Exploring the CAM4 Data Breach: Security Vulnerabilities and Response Strategies

Jacob Sorn*, Patrick Carroll*, Zachary Pang*, Suman Bhunia*, Mohammad Salman†, Paulo A Regis‡
*Department of Computer Science and Software Engineering, Miami University, Oxford, Ohio, USA
†University of Anbar, Anbar, Iraq
‡ Department of Computer Science, Southeastern Louisiana University–Hammond, Louisiana
Email: {sornjj, carrolp4, pangk2, bhunias}@miamioh.edu, mohammed_salman@uoanbar.edu.iq, pregis@southeastern.edu

***Abstract*—The CAM4 data breach was a significant security incident involving the compromise of nearly 11 billion records from CAM4, a service focused on adult content. This paper delves into the causes of this extensive data exposure, particularly highlighting the role of a misconfigured search engine within CAM4's infrastructure. The impact and potential ramifications of this breach are examined by comparing it with other cyber attacks on similar adult-oriented websites. Notably, the exposure was identified by security researchers rather than malevolent entities. The paper discusses how the absence of default authentication measures in ElasticSearch, coupled with human error, were key factors in this breach. Granity Entertainment's prompt response in decommissioning the compromised server is also covered. The paper examines potential consequences of the breach, including economic losses, personal risks like sextortion, and erosion of trust. It also explores various motivations behind cyber attacks in the context of the CAM4 incident. The study concludes by emphasizing the relatively positive outcome, where minimal legal or financial repercussions occurred since the breach was detected by researchers, not by hostile actors.**

***Index Terms*—CAM4, ElasticSearch, Granity Entertainment, Shodan, NoSQL, Authentication.**

I. Introduction

This paper aims to investigate the causes and implications of the CAM4 Data Breach. CAM4, an adult content website, experienced a significant data leak. This study will explore the detection of the leak, identify the parties who discovered it, and analyze the subsequent response. Additionally, the paper will examine similar incidents on other adult websites and their aftermath. The methodology of hacking the CAM4 database will also be discussed. The paper will provide an overview of the evolving landscape in data and IT systems within the adult entertainment industry. Finally, it will propose preventative strategies for companies to avoid future breaches.

In the evolving digital landscape, the CAM4 data breach exemplifies the critical need for robust cyber security in handling sensitive data. This incident underscores the growing complexity and frequency of cyber threats in an era increasingly reliant on digital platforms. This paper will also consider the ethical and legal challenges arising from such breaches, reflecting on the need for stringent data protection and privacy measures. This discussion sets the foundation for a comprehensive analysis of security vulnerabilities and strategies for safeguarding against similar incidents, particularly in platforms handling highly sensitive personal and financial information.

In summary, the main contributions of this paper are:

- Investigating the causes and implications of the CAM4 data breach.
- Analysis of the breach detection, identifying parties involved, and their response.
- Comparing this breach with similar incidents in the adult website industry.
- Discussion on the methodology behind the CAM4 database hack.
- Overview of changes in data and IT systems within the adult entertainment industry.
- Proposing preventative strategies to avoid future data breaches.

II. Background

A. What is CAM4?

CAM4 is an adult web streaming platform where adult users can upload and consume explicit videos. The website requires users' personal information for account creation and maintenance, allowing them to log in for both consuming and producing content. Additionally, financial information is collected, as users can purchase tokens to tip webcam performers. The sensitive nature of the site and the data it holds makes any breach particularly concerning for its users [1].

Regarding the data breach, CAM4 had 11 billion records from its users' information exposed [2]. This compromise was initially discovered by a research group, and in response, the company immediately took down the affected database [2]. The exposed information ranged from full names and payment logs to email addresses. Users have been cautioned against clicking suspicious links due to risks of sextortion scams and ransom demands [2]. However, a WIRED article reported that although these records were exposed, it does not necessarily mean they were accessed [3]. Furthermore, it appears that only 93 people could actually be identified from the payment information in this exposure [3]. The article suggests that this incident likely resulted from a misconfiguration of a production database [3].

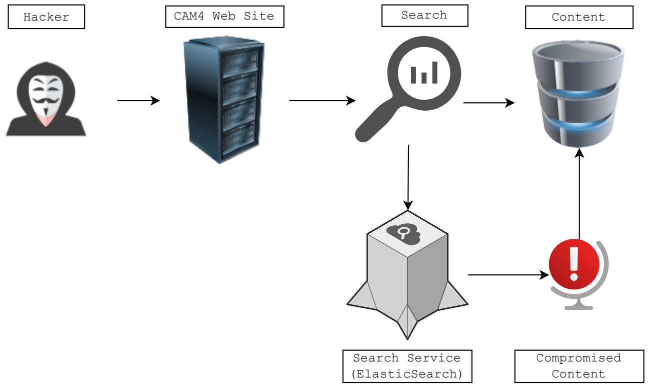

Figure 1: Diagram of services and how ElasticSearch has compromised the data.

B. ElasticSearch

According to an article by TeamPassword, a misconfiguration in Elasticsearch was the critical vulnerability that allowed the security firm to access sensitive data on CAM4 [4]. The breach led to the exposure of approximately 11 billion records, which included comprehensive user information such as first and last names, password hashes, email accounts, and countries of origin. Additionally, the leak revealed more specific data like transcripts of email correspondences, communications with other users and CAM4 support, token information, IP addresses, and logs related to fraud and spam activities.

[4].

Restating the all of this data was leaked to a misconfigured Elasticsearch database, it would be beneficial to know what a Elasticsearch database is and how it is used. According to Elastic's website, Elasticsearch is a free and open search and analytics engine for all types of data. This includes textual, numerical, geospatial, structured, and unstructured data. [5] CAM4 was most likely using Elasticsearch as an internal search engine that can be used by employees to track user activity. This data could then be used in data analysis to help build a better product.

Leaks due to misconfigured Elasticsearch databases are not uncommon. After a quick search we found countless instances of data being left wide open. One example of this is StoreHub. Researchers at a security recommendation service Safety Detectives claimed to have found roughly a million records being left defenseless on a Elasticsearch server ran by Malaysian point-of-sale software vendor StoreHub. [6].

C. Ashley Madison Data Breach

When looking at a leak from an environment that is fairly taboo like CAM4, many people may look back at the chaos that ensued after Ashley Madison had a data breach in 2015. Ashley Madison is a commercial website known for enabling extramarital affairs. [7] A group calling themselves *The Impact Team* stole user data from Ashley Madison and threatened to leak personal identities if the website was not shut down within 30 days. [8] Exactly 30 days later the group released 60 gigabytes of user data onto the dark web.

The impact of this user data being released is that they soon became targets of extortion and public shaming. This attack by the group was initiated by their dislike of the full delete service offered by the website. [8] Individuals who had accounts made using their information without their permission either as a prank or malicious could have their accounts fully deleted and removed from their databases. However, this was revealed to be false by the data breach as many individuals who paid for the service found their details in the leaks. [7] Ashley Madison made millions off of the full delete service and did not fill their promise of deletion from their databases.

Even though there is no proof that any malicious individual accessed the data from CAM4, the Ashley Madison data breach goes to show the possible aftermath of this type of data being leaked.

III. Attack Methodology

CAM4 database exposed nearly 10.88 billion records which are approximately seven terabytes of information. These include users' full names, email addresses, and other personally sensitive information. [4] It is worth mentioning that there is no evidence that suggests the website was maliciously broken by unauthorized users. However, the database remains vulnerable, as the company has no choice but to take down the entire server in order to fix the issue.

The vulnerable database was initially found by a security review site "Safety Detectives." [9] When browsing through the Shodan search engine, the team noticed that CAM4's ElasticSearch production database is not protected by a password, which means technically anyone could access the firm's private server.

A. Elasticsearch Database

Elasticsearch database is a NoSQL, JSON-based data store system. However, Elasticsearch itself is not a database but a search engine. Organizations and firms use this cloud-based search engine to gather vast amounts of data and, at the same time, use the platform's depositories to store their information without further consideration. Not surprisingly, security technicians might forget to set authentication or give a simple, easy-to-guess password to the database. Internet-security researchers estimate eight hours for a hacker to do reconnaissance, scanning, accessing, escalation, and covering their tracks to a vulnerable database. [10] Most importantly since Elasticsearch is an open-source engine based on Apache Lucene, which was first released back in 2010, there are increasingly abundant information and tutorials that are available online for people to learn how to breach into the system.

B. Pentesting Elasticsearch

1) Reconnaissance and Footprinting: Elasticsearch database works like the JSON database in which each file contains many values with their unique header(keys). An inverted index, also known as an inverted file, connects all

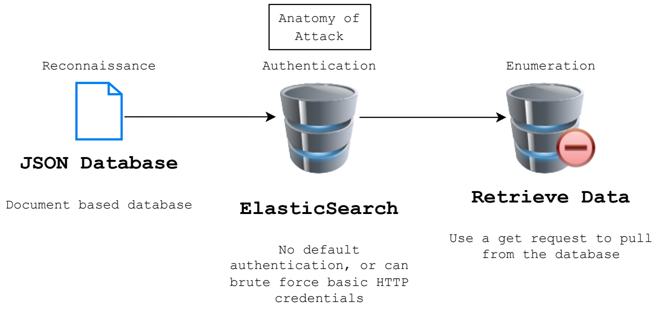

Figure 2: Diagram of how pentesting occured on the Elastic-Search component of the CAM4 stack.

the files together to ensure a fast full-text search in near-real time.

2) Authentication: Elasticsearch does not have a default authentication, which allows unauthorized users to access the data. This process could be tested by the command “curl -X GET ”ELASTICSEARCH-SERVER:9200/_xpack/security/user“, and verified by an ”error... Enable security by setting...“. However, if the returning statement is ”error... missing authentication,“ which suggests that the HTTP basic authentication is turned on, then any Bruteforce HTTP basic auth such as ”hydra -L /usr/share/brutex/wordlists/simple-users.txt -P /usr/share/brutex/wordlists/password.lst sizzle.htb.local http-get /certsrv/“ and ”curl -X GET http://user:password@IP:9200/“.

3) Enumeration and Elastic Info: To enumerate data, the command is curl -X GET ”ELASTICSEARCH-SERVER:9200”, follows the desired command, for instance “/_security/user” or “/_cat/tasks”. [11]

IV. Impact

A. Impact on the Company

Luckily for Granity Entertainment who is CAM4’s parent company, there was little public outcry due to the data breach. [4] This may be due to the nature of the website and users wanting to stay anonymous or that there was no sign of the data actually being accessed by a malicious party. Although this is seen as one of the largest data leaks [12] they had a very fast response to the report unlike many of the companies we will look at later. Within 30 minutes of the report Granity Entertainment had taken the database offline to repair the compromise. [4] Although this leak was their fault, nothing appears to have been accessed by a malicious party and they had a very fast response to the breach. This may contribute to the reason why there was little outcry about this situation. However, due to the fact that it is an adult website, it is hard to find critiques about the website one way or the other. For this reason we can only infer whether or not they lost users due to this incident. With there being many websites that offer similar services, people very easily could of decided to go elsewhere and feel more secure about there information being stored there.

Table I: Potential Attacks

Information Leaked	Potential Attacks
First/Last Names	Phishing, Social Engineering, Identity Theft, Sextortion
Emails	Phishing, Social Engineering, Sextortion
IP Address	DDoS, Port Scan, IP Spoofing
Payment Information	Identity Theft
User Conversation	Phishing, Social Engineering
Usernames/Password Hashes	Account Stuffing, Extortion

B. Potential Impact

Even though it is assumed that none of the data was actually accessed by an outside source, it is a good idea to look at the potential consequences that could take place if a malicious party were to have accessed that data. According to an article by HackNotice, information such as first/last names, emails, IP addresses, payment information, user conversation, username/password hashes, and more were in the database that could of been accessed. [13] Looking at table 1, we can see a few of the possible attacks a hacker could launch if they were to access this type of user information. For example, getting access to someone first and last names as well as their respective email addresses could make them subject to phishing, social engineering, identity theft, and sextortion. Both phishing and social engineering are similar in the way they would be launched using this type of information. The hacker would be able to disguise themselves as the person with exposed data and may be able to receive sensitive data from friends/colleagues. Almost everyone knows someone that has faced identity theft in some way. The hacker can act like them and try to buy things under their name. This can impact credit score and more. [14] By allowing this type of data to be exposed, Granity Entertainment opened itself up to face many lawsuits were this type of thing happen. For this reason this paper will go more in-depth of similar situations that other adult websites went through and the backlash they received from it.

C. Revisiting Ashley Madison

The Ashley Madison Data Breach was already covered in the background section of this paper so here we will just review the aftermath and legal issues they went through in consequence of the breach. When it comes to the legal consequences of what came to light due to the breach, Ashley Madison’s parent company paid 11.2 million dollars in a settlement with over two dozen data breach lawsuits. [15] One-third of this settlement went to attorneys’ fees, 500,000 was put to the side to help administer the remaining 7 million to affected Ashley Madison members.

Bots, fake accounts, and a faulty full delete service were all exposed during this incident. People could pay money to interact with *engagers* on their website. Over 70,000 of these were bots with fake female profiles. The company then announced they would refund up to 500 dollars to any member that submits a valid claim proving they spent money on one of these bots. [15]

When it comes to the public backlash from this breach, it is fairly extreme. Compared to CAM4 which is considered a porn site which is still taboo, Ashley Madison offers extramarital affairs. Many people understandably have very strong feelings toward people who commit adultery and them being exposed like this can effectively ruin their life. There were quite a few people who took the side of the hackers and believed they were doing everyone a service by exposing the people having affairs. Promising to offer anonymity and not being able to follow through with it can cause many people to shy away from your services.

Surprisingly, Ashley Madison is still going strong today if not stronger than it was in 2015 before the incident. There may be many factors in this such as it still being the largest provider of the certain service it has to offer. It may also be due the the cultural shift towards hookups that is continuing to bloom. You can not say for sure what may cause Ashley Madison to still be popular today but many people still look down on it due to it's nature and the data breach.

D. Luscious Data Breach

The Luscious data breach is where almost 1.2 million users had their private information exposed. One of the main concerns from this data breach was the use of identifiable emails. Many users form countries such as France, Germany, Russia, Brazil, Italy, Canada and Poland all used government associated emails. Issues can arise with valid user emails such that open source intelligence gathering can link these unique emails to specific individuals. Although more a discussion of defense, users on taboo websites can protect anonymity by using personal/anonymous emails when registering for personal services. Furthermore, exploiters of this information take advantage of the social implications of association with porn sites. Relationship damage, social judgment, and surfacing of private life all lead to victims more likely to paying anybody blackmailing them. In order to combat these issues the company has recommend that users do not use identifying information when registering for such accounts. In addition, the research group recommends that you make usernames and passwords that re unrelated to your name/email so that they cannot be easily associated in the incidence of a breach. [16]

E. Adult Friend Finder Breach

In 2016, six databases were breaches which were all owned by Friend Finder Networks Inc. This breach exposed 412 million records from the self proclaimed "world's largest sex and swinger community". This breach endured large criticism as the company has had previous data integrity issues. The data included sexual preference, purchases and usernames/passwords. One of the largest takeaways from this hack is how a company responds. Friend Finder was made aware of the vulnerability on October 18th while the breach occurred on October 20th. This indicates the company was slow if not negligent in their response. Furthermore, this is the company's second data breach, which greatly reduces the trust consumers will have with this sight.

Regarding the data, many of the emails exposed could lead to extortion or identify specific individuals. These include the 78,000+ military emails and the 5,000+ .gov emails. Due to the freedom of information act and government websites these individuals are at a very high risk to be individually identified. It would be advisable based on this leak and the others discussed to use personal emails or more vague emails to limit identification if a leak were to happen. In addition these people are at a higher risk of identity theft if a criminal can attach the email to a name, address, job, and other possible points of identification. [17]

F. Retrospective

After looking at the aftermaths of some of the other major adult website data breaches, CAM4 was lucky that nothing came from their data leak. The legal reprimands and social backlash that can be seen from some of these cases are terrifying for a business to imagine. Next we can look at the defenses that can be put in place and what can change in order for the sensitive data to be secured.

V. Defense Solution

A catastrophic data breach could impact thousands if not millions of users. To emphasize the situation we discussed in this article. We will be focusing on data breaches caused by human faults. The use of malicious and hardware devices will also be mentioned later in this section.

A. Human Error

According to an article published in 2021, human error is the number one cause of data breaches. [18] Among all data breach cases, all most one-third of incidents were due to the result of employee negligence. In contrast, data losses due to external cyber attacks only occupied 22 percent. The report also indicated that more than 25 percent of employees do not lock their computers with passwords. [19] Workers with no computer science or online safety knowledge might simply not understand the importance of database security. Even with knowledgeable employees, there is always the possibility that an error will happen due to a lack of focus. Fortunately, many procedures and rules are already designed and optimized in order to prevent employees make mistakes.

Social engineering is a commonly used hacking method which often deployed by black hat hackers. It refers to malicious activities accomplished through human interactions by using psychological manipulation. [20] For example, a most commonly used method for social engineering attacks is suspicious emails sent by hackers. Since company email addresses and employees' email addresses could be easily found by using LinkedIn or by using "whois", hackers could send emails that pretend to be a superior or fellow co-worker who uses their private email address to contact the employee. They could ask for the company server username and password or want authorization to the company's sensitive database. To prevent data leaks due to social engineering attacks, there are some rules and boundaries people could follow.

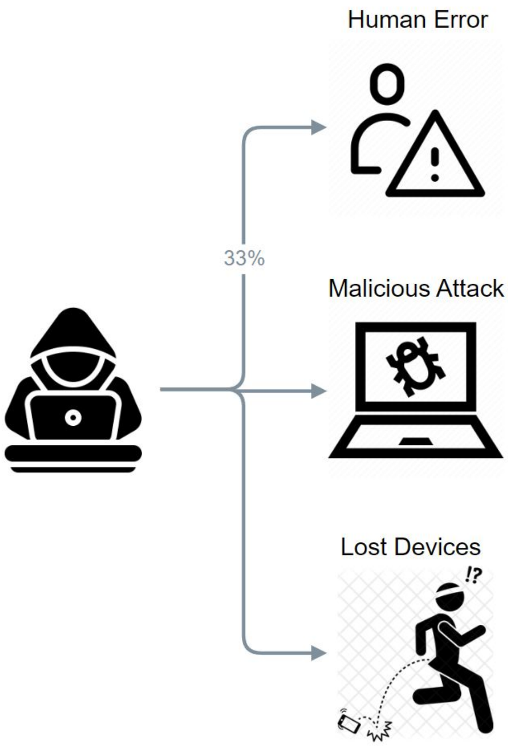

Figure 3: Diagram of corresponding defense method 1 according to the attack methods.

First of all, works related information should only be discussed and sent by a secured email under the company's domain. That means if there is a database authentication update request that an employee needs from his supervisor, he should make sure his request is only sent and received under the company's private email address. Moreover, any workers should avoid opening emails or attached files from suspicious addresses in order to minimize the risk.

Secondly, workers should often update and maintain computer antivirus software. It is not very common to see, but sometimes, a malicious email sent by a hacker could contain small pieces of software that break into the computer system. Luckily most modern antivirus scanners could prevent that from happening.

Besides, raising employees' security awareness could be highly effective. Companies could offer online security workshops or meetings that discuss what potential malicious attacks look like. For example, if an employee gets an email offering a disturbingly large amount of money, he or she should know something is not right. Or, if a non-employee shows up at the front desk asking for access to the company's server room, people should know the person is up to something. Most importantly, the company should advise all employees to set up a secure password to accounts related to the company's data.

Official Contact
Security Trainning
Personal Awareness
Update Patches
UPDATE
Security Scan
Recuit Experts

Figure 4: Diagram of corresponding defense method 2 according to to attack methods.

In addition, large companies should also consider using multifactor authentication systems to help ensure the safety of personal accounts in the event of system compromise. There are a lot of security companies on the market to offer various products, including Imperva, Pii-Tools, Duo, and so on. These authentication systems will add additional work to the employees; however, they are effective in increasing the company's data security.

B. Malicious Attack

Malicious attacks could be tricky to prevent. Sometimes, data and information may be already leaked without any warning or notification. However, there is still something people could do in order to minimize the risks. First of all, many malicious attacks target certain vulnerabilities of a software or system. It could be a badly configured AWS or Github server, exploited software patches, and many more. For example, the 2017 WannaCry ransomware worm, which impacted more than 200,000 users, was caused by a vulnerability in an old version of Microsoft Windows. [21]

To prevent this happened again. A company should advise its employees to regularly update their software or allows internet security personal to install the latest software patches. Most issues and software vulnerabilities are fixed in a short amount of time after it is been discovered.

On top of that, users should also avoid using default account names and passwords. For example, usernames "Administrator" or "Admin" and passwords "12345" or other simple coherent numbers should be avoided.

Companies with a certain amount of employees should consider recruiting a team of security specialists. A team of experts could offer a lot more insightful suggestions to the company, and more importantly, they could maintain the company's vulnerable data and keep malicious actions away.

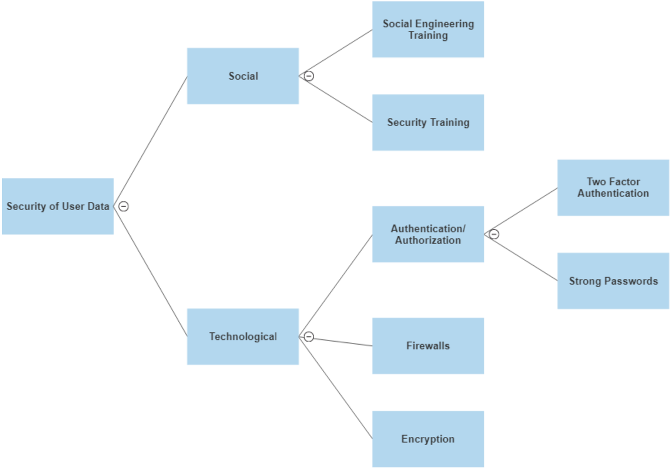

Figure 5: Tree Diagram of Options to Improve Security of User Data

C. Lost or Stolen Devices

Security breaching from a hardware level is very rare. On the one hand, we have seen some software vulnerabilities caused by hardware design flaws, but large data leaks are very rare. It not only requires hackers to have outstanding intruding skills but also needs perfect social engineering skills in order to have access to the database server itself or get an authorized device that connects to the server.

VI. Conclusion

The CAM4 data leak underscores the critical need for robust data protection in services where user anonymity is expected. Similar breaches, like those at Ashley Madison and Adult Friend Finder, demonstrate the recurring issue of inadequate security measures. In CAM4's case, the data was left virtually unprotected, posing a significant risk had malicious actors accessed it. This incident should serve as a wake-up call for implementing stringent security standards, including employee training, encryption, and two-factor authentication, to safeguard users' personal information and prevent future breaches. Security in such environments must be proactive, not reactive.

References

[1] Wikipedia, "CAM4." https://en.wikipedia.org/wiki/CAM4.

[2] "CAM4 data exposure leaks billions of records from adult streaming website." https://www.idtheftcenter.org/post/cam4-data-exposure-leaks-billions-of-records-from-adult-streaming-website/#:~:text=A%20team%20of%20researchers%20uncovered,email%20addresses%20and%20payment%20logs journal=ITRC, publisher=Identity Theft Resource Center.

[3] B. Barrett, "An adult cam site exposed 10.88 billion records." https://www.wired.com/story/cam4-adult-cam-data-leak-7tb/.

[4] "What happened with the CAM4 Data Leak?." https://teampassword.com/blog/what-happened-with-the-cam4-data-leak.

[5] "What is Elasticsearch?." https://www.elastic.co/what-is/elasticsearch.

[6] S. Sharwood, "What is Elasticsearch?." https://www.theregister.com/2022/06/16/storehub_data_leak/.

[7] Wikipedia, "Ashley Madison data breach." https://en.wikipedia.org/wiki/Ashley_Madison_data_breach.

[8] "A Retrospective on the 2015 Ashley Madison Breach." https://krebsonsecurity.com/2022/07/a-retrospective-on-the-2015-ashley-madison-breach/.

[9] A. Sen, "Live streaming adult site leaves 7 terabytes of private data exposed," May 2020.

[10] A. Russell, "What is elasticsearch and why is it involved in so many data leaks?," Oct 2020.

[11] "Basic information." https://book.hacktricks.xyz/network-services-pentesting/9200-pentesting-elasticsearch.

[12] D. Chester, "10 biggest data breaches in 2020-2021. Leakage protection.." https://cooltechzone.com/threats/malware-removal/biggest-data-breaches.

[13] HackNotice, "CAM4 Data Leak Exposes Personal Data of Millions of Users – Security Boulevard." https://hacknotice.com/2020/05/05/cam4-data-leak-exposes-personal-data-of-millions-of-users-security-boulevard/.

[14] "4 Lasting Effects of Identity Theft." https://lifelock.norton.com/learn/identity-theft-resources/lasting-effects-of-identity-theft.

[15] "Lawyers score big in settlement for Ashley Madison cheating site data breach." https://arstechnica.com/tech-policy/2017/07/sssshhh-claim-your-19-from-ashley-madison-class-action-settlement/.

[16] C. Jones, "Adult site Luscious data breach affects more than a million users." https://www.itpro.com/data-breaches/34239/adult-site-luscious-data-breach-affects-more-than-a-million-users.

[17] "How did the adult friend finder hack happened," Aug 2020.

[18] C. Pruden, "Employee negligence and data leak prevention — pii tools," Aug 2021.

[19] E. Rosen, "Human error biggest cause of data breach: Survey," May 2015.

[20] I. Team, "What is social engineering: Attack techniques &; prevention methods: Imperva," Dec 2019.

[21] A. T. Tunggal, "How to prevent data breaches in 2022: Upguard," Aug 2022.