Association for Information Systems

AIS Electronic Library (AISeL)

ICIS 2022 Proceedings

Cybersecurity, Privacy and Ethics in AI

Dec 12th, 12:00 AM

Weaponizing the GDPR: How Flawed Implementations Turn the Gold Standard for Privacy Laws into Fool's Gold

Alexander Gladis
RWTH Aachen University, gladis@time.rwth-aachen.de

Nicole Janine Hartwich
RWTH Aachen University, hartwich@time.rwth-aachen.de

Oliver Salge
RWTH Aachen University, salge@time.rwth-aachen.de

Follow this and additional works at: https://aisel.aisnet.org/icis2022

Recommended Citation

Gladis, Alexander; Hartwich, Nicole Janine; and Salge, Oliver, "Weaponizing the GDPR: How Flawed Implementations Turn the Gold Standard for Privacy Laws into Fool's Gold" (2022). *ICIS 2022 Proceedings*. 7.
https://aisel.aisnet.org/icis2022/security/security/7

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICIS 2022 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

Weaponizing the GDPR: How Flawed Implementations Turn the Gold Standard for Privacy Laws into Fool's Gold

Completed Research Paper

Alexander Gladis, Nicole Hartwich, Torsten-Oliver Salge
Institute for Technology and Innovation Management, RWTH Aachen University
https://www.rwth-aachen.de/digital-responsibility-lab
{gladis, hartwich, salge}@time.rwth-aachen.de

Abstract

Despite its ambitious goals of protecting personal data and generally being well-received, the General Data Protection Regulation (GDPR) can be exploited for identity theft by weaponizing subject access requests (SARs). To understand this threat and investigate the impact of victims' privacy awareness and public exposure on its effectiveness, we selected three victims – highly privacy aware person, average user, and semipublic figure – and tasked six realistic attackers with stealing their personal data. Based on 718 submitted SARs, we provide novel insights from a realistic case study of a law being weaponized and advance the understanding of GDPR-based identity theft by demonstrating its practical viability. Further, we derive patterns from common flaws observed in SAR handling processes, and explore threat mitigation options for individuals, organizations, and lawmakers. Generalizing our findings, we uncover approaches for cybersecurity researchers to probe further laws for flaws.

Keywords: GDPR, subject access request, social engineering, identity theft, cybersecurity

Introduction

The General Data Protection Regulation (GDPR) law (European Union, 2016) set a precedent as the most ambitious attempt at regulating the collection, recording, storage, and processing of personal data not just in the European Union (EU) but worldwide. Guiding how organizations approach privacy and cybersecurity, the law is regarded as the current gold standard for privacy laws by many (e.g., Albrecht, 2016; Andrew & Baker, 2021). It requires most organizations, including extraterritorial ones, that target European citizens to appoint a data protection officer (DPO), rethink their data collection practices, and adhere to a range of obligations with regards to data security and data processing transparency.

Paradoxically, despite these ambitious goals of protecting personal data and generally being well-received, there is initial evidence that the GDPR can be abused as a weapon for identity theft. Previous research indicates that flawed implementations of the law by organizations can lead to unintended cybersecurity implications, potentially allowing access to personal data without authorization (e.g., Bufalieri et al., 2020; Di Martino et al., 2019). The primary focus of these studies was the right of access (Art. 15), granting EU citizens (*data subjects*) the right to submit a so-called subject access request (SAR) to any organization (*data controller*) affected by the GDPR. In doing so, a data subject is granted the right to request any personal data stored about themselves from the data controller, who in turn is obliged to verify the data subject's identity. By impersonating their victim and sending spoofed SARs to organizations with flawed implementations of this identity verification, the researchers in the aforementioned studies were able to exfiltrate personal data without breaching the targeted organizations' technical cybersecurity systems. Such an attack is classified as "social engineering" in cybersecurity terminology (Wang et al., 2020).

While these studies serve as a promising proof of concept for SAR identity theft, they either focus on isolated components of such an attack rather than the whole process, or provide evidence for its practical viability only within certain restrictive assumptions. Additionally, they leave a detailed analysis of patterns regarding especially successful attacker strategies as well as typical flaws in processes implemented by organizations for future research. We extend this important line of research by conducting three extensive case studies of simulated SAR identity thefts under realistic conditions. This way, we answer the following research questions: Is SAR identity theft feasible under real-world conditions, e.g., by an attacker that has no prior knowledge about the victim? If so, how is the attack affected by the victim's privacy characteristics, what damage could the malicious actor inflict, and what are prominent attack strategies and response patterns?

In order to understand the impact of privacy awareness and preferences as well as public exposure on the effectiveness of SAR identity theft, we selected three structurally distinct victims to be attacked, representing a highly privacy aware person, an average user, and a semipublic figure. A team of six attackers was tasked with aggregating as much personal data as possible about their three victims (all German citizens) within a three-month window. Without any practical experience in social engineering, possessing no prior knowledge about their victims except name and workplace, and bound by certain ethical and legal considerations, our attackers were as weak as reasonably assumable in a realistic scenario – meaning that just about anyone is capable of replicating their attacks. Yet, even under these circumstances, they were able to exfiltrate a broad range of sensitive personal data on our three victims, including home address, phone numbers, utility bills, national identity card and bank account information, as well as loan financing and insurance data.

This work contributes to social engineering research by providing novel insights from a realistic case study of a law being used as a sword rather than the shield it was supposed to be. Additionally, we advance the understanding of GDPR-based identity theft attacks not only by demonstrating their practical viability, but also by deriving patterns from common flaws in SAR handling processes that we observed. Generalizing these observations on "law hacking", we uncover approaches for cybersecurity researchers to probe further laws for flaws. For example, we discover a "weakest link" effect as a systemic weakness in the GDPR resulting from attacking the system of organizations as a whole rather than individual organizations in isolation.

We contribute to practice by establishing a lower boundary for the damage potential (and hence threat) of SAR identity theft in the real world, thus overcoming limitations of prior studies and uncovering immediate need for action by lawmakers and organizations alike. Based on the patterns and effects observed in our study, we explore options for how organizations can improve processes and reduce susceptibility to such attacks. From a data subject's perspective, we investigate how privacy characteristics affect the feasibility and impact of SAR identity theft, providing insights into how individuals can (partially) mitigate this threat.

Conceptual Background

The GDPR was adopted by the EU in 2016 and came into effect in May 2018, replacing the 1995 Data Protection Directive (European Union, 2016). It has influenced the design of privacy laws worldwide, such as the California Consumer Privacy Act (CCPA) of 2018. Among others, the law grants all European citizens broad control over how their personal data is collected and processed, including but not limited to the right of access (Art. 15), the right to rectification (Art. 16), and the right to erasure (Art. 17).

In our study, the right of access will be weaponized through means of social engineering techniques in order to fool organizations into disclosing personal data to an unauthorized adversary. Social engineering can be defined as "a type of attack wherein the attacker(s) exploit human vulnerabilities by means of social interaction to breach cybersecurity, with or without the use of technical means and technical vulnerabilities" (Wang et al., 2020). It is an increasingly widespread tool for malicious actors in general, being experienced by 85% of organizations in 2018 (Accenture Security, 2019). Popularity and effectiveness of such attacks were boosted even further by the COVID-19 pandemic, for example because many knowledge workers worked from home and thus relied more on digital rather than face-to-face communication (Naidoo, 2020).

The most well-known and well-researched representative of social engineering attacks is phishing, an omnipresent threat to organizations especially when knowledge workers are targeted (Krombholz et al., 2015). Studies show that susceptibility to phishing attacks is mainly determined by two key factors: On the one hand, it is affected by innate traits of the victim, such as personality factors (Moody et al., 2017) or behavioral factors (Wright & Marett, 2010). On the other hand, attackers can increase susceptibility through social engineering approaches such as impersonating a credible source (Algarni et al., 2017), appropriately contextualizing their interaction with the victim (Goel et al., 2017), or making use of influence techniques (Wright et al., 2014). Jaeger and Eckhardt (2021) highlight that susceptibility to phishing is not static. Rather, the victim's situational information security awareness on a case-by-case basis must also be taken into account.

Similar to phishing attacks, recent research indicates that social engineering techniques could also be used to exploit flawed implementations of the GDPR right of access, manifested in the process of handling SARs, to access personal data stored by organizations without authorization. For example, Cagnazzo et al. (2019) managed to fool 10 out of 14 companies into leaking personal data by sending spoofed SARs pretending to be another data subject. In a broader study, Bufalieri et al. (2020) found that more than half of over 300 data controllers had flawed authentication or data exchange procedures in place when handling SARs. As shown by Pavur and Knerr (2019) and Di Martino et al. (2019), a malicious actor could theoretically exploit these flaws for a novel form of identity theft which, paradoxically, is only possible because of the GDPR.

However, to the best of our knowledge, no study has yet demonstrated the viability of such identity theft attacks under real-world conditions. For example, whilst providing valuable insights into how identity verification is implemented by organizations, Di Martino et al. (2019) knew in advance which organizations stored data on their victim and submitted SARs only to such organizations. This simplification does not reflect reality, where a malicious actor committing the identity theft would likely lack this prior knowledge.

Further, a more comprehensive understanding of typical patterns in attacker behavior and organizational processes facilitating SAR identity theft is required in order to work towards mitigating the threat. Additionally, evaluating how structural differences in the victims' privacy characteristics affect such attacks might yield insights into how individuals can protect themselves against them.

Methods

To explore the phenomenon of SAR identity theft and answer our research questions, we embraced a multiple case study design according to Yin (2009). For this purpose, we tasked a team of six attackers with executing such attacks on three victims by weaponizing the GDPR through social engineering.

Research Design

The primary goal of our case study was to simulate identity theft through weaponizing SARs – impersonating the victim and sending illegitimate ones in their name –, resembling a real-world attack scenario as closely as possible. This allowed us to study the handling of SARs by organizations and to investigate the potential damage an attacker could have caused. Further, simulating the attack under realistic conditions facilitated mitigating potential biases, such as the risk of inadvertently deploying a-priori knowledge due to the victim being well-known to the adversary.

In order to understand if and how a victim's data privacy awareness and attitude affect the effectiveness of such attacks and hence derive potential mitigation mechanisms for data subjects, three individuals with structurally distinct privacy characteristics volunteered as victims for our case study. A team of six attackers was tasked with gathering as much sensitive personally identifiable information on these victims as possible within the time frame and operational constraints of the case study. This attack was simulated in four stages, taking place from November 2020 until February 2021. During the first stage, the attackers gathered initial data on their victims from openly available sources (e.g., social media). Afterwards, they went through three iterations of submitting spoofed SARs and evaluating the responses.

Realism of the Attack Scenario

With our research goal in mind of understanding if and what damage a real-world attacker could potentially cause through SAR identity theft, we tailored the design of our case study to replicate the capabilities of a realistic, yet severely constrained and weak attacker in order to establish a lower boundary for their threat.

In our study, the attackers and their victims were strangers prior to conducting the simulation, such that the simulated attackers had the same knowledge constraints that real ones would have. Similar to a real-world attack, the attackers were able to freely select targeted organizations, with restrictions only on the healthcare sector due to ethical concerns, as presented later. Furthermore, the victims took a passive role and did not provide any assistance or additional information to the attackers throughout the study. Due to legal and ethical constraints, the attackers were not informed about letters mailed to the victims and had no ability to intercept them, which we deemed likely to be the case in a real identity theft scenario as well. Similarly realistic, the attackers had no access to the victims' real email accounts or phone numbers at any time.

However, also because of legal considerations, the attackers were unable to falsify documents or scans and could not impersonate their victims in phone calls (from a fake number) for identity verification. Unlike our simulated attackers, a real-world adversary already in the process of committing identity theft might not hesitate to engage in such criminal activities. In their study, Di Martino et al. (2019) found that 8 out of the 15 organizations that fell for malicious SARs in total did so because the adversary provided an altered identity card. Hence, we believe that a significant percentage of the organizations from our sample could be fooled by forging a (redacted) scan of the data subject's passport or identity card, too. However, we decided that investigating this is not within the scope of this work and instead subject to future research[1].

A further operational constraint for the attackers was that their frequency of interaction with organizations was limited to once every 30 days by the iterative design of our SAR process, which will be presented later. As this limitation would not exist in a real-world setting, a malicious adversary would be able to react to responses by organizations more quickly and more often, likely improving their success chances.

Legal and Ethical Considerations

Aiming to protect the victims, the attackers, and the organizations targeted in our case study as well as the persons handling our SARs from harm, we derived a set of operational constraints for the design of our attack simulation. These were guided by legal restrictions and established standards for ethical research.

First, we required that the simulated victims (data subjects) were kept informed about the state of the attack, retained control over their personal data, and were able to withdraw their consent at all times. The simulated attackers (data requesters) required a legally binding guarantee that they could not be held liable for their attack as long as they followed rules mutually agreed upon with their victims a priori. Given that the attackers, by design, were unfamiliar with their victims prior to the simulated attack, we asked both parties to sign a contract outlining the case study design and establishing rules. Additionally, as doing so may violate German law, the attackers were prohibited from forging any documents, impersonating their victim in phone calls, or attempting to intercept mailed letters for the purpose of identity verification.

Second, we ensured that our study causes no harm to the organizations (data controllers) addressed in the adversaries' SARs. To prevent organizations from becoming liable to legal penalties according to the GDPR, we designed our SAR submission procedure in such a way that the data controllers technically did not transmit any data to an unauthorized third party, even if they would have done so in a real attack due to insufficiently verifying the data subject's identity. Furthermore, no incidents or flawed verification processes were reported to governmental data protection agencies. For the purpose of avoiding reputational damage to the affected organizations in our sample, we do not disclose their names nor give descriptions detailed enough to deduce their identity. Additionally, we took care to not disrupt any organization nor waste an overproportional amount of organizational resources. For example, the total number of SARs sent to a small sports club would have been restricted to one across all attack victims, whereas a large corporation could have received one request for each victim. SARs to healthcare professionals (e.g., local doctor's offices) were limited to a few instances to avoid overburdening their resources already strained by the COVID-19 pandemic.

[1] For a preliminary investigation of this assumption, we made a genuine black-and-white scan of one victim's German national identity card – using a low image resolution on purpose – and redacted all information except for name, date of birth, and address. Due to the poor scan quality and heavily redacted information, this image would have been trivial to forge even for an unsophisticated adversary. Furthermore, all visible data had previously been gathered from interactions with multiple organizations by our simulated attackers. Yet, this document sufficed to persuade an organization that had initially denied the SAR for identity verification reasons, even though that request had already contained all information not redacted in the scan. Whilst this is an interesting point of reference for future studies, the interaction took place after our simulated attacks had concluded and was thus not considered in our analysis.

Third, the individuals that improperly handled our attackers' SARs needed to be protected from repercussions (e.g., being penalized by their employer). From our outside perspective, we were unable to judge if flawed organizational policies or individual errors were at fault for a successful attack in some instances. Hence, we decided on a case-by-case basis to responsibly disclose vulnerabilities only if we were able to rule out individual error. When doing so, we emphasized constructive advice on improving policies rather than putting blame on the individuals that execute them. For organizations with an external DPO, we chose to address our disclosure to the external entity rather than the appointing organization. This way, they could improve their identity verification process without being at risk of having their appointment revoked.

Victim Personas and Privacy Characteristics

Our victims were selected based on structural differences in their privacy characteristics. We restrict our use of identifying information in this paper to protect their anonymity, e.g., by using gender-neutral pronouns.

VictimA: As a university professor, this person has interacted with a large number of organizations throughout their professional career and their private life. While aware of the resulting privacy implications, VictimA tends to disclose their real name, date of birth, and more information when signing up on websites. They use their work email and phone number for many such interactions with organizations, do not follow the principle of data minimization, and have little regard for recommended cybersecurity practices such as using a password manager and unique randomly generated passwords for each website. A variety of key identifiers (e.g., date and place of birth) as well as a detailed CV are publicly accessible on the Internet. Additionally, their research interests and parts of their professional network can be identified based on publications. VictimA has accounts on multiple social media platforms, however most information shared on there is not publicly visible. In our case study, they represent a semipublic figure with limited privacy awareness.

VictimB: Similar to VictimA, this person typically uses non-pseudonymized data when interacting with organizations and does not strictly practice data minimization. Hence, numerous organizations store VictimB's personal data. As a research associate, however, they are less publicly exposed and have published fewer key identifiers in their public CV. Despite being mostly non-public, their social media presences disclose some data such as their age (implying year of birth) and a few hobbies. For our case study, VictimB represents an average user with some privacy awareness, such as making most social media content private.

VictimC: Being proficient in cybersecurity and working as a research associate, this person is highly aware of their digital footprint and privacy, trying to minimize interactions with organizations that require disclosing personal data. VictimC uses pseudonymized data for interactions with organizations whenever possible and maintains multiple personal email addresses for creating accounts on websites, e.g., one email when it is necessary to give their real name, and a different one when using pseudonymized data. Further, they use random passwords and a password manager, and keep track of reported data breaches to react accordingly. When no longer interested in further interaction with an organization, VictimC makes use of the right to erasure (Art. 17 GDPR) to force the organization to delete their personal data. Their use of social media is limited to professional networks, where they take care to minimize disclosure of personal data. However, some details such as a profile picture and their higher education are publicly visible on social media and the website of their employer. Within our case study, VictimC represents a highly privacy aware person.

Initial Knowledge and Open-Source Intelligence

The attackers were provided with only their victims' name and workplace, reflecting a bare minimum set of identifying information a real-world adversary would possess. Prior to submitting SARs, they expanded their knowledge through open-source intelligence (OSINT), i.e., "intelligence that is produced from publicly available information and is collected, exploited, and disseminated in a timely manner to an appropriate audience for the purpose of addressing a specific intelligence requirement." (U.S. House. 109th Congress, 2006, Sec. 931). In doing so, the attackers scoured the victims' social media presences[2], their employers' websites, newspaper archives, and more for exploitable information. Data validity and integrity were established by conservatively filtering out ambiguous findings based on a combination of reasonable assumptions

[2] In order to not risk alarming the victims, this reconnaissance step was purely passive. For example, no friend requests were sent with the goal of gaining access to more sensitive data. Instead, only publicly visible information on social media profile pages was gathered.

(e.g., that the victim lives in a city close to their workplace) and cross-referencing data from multiple sources.

SAR Process

The design of our case study was guided by striking a balance in the trade-off between our goal of simulating a realistic attack by a malicious adversary on the one hand, and fulfilling the previously outlined legal and ethical requirements on the other hand. In doing so, we developed an iterative process for coordinating the SARs between attackers and victims, as depicted in Figure 1 and described in the following.

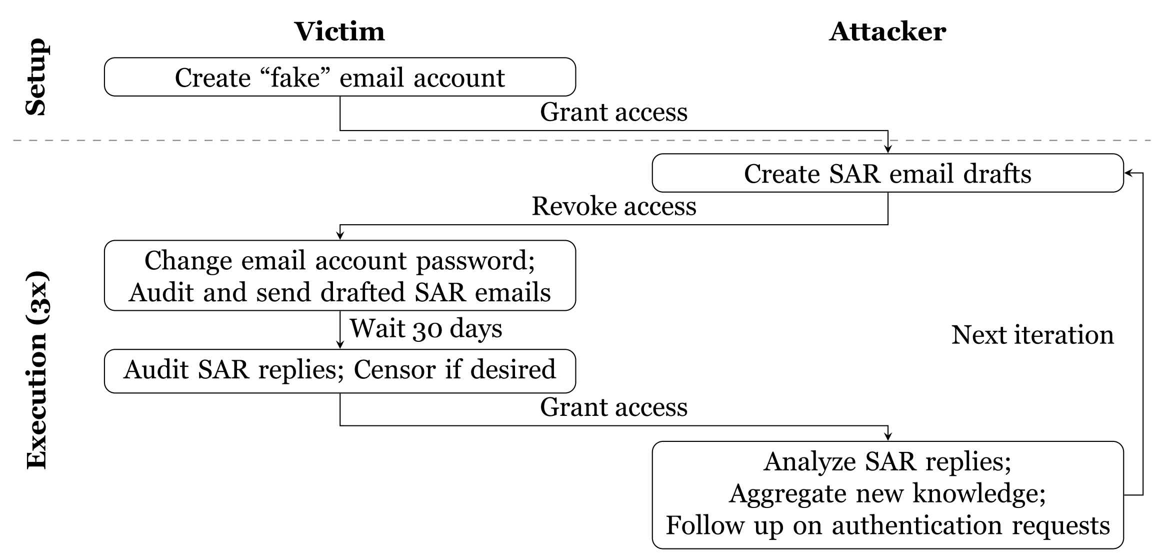

Figure 1. SAR Process of the Case Study

Before starting the first iteration of our simulated attack, each victim created a "fake" email account in their own name at a popular free email provider[3]. Having the victims create an email account themselves helped avoid potential legal implications for organizations, as any reply from targeted organizations (including those leaking data without properly verifying the identity) would be received by an account that, on paper, belongs to the true data subject. In a real identity theft scenario, the attackers would have created this email account in order to impersonate their victim. As a strategy commonly deployed in highly targeted phishing attacks, this helped deceive targeted organization into believing the legitimacy of SARs. Furthermore, this approach also provided the attackers with a plausible pretext for not having access to the victim's real email accounts, which might be known to the targeted organization and hence could be used for reasonably secure proof of identity. For example, if the addressed data controller sent a password to the victim's real email upon receiving the SAR, the adversaries – impersonating the victim – could ask for the password to be sent to the "new" email instead, pretending to have lost access to that account due to having been hacked.

At the start of each iteration, the victims granted the attackers access to their "fake" email, but never to their real email accounts. This does not negatively impact realism, as the attackers would have created this email account themselves in a real scenario and thus have full access to it. The attackers then had 24 hours to create drafts for SAR emails to be sent to organizations of their choosing, without being allowed to send any emails themselves. Organizations were selected using multiple criteria and approached in different ways, based on individual strategies developed by the attackers to best fit their current knowledge about each victim. Given the similarity between phishing and our simulated attack, the attackers (who had no practical experience in social engineering but read up on the basics) formulated their SARs using simple techniques shown to boost susceptibility to phishing, such as credible impersonation (Algarni et al., 2017), contextualization (Goel et al., 2017), and influence techniques (Wright et al., 2014). For example, they instilled a sense of urgency by emphasizing that replying within a month is mandatory. The attackers emailed SARs either in German or English, depending on the country of origin of the targeted organization, and addressed them to the organization's data protection inquiry email or, as a fallback solution, generic customer support address.

[3] For example *FirstName.LastName@gmail.com* or *LastName_FirstCharacterOfFirstName@outlook.com*

Subsequently, the victims revoked the attackers' access to their "fake" email by changing the account password. Then, they sent out the attackers' drafts without being allowed to alter them. This proxied design was crucial to fulfilling our legal and ethical requirements, since it ensured that all SARs were technically submitted by the true data subject (the victim) rather than an impersonator (the attackers), eliminating potential legal issues for all involved parties, as laid out earlier. Further, this design also guaranteed that the victims reaffirm their consent to each individual SAR being sent out.

Afterwards, the victims waited for the (with some exceptions) maximum legally allowed response time of one month, throughout all of which the attackers had no access to the "fake" email account. Then, the victims audited all received responses and had the opportunity to censor any personal data from reaching the attackers[4] by deleting individual received emails, again in line with our ethical guidelines.

Finally, the attackers were again granted access to the "fake" email account. They would then analyze all replies, aggregate newly gained knowledge, decide if and how to follow up on proof of identity requests by data controllers, and prepare a list of new organizations to contact in the next iteration.

Data Collection

Our data sample comprehensively depicts the entire chronology of the simulated attacks, comprising 718 SARs submitted by the attackers in total. Throughout the study, we systematically recorded the attackers' strategies, including their reasoning for choosing them, as well as how organizations reacted to submitted SARs. We acquired reports about the attackers' initial OSINT findings and their discussions regarding attack strategy, target selection, as well as development of highly targeted SARs sent to specific organizations. Further, we recorded all communication with organizations (e.g., received emails and letters) with detailed metadata such as request and response timestamps, recipient (e.g., a response sent to the victim's true email address rather than the attackers' fake one), and whether an external DPO handled the request.

Data Analysis

We analyze our data sample in a two-stage process. First, we focus on the individual victims as isolated cases in a within-case analysis. In doing so, we investigate what data was leaked by organizations and how, as well as especially effective attack strategies and key challenges for the attackers. Afterwards, we compare the three cases in a cross-case analysis. From the similarities and differences identified this way, we derive commonly observed patterns with regards to flaws in SAR handling processes in organizations and attacker behavior. Further, we investigate the effect of the victims' distinct privacy characteristics on the attack by contrasting scope and sensitivity of leaked personal data acquired in each case by the attackers.

Within-Case Analysis

Starting with varying levels of OSINT knowledge gathered about their victims, the attackers pursued different strategies for the identity theft. For example, they would send a highly-targeted SAR email to specific organizations known to store personal data, such as the victim's former school known from OSINT. In other cases, they would prefer submitting a generic SAR email similar to those sent by Pavur and Knerr (2019) to a broad range of organizations, guided by educated guessing, in hopes of reaching some that store data on their victim. The attackers' strategies were adapted after each iteration to account for newly acquired and now deployable information. In the following, we will present their approaches, successful attacks resulting in important data leaks, as well as other notable incidents in an anecdotal manner on a case-by-case basis.

VictimA

Prior to the first iteration of our case study, the attackers had already collected some key data commonly used for weak forms of knowledge-based authentication about VictimA through OSINT, such as date and place of birth as well as a likely home address. Through educated guesses based on the aggregated data

[4] Made use of only once. The victim concealed sensitive data about their childhood, family, religion, and more from the attackers. As it was leaked in the final iteration of our study, this had no effect on its outcome because there were no subsequent SARs for the data to be deployed in. The victim precisely described what kind of data was leaked so this incident could be included in our evaluation.

sources, in particular the victim's public CV, they were able to reconstruct a coarse timeline of places that VictimA had lived in since birth. For example, given that the victim was born in the same city that they went to secondary school in, it is likely that they and their immediate family lived in or close to that city from birth until, at least, graduating secondary school.

First Iteration

Utilizing this a-priori knowledge, the attackers decided to pursue a dual-track strategy, submitting SARs to a total of 118 organizations in the first iteration.

On the one hand, they submitted highly-targeted SARs to select organizations that were certainly known to store data about VictimA (e.g., the former school) or were likely to do so (e.g., sports clubs matching the victim's hobbies known from OSINT in cities they have likely lived in). For example, when approaching the victim's former secondary school, the attackers weaponized data from their public CV, summarizing their professional career and referencing the year of graduation in order to convey legitimacy by disclosing supposedly non-public knowledge. Additionally, they attackers gave a plausible pretext motivating the SAR submission by mentioning an academic research project on the GDPR, which – ironically – was not far from the truth. This approach successfully tricked the school principal handling the request, who replied scanned documents containing a variety of sensitive data from VictimA's childhood, such as their parents' names and address, their religion, and the name of their primary school. Any of these could have been asked for by the principal for further proof of identity, which would have rendered the attack unsuccessful in this case.

On the other hand, the attackers selected certain industries that (almost) everyone interacts with but that also have relatively few key players, such as airlines or insurers operating in Germany, and sent a generic SAR to all organizations within. As the GDPR requires organizations to respond to SARs even if they store no data on the requester, the attackers hoped to find out which organizations store data on their victim by principle of exclusion – if all organizations except one reply that they have not interacted with the victim, it is highly likely that this one organization has. Through such "organization enumeration attacks", the attackers could acquire knowledge that helped them narrow down the scope of their subsequent efforts, without necessarily being able to fulfill potential identity verification requests by the organizations of interest yet.

Using this dual-track strategy, the attackers succeeded in causing four significant data leaks already during the first iteration. The knowledge acquired this way confirmed previous assumptions on key information usable for proof of identity in subsequent iterations, such as VictimA's current home address. Further, the attackers gathered new key identifiers such as a private email address used by the victim and learned a variety of information specific to individual organizations leaking data (e.g., customer numbers). Whilst potentially useful to a real-world adversary engaging in further attacks beyond submitting SARs, organization-specific data was of little use to our attackers as it could not be used for authentication with other organizations.

Despite their successes, our attackers also inadvertently caused the victim to quickly become aware of the simulated identity theft. Some of the small organizations they targeted based on educated guesses, such as a local language club and a local general practitioner, became suspicious upon receiving the SAR. This was because they had not had any interaction with VictimA and hence did not know them. Further, given the local scope and size of these organizations, receiving even a legitimate SAR from a known individual is presumably an atypical scenario for them. Whereas some DPOs of such organizations simply ignored the SAR or replied that the requester is unknown, in two instances they correctly identified our attack as a potential identity theft and tried to warn VictimA. One DPO did so by sending an email to the victim's work email address and calling the work phone number given in the SAR, both of which are verifiable on their employer's website, and additionally by contacting several colleagues listed there. Another DPO went even further and filed a legal complaint against persons unknown with the local police, which was dismissed after VictimA explained our case study to a detective who contacted them to warn about the potential identity theft. The attackers were not informed of these events, as would be the case in a real-world scenario.

Second Iteration

In the first iteration, the attackers had already exhausted all organizations known to them to have interacted with VictimA. Hence, they focused more on educated guesses for new organizations and following up on requests for additional proof of identity using newly acquired knowledge during the second iteration, submitting 126 SARs to additional organizations in total. Among other organizations, they succeeded in exfiltrating personal data from a multinational conglomerate in the furniture industry, a multinational car rental company, a comparison shopping website, and multiple scientific publishers. In addition to numerous data specific to individual organizations such as customer numbers or purchase histories, they acquired more key identifiers such as VictimA's private mobile phone number, their former home address, their national identity card and driver's license numbers, and a personal bank account number.

Further, they gained some insights into the victim's insurances and a building loan, presumably for their current home. An improperly secured data exchange allowed the attackers to access names, but not contents, of files sent in the SAR reply. These file names, however, contained relevant metadata such as dates and descriptions of individual insurance policies or loan financing plans, e.g., "20170821_lifeinsurance.pdf".

Third Iteration

Previously, the attackers had obtained a broad range of personal data through highly-targeted attacks as well as by exhausting organizations from industries that people matching their profile of VictimA are likely to interact with. Attempting to acquire more data on VictimA in the third iteration, they decided to target a total of 54 more specialized organizations, e.g., government agencies from the victim's hometown, a winery, and a lottery. Despite some of them storing data on the victim, no further leaks could be achieved.

VictimB

In the case of VictimB, the attackers were unable to find certain key data commonly used for proof of identity, such as date and place of birth or home address, through OSINT prior to the beginning of our case study.

First Iteration

Afraid of being unable to fulfill a potential request for proof of identity due to their limited knowledge, they decided to not immediately contact the victim's former school (as known from a public CV) in the first iteration. Instead, since their knowledge was mostly related to the victim's professional career, they decided to focus on submitting SARs to 26 organizations, which they deemed likely to have had interactions with a person from VictimB's line of work. For example, they contacted select European railway companies as well as hotel chains popular for work-related travel throughout Germany and Europe. Out of these organizations, one hotel chain provided them with the victim's home address and entire booking history.

Second Iteration

The newly acquired home address was subsequently utilized for a total of 126 additional SARs as well as following up on requests for proof of identity during the second iteration. Given their limited success in the first iteration, the attackers changed their targeting strategy to a combination of educated guesses and broader organization enumeration attacks, similar to how they approached attacking VictimA. This change proved effective, resulting in six major leaks yielding key information including VictimB's date of birth, former and current private mobile phone number, former home address, and personal bank account number – all obtained from multiple organizations, thus strengthening data validity through cross-references. All of these data leaks likely resulted from flawed processes for handling SARs established by the affected organizations.

For example, an automobile association with millions of members responded with an encrypted ZIP file attached to an email disclosing to the attackers that VictimB's date of birth was used as password. This implies that the association considered knowledge of the data subject's date of birth, in addition to identifiers such as name and work email matching their records, to be sufficient for accessing data such as bank account numbers. While questionable whether this is compliant with the GDPR, their approach also exhibits a significant vulnerability rendering the requirement to know the data subject's date of birth void: Using any date of birth as key for encrypting a file archive transmitted to the attackers (implying losing the ability to rate-limit attempts at cracking the password) provides no security. The attackers simply tried out all valid dates of birth[5] within milliseconds using a computer program. The data archive leaked by this organization contained key personal data such as VictimB's date of birth, home address, current private phone number, and personal bank account number, all of which proved useful to the attackers in the third iteration.

As another example, a former insurer of VictimB initially replied to the SAR in a letter mailed to the stored home address. Given that intercepting postal mail was beyond our attackers' capabilities, this approach would be considered secure in terms of authentication and data exchange. However, as the stored address was not up to date, the letter was returned as undeliverable. In an act of helpfulness, the employee in charge contacted the attackers via their fake email and provided all sensitive personal data electronically without encryption or proof of identity. This incident highlights that even if a secure process for handling SARs exists for everyday cases, the security concept is at risk of falling apart in extraordinary situations. A better way to handle this situation would have been requesting the former home address for identity verification.

In addition to the aforementioned leaks of VictimB's data, our attackers also received personal data (personal email and phone number, home address, passport number, and more) on a different person bearing the same name. This incident is particularly noteworthy not only because the leaking organization is a multinational online travel agency with over a million customers as of 2021, but also because all data they transmitted to the attackers, other than the person's name, mismatched the corresponding data supplied in the SAR. Hence, only first and last name were used for identity verification, certainly violating the legal requirements.

Third Iteration

Building upon their previous successes, the attackers decided to proceed with their established targeting strategy and contacted 54 new organizations in the third iteration. Additionally, they submitted SARs to organizations that they identified as candidates for systematically flawed SAR handling processes during attacks imitating the other two victims in previous iterations, such as the multinational car rental company that leaked VictimA's data in iteration two. In analogy to that interaction, the company also transmitted VictimB's private phone numbers, date and place of birth, national identity card and driver's license numbers, and bank account number without requesting proof of identity.

Further, now confident in being able to authenticate themselves as VictimB, the attackers sent a SAR to the victim's former school. Without questioning their identity, the appointed DPO fulfilled the request and supplied data on the victim's family, religion, as well as grades and courses throughout their school years.

Having found a photograph of VictimB wearing glasses via OSINT, the attackers also targeted optical store chains. Despite storing the victim's personal email and phone number, which could have been used for a secure authentication given our attackers' capabilities, one optical store chain (with a significant share of the German eye-wear market) disclosed the victim's visual acuity measurements and entire purchase history.

VictimC

Out of the three victims in our study, the attackers collected the least amount of data on VictimC via OSINT. More importantly, they were unable to find any key identifiers, such as date of birth or personal email address, beyond work-related ones published on the website of the victim's employer. Further, they could not find any hints regarding VictimC's hobbies or affiliations with organizations other than a former school.

First Iteration

Given the limited amount of work-related information the attackers knew about VictimC, they chose to pursue only organization enumeration attacks in the first iteration, submitting SARs to 99 organizations (e.g., airlines or supermarket chains) in total. Even if they were likely unable to follow up on requests for proof of identity, they hoped to gain some insights into what organizations stored data about VictimC this way. However, their efforts were of limited success, as none of the targeted organizations leaked personal data and only three organizations revealed that they stored data on the victim in their authentication requests. Unlike the schools contacted impersonating VictimB and VictimA, a scanned national identity card was requested

[5] Assuming a reasonable range for dates of birth from 1900/01/01 to 2020/01/01, only approximately 45,000 passwords must be tested.

for verification by the employee handling the SAR submitted to VictimC's secondary school.

A noteworthy observation during this iteration is that several of the contacted organizations denied storing any data on VictimC, despite doing so. We suspect that these organizations look up data for SARs by email address rather than name, as the victim had used only their personal email address – which was unknown to the attackers at this point and hence not included in the SAR – for previous interactions with them.

Second Iteration

Having gained no new knowledge on VictimC, the attackers targeted 94 organizations in the second iteration without changing their strategy. Contrary to the first iteration, they succeeded in fooling a local energy provider operating in the vicinity of VictimC's workplace into disclosing a broad spectrum of personal data, such as former and current home address, personal email address, bank account number, customer number, and electricity bills throughout the past decade. None of the key identifiers supplied by the attackers in their SAR, such as work email or work phone number, were known to the energy provider. This implies that the only data their DPO could have considered for proof of identity was the victim's name, which does not comply with the requirements imposed by the GDPR.

The attackers hence identified this organization as a likely candidate for a systemically flawed SAR handling process, given that the energy provider had a range of secure options for proof of identity at their disposal. For example, they could have requested the victim's current or former home address, customer number, and information from their latest electricity bill for a reasonably secure form of knowledge-based authentication. Alternatively, given that they stored the victim's personal email address in their database, they could have sent the SAR reply to that address, or they could have requested proof of ownership of that email account. Even better, they could have used the victim's current home address for identity verification via postal mail, for example by mailing a letter containing the password to a web portal hosting the SAR reply. Despite all these reasonably secure options, no identity verification beyond looking up VictimC's name was performed. This is especially noteworthy as the SAR was handled by an externally appointed DPO, who specializes in providing GDPR-compliant services to a variety of medium and large companies.

Third Iteration

In the third iteration, the attackers followed up on numerous SARs sent in the previous iterations, providing newly acquired key identifiers such as VictimC's personal email and home address. Additionally, they submitted SARs to further 21 organizations, including a large multinational video game and consumer electronics retailer who disclosed VictimC's date of birth. While this organization could have requested proof of ownership of the victim's personal email address as a means for secure authentication, the SAR was processed without any further identity verification. Even worse, the attackers' fake email address was automatically added as a legitimate alternative email to the victim's user account.

Most of these newly contacted organizations had been classified as potentially systematically vulnerable in previous iterations of attacks on the other victims, for example the multinational car rental company that leaked data for VictimA and VictimB. However, the attackers were unable to obtain more personal data this way because none of these organizations stored any data on VictimC.

Cross-Case Analysis

Despite our attackers' limited capabilities, they successfully exfiltrated personal data on all three victims from different organizations by submitting spoofed SARs. In doing so, they uncovered a broad spectrum of different SAR handling processes implemented by organizations with little, if any, standardized behavior being observable within an industry or shared by organizations exhibiting similar attributes (e.g., typical communication channels with their customers). However, we observed certain patterns that emerged when analyzing the organizations' reactions to the 718 SARs submitted throughout our case study.

Leaked Data and Potential Damage by the Attackers

One of the most prevalent uses of stolen identities by criminals is credit card fraud or similar activity to steal money (Willox Jr et al., 2004). Such goals could have been accomplished using only the personal data exfiltrated in our case studies. For example, using the victims' bank account data, an attacker could have purchased goods in online shops via direct debit. Knowing VictimA's national identity card number in addition to that could have sufficed to register a credit card or mobile phone subscription in their name.

A more sophisticated attacker might have even been capable of forging a national identity card using VictimA's data with a picture of another person. This way, the stolen data could have been abused by criminals to, as an example, cross state borders under a false identity for the purpose of drug trafficking.

Privacy Characteristics, A-priori Knowledge, and Convergence of Leaked Data

The three victims of our case study were selected to be structurally distinct in their public exposure as well as privacy awareness and preferences. For example, whereas VictimA is a semipublic figure with little regard for their digital footprint, VictimC tries to keep publicly available personal data to a minimum and uses pseudonymized data whenever possible in interactions with organizations. This behavior of the latter victim reduced the effectiveness of our simulated SAR identity theft in two ways. On the one hand, the attackers were unable to find key identifiers such as VictimC's date of birth via OSINT while preparing their attack. On the other hand, the number of organizations that could have leaked information on this victim was significantly lower than, for example, the number of those that could have leaked VictimA's personal data, because VictimC had interacted with fewer organizations using their real identity.

	VictimA	VictimB	VictimC
First & last name	provided	provided	provided
Workplace	provided	provided	provided
Workplace email address	OSINT	OSINT	OSINT
Workplace phone number	OSINT	OSINT	OSINT
Home address	OSINT	1	2
Date of birth	OSINT	2†	3
Place of birth	OSINT	3	-
Personal email address	1	2	2
Personal phone number	2	2	-
Bank account number	2	2	2
National identity card number	2	3	-
Driver's license number	2	3	-

† Year of birth known from OSINT prior to the first iteration.

Table 1. Iterations Needed to Acquire Victims' Key Identifiers

However, as can be seen in Table 1, even though they engaged in such privacy efforts, VictimC was unable protect themselves against SAR identity theft. Rather than fully mitigating this attack, we observed that an increased level of privacy resulted in a time shift (measured in iterations) of data known to the attackers. For example, multiple data leaks across three iterations were required for the attackers' knowledge on VictimC to reach a level slightly above what they knew from OSINT about VictimA prior to the first iteration of our case study. In analogy, it took them only two iterations for VictimB (representing an average user) to reach that same level of knowledge. Based on our observations, we hypothesize that this convergence of known data would have continued throughout subsequent iterations, ultimately nullifying VictimC's advantage.

Key Identifiers Requested by Organizations

Throughout the attackers' interactions with targeted organizations, we observed that certain types of data are commonly requested either for knowledge-based proof of identity or for identifying the data subject in the organization's data records. For example, in the first iteration of attacks on VictimC, several organizations that stored data about the victim did not recognize them because the victim's personal email address was not included in the SAR. These key identifiers can be split into two groups: organization-specific data (e.g., customer number) and generic personal information (e.g., name or date of birth).

In case of a data leak, organization-specific data typically cannot be exploited for SARs submitted to further organizations. As doing so is the core idea of the attack scenario simulated in our case study, such data was of little value to our adversaries. Further, organizations insisting on requesting such data for identity verification, rather than being content with generic personal information, constituted a roadblock given the capabilities of our simulated attackers. While we consider this approach an improvement over requesting generic personal information, it is not necessarily secure in a real-world scenario. A more sophisticated attacker might be able to obtain organization-specific knowledge through additional (potentially illegal) means.

Every bit of generic personal information, however, was highly useful to our attackers as they could weaponize it for future SARs or for following up on proof of identity requests by organizations. We observed that even small leaks of such data, seemingly unimportant when seen in isolation, played a significant role when accumulated in a large-scale attack targeting hundreds of organizations. Assume, for example, an organization recording only name, home address, as well as date and place of birth of their customers. By itself, it would appear reasonably secure for that organization to request name, home address, and date of birth for proof of identity. However, doing so would disclose the subject's place of birth to the attackers. In a large-scale identity theft scenario, this newly acquired data boosts the attackers' chances of convincing further organizations that the spoofed SARs are legitimate, thus cascading into more leaks. As a consequence of this observation, knowledge-based authentication using generic personal information such as date of birth cannot be considered sufficiently secure for verifying a SAR data subject's identity, despite its widespread use.

Further, we identified certain "critical mass" thresholds of generic personal data knowledge. Once such a threshold was crossed, we observed a significant increase in the number of organizations able and willing to process the attackers' requests. For example, knowing a broad range of personal contact information (e.g., home address, personal email address, and personal phone number) enabled most organizations, which the victim had privately interacted with, to identify the victim's records in their database. Based on the observations by Di Martino et al. (2019), we believe that an even more significant critical mass effect can be observed once the attackers are capable of convincingly forging a national identity card scan. While our attackers managed to acquire the data necessary for such an endeavor, such as the victim's home address and national identity card number, they did not attempt to falsify any documents for legal and ethical reasons.

Systematically Flawed SAR Handling Processes

The majority of data leaks throughout our case study appeared to result from systematic flaws in the processes and policies implemented by organizations for handling SARs, rather than from individual errors made by the employees handling our attackers' requests. Whenever the attackers suspected such flaws to be systematic, they flagged the organization as a candidate for further SAR submissions impersonating the other two victims in the subsequent iteration. This approach was often met with success, providing further evidence for systematic flaws. These flaws can be clustered into two groups, as explained in the following.

Insufficient Identity Verification

Certain organizations implemented an insecure process for authenticating the data requester's identity, or had no such process at all. For example, some repeatedly disclosed personal data despite confirming the legitimacy of the SAR only through relatively easy to acquire key identifiers (e.g., date of birth).

Even more severe, the local energy provider that disclosed a broad set of personal data on VictimC without any authentication also leaked data on VictimB in a similar manner. In order to validate that lack of authentication, one of the attackers – coincidentally also customer of that organization – submitted a SAR supplying nothing but their own first and last name. Without further proof of identity, the energy provider disclosed that attacker's home address, phone number, bank account number, and more. Given that these leaks were reliably reproducible across multiple unrelated data subjects, we deem it unlikely that they can be attributed to an individual employee's failure to comply with established secure policies.

Insecure Data Exchange

Some organizations failed to implement a secure way of exchanging data with the requester in response to SARs, causing multiple data leaks throughout our case study. As the GDPR mandates that organizations ensure data confidentiality, such flaws can constitute a noncompliant implementation of the law. For example, our attackers received data from some organizations as plain text in an unencrypted email, implementing no security measures to protect the data exchange against interception by a malicious third party.

Other organizations took measures in an attempt to securely transmit personal data to the SAR subject, but failed to do so successfully. A common approach was to transmit data contained in an encrypted ZIP archive via email, and disclose the corresponding password outside of that email. For example, one organization indicated that the file was encrypted using the data subject's date of birth as password, which provides no security because all valid passwords can be tried by a computer program within seconds. Another organization generated a random password using eight alphanumeric characters, which was mailed to the data subject's home address recorded in their database. Again, all possible combinations can be tested by deploying a few compute hours on modern hardware, rendering this effort at securing the data void.

While sending the password via mailed letter provides some level of security assuming it is sufficiently long, transmitting encrypted ZIP files in response to an unauthenticated SAR request should generally not be considered a state-of-the-art mechanism for secure data exchange. This is because only the contents of files in ZIP archives can be encrypted, but not their metadata such as file names, which can disclose valuable information to an adversary as was the case with VictimA. Further, all organizations in our case study encrypted the transmitted ZIP files using the default PKZIP cipher, which is shown to exhibit weaknesses such as a known-plaintext attack likely exploitable in our scenario (e.g., Jeong et al., 2012).

Even files encrypted using an algorithm currently considered to be secure should not be transmitted to a potentially unauthorized entity. Despite being incapable of decrypting such files immediately, a sufficiently determined adversary might decide to store them for years until advancements in technology (e.g., growth of computation power or quantum computing) compromise their security. Whereas certain types of data such as a home address might be outdated by then, others such as place and date of birth, religion, chronic diseases, or sequenced DNA could still be valuable to them even after decades.

Discussion

Implications for Research

Operating at the intersection between cybersecurity and cybercrime, lawmaking, and management, our study contributes to research in three meaningful ways. First, we integrate insights and established techniques from social engineering research (e.g., Jaeger & Eckhardt, 2021; Wright et al., 2014) to demonstrate the real-world viability – and hence threat – of SAR identity theft and provide a deeper understanding of the phenomenon. In doing so, we raise awareness among cybersecurity researchers and provide a solid foundation for future studies that, for example, propose and evaluate remediation options for the novel threat.

Second, our study contributes to cybersecurity (e.g., Willox Jr et al., 2004) as well as social engineering (e.g., Algarni et al., 2017; Orgill et al., 2004) research by identifying a novel domain as testbed for evaluating the generalizability of insights gained in studies on phishing. For example, based on theory from phishing literature, a plausible explanation for DPOs falling for spoofed SARs might be lack of (situational) awareness (Jaeger & Eckhardt, 2021; Nguyen et al., 2021), susceptibility due to personality (Moody et al., 2017) or behavioral factors (Wright & Marett, 2010), convincing use of influence techniques (Wright et al., 2014) and contextualization (Goel et al., 2017) by the attackers, or a combination thereof.

Third, we advance research on law hacking (e.g., Di Martino et al., 2019) by inferring two key effects from our findings, which provide insights helpful for cybersecurity researchers to evaluate the security implications of existing as well as future laws. The first key effect is the *weakest link and cascading leaks effect*. Every leak of personal data not specific to a particular organization contributed to successfully fooling further organizations. We observed such a cascade of data leaks for all three victims, regardless of privacy characteristics. Some of those leaks were caused by individual employees falling for spoofed SARs, emphasizing the need for further research into employee awareness training similar to that against phishing (e.g., Nguyen et al., 2021). However, even organizations with a – if seen in isolation – reasonably secure and GDPR-compliant identity verification process (e.g., storing date and place of birth, and requesting only date of birth) inadvertently contributed to this effect. Therefore, in analogy to IT infrastructure, the security level of the whole system (i.e., the entirety of organizations) is not determined by the average resilience of the organizations, but rather by that of its weakest link (i.e., organizations with flawed identity verification). However, this requirement to secure the system of all organizations as a whole evidently conflicts with the GDPR generically regulating only organization-level security and leaving implementation details to DPOs, which manifested in the broad spectrum of different SAR handling processes we observed. The importance of taking the whole system into account is further emphasized by our organization enumeration attacks, where mandatory replies that no data is stored resulted in revealing which organizations the victim had interacted with, hinting at a systemic information leak rooted in the GDPR. On a more abstract level, this insight that vulnerabilities arose from a shift of scope might contribute to cybersecurity research as a tool for discovering flaws in other laws.

The second key effect we infer is the *knowledge convergence and scalability effect*. As an immediate consequence of the weakest link and cascading leaks effect, we observed that the advantage of an increased privacy awareness eroded with an increasing number of SARs submitted. This effect was strengthened by the lack of restrictions with regards to submitting SARs to a large number of organizations. Contacting more organizations resulted in an increased likelihood of achieving more data leaks and enabled organization enumeration attacks, without the attackers risking any repercussions other than alarming their victim of the ongoing attack. Thus, in analogy to IT security, our study contributes to cybersecurity research by emphasizing the need for protecting against “brute-force” attacks in the context of lawmaking.

Implications for Practice

Our study has immediate implications for practice, as we expect malicious actors to observe similar patterns and gain the same insights as we did in real attack scenarios. Having demonstrated the real-world feasibility of SAR identity theft and the extent of damage attackers could have caused by weaponizing the exfiltrated personal data in our simulated attacks, it is crucial to recognize that such findings can and will be used to optimize attacker behavior. For example, an experienced adversary might compile a list of organizations with systematically flawed SAR handling processes to be tried first in an attack, or could adjust their strategy to reduce the likelihood of the victim becoming aware of the ongoing attack.

In addition to highlighting the need for action by all involved parties to mitigate this growing threat, we contribute to this effort by advancing the understanding of SAR identity theft attacks. Based on our findings, individuals as data subjects do not have the ability to fully mitigate the attack, especially due to the weakest link and knowledge convergence effects. However, minimizing one’s digital footprint and using pseudonyms whenever possible complicates the identity theft for attackers, requiring more leaks to unlock access to more sensitive data. This insight should be considered by individuals – particularly in their use of (professional) social media, where important key identifiers (e.g., date of birth) are often shared on public CVs.

Organizations, on the other hand, should reevaluate their SAR handling processes in order to eliminate potential flaws. Further, they should reconsider their use of knowledge-based authentication in general. As evidenced by our study, commonly used key identifiers such as date of birth or home address must be considered insecure for identification purposes not just in the context of SARs, but also on customer support hotlines, for example. Despite having to strike a balance between customer convenience and security, they should request organization-specific identifiers (e.g., customer number) or require proof of ownership (e.g., emailing a one-time code to a stored address) instead. In addition to protecting their customers’ data, this would also contribute to breaking the chain of leaks resulting in the weakest link effect.

In a further effort to mitigate this effect, lawmakers might want to consider prohibiting the use of data that is not organization-specific for SAR identity verification. Additionally, they could introduce means to limit the number of SARs each data subject is allowed to submit within a given time frame. For example, a centralized instance – accessed by EU citizens using government-issued credentials – could grant a limited number of one-time tokens to be submitted alongside a SAR and consumed by the receiving organization. This would weaken the scalability effect and complicate the organization enumeration attacks executed in our study.

On the bright side, we uncovered certain limitations to the practical viability of SAR identity theft, despite the successful attacks on all three victims. Attackers without knowledge about which organizations store data on their victim need to submit SARs to a large number of organizations. While the process of submitting them via email can be automated, the replies our attackers received were mostly unstructured and, particularly those from smaller organizations, highly individualized. Further, such replies sometimes implied hints for future educated guessing, even when not disclosing any personal information. For example, some organizations securely mailed the victim's personal data via letter, but also replied to the attackers' SAR email stating that a letter was mailed to the stored home address – implying that the organization stores data on the victim. Based on these observations, we consider it infeasible to (fully) automate evaluating the replies. This results in limited scalability, meaning that even though SAR identity theft is feasible for targeted attacks on select individuals, it is unlikely to become an automated large-scale threat such as phishing emails.

Further, all three victims had quickly become aware of the attack due to most organizations handling identity verification securely. For example, some organizations responded via email sent to a known address from their database, or in a letter mailed to the stored home address – alarming the victim in the process. Some organizations, such as a local language club in the case of VictimA, even recognized the identity theft and attempted to actively protect the victim by notifying them or by reporting their suspicion to the police.

Limitations and Future Work

A primary goal of our study was to establish a lower boundary for the threat of SAR identity theft. Hence, we designed it so that the simulated attackers represent the weakest ones reasonable in a real-world scenario. Future studies might want to simulate the attack with fewer constraints on the adversaries' capabilities, e.g., allowing them to simulate forging national identity card scans or to interact with organizations through additional communication channels such as phone calls. This would yield further insights into the damage a more determined and sophisticated attacker with no regard for the law could cause.

Our case study took place between November 2020 and February 2021, within the second wave of the COVID-19 pandemic in Germany, which peaked in December 2020. In order to not overburden the healthcare sector, we decided to carefully submit only a few SARs to organizations like general practitioners. However, such organizations constitute an interesting target for malicious actors in our scenario, as they store highly sensitive personal data that could significantly impact a data subject when stolen, yet are small and likely inexperienced in handling SARs. Further, as observed in recent research, threat and effectiveness of social engineering attacks in general increased due to the COVID-19 pandemic and the accompanying restrictions (Naidoo, 2020). This might have facilitated our attacks and boosted their success rates, possibly resulting in an effect adverse to our goal of simulating a weakest reasonably assumable attacker. Hence, we propose a post-pandemic repetition of our study to overcome these limitations.

Further, the focus of our study lies on the attackers' perspective of SAR identity theft. While we were able to identify some patterns in how organizations handle SARs, a complementary study from the organizational perspective would contribute to a more complete understanding of the phenomenon. For example, DPOs could be interviewed in an effort to investigate not only their organization's policies for handling SARs, but also how and by whom they were developed. This way, root causes for systematic flaws could be identified.

Acknowledgements

We are grateful to Ekaterina Korneeva for her immensely helpful support during the planning and execution of our case study. We would also like to thank the nine anonymous case study participants, who sacrificed their valuable time to make this research project possible.

References

Accenture Security. (2019). *Ninth Annual Cost of Cybercrime Study*. https://www.accenture.com/_acnmedia/PDF-96/Accenture-2019-Cost-of-Cybercrime-Study-Final.pdf

Albrecht, J. P. (2016). How the GDPR will change the world. *European Data Protection Law Review*, *2*, 287. https://doi.org/10.21552/EDPL/2016/3/4

Algarni, A., Xu, Y., & Chan, T. (2017). An empirical study on the susceptibility to social engineering in social networking sites: the case of Facebook. *European Journal of Information Systems*, *26*(6, SI), 661–687. https://doi.org/10.1057/s41303-017-0057-y

Andrew, J., & Baker, M. (2021). The General Data Protection Regulation in the Age of Surveillance Capitalism. *Journal of Business Ethics*, *168*(3), 565–578. https://doi.org/10.1007/s10551-019-04239-z

Bufalieri, L., Morgia, M. L., Mei, A., & Stefa, J. (2020). GDPR: When the Right to Access Personal Data Becomes a Threat. *2020 IEEE International Conference on Web Services (ICWS)*, 75–83. https://doi.org/10.1109/ICWS49710.2020.00017

Cagnazzo, M., Holz, T., & Pohlmann, N. (2019). GDPiRated – Stealing Personal Information On- and Offline. *Computer Security – ESORICS 2019*, 367–386. https://doi.org/10.1007/978-3-030-29962-0_18

Di Martino, M., Robyns, P., Weyts, W., Quax, P., Lamotte, W., & Andries, K. (2019). Personal Information Leakage by Abusing the GDPR 'Right of Access'. *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*, 371–385.

European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). *Official Journal L119*, *59*, 1–88.

Goel, S., Williams, K., & Dincelli, E. (2017). Got Phished? Internet Security and Human Vulnerability. *Journal of the Association for Information Systems*, *18*(1), 22–44. https://doi.org/10.17705/1jais.00447

Jaeger, L., & Eckhardt, A. (2021). Eyes wide open: The role of situational information security awareness for security-related behaviour. *Information Systems Journal*, *31*(3), 429–472. https://doi.org/10.1111/isj.12317

Jeong, K. C., Lee, D. H., & Han, D. (2012). An improved known plaintext attack on PKZIP encryption algorithm. *Information Security and Cryptology - ICISC 2011*, 235–247. https://doi.org/10.1007/978-3-642-31912-9_16

Krombholz, K., Hobel, H., Huber, M., & Weippl, E. (2015). Advanced social engineering attacks [Special Issue on Security of Information and Networks]. *Journal of Information Security and Applications*, *22*, 113–122. https://doi.org/https://doi.org/10.1016/j.jisa.2014.09.005

Moody, G. D., Galletta, D. F., & Dunn, B. K. (2017). Which phish get caught? An exploratory study of individuals' susceptibility to phishing. *European Journal of Information Systems*, *26*(6, SI), 564–584. https://doi.org/10.1057/s41303-017-0058-x

Naidoo, R. (2020). A multi-level influence model of COVID-19 themed cybercrime. *European Journal of Information Systems*, *29*(3, SI), 306–321. https://doi.org/10.1080/0960085X.2020.1771222

Nguyen, C., Jensen, M., & Day, E. (2021). Learning not to take the bait: a longitudinal examination of digital training methods and overlearning on phishing susceptibility. *European Journal of Information Systems*. https://doi.org/10.1080/0960085X.2021.1931494

Orgill, G. L., Romney, G. W., Bailey, M. G., & Orgill, P. M. (2004). The urgency for effective user privacy-education to counter social engineering attacks on secure computer systems. *Proceedings of the 5th conference on Information technology education*, 177–181. https://doi.org/10.1145/1029533.1029577

Pavur, J., & Knerr, C. (2019). GDPArrrrr: Using Privacy Laws to Steal Identities. *CoRR*, *abs/1912.00731*.

U.S. House. 109th Congress. (2006). *H.R.1815 - National Defense Authorization Act for Fiscal Year 2006*.

Wang, Z., Sun, L., & Zhu, H. (2020). Defining social engineering in cybersecurity. *IEEE Access*, *8*, 85094–85115. https://doi.org/10.1109/ACCESS.2020.2992807

Willox Jr, N. A., Gordon, G. R., Regan, T. M., Rebovich, D. J., & Gordon, J. B. (2004). Identity fraud: A critical national and global threat. *Journal of Economic Crime Management*, *2*(1), 3–48. https://doi.org/10.21552/EDPL/2016/3/4

Wright, R. T., Jensen, M. L., Thatcher, J. B., Dinger, M., & Marett, K. (2014). Influence Techniques in Phishing Attacks: An Examination of Vulnerability and Resistance. *Information System Research*, *25*(2), 385–400. https://doi.org/10.1287/isre.2014.0522

Wright, R. T., & Marett, K. (2010). The Influence of Experiential and Dispositional Factors in Phishing: An Empirical Investigation of the Deceived. *Journal of Management Information Systems*, *27*(1), 273–303. https://doi.org/10.2753/MIS0742-1222270111

Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). SAGE.