

Analyzing Third-Party Data Leaks on EU Healthcare Websites

Sammani Rajapaksha^{1,*}, Timi Heino¹, Panu Puhtila¹ and Sampsa Rauti^{1,*}

¹University of Turku, Vesilinnantie 5, 20500 Turku, Finland

Abstract

In the present-day web-based health services, users often reveal sensitive data concerning their health status. Specifically, this is often the case when using the search function in various online services. Users trust that their data stays confidential and private when using websites. However, at the same time, many online health services use third-party web analytics and other third-party services and libraries, which may put users' sensitive data in jeopardy. In this study, we analyze 480 web-based health services in the EU area. We conduct a network traffic analysis of the data sent out to third-party services when using the studied health websites and provide an analysis of data leaks. We found that 60.2% of the studied websites leaked URLs without consent from the user. Moreover, 58.9% of the websites that had search functionality leaked search terms to third parties. Our study also highlights some regional disparities in website privacy. Our findings are a stark reminder of the current challenges in protecting users' personal data in online health services. They highlight the urgent need for web developers and health website maintainers to reassess the used third-party services and fix the privacy issues.

Keywords

Medical websites, data leaks, data concerning health, web privacy, third-party services

1. Introduction

As technology keeps advancing, the significance of web-based health services has grown in recent years. The demand for accessible healthcare continuously increases, and e-health solutions online offer new opportunities for health and social care, as well as easy access to health information. Digitization of healthcare services increases the need to protect the privacy and security of users' data through several regulatory data protection frameworks such as GDPR and HIPAA [1]. Users' sensitive personal data is processed by many web-based healthcare services, such as medical centers' websites and hospital websites. Because this data is extremely sensitive and often concerns website visitors' health, it is important to make sure this data remains confidential and does not leak to third parties [2]. In GDPR (General Data Protection Regulation), health data is classified as special category data and requires strict protection due to its sensitive nature.

Earlier work on website privacy and third-party data leaks has shown that even in critical online services, third-party tools and components, such as web analytics, are regularly employed by web developers [3, 4, 5]. Business objectives can be tracked, and user experience can be improved with these tools, but there is also a huge risk of leaking the website visitor's sensitive health information to third-party servers. Unfortunately, this often happens without users' knowledge or consent [6], and even developers and maintainers are frequently unaware of leaking personal data [7].

In the current study, we examine the privacy of 480 web-based health services in the EU area. Specifically, we address third-party data leaks taking place on these websites, as the user visits different pages in the web-based healthcare services and uses the search functionality. By studying websites from 24 EU countries, we present an overview of health data leaks, an issue that is likely to affect a much larger group of web-based health services than covered in this study. Therefore, in the current study, we address the following research question: *How prevalent is the leakage of sensitive health-related data*

TKTP 2025: Annual Doctoral Symposium of Computer Science, 2.–3.6.2025 Helsinki, Finland

*Corresponding author.

✉ syraja@utu.fi (S. Rajapaksha); tdhein@utu.fi (T. Heino); papuht@utu.fi (P. Puhtila); sjprau@utu.fi (S. Rauti)

🆔 0000-0003-4647-3885 (S. Rajapaksha); 0009-0008-4798-5261 (T. Heino); 0009-0004-6418-1063 (P. Puhtila);

0000-0002-1891-2353 (S. Rauti)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

through URLs and search terms in web-based healthcare services in the EU area? We also analyze the differences in privacy risks between the studied EU countries and discuss the privacy threats when integrating third parties in web-based health services.

The primary scientific contribution of this study is that, to the best of our knowledge, it is the first study addressing third-party data leaks on medical websites in the whole EU area. While certain other studies such as Yu et al. [3] have conducted very large-scale surveys of medical websites, their focus and approach has been different from ours.

The remainder of the paper is structured as follows. Section 2 presents the existing work on the third-party data leaks on medical websites. Section 3 covers the setting and the method of the current study, describing the selection process of the studied websites, and network traffic analysis using an automatic tool. Section 4 discusses the results of our network traffic analysis and explores the found data leaks. Section 5 presents the key findings, the limitations of the study, and the implications of our results for software developers and users of medical websites. Finally, section 6 concludes the paper.

2. Related Work

Privacy violations in medical online platforms and their implications have been studied widely. Already in 2013, Huesch [8] raised concerns about what potential privacy-infringing effects, such as becoming profiled and targeted with advertising, searching of health information from free online sources has. Libert [9] studied what implications leaking health data in URL addresses has. Kes et al. [10] argued in their paper that the collection of user data in medical contexts is actually a positive practice, as it allows for the online medical providers to know their patients better and thus provide them better services.

Hwang and Yun [11] conducted a survey of patient concerns regarding potential privacy violations associated with the use of electronic medical records in Taiwan and found out that there is high demand for hospitals and other medical providers to communicate their privacy policies better to alleviate the concerns that were raised by their respondents. Zheutlin et al. [4] investigated how user data is tracked in USA-based health-sector websites. Wesselkamp et al. [12] researched the privacy situation in 385 medical websites and found severe privacy issues on majority of them. Their findings revealed that 62% of the websites engaged in data collection even before consent had been granted by the user, and 15% even after all data collection had been rejected.

Yigzaw et al. [13] discussed central cybersecurity and privacy risks associated with medical websites, central techniques such as cryptography and data de-identification that can be used to manage these risks and how they should be implemented in medical contexts. Yu et al. [3], using web-crawling software conducted a very large scale privacy analysis on hospital websites across the globe, finding that 53.5% of them used some kind of data analytical tool that harvested user data. Huo et al. [14] conducted a study of 459 health-sector web portals, which revealed that 9 of them leaked very sensitive information, such as information on medical prescriptions and test results to third parties.

Friedman et al. [15] inspected websites of abortion clinics, and came to a conclusion that almost all of them (99.1%) collected user data through website analytical tools. Friedman et al. [16] also published an article discussing what kind of risks for medical providers are associated with using the analytical tools in hospital websites, pointing out that this may put them in legal jeopardy. Surani et al. [17] studied privacy policies of online health service providers, and revealed obvious problems in how they presented their data collection.

3. Study Setting and Method

The websites were collected by searching with the keyword "hospitals in (country)" or more specific queries like "popular hospitals in (country)" on the Google search engine. Out of the search results, articles listing hospitals, Wikipedia lists and government provided lists were used to collect the data set. We found more than 500 websites initially but later filtered them after checking the website availability and relevance.

To make the results consistent, we removed the countries that did not appear to have 20 medical websites from the study. The three smallest EU member states, Malta, Luxembourg and Cyprus, were left out. This left us with 24 EU member states. From each of these countries, we chose 20 websites of hospitals and medical centers, both from the private and public sector. The websites were selected randomly. However, we made an effort to avoid selecting several websites owned by the same company in one country, keeping the dataset diverse. In total, our dataset includes 480 healthcare websites. While the exact number of health websites in the studied 24 EU countries is difficult to determine, based on initial searches and health directories, we estimate the total number to be in the range of several thousands. This makes our sample a relevant subset, although it is not exhaustive.

In the current study, we only discuss privacy issues on the websites at a general level and avoid singling out the studied health service providers in a negative way. To follow ethical research practices, we do not use the actual names of the selected health websites, but we only present the general phenomenon and use aggregated data to ensure anonymity.

To analyze network traffic on all 480 websites, the analysis was mostly carried out with an automatic tool specifically built to analyze URL address and search term leaks to third-party servers. We describe the design and implementation of the tool in more detail in [18]. The tool is written in Python and uses Selenium WebDriver library for automating browsing and mimicking a human user. For the small set of websites on which the tool failed to function properly, the analysis was carried out manually with Google Chrome's Developer Tools. To assess our tool's reliability, we also cross-validated a sample of its outputs by double-checking them manually. In the rare cases where the tool failed, this happened for example due to some websites having pop-up elements (e.g. cookie consent banners) overlaid with the search functionality, preventing our tool from interacting with the search button.

In our manual network traffic analysis sequence, the browser cache was first cleared, and cookies were deleted. We then first accessed the front page of the health service. On the front page, all cookies and data collection were rejected. If the website had a search function, we then made a search to see whether the search term leaked. When browsing the health website, all generated network traffic was recorded using Google Chrome browser developer tools (DevTools). The network traffic was logged and saved as HAR files (HTTP Archive) for further analysis. The log files were then manually examined, searching through the HTTP request payloads, and all instances of leaking URLs and search terms were meticulously documented.

In the current study, we focus on two different categories of data leaks:

- *URL leaks.* A URL leak happens when the URL address of the page the user is currently visiting leaks. On medical websites, the exact URL address can reveal information about diseases or treatments the user is interested in, for example. While the URLs of pages the user have accessed may seem like innocent information at first, they tell a lot about health-related topics the user wants to get information about. The health-related keywords can be included either directly in the URL, or at the very least the topic of the page in question is revealed by visiting the leaked URL.
- *Search term leaks.* These leaks happen when a search term inputted by a user leaks. Search terms on medical websites can contain information about diseases or symptoms, which is extremely sensitive and private information and can reveal a lot about the status of a user's health. What makes search terms especially dangerous as personal information is the fact they are freely chosen by the user. Search terms usually leak as a part of an URL address that is sent to a third party. However, sometimes the HTTP request payload also contains search terms, independent of the URL address.

We focus on URL and search terms leaks specifically because of their high sensitivity and potential to reveal information about the user's state of health. They are also easy to detect with an automatic tool compared to many other personal data items, such as diverse pieces of information that may leak in the context of booking a doctor's appointment.

The fact that contextual data like URLs and search terms leak to third parties would not be very problematic by itself if it was not possible to identify specific users. However, what makes data leaks

dangerous is the fact they combine both identifying and contextual data. A user can usually be identified by e.g. their IP address, and this identifying data is then combined with sensitive contextual data such as details on medical conditions the user has searched for. This enables third parties to infer the user's potential medical conditions, for example.

Although identifying personal data such as an IP address cannot always be linked to a person's identity (real name), large technology companies such as Google and Meta can often accurately identify the user. This is because users may use the same device to login to the other services run by these companies, and also because there are often persistent cookies stored on the user's computer containing unique identifiers.

Aside from IP addresses and cookies, the user may be identified by third-party analytics companies using browser fingerprinting. This is a technique used to track and identify users based on their unique browser and device characteristics. These pieces of technical information can include browser type and version, operating system, screen resolution, installed browser extensions etc. When combined, these technical details can single out users with good accuracy.

It is also important to understand the significance of the time here. When the same third-party keeps collecting bits and pieces of a user's sensitive health data over time and links it all together, a more and more comprehensive profile could be built on the user's medical history. Large technology companies can also collect data across several websites. These considerations together make data collection a serious privacy risk.

4. Results

Our results show leaking potentially sensitive data even without the users' consent is relatively common on the studied health websites. Out of the studied websites, 60.2% (289/480) leaked the visited URLs to a third party. A search term was leaked by 58.9% (235/399) of the websites that had a search function.

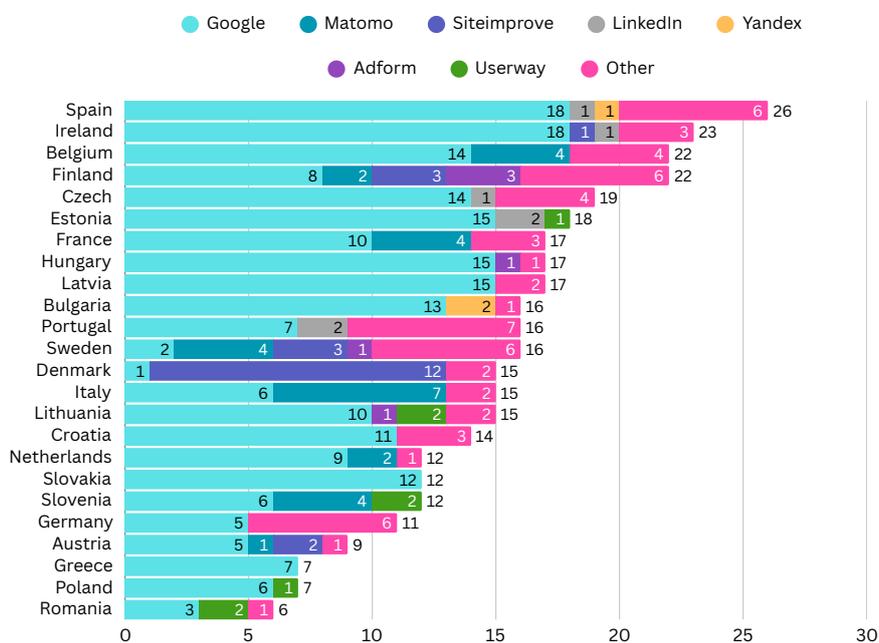


Figure 1: URL leaks in the studied EU countries and the most frequent third parties receiving leaks.

URL leaks in the studied EU countries and the most frequent third parties receiving potentially sensitive data are shown in Figure 1. Looking at the URL leaks in different countries (one third party is always counted only once per one studied website), Spain comes on top with 26 URL leaks on its 20 websites, meaning there are approximately 1.3 URL leaks per website even without consent. Ireland

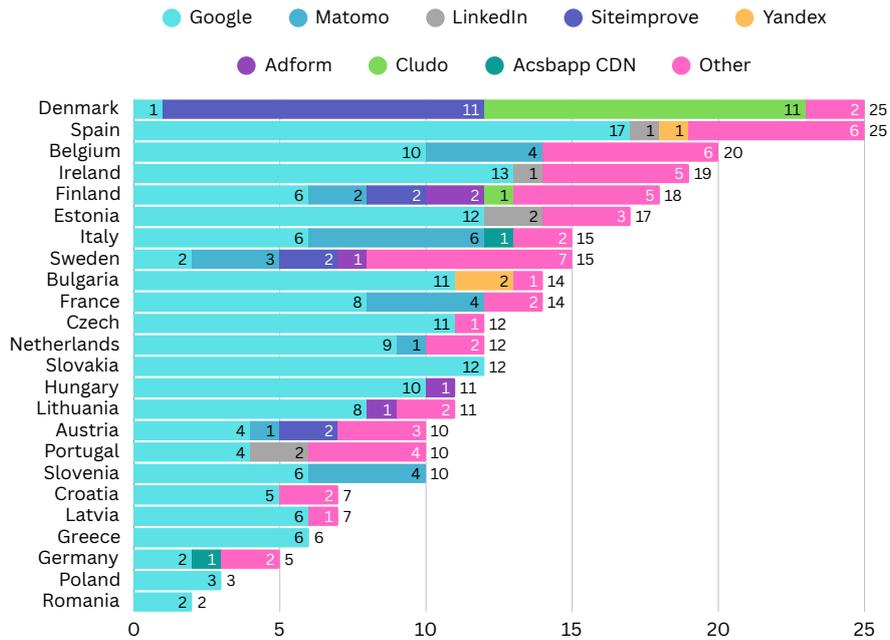


Figure 2: Search term leaks in the studied EU countries and the most frequent third parties receiving leaks.

comes second with 23 URL leaks, followed by Belgium and Finland (22 leaks both). Romania (6), Poland (7) and Greece (7) have the smallest number of leaks. Perhaps a little surprisingly, Austria (9) and Germany (11), known for their strict stance on online privacy, are not doing as well as one might expect in this study.

It is worth noting, however, that the number of leaks is not the only aspect that should be used in assessing the level of privacy in different countries. For example, the Danish health websites mostly have Siteimprove analytics as a third party. This analytics service provider is based in Copenhagen, Denmark and is at least marketed as a highly privacy-aware and GDPR-compliant solution. Most Danish health services have systematically opted for this analytics service instead of Google Analytics and there was only one occurrence of Google Analytics on studied Danish websites. Likewise, Swedish websites often opted for privacy-friendly options such as Matomo and Siteimprove.

An obvious observation from the diagram is also that Google is present in every single studied country, observing visited pages. For example, in 12 (50%) out of 24 studied EU countries, Google is collecting visited URLs on 10 or more websites. This means that in half of the countries studied, the change of visited pages leaking to Google without the user's consent seems to be 50% or higher. While such observations cannot be reliably generalized by only observing 20 websites from each country, this finding still shows that there is a great reason for concern especially when it comes to Google's widespread collection of sensitive personal data on European health websites.

Third parties with fewer than 5 occurrences have been grouped under the 'Other' category, but we have highlighted some less common but still significant third parties in this study. One particularly noteworthy third party is Yandex. Yandex is a Russian third-party web analytics service which may collect user data from websites. Especially in the current geopolitical situation, potentially sending users' sensitive health data from Europe to Russia without consent is not acceptable and can pose data privacy risks.

Figure 2 shows the search term leaks in the studied EU countries and the most frequent third parties receiving potentially sensitive health searches. In this comparison, too, Spain is on top with 25 search terms leaks. However, this top place is now shared with Denmark, since the Cludo website search tool receives the search term on 11 Danish websites. However, besides the USA, Cludo also has a headquarters in Denmark, potentially ensuring better compliance with GDPR for data protection and

privacy. Just like with URL leaks, Ireland, Belgium and Finland are other significant countries leaking search terms. Romania (2), Poland (3), and Germany (5) have the smallest numbers of search term leaks.

It is easy to see in Figure 2 that Google dominates the search term leaks as well and is again present in every single country. Likewise, we observe that Yandex is still there receiving search term leaks, potentially collecting users' health-related queries and storing them on Russian soil. Beside Yandex, another controversial service collecting search terms we have highlighted in the figure is AccessiBe, or Acsbapp. While meant to be a solution improving accessibility, some privacy concerns have been voiced regarding this service¹.

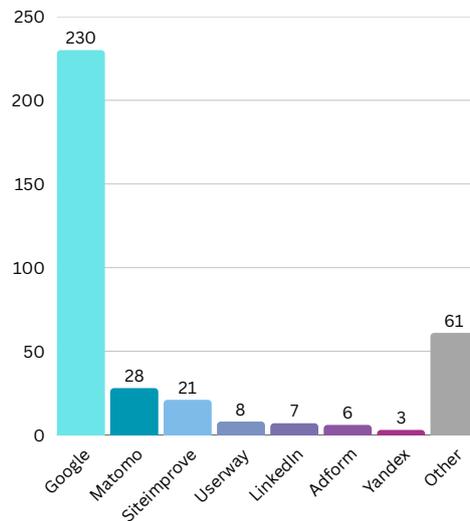


Figure 3: The third parties most frequently receiving URL leaks.

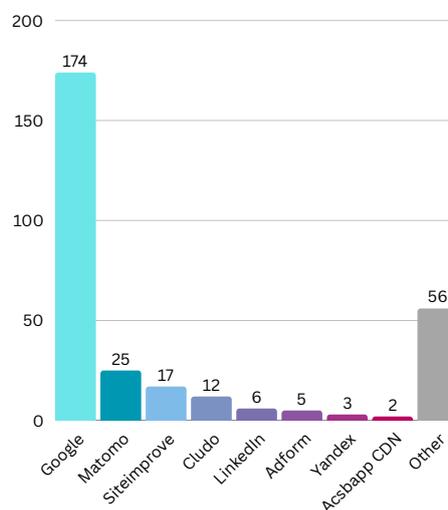


Figure 4: The third parties most frequently receiving search term leaks.

Figures 3 and 4 illustrate the most common receivers of URL and search term leaks, respectively. These figures also show very clearly how dominant Google is among the found third-party services. This becomes even more prominent when we note that Matomo and Siteimprove analytics, which are in the second and third place in the figures, are usually considered privacy-friendly options [19, 20].

¹<https://tink.uk/accessibe-and-data-protection/>

Meanwhile, there have been concerns about data collection and privacy compliance of Google Analytics in the EU area [21].

Matomo is a privacy-conscious analytics service that can be deployed locally. In the current study, we have only included third-party versions of Matomo analytics (often hosted within the domain matomo.cloud). While Matomo as a third party might not present a significant risk, it is still technically considered a third party.

Compared to some earlier studies (such as [14]), we observed a higher number of data leaks and increased prevalence of google Analytics, for example. The differences are partly explained by different collection methods and different data leaks types, as we concentrated on URL and search term leaks specifically. However, while [14] found Google Analytics on only 14% of the medical websites they studied, we found data leaks to Google on almost 50% of the studied websites. Besides studying a different set of websites (USA vs EU), the growing prevalence of Google Analytics is most likely explained simply by its increased popularity and market share.

5. Discussion

5.1. Key Findings

The key findings of our study can be summarized as follows:

- **High numbers of leaks.** While the studied websites did not have a large number of different third parties per website on average, numbers can still be considered high. 60.2% of the studied websites leaked URLs and 58.9% of the websites with search function leaked search terms. These are not good results considering that we are talking about health websites and data transfers happening without the user's consent.
- **Geographic disparities in website privacy.** The current study shows that several countries such as Spain and Ireland have a pressing need to clearly improve the privacy of their websites both in terms of number of data leaks and third-party services they choose (for example, 90% of both Spanish and Irish websites leaked visited URLs to Google).
- **Strong involvement of Google.** The vast majority of the detected data leaks had to do with Google's services, mostly Google Analytics. Google's services were found collecting sensitive personal data in every single studied country. Almost 50% of the studied websites leaked URLs to Google.
- **Problematic third parties.** Besides Google, there were some third parties that do not belong to health-related websites in Europe. One such third party is Russian analytics service Yandex. Especially in the current geopolitical situation, using this service puts the online privacy of European patients in jeopardy.

5.2. Implications for Developers

There is no sound reason for the use of third-party analytics in web-based health services. It is likely that the analyzed health services have not leaked sensitive personal data on purpose. However, while third parties may not abuse it, the data they receive is identifiable [14], and the fact this data is sent to third parties is already a significant privacy risk. As software development is a risk-based discipline [22], these kinds of threats should be avoided. In what follows, we will discuss multiple proactive measures that web developers should employ to prevent leaks of sensitive health data.

The most obvious and easy solution to the privacy problem caused by third-party services would be to omit these services from websites entirely. Locally hosted services like Matomo [21] should be used if web analytics are deemed absolutely necessary. This solution enables the health service provider to fully control the collected data and it does not unnecessarily fall into the hands of any third party.

If third-party services are still deemed to be really necessary and developers want to integrate them in a web service, a careful evaluation should take place. The justification of using each selected external

service should be carefully documented, and every service should be thoroughly vetted for privacy. While the use of third-party analytics is difficult to justify for health websites, some uses for trusted third-party services may be necessary [5]. For example, appointment booking systems and chat services may be essential for the core functionality of a medical website.

A rigorous assessment of data a critical website sends to third parties should be conducted. Such network traffic analysis could be carried out using the method we followed in the current study. As the outgoing network traffic is examined, special attention should be paid to the critical function in the web-based health service such as the search function. By employing network traffic analysis, developers can get a detailed overview of what kinds of data third parties collect and which third-parties cause privacy concerns and should be removed from the website. Because many platforms such as content management systems often offer easy integration options for third-party analytics or even include these data-collecting services by default, network traffic analysis is essential to give the website maintainers a good understanding of what data actually flows out of their website.

Domain area knowledge, such as understanding the specific privacy requirements in the healthcare sector, is very important [23]. The developers should acquire at least some level of knowledge about the privacy regulations and threat models concerning this specific industry. The development team has to effectively communicate with stakeholders in order to design a system where sensitive health data is appropriately protected. Critical online resources such as hospital websites should also always undergo a rigorous external privacy audit, which does not seem to currently be the case even in the strictly regulated EU area.

5.3. Implications for Users

For an average user, leaking personal data concerning health can have significant consequences. If the user's personal data is leaked and somehow becomes public, the user can be stigmatized or discriminated against. When users become aware of these kinds of serious data leaks, they can also completely lose their trust in the online health services. As we have seen in the current study, data leaks can happen without users knowing about them and without any consent. It is also clear most users do not have the capacity to investigate how third-party analytics services collect data on them.

The risk of a data breach, selling personal data, or handing over personal data becomes more pronounced when the data is stored in multiple servers around the world. Even when the third-party analytics service provider stores the data securely and no leak occurs, the user may still find this data collection unethical and should be able to control whether it is happening. For example, it is worth asking why health data of EU citizens should be collected by U.S.-based large technology companies to begin with.

6. Conclusion

The findings of this study emphasize the need for software developers and data protection officers around the EU area to carefully evaluate the third-party services used in their web services. Following the privacy-by-design approach and adopting fair data processing practices when handling sensitive customer data is essential. Our results also highlight the importance for the data protection authorities in the EU to enforce privacy regulation more strictly especially when it comes to health-based websites.

Users also have to be transparently informed about what kind of personal data on them is being shared and with whom. Such sharing has to happen with the users' consent, unlike what was happening on the websites analyzed in this study. In web-based health services, external services that collect sensitive data should not be used unless they are fully trusted. Checking their reputation and data processing practices thoroughly and conducting a risk analysis is important. Failing to carry out such evaluation increases the vulnerability of several user groups that are already vulnerable. After all, it would be important to build web-based health services that make users feel as secure and assured as they would in a traditional, in-person healthcare setting.

Acknowledgments

This research has been funded by Academy of Finland project 327397, IDA – Intimacy in Data-Driven Culture.

Declaration on Generative AI

The authors have not employed any generative AI tools.

References

- [1] E. Isibor, Regulation of healthcare data security: Legal obligations in a digital age, Available at SSRN 4957244 (2024).
- [2] S. Saha, C. Chowdhury, S. Neogy, A novel two phase data sensitivity based access control framework for healthcare data, *Multimedia Tools and Applications* 83 (2024) 8867–8892.
- [3] X. Yu, N. Samarasinghe, M. Mannan, A. Youssef, Got sick and tracked: Privacy analysis of hospital websites, in: *2022 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, IEEE, 2022, pp. 278–286.
- [4] A. R. Zheutlin, J. D. Niforatos, J. B. Sussman, Data-tracking on government, non-profit, and commercial health-related websites, *Journal of general internal medicine* (2021) 1–3.
- [5] S. Rauti, R. Carlsson, S. Mickelsson, T. Mäkilä, T. Heino, E. Pirjatanniemi, V. Leppänen, Analyzing third-party data leaks on online pharmacy websites, *Health and Technology* (2024) 1–18.
- [6] T. Kriecherbauer, R. Schwank, A. Krauss, K. Neureither, L. Remme, M. Volkamer, D. Herrmann, Is personalization worth it? notifying blogs about a privacy issue resulting from poorly implemented consent banners, in: *Proceedings of the 19th International Conference on Availability, Reliability and Security*, 2024, pp. 1–7.
- [7] C. Utz, S. Amft, M. Degeling, T. Holz, S. Fahl, F. Schaub, Privacy rarely considered: Exploring considerations in the adoption of third-party services by websites, *arXiv preprint arXiv:2203.11387* (2022).
- [8] M. D. Huesch, Privacy threats when seeking online health information, *JAMA Internal Medicine* 173 (2013) 1838–1840.
- [9] T. Libert, Privacy implications of health information seeking on the web, *Communications of the ACM* 58 (2015) 68–77.
- [10] I. Kes, D. Heinrich, D. M. Woisetschlager, Behavioral targeting in health care marketing: Uncovering the sunny side of tracking consumers online, in: *Let's Get Engaged! Crossing the Threshold of Marketing's Engagement Era: Proceedings of the 2014 Academy of Marketing Science (AMS) Annual Conference*, Springer, 2016, pp. 297–297.
- [11] H.-G. Hwang, Y. Lin, Evaluating people's concern about their health information privacy based on power-responsibility equilibrium model: A case of taiwan, *Journal of Medical Systems* 44 (2020) 112.
- [12] V. Wesselkamp, I. Fouad, C. Santos, Y. Boussad, N. Bielova, A. Legout, In-depth technical and legal analysis of tracking on health related websites with ernie extension, in: *Proceedings of the 20th Workshop on Workshop on Privacy in the Electronic Society, WPES '21*, Association for Computing Machinery, New York, NY, USA, 2021, p. 151–166.
- [13] K. Y. Yigzaw, S. D. Olabariaga, A. Michalas, L. Marco-Ruiz, C. Hillen, et al., Health data security and privacy: Challenges and solutions for the future, in: E. Hovenga, H. Grain (Eds.), *Roadmap to Successful Digital Health Ecosystems*, Academic Press, 2022, pp. 335–362.
- [14] M. Huo, M. Bland, K. Levchenko, All eyes on me: Inside third party trackers' exfiltration of phi from healthcare providers' online systems, in: *Proceedings of the 21st Workshop on Privacy in the Electronic Society, WPES'22*, Association for Computing Machinery, New York, NY, USA, 2022, p. 197–211.

- [15] A. B. Friedman, L. Bauer, R. Gonzales, M. S. McCoy, Prevalence of third-party tracking on abortion clinic web pages, *JAMA Internal Medicine* 182 (2022) 1221–1222.
- [16] A. B. Friedman, R. M. Merchant, A. Maley, K. Farhat, K. Smith, J. Felkins, R. E. Gonzales, L. Bauer, M. S. McCoy, Widespread third-party tracking on hospital websites poses privacy risks for patients and legal liability for hospitals, *Health Affairs* 42 (2023) 508–515.
- [17] A. Surani, A. Bawaked, M. Wheeler, B. Kelsey, N. Roberts, D. Vincent, S. Das, Security and privacy of digital mental health: An analysis of web services and mobile apps, in: *Conference on Data and Applications Security and Privacy*, 2023.
- [18] R. Carlsson, P. Puhtila, S. Rauti, Towards an automatic tool for detecting third-party data leaks on websites, *Proceedings <http://ceur-ws.org> ISSN 1613 (2023) 0073*.
- [19] J. Gamalielsson, B. Lundell, S. Butler, C. Brax, T. Persson, A. Mattsson, T. Gustavsson, J. Feist, E. Lönroth, Towards open government through open source software for web analytics: The case of matomo, *JeDEM-eJournal of eDemocracy and Open Government* 13 (2021) 133–153.
- [20] S. Panda, R. Chakravarty, Evaluating the web accessibility of iit libraries: a study of web content accessibility guidelines, *Performance Measurement and Metrics* 21 (2020) 121–145.
- [21] D. Quintel, R. Wilson, Analytics and privacy, *Information Technology and Libraries* 39 (2020).
- [22] B. W. Boehm, T. DeMarco, Software risk management, *IEEE software* 14 (1997) 17.
- [23] L. Sion, D. Van Landuyt, W. Joosen, Leveraging the domain experts: specializing privacy threat knowledge, in: *European Symposium on Research in Computer Security*, Springer, 2024, pp. 534–541.