part of the session, we focused on ideation for countermeasures against such attacks. We asked the participants to “*think of any design idea to improve the privacy of account owners.*” We let the participants know that they could think of any out-of-the-box ideas to address such attacks. Participants discussed their ideas within the group and then documented or sketched them. Finally, we asked a volunteer from each group to come on stage and explain their idea. At the end of the session, the facilitators thanked the participants and compensated them for their two hours of participation. Before leaving the session, participants signed the payment form and received a cash payment of CHF 40 (~USD 45). This study has no particular ethical implications. We received approval from our institution’s ethics committee and followed standard protocols such as participant anonymity and data deletion.

## 6.2 Focus Group Limitations

We acknowledge the potential influence of participants’ cultural backgrounds, ages, and proficiency levels on their perceptions and responses. The focus group sessions involved a diverse set of participants from various national backgrounds; however, the findings are not intended for broad generalization. Additionally, the study primarily recruited students, which may affect the generalizability of the results to a wider population. Despite these limitations, the study provides valuable qualitative insights into user perspectives on account enumeration attacks.

## 6.3 Findings

Seven participants were women and six were men. The average age of the participants was 20.4 years ( $SD = 1.6$ ). Six participants were from hybrid regime governments, including Morocco, Tunisia, and Turkey. Four were from authoritarian governments, including Egypt, Cameroon, and Lebanon. Two participants were from governments with flawed democracies (the United States and Hungary), and only one was from governments with full democracies (France).

**RQ3a.** When we asked participants to think about sensitive cases where an attack could make them unhappy, they mentioned having an account on adult video streaming sites, dating services, gambling services, and online social networks. For social networks, they mainly discussed the example of having a fake account to bully or stalk other users, trolling on social media, or promoting another account, and the fact that account enumeration attacks can embarrass those account owners: “*If you have an account on social media where you post mean stuff, you don’t want other people to know who you are.*” [P5]. One participant also mentioned the case of the dark web, which is beyond the

focus of our work:<sup>37</sup> *“If people have an account on a site where they can buy drugs or guns [...]”* [P2]. When we asked participants who are the people that the account owner might feel uncomfortable with after such attacks, they mentioned friends, family, and social media followers.

We asked participants about the severity or seriousness of such attacks, and they mostly agreed that *“it depends on the situation”* [P10]. P12 discussed that if the services are considered “taboo,” it could have more severe consequences, adding that *“I think in this country there is diversity. We have people from all over the world. So I think it is less sensitive.”* P2 agreed and added: *“If you come from a conservative country, having an account on a gambling or adult website is definitely problematic. I come from Morocco. If someone in my family had an account on those sites, it would be a big shame.”* And P12 replied, *“The difference with Western countries is that people in my country [Egypt] refrain from saying it out loud [...]”* The participants also agreed that even if they do not live in a conservative country, the attack can cause problems in their destination country when they travel. **The results showed that the severity of account enumeration attacks depends on the cultural context and the perceived sensitivity of the services.** A participant also mentioned the potential of being discriminated against in job interviews after the attack.

When we asked participants to think about possible ways to determine if someone has an account on a particular service, they thought about password reset, account creation, and friend finder attack vectors.<sup>38</sup> Additionally, two participants expressed that there may be a dedicated website or software to perform such attacks: *“You put the e-mail up, and they give you all the accounts associated with that e-mail.”* Finally, P5 mentioned the case of having physical access to the account owner’s device, such as unlocking someone’s phone and checking their password manager. This is beyond the scope of our study. Overall, the discussion showed that **the attack vectors are straightforward enough to be understood by non-expert users and that even lay users might consider performing such attacks.** Lastly, two participants *mistakenly* thought that account enumeration attacks were always non-stealthy: *“If the attacker uses the ‘forget password’ button, then the user gets an e-mail.”* [P9]. In fact, if there is no account in that service, the e-mail owner usually does not receive an e-mail.

**RQ3b.** When we asked participants how they would protect themselves, the majority agreed that they would use a different e-mail address that is unknown to others, is not used for daily communication, and cannot be guessed (e.g., not using their real identity in the e-mail address). One participant mentioned using disposable e-mail. These findings are consistent with our previous findings (see Figure 5). In addition, one participant mentioned the Hide My E-mail feature available to Apple and iCloud+ users for SSO.<sup>39</sup> This feature creates a random and anonymous e-mail address during account creation and forwards e-mail messages to the user’s primary e-mail address. The feature is different from SSO because it creates an alias e-mail. We asked participants if they had ever taken any action to protect themselves. Two participants (P2 and P12) mentioned using a different e-mail address, and one (P8) shared that he had stopped using “shady websites.” Our participants suggested four main ideas to enhance account privacy. Next, we explain each concept and briefly discuss them.

First, most participants argued that the **attacks should not be stealthy** (G1, G3, G4).<sup>40</sup> While some participants (G4) discussed that the account owners should be informed about the attack, others (G1, G3) believed that the e-mail owners (i.e., those who may not have an account) should

<sup>37</sup>Note that the website on the dark web may have account creation and login features, but the level of security and anonymity may be higher, making it more difficult to perform account enumeration attacks.

<sup>38</sup>To read about the friend finder, see the ‘anonymous account feature’ in the findings related to RQ3b.

<sup>39</sup>See [support.apple.com](https://support.apple.com), last visited: Feb. 2024.

<sup>40</sup>G1 stands for participants from Group 1. The participants of the focus group were separated into four different groups.



also be informed about their e-mail addresses being used to execute the attack. Most participants (G3, G4) mentioned that the service providers should send these e-mails. This idea does not prevent the attacks, but it is a good suggestion to inform the participant about information leakage—a posteriori. Some participants (G1) suggested a global service that can detect all login, password reset, and account creation attempts on the Internet and notify e-mail owners. Such a feature can empower users to monitor and protect their accounts. However, it would require a distributed and extensive network of monitoring nodes and advanced algorithms to detect suspicious activity. Given the complexities and challenges involved, implementing such a global service would require a collaborative effort among multiple stakeholders, including private companies, government agencies, and Internet service providers.

Second, several participants (G3, G4) suggested **the anonymous account feature** in online services to hide the account and prevent friend finder attempts: when users of an online service can search for their friends’ e-mail to see if they use the same service (e.g., when searching for friends on Facebook or collaborators on Overleaf). Online services should give users the choice to opt in for anonymity. This feature can be enhanced by using an alias e-mail to allow users to share their accounts with their friends using the alias e-mail. However, it is not known whether service providers would be willing to use such a feature. For example, social media services that rely on friend connections and networking may hesitate to adopt this idea because it may disrupt their core functionality. So, the feasibility of implementing this idea might vary depending on the type of online service.

Third, some participants (G4) also proposed the idea of **MasterApp** to be used during account creation and account login. This idea is very similar to SSO. However, SSO is mostly available for Google, Apple, and Meta users. With MasterApp, users can use SSO without having an account in one of these companies. The MasterApp acts as a diversified SSO service, allowing users to be independent of tech giants.

Fourth, several participants (G2, G3, G4) suggested that **oblivious messages** be displayed when login, password reset, and account creation fail. As proposed in [Section 4.1](#), such messages do not reveal the existence of the account and are an effective strategy because attackers will not be able to determine the existence of an account. The other positive side of oblivious messages is that they do not require significant implementation effort to deploy. However, it is important to consider the potential impact of oblivious messages on the user experience. For example, if a user forgets their username and/or password and tries several different passwords during the login process, receiving the same generic “Invalid username or password” message for all attempts can be frustrating. Thus, future work should carefully consider how to minimize such negative effects.

## 7 DISCUSSION

In this section, we discuss how the attacks considered in this paper can become an even more serious threat if fully automated. Then, we provide recommendations for mitigating the attacks. Finally, we conclude with legal considerations as well as limitations.

### 7.1 Attack Automation

There are two main reasons for which one would want to automate the attack: scaling the number of targeted services and scaling the number of targeted users (i.e., e-mail addresses). Automating the attack vectors *across* online services is non-trivial since each service tends to have its own variants of login, account creation, and password reset forms. While such variations can be easily managed by a human being, scripting against them could be tricky. However, when it comes to automating these attack vectors for a *specific* online service, the task turns out to be rather easy,

especially in the absence of CAPTCHAs. Indeed, once the structure of a given form is known, feeding the forms with a large number of e-mail addresses can be easily automated (it is even easier if the service offers an API to verify the existence of an account). In order to demonstrate the feasibility of attack automation, we implemented a proof-of-concept script for doing so for several online services, including Netflix and an online retailer. We provide its source code on the OSF repository. Note that it takes only a dozen lines of code (i.e., information about the form to be submitted and a function to determine the existence of the account from the returned HTML page) to extend our script to a new online service. The script is written in Python and relies on Selenium for automating user interactions (i.e., input text—such as usernames and passwords—and submit forms) with the browser. The script takes as input a list of e-mail addresses and performs the attack by simulating interactions with the websites of the targeted services and by using the results of our study to determine the existence of accounts associated with the targeted e-mail addresses. We also provide a demo video. As for online services that block too many form submissions from the same IP address, they can be fooled by relying on VPNs, proxies, or botnets.

In order to further stress this (automated) approach in real conditions, we tested it on the website of the aforementioned online retailer. On this website, the account creation webpage informs the user of the existence of an account associated with the e-mail address provided without the need to actually submit the form (asynchronous requests). We tried, automatically (with Selenium), different e-mail addresses at regular intervals. We started with one address every 10 minutes and increased the pace progressively (by 10%, every 10 minutes) until we reached a pace of one attack every 10 seconds. The attack ran flawlessly, and we stopped it after 24 hours. At such a (modest) pace, it is possible to enumerate millions of accounts in a few months from a *single* IP address.

## 7.2 Recommendations and Countermeasures

To thwart the threat raised by the considered attack, some rather simple countermeasures could and should be implemented. It is important to note that the focus of these countermeasures is primarily on e-mail-based attacks (not phone-based ones), which aligns with the main scope of this research.

On the service provider side, a simple countermeasure is to display *oblivious* messages (i.e., that do not reveal whether an account exists) for all operations related to account management (including log in, password reset, and account creation). For instance, upon password reset, a suitable oblivious message reads “*if an account associated with the e-mail address exists, a password-reset link has been sent to the e-mail address*” (see Figure 1c). For account creation, a suitable oblivious message reads “If no account associated with the e-mail address already exists, one has been created and a confirmation e-mail has been sent. [optional: Otherwise, a password-reset link has been sent to the e-mail address]”. However, it is essential to consider the usability implications of such messages, as they may introduce some level of frustration for users. For example, account creation forms usually require users to input large amounts of data (first name, last name, address, date of birth, *etc.*). Therefore, notifying users that an account already exists *before* they input all this data presents obvious usability benefits. A potential usability enhancement could involve a two-stage account creation process, as illustrated in Figure 7: (1) the user inputs their e-mail address and submits the form. (2) If no account exists, an e-mail with a link to finalize the account creation (including the entry of the remaining data) is sent. The usability of such an account-creation process should be further studied. In any case, service provider should totally avoid the use (and provision) of an *open* API for determining the existence of an account. Future research could conduct usability evaluations or surveys to assess the practicality and user-friendliness of these countermeasures in real-world scenarios.

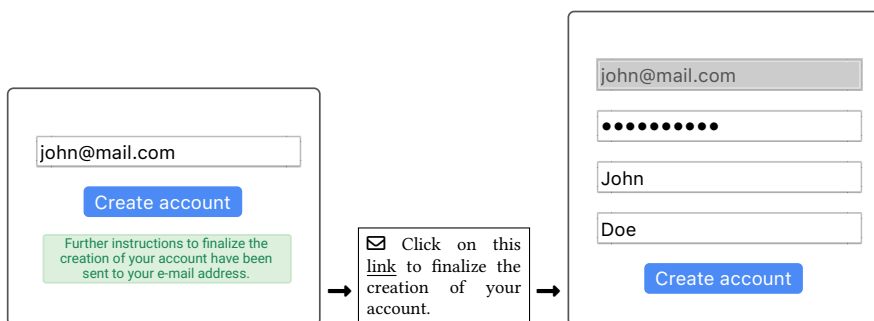


Fig. 7. Recommended design for a 2-stage account creation process.

Another (partial and controversial, as explained above) countermeasure is to offer the possibility to register and log in via a third-party SSO service. However, the practicality and impact of SSO on security and usability should be thoroughly assessed. Moreover, service providers should offer an easy online procedure to delete one’s account, rather than going through a tedious e-mail-based procedure, making it more user-friendly and efficient. Future research can explore the design and implementation of streamlined account deletion processes. Finally, service providers could rely on usernames instead of e-mail addresses as identifiers. The impact of all the aforementioned countermeasures would be even higher if they were implemented in user-management frameworks such as UserFrosting (PHP, <https://www.userfrosting.com>) and flask-base (Python, <https://github.com/hack4impact/flask-base>).

Finally, service providers could put in place mechanisms for *detecting* such attacks (e.g., when many requests associated with different e-mail addresses come from a limited set of IP addresses) and for *warning* users about requests related to their e-mail addresses, regardless of whether they have an account, as done by LEGO<sup>®</sup> and Zoom. Naturally, service providers should strike a balance between security/privacy of online accounts and users’ information load. Such mechanisms would make the considered attacks non-stealth, thus deterring covert adversaries from carrying these attacks. On the user side, simple protective measures could also be implemented (yet, we believe that, to concretely counter these attacks, it should be the service provider who implements protective measures). For instance, a user could use distinct e-mail addresses for different online services (or use disposable e-mails), at least for those that are critical from a privacy viewpoint. A relatively simple way to do so is to rely on the alias feature offered by e-mail service providers. For instance, Gmail users can use aliases of the form `username+alias@gmail.com` (the “+” symbol denotes the “+” ASCII character, not a concatenation operation) for generating an arbitrary number of different e-mail addresses while still receiving the e-mails sent to these addresses in the mailbox of their address `username@gmail.com`.<sup>41</sup> For such a countermeasure to be efficient, however, users should make sure that the aliases they use are not too easy to guess. For instance, alias `username+servicename@gmail.com` is relatively easy to guess. Another simple countermeasure is to systematically delete accounts on services that the user no longer uses.

While the countermeasures discussed primarily focus on e-mail-based attacks, it is essential to consider their applicability to phone-based attacks. Phone numbers play a significant role in account registration, especially in mobile applications, and may be vulnerable to enumeration attacks. Unlike e-mail addresses, phone numbers are often more personal, less frequently changed, and may require additional steps or fees to obtain new ones. The effectiveness and practicality

<sup>41</sup>See [blog.101domain.com](https://blog.101domain.com), last visited: Feb. 2024.

of applying the same countermeasures to phone-based attacks require further investigation. For example, displaying oblivious messages for phone-based operations related to account management could have different usability implications than for e-mail-based operations. Similarly, while a two-step account creation process with e-mail verification is common, similar verification methods can be applied to phone numbers. Future research should explore the specific challenges and nuances of countering phone-based enumeration attacks and evaluate the usability and security implications of applying similar countermeasures.

### 7.3 Legal Analysis

The vast majority of the services considered in our analysis are available to residents of the European Union (EU), who are protected by the General Data Protection Regulation (GDPR). Under the GDPR, an e-mail address is considered as personal data (Article 4(1)) and, by extension, so is the existence of the corresponding service account data associated with this e-mail address. In its capacity of data controller, the service provider is obliged to protect the data of its users (the data subjects, identified by at least their e-mail addresses), including the existence of their accounts on the service. Therefore, revealing through different features of an online service the existence of an account associated with the given e-mail address constitutes a data breach. To avoid this, the service provider could implement the aforementioned recommendations, such as displaying the same response to requests, regardless of whether an account associated with an e-mail address exists.

## 8 CONCLUSION

In this paper, we have studied three well-known basic attacks that enable an adversary to infer the existence of an online account associated with a given e-mail address, on a specific service. Our experimental results show that the considered attacks are quite successful and can often be carried out stealthily. Our user survey then shows that users find this information sensitive, but only a small fraction of them take (effective) measures to conceal it. As part of future work, we intend to further research the feasibility of automating the attack across websites and to study other known attack vectors (e.g., based on timings) for accessing the same information and to identify new ones. We also intend to study user behavior (instead of attitude) through in-the-wild experiments and to study the usability and user perception of the proposed countermeasures (e.g., two-stage account creation process). Finally, we intend to collaborate with service providers and developers of user-management frameworks to fix the vulnerabilities exploited by the considered attack and with data protection authorities to increase compliance regarding the protection of account databases of online services.

## ACKNOWLEDGEMENTS

The authors are grateful to Vincent Vandersluis for his great editing job, to Bertil Chapuis, Nicolas Frey, and Christophe Hauert for their feedback on the attack, to Valérie Junod and Sylvain Métille for their feedback on the legal aspects, and to Arnaud Salvadori for his help in the preliminary investigation. The work was partially funded with grant #22018 from the Hasler Foundation.

## REFERENCES

- [1] Ayako Akiyama Hasegawa, Takuya Watanabe, Eitaro Shioji, Mitsuaki Akiyama, and Tatsuya Mori. 2020. Addressing the Privacy Threat to Identify Existence of a Target’s Account on Sensitive Services. *Journal of Information Processing* 28, 0 (2020), 1030–1046. <https://doi.org/10.2197/ipsjip.28.1030>
- [2] Yonatan Aumann and Yehuda Lindell. 2007. Security Against Covert Adversaries: Efficient Protocols for Realistic Adversaries. In *Proc. of the Theory of Cryptography Conference (TCC)*, Vol. 4392. Springer, 137–156. [https://doi.org/10.1007/978-3-540-70936-7\\_8](https://doi.org/10.1007/978-3-540-70936-7_8)





- [3] Marco Balduzzi, Christian Platzer, Thorsten Holz, Engin Kirda, Davide Balzarotti, and Christopher Kruegel. 2010. Abusing Social Networks for Automated User Profiling. In *Proc. of the Int'l Workshop on Recent Advances in Intrusion Detection (RAID)*, Vol. 6307. Springer, 422–441. [https://doi.org/10.1007/978-3-642-15512-3\\_22](https://doi.org/10.1007/978-3-642-15512-3_22)
- [4] Lujo Bauer, Cristian Bravo-Lillo, Elli Fragkaki, and William Melicher. 2013. A comparison of users' perceptions of and willingness to use Google, Facebook, and Google+ single-sign-on functionality. In *Proc. of the ACM Workshop on Digital identity management (DIM). Colocated with the ACM Conf. on Computer and Communications Security (CCS)*. ACM, Berlin, Germany, 25–36. <https://doi.org/10.1145/2517881.2517886>
- [5] Andrew Bortz and Dan Boneh. 2007. Exposing private information by timing web applications. In *Proc. of the Int'l Conf. on World Wide Web (WWW)*. ACM, Banff, Alberta, Canada, 621. <https://doi.org/10.1145/1242572.1242656>
- [6] Rahul Chatterjee, Periwinkle Doerfler, Hadas Orgad, Sam Havron, Jackeline Palmer, Diana Freed, Karen Levy, Nicola Dell, Damon McCoy, and Thomas Ristenpart. 2018. The Spyware Used in Intimate Partner Violence. In *Proc. of the IEEE Symp. on Security and Privacy (S&P)*. IEEE, San Francisco, CA, 441–458. <https://doi.org/10.1109/SP.2018.00061>
- [7] Eugene Cho, Jinyoung Kim, and S. Shyam Sundar. 2020. Will You Log into Tinder using your Facebook Account? Adoption of Single Sign-On for Privacy-Sensitive Apps. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI) - Extended Abstracts*. ACM, Honolulu, HI, USA, 1–7. <https://doi.org/10.1145/3334480.3383074>
- [8] Yana Dimova, Tom Van Goethem, and Wouter Joosen. 2023. Everybody's Looking for SSOmething: A large-scale evaluation on the privacy of OAuth authentication on the web. *Proceedings on Privacy Enhancing Technologies* (2023). <https://petsymposium.org/popets/2023/popets-2023-0119.php>
- [9] Verena Distler. 2023. The Influence of Context on Response to Spear-Phishing Attacks: an In-Situ Deception Study. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI) (CHI '23)*. ACM, New York, NY, USA, 1–18. <https://doi.org/10.1145/3544548.3581170>
- [10] D. Dittrich and E. Kenneally. 2012. *The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research*. Technical Report. U.S. Department of Homeland Security.
- [11] Vincent Drury and Ulrike Meyer. 2020. No Phishing With the Wrong Bait: Reducing the Phishing Risk by Address Separation. In *Proc. of the IEEE European Symp. on Security and Privacy Workshops (EuroS&PW)*. 646–652. <https://doi.org/10.1109/EuroSPW51379.2020.00093>
- [12] Serge Egelman. 2013. My profile is my password, verify me!: the privacy/convenience tradeoff of facebook connect. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI)*. ACM, Paris, France, 2369–2378. <https://doi.org/10.1145/2470654.2481328>
- [13] Serge Egelman, Marian Harbach, and Eyal Peer. 2016. Behavior Ever Follows Intention?: A Validation of the Security Behavior Intentions Scale (SeBIS). In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI)*. ACM, San Jose, CA, USA, 5257–5261. <https://doi.org/10.1145/2858036.2858265>
- [14] Diana Freed, Sam Havron, Emily Tseng, Andrea Gallardo, Rahul Chatterjee, Thomas Ristenpart, and Nicola Dell. 2019. "Is my phone hacked?" Analyzing Clinical Computer Security Interventions with Survivors of Intimate Partner Violence. *Proceedings of the ACM on Human-Computer Interaction* 3, CSCW (Nov. 2019), 1–24. <https://doi.org/10.1145/3359304>
- [15] Mohammad Ghasemisharif, Chris Kanich, and Jason Polakis. 2022. Towards Automated Auditing for Account and Session Management Flaws in Single Sign-On Deployments. In *Proc. of the IEEE Symp. on Security and Privacy (S&P)*. IEEE, San Francisco, CA, USA, 1774–1790. <https://doi.org/10.1109/SP46214.2022.9833753>
- [16] Hackspaining. 2022. Avoiding User Enumeration. (2022). <https://www.hackspaining.com/prevention/user-enumeration> \lastvisited.
- [17] Ayako Akiyama Hasegawa, Takuya Watanabe, Eitaro Shioji, and Mitsuaki Akiyama. 2019. I know what you did last login: inconsistent messages tell existence of a target's account to insiders. In *Proc. of the Annual Computer Security Applications Conference (ACSAC)*. ACM, San Juan, Puerto Rico, USA, 732–746. <https://doi.org/10.1145/3359789.3359832>
- [18] Sam Havron, Diana Freed, Rahul Chatterjee, Damon McCoy, Nicola Dell, and Thomas Ristenpart. 2019. Clinical computer security for victims of intimate partner violence. In *Proc. of the USENIX Security Symposium (USENIX Security)*. USENIX Association, Santa Clara, CA, 105–122. <https://www.usenix.org/conference/usenixsecurity19/presentation/havron>
- [19] Hang Hu, Peng Peng, and Gang Wang. 2019. Characterizing Pixel Tracking through the Lens of Disposable Email Services. In *Proc. of the IEEE Symp. on Security and Privacy (S&P)*. IEEE, San Francisco, CA, USA, 365–379. <https://doi.org/10.1109/SP.2019.00033>
- [20] Ruogu Kang, Laura Dabbish, Nathaniel Fruchter, and Sara Kiesler. 2015. "My Data Just Goes Everywhere:" User Mental Models of the Internet and Implications for Privacy and Security. In *Proc. of the Symp. On Usable Privacy and Security (SOUPS)*. 39–52. <https://www.usenix.org/conference/soups2015/proceedings/presentation/kang>
- [21] Jinwoo Kim, Kuyju Kim, Junsung Cho, Hyoungshick Kim, and Sebastian Schrittwieser. 2017. Hello, Facebook! Here Is the Stalkers' Paradise!: Design and Analysis of Enumeration Attack Using Phone Numbers on Facebook. In *Proc. of the Int'l Conf. on Information Security Practice and Experience (ISPEC)*, Vol. 10701. Springer International Publishing, Cham,


























































- 663–677. [https://doi.org/10.1007/978-3-319-72359-4\\_41](https://doi.org/10.1007/978-3-319-72359-4_41)
- [22] K. Krombholz, K. Busse, K. Pfeffer, M. Smith, and E. von Zezschwitz. 2019. "If HTTPS Were Secure, I Wouldn't Need 2FA" - End User and Administrator Mental Models of HTTPS. In *Proc. of the IEEE Symp. on Security and Privacy (S&P)*. IEEE, San Francisco, CA, USA, 1138–1155. <https://doi.org/10.1109/SP.2019.00060>
- [23] Leona Lassak, Philipp Markert, Maximilian Golla, Elizabeth Stobert, and Markus Dürmuth. 2024. A Comparative Long-Term Study of Fallback Authentication Schemes. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI) (CHI'24)*.
- [24] Karen Levy and Bruce Schneier. 2020. Privacy threats in intimate relationships. *Journal of Cybersecurity* 6, 1 (Jan. 2020), tyaa006. <https://doi.org/10.1093/cybsec/tyaa006>
- [25] Naresh K. Malhotra, Sung S. Kim, and James Agarwal. 2004. Internet Users' Information Privacy Concerns (IUIPC): The Construct, the Scale, and a Causal Model. *Information Systems Research* 15, 4 (Dec. 2004), 336–355. <https://doi.org/10.1287/isre.1040.0032>
- [26] Ioana Andreea Marin, Pavlo Burda, Nicola Zannone, and Luca Allodi. 2023. The Influence of Human Factors on the Intention to Report Phishing Emails. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI) (CHI '23)*. ACM, New York, NY, USA, 1–18. <https://doi.org/10.1145/3544548.3580985>
- [27] Philipp Markert, Leona Lassak, Maximilian Golla, and Markus Dürmuth. 2024. Understanding Users' Interaction with Login Notifications. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI) (CHI'24)*. <https://doi.org/10.1145/3613904.3642823> arXiv:2212.07316 [cs].
- [28] Tenga Matsuura, Ayako A. Hasegawa, Mitsuaki Akiyama, and Tatsuya Mori. 2021. Careless Participants Are Essential for Our Phishing Study: Understanding the Impact of Screening Methods. In *Proc. of the European Symp. on Usable Security (EuroUSEC) (EuroUSEC '21)*. ACM, New York, NY, USA, 36–47. <https://doi.org/10.1145/3481357.3481515>
- [29] Tara Matthews, Kathleen O'Leary, Anna Turner, Manya Sleeper, Jill Palzkill Woelfer, Martin Shelton, Cori Manthorne, Elizabeth F. Churchill, and Sunny Consolvo. 2017. Stories from Survivors: Privacy & Security Practices when Coping with Intimate Partner Abuse. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI)*. ACM, Denver, CO, USA, 2189–2201. <https://doi.org/10.1145/3025453.3025875>
- [30] Allison McDonald, Carlo Sugatan, Tamy Guberek, and Florian Schaub. 2021. The Annoying, the Disturbing, and the Weird: Challenges with Phone Numbers as Identifiers and Phone Number Recycling. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI)*. ACM, Yokohama Japan, 1–14. <https://doi.org/10.1145/3411764.3445085>
- [31] Gregory D. Moody, Dennis F. Galletta, and Brian Kimball Dunn. 2017. Which phish get caught? An exploratory study of individuals' susceptibility to phishing. *European Journal of Information Systems* 26, 6 (Nov. 2017), 564–584. <https://doi.org/10.1057/s41303-017-0058-x>
- [32] Srivathsan G. Morkonda, Sonia Chiasson, and Paul C. van Oorschot. 2023. Influences of Displaying Permission-related Information on Web Single Sign-On Login Decisions. (Aug. 2023). <https://doi.org/10.48550/arXiv.2308.13074> arXiv:2308.13074 [cs].
- [33] Srivathsan G. Morkonda, Sonia Chiasson, and Paul C. van Oorschot. 2023. "Sign in with ... Privacy": Timely Disclosure of Privacy Differences among Web SSO Login Options. (Aug. 2023). <https://doi.org/10.48550/arXiv.2209.04490> arXiv:2209.04490 [cs].
- [34] pagvac. 2007. Username Enumeration Vulnerabilities. (April 2007). <https://www.gnucitizen.org/blog/username-enumeration-vulnerabilities/> \lastvisited.
- [35] Elissa M. Redmiles, Ziyun Zhu, Sean Kross, Dhruv Kuchhal, Tudor Dumitras, and Michelle L. Mazurek. 2018. Asking for a Friend: Evaluating Response Biases in Security User Studies. In *Proc. of the ACM Conf. on Computer and Communications Security (CCS)*. ACM, Toronto, ON, Canada, 1238–1255. <https://doi.org/10.1145/3243734.3243740>
- [36] Kristen Roggemann, Galia Nurko, and Alexandra Tyers-Chowdhury. 2020. *User Perceptions of trust and Privacy on the Internet*. Technical Report. DAI's Center for Digital Acceleration. 48 pages.
- [37] Johnny Saldaña. 2021. *The coding manual for qualitative researchers* (4th ed ed.). SAGE Publishing, Thousand Oaks.
- [38] Dafydd Stuttard. 2007. Preventing username enumeration. (April 2007). <https://portswigger.net/blog/preventing-username-enumeration> \lastvisited.
- [39] San-Tsai Sun, Eric Pospisil, Ildar Muslukhov, Nuray Dindar, Kirstie Hawkey, and Konstantin Beznosov. 2013. Investigating Users' Perspectives of Web Single Sign-On: Conceptual Gaps and Acceptance Model. *ACM Transactions on Internet Technology* 13, 1 (Nov. 2013), 2:1–2:35. <https://doi.org/10.1145/2532639>
- [40] Anne Clara Tally, Jacob Abbott, Ashley M Bochner, Sanchari Das, and Christena Nippert-Eng. 2023. Tips, Tricks, and Training: Supporting Anti-Phishing Awareness among Mid-Career Office Workers Based on Employees' Current Practices. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI) (CHI '23)*. ACM, New York, NY, USA, 1–13. <https://doi.org/10.1145/3544548.3580650>
- [41] Emily Tseng, Rosanna Bellini, Nora McDonald, Matan Danos, Rachel Greenstadt, Damon McCoy, Nicola Dell, and Thomas Ristenpart. 2020. The Tools and Tactics Used in Intimate Partner Surveillance: An Analysis of Online Infidelity

Forums. In *Proc. of the USENIX Security Symposium (USENIX Security)*. 18.

- [42] Rui Wang, Shuo Chen, and XiaoFeng Wang. 2012. Signing Me onto Your Accounts through Facebook and Google: A Traffic-Guided Security Study of Commercially Deployed Single-Sign-On Web Services. In *Proc. of the IEEE Symp. on Security and Privacy (S&P)*. IEEE, San Francisco, CA, USA, 365–379. <https://doi.org/10.1109/SP.2012.30>
- [43] Rick Wash, Emilee Rader, and Chris Fennell. 2017. Can People Self-Report Security Accurately? Agreement Between Self-Report and Behavioral Measures. In *Proc. of the ACM Conf. on Human Factors in Computing Systems (CHI) (CHI '17)*. ACM, New York, NY, USA, 2228–2232. <https://doi.org/10.1145/3025453.3025911>

## A RAW EXPERIMENTAL DATA

Table 1. Raw experimental data. The  icon denotes the stealthiness of the (successful) attack and the  icon denotes the presence of CAPTCHAs for the corresponding form. The possibility to delete one’s own account is denoted with icon ; the  icon denotes the fact that it can be done only by e-mail (no online form). The presence of the “Login with...” (i.e., single sign-on) feature is denoted with the icon of the corresponding identity provider (i.e., **f**, **G**, or **🍏**).

service	login			password reset			account creation				f	G	🍏
 20min	✓	✓	✗		✓	✓	✗		✓	✗	✓	✓	✓
 Airbnb	✓	✓	✗		✓	✗	✗		✓	✓	✗	✓	✓
 Amavita	✓	✗	✗		✓	✗	✓		✗		✓	✓	✓
 Amazon	✓	✓	✗		✓	✓	✗		✓	✗	✓	✓	✗
 Autoscout24	✓	✗	✗		✓	✗	✗		✓	✗	✓	✓	✗
 Badoo	✗	✗	✗		✓	✗	✓		✓	✗	✓	✓	✗
 BettyBossi	✗	✗	✗		✓	✗	✗		✓	✗	✗	✗	✗
 Blick	✓	✓	✗		✓	✗	✗		✓	✗	✗	✗	✗
 Bon Prix	✓	✗	✓		✗	✗	✗		✓	✗	✗	✗	✗
 Booking	✓	✓	✗		✓	✗	✗		✓	✗	✓	✓	✓
 Comparis	✓	✗	✗		✓	✗	✗		✓	✗	✓	✓	✓
 Coop At Home	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
 DPD	✓	✗	✗		✓	✗	✗		✓	✗	✓	✓	✓
 Decathlon	✓	✓	✗		✓	✗	✗		✓	✗	✓	✓	✓
 DeepL	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
 Doctissimo	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
 Doodle	✗	✗	✗		✗	✗	✗		✓	✗	✓	✓	✗
 Facebook	✓	✓	✗		✓	✓	✗		✓	✗	✗	✗	✗
 Flickr	✓	✓	✗		✓	✓	✗		✓	✗	✗	✗	✗
 Galaxus	✗	✗	✗		✗	✗	✗		✓	✗	✓	✓	✓
 Hornbach	✓	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
 IMDB	✓	✓	✗		✓	✗	✓		✓	✗	✓	✓	✓
 Ifolor	✗	✗	✗		✓	✗	✗		✓	✗	✗	✗	✗
 Ikea	✓	✓	✗		✗	✗	✓		✓	✗	✓	✓	✓
 Immoscout24	✓	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
 InshAllah	✓	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
 Instagram	✓	✓	✗		✓	✗	✗		✓	✗	✓	✓	✗
 JW	✗	✗	✗		✗	✗	✓		✓	✗	✓	✓	✗
 Jeuxvidéo	✗	✗	✓		✗	✗	✓		✓	✗	✗	✗	✗
 Jobs	✓	✓	✗		✓	✗	✗		✓	✗	✗	✗	✗
 Jobup	✓	✓	✗		✓	✗	✗		✓	✗	✗	✗	✗
 Journal des Femmes	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
 JustEat	✗	✗	✗		✗	✗	✗		✓	✗	✓	✓	✗
 Landi	✓	✓	✗		✗	✗	✗		✓	✗	✗	✗	✗
 Loterie Romande	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
 Migros Online	✗	✗	✗		✓	✗	✗		✓	✗	✗	✗	✗
 Moneyhouse	✓	✗	✗		✓	✗	✗		✓	✗	✗	✗	✗
 Netflix	✓	✓	✗		✓	✗	✗		✓	✗	✓	✓	✓
 Obi	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
 Ochsnersport	✗	✗	✗		✓	✗	✗		✓	✗	✗	✗	✗
 Orel Füssli	✗	✗	✗		✓	✗	✗		✓	✗	✗	✗	✗
 Outdooractive	✓	✓	✗		✓	✗	✗		✓	✗	✓	✓	✓
 Paypal	✗	✗	✗		✓	✗	✗		✓	✗	✗	✗	✗
 Playstation	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
 Pornhub	✗	✗	✗		✗	✗	✓		✓	✗	✗	✗	✗
 Post.ch	✗	✗	✗		✗	✗	✓		✓	✗	✗	✗	✗
 Qualipet	✗	✗	✗		✗	✗	✓		✓	✗	✗	✗	✗
 RTS	✗	✗	✗		✓	✗	✓		✓	✗	✓	✓	✓
 Ricardo	✗	✗	✗		✓	✗	✓		✓	✗	✗	✗	✗
 SBB	✗	✗	✗		✓	✗	✗		✓	✗	✗	✗	✗
Shop Apotheke	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
Shutterstock	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
Spotify	✗	✗	✗		✓	✗	✓		✓	✗	✓	✓	✓
Statista	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
Swisslos	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
Tripadvisor	✗	✗	✗		✗	✗	✗		✓	✗	✓	✓	✗
Twitch	✗	✗	✗		✗	✗	✓		✓	✗	✓	✓	✗
Twitter	✓	✓	✗		✓	✗	✓		✓	✗	✗	✓	✓
Wikipedia	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
Wuerth	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗
Xhamster	✓	✓	✗		✓	✗	✗		✓	✗	✗	✓	✗
Zalando	✗	✗	✗		✓	✗	✗		✓	✗	✗	✗	✗
Zooplus	✗	✗	✗		✗	✗	✗		✓	✗	✗	✗	✗



## B SURVEY QUESTIONNAIRE

Here we provide the complete transcript of the survey questionnaire.

Note: Coding rules are marked in gray (not visible to respondents)

### Sec. A: Consent Form

You are invited to participate in the **Survey on Online Accounts**, research conducted by the by the Information Security and Privacy Lab at the University of Lausanne (UNIL) for **improving the security of online services**. Thank you for your participation!

You will need about **7-10 minutes** to complete the questionnaire.

Your participation is entirely voluntary and you can withdraw your participation at any time during the experiment without specifying a reason: Your data will be deleted but you will not be paid if you stop your participation during the experiment. Your responses will be collected by the researchers in a confidential and anonymous way. If the results of this study will be published, we will do so in an anonymous manner. If you have any questions about the research, do not hesitate to email us at: [ux-research@unil.ch](mailto:ux-research@unil.ch) By signing this (By clicking on "I agree"), you declare that you have read the preceding information and agree to participate in this study.

- (1) [Single Selection.] Do you agree with the terms above?
- (a) I agree
  - (b) I do not Agree [Terminate.]

### Sec. B: Services

- (2) [Multiple selection. Order randomized.] Please select all the online services on which you have an account.
- (a) Netflix
  - (b) Immoscout24
  - (c) Instagram
  - (d) Galaxus
  - (e) Migros Online
  - (f) Ricardo
  - (g) IKEA
  - (h) Zalando
  - (i) Loterie Romande
  - (j) Doctissimo
  - (k) Pornhub
  - (l) Badoo
  - (m) Magic X Erotic Shop
  - (n) Grindr
  - (o) UDC (Union démocratique du centre)
  - (p) None of the above
- (3) [Single Selection.] How many different e-mail addresses do you have?
- (a) 1
  - (b) 2
  - (c) 3 or more

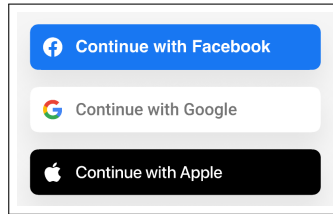
[Display Q4 if Q3\$!=1]

- (4) [Single Selection.] How many different e-mail addresses do you **use to register and login to different online services** (e.g., using your Swisscom e-mail address to register and login to IKEA and your Hotmail address for Netflix)?

- (a) 1
- (b) 2 [Display this choice if Q3\$=2 or Q3\$=3 or more]
- (c) 3 or more [Display this choice if Q3\$=3 or more]

[Display Q5 if Q4\$=2 or Q4\$=3 or more]

- (5) [Open-ended.] Why do you register to different online services with different e-mail addresses?  
[ ](text field)
- (6) [Single Selection.] In the last 6 months, how many times did you use the "Sign in with ..." (e.g., Apple, Facebook, Google, Microsoft, Twitter) feature to register online accounts?
- (a) 0
  - (b) 1–2
  - (c) 3–4
  - (d) 5–6
  - (e) 7+



### Sec. C: Privacy

- (7) [Grid question. Row order randomized.] From a privacy perspective, how concerned would you be if the following entities could know the list of the online services on which you have an account?

Row options:

- Companies
- Governments
- Friends
- Colleagues
- Employer
- Relatives
- Intimate partner

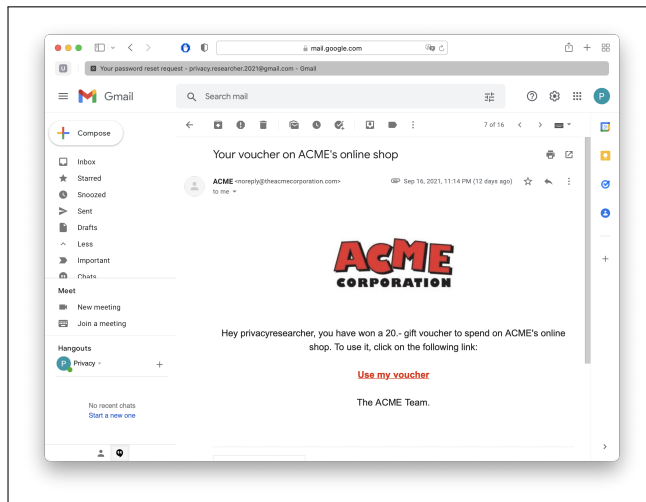
- (a) Not at all concerned
- (b) Slightly concerned
- (c) Moderately concerned
- (d) Very concerned
- (e) Extremely concerned

- (8) [Single Selection.] Do you use any strategy for preventing others to know about the existence of (some of) your online accounts?
- (a) Yes
  - (b) No
  - (c) Not sure

[Display Q9 if Q8\$=Yes]

- (9) [Multiple selection. Order randomized.] Please select all the strategies that you use for preventing others to know about (some of) your online accounts.
- (a) Use an e-mail address that nobody knows
  - (b) Use a disposable e-mail address (e.g., YOPmail)
  - (c) Use a proxy
  - (d) Delete the account
  - (e) Connect only from a personal device (e.g., computer, smartphone, tablet)
  - (f) Other (please specify) - [ ](text field)

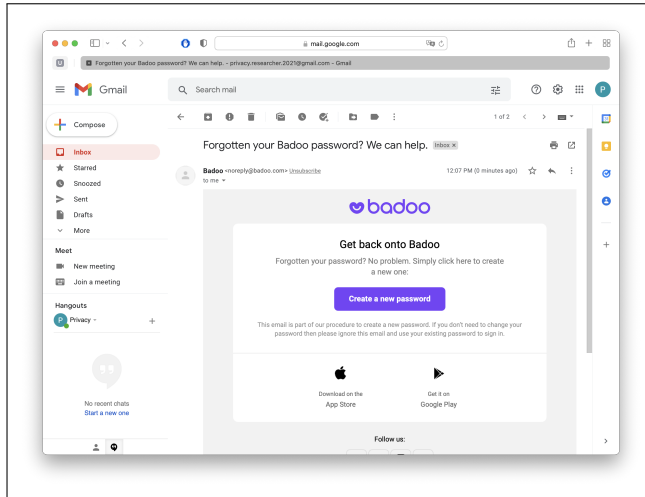
**Sec. D: Phishing**



Please look at the screenshot above. Imagine that you received this e-mail today.

- (10) [Single Selection.] How likely would you click on the link provided in this e-mail **if you do have an account** on this online service?
- (a) Very unlikely
  - (b) Unlikely
  - (c) Somewhat unlikely
  - (d) Neutral (neither likely nor unlikely)
  - (e) Somewhat likely
  - (f) Likely
  - (g) Very likely
- (11) [Single Selection.] How likely would you click on the link provided in this e-mail **if you do not have an account** on this online service?
- (a) Very unlikely
  - (b) Unlikely
  - (c) Somewhat unlikely
  - (d) Neutral (neither likely nor unlikely)
  - (e) Somewhat likely
  - (f) Likely

(g) Very likely



[Display this image if Q2\$!=Badoo] Please look at the screenshot above. Imagine that **you do have an account on this service** and that you received this e-mail today **even though you did not make a request to reset your password.**

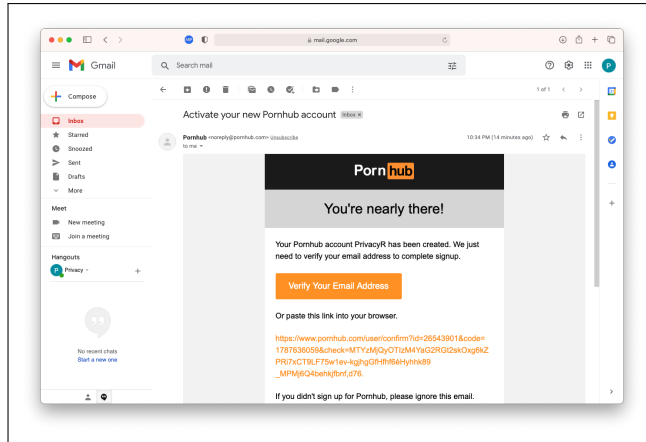
[Display this image if Q2\$Badoo] Please look at the screenshot above. Imagine that you received this e-mail today **even though you did not make a request to reset your password.**

- (12) [Open-ended.] In your opinion, what could have led to the sending of this e-mail?  
[ ](text field)
- (13) [Single selection.] Would you take any action after you have received this e-mail?  
(a) No  
(b) Yes (please specify) - [ ](text field)
- (14) [Single Selection.] In the last 6 months, how many times did you receive an **unsolicited** e-mail similar to the one above for one of your existing accounts?  
(a) 0  
(b) 1–2  
(c) 3–4  
(d) 5–6  
(e) 7+

[Display Q15 if Q4\$!=1]

- (15) [Single Selection.] How likely would you check whether the e-mail above was sent to the e-mail address that you have used to register your existing account?  
(a) Very unlikely  
(b) Unlikely  
(c) Somewhat unlikely  
(d) Neutral (neither likely nor unlikely)  
(e) Somewhat likely  
(f) Likely

(g) Very likely



[Display this image if Q2\$!=Pornhub] Please look at the screenshot above. Imagine that you received this e-mail today **even though you did not make a request to create an account.**

[Display this image if Q2\$Pornhub] Please look at the screenshot above. Imagine that you **do not have an account on this service** and that you received this e-mail today **even though you did not make a request to create an account.**

- (16) [Open-ended.] In your opinion, what could have led to the sending of this e-mail?  
[ ](text field)
- (17) [Single selection.] Would you take any action after you have received this e-mail?  
(a) No  
(b) Yes (please specify) - [ ](text field)
- (18) [Single Selection.] In the last 6 months, how many times did you receive an **unsolicited** e-mail similar to the one above on one of your e-mail addresses?  
(a) 0  
(b) 1–2  
(c) 3–4  
(d) 5–6  
(e) 7+

**Sec. E: IUIPC**

- (19) [Grid question. Row order randomized.] Please rate the following statements.

Row options:

- All things considered, the Internet would cause serious privacy problems.
- Compared to others, I am more sensitive about the way online companies handle my personal information.
- To me, it is the most important thing to keep my privacy intact from online companies.
- I believe other people are too much concerned with online privacy issues.
- Compared with other subjects in my mind, personal privacy is very important.
- I am concerned about threats to my personal privacy today.

- (a) Strongly disagree
- (b) Somewhat disagree
- (c) Neither disagree nor agree
- (d) Somewhat agree
- (e) Strongly agree