Security & Privacy Awareness & Concerns of Computer Users Posed by Web Cookies and Trackers

Smah Almotiri

A thesis submitted to the Faculty of Design and Creative Technologies
Auckland University of Technology

In partial fulfilment of the requirements for the degree of Master of Information Security and Digital Forensics

School of Engineering, Computer and Mathematical

Sciences

Auckland, New Zealand, 7 October 2022

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgments), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

.

Signature: Date: 7 October 2022

Acknowledgement

I would like to express my gratitude to Dr. Alastair Nisbet, for giving me the chance to incorporate my interests in coding and data privacy into a master's thesis. I also want to acknowledge him for sharing his knowledge with me and guiding me through the different stages of this research. I additionally want to thank my family for their unwavering support, for being there for me whenever I needed them, and for their words of encouragement. Finally, I would like to express my appreciation to my husband for his constant support while I studied toward my master's degree and to my son, Yahya, for his tolerance throughout this time.

Abstract

Web-cookies are an essential element across every website. Despite its importance for the growth and development of websites, cookies are associated with severe security and privacy risks. In numerous empirical studies, browser cookies have been proven to be insecure and vulnerable to cyber-attacks, resulting in compromised user data. Additionally, cookies breach users' online privacy in violation of the rules and laws that are in place to protect information privacy. The objectives of this study are to evaluate users' awareness of the threats of HTTP cookies and web trackers to data privacy and security as well as to determine the level of end-user privacy concerns regarding violations of this technology. An anonymous online survey addressed to adult Internet users on an international level has been developed. 34 questions, including one of which required accessing a website set up for this research, were completed by 471 respondents in order to evaluate respondents' levels of awareness of the risks of web tracking for the security and privacy of their personal data as well as determining their level of privacy concern. The findings of the survey analysis revealed that users are highly aware of the dangers that cookies represent to the security and privacy of personal data. Further, the findings demonstrated that there is a significant level of concern among respondents regarding the collection, usage, and storage of personal data via web tracking and cookies. These findings lead to the conclusion that local and international legislation should work to develop a framework with unified standards on an international level for governance and protection of privacy of users' data in the context of cookies and tracking in response to the concerns of users from different regions around the world about the practices of web trackers on their data that pose security risks and violate privacy

Table of Contents

List of Figures

List of Tables

List of Abbreviations

IPA	Information Privacy Awareness
ISA	Information Security Awareness
IPC	Information Privacy Concern
OBA	Online Behavioral Advertising
KPI	Key Performance Indicators
HTTP	Hyper Text Transfer Protocol
IP	Internet Protocol
CIA	Confidentiality, Integrity, and Availability
ML	Machine Learning
XSS	Cross-Site Scripting
XSRF	Cross-Site Request Forgery
CSRF	Cross-Site Request Forgery
MITM	Man-In-the-middle
PII	Personally Identifiable Information
SPI	Sensitive Personal Information
IPPs	Information Privacy Principles
Tor	The Onion Router
GDPR	General Data Protection Regulation
TPT	Tracking Prevention Tool
IUPAC	Internet Users' Information Privacy Concerns

OPL	Online Privacy Literacy
EDA	Exploratory Data Analysis
IoT	Internet of Things
SNSs	Social Network Sites

CHAPTER 1: INTRODUCTION

1.1 Background and Motivation

Individuals utilize the World Wide Web to browse leaving a huge number of digital fingerprints from users on countless websites across the globe. The most prominent industry to capitalize from such digital fingerprints that users record on websites is web tracking. Web tracking refers to the practice of tracking a user across several websites, such as advertising agencies (Sprankel, 2011). The function of tracking is to gather data about the activities, interests, and locations of users (Boerman et al., 2017; Bujlow et al., 2015; Kumar et al., 2012). In order to comprehend visitors' preferences and behaviours and deliver them with personalised content, web tracking makes it possible to identify visitors to websites as well as the Web pages they have viewed (Hamed & Ayed, 2015; Kumar et al., 2012). When an end-user accesses a website, their behaviours, such as mouse clicks and button pushes are automatically recorded and provided to different monitoring services by the website in an anonymous manner. Such functions can be performed via web cookies using the HTTP Internet protocol (Sipior et al., 2011). Cookies are arbitrary strings that are kept on the user's computer and are owned by a website. Its goal is to enable stateful browser-server communication using a stateless protocol (Sipior et al., 2011; Sprankel, 2011). Cookies and other tracking technologies that enable third parties to uniquely identify users can be used in conjunction with the "referrer" header, which informs the third party which first-party site the user is now visiting, to get users' browsing histories (Englehardt & Narayanan, 2016). Pages on web browsers are progressively made up of material from several undisclosed third-party services in the networking sites, commercial advertisement, predictive modelling, and other industries (Mayer & Mitchell, 2012). Hence, web tracking is of considerable interest to marketing firms (Sprankel, 2011). Such industries have significantly advanced web tracking over time. Much of the earliest analysis of online tracking is provided by Krishnamurthy and Wills (2009), who tracked the increase of the major third-party businesses from 10% to 20–60% of top websites between 2005 and 2008. Studies from the next years reveal a continuous rise in third-party tracking as well as a range of tracking methods (Acar et al., 2013; Mayer & Mitchell, 2012; Roesner et al., 2012). Libert (2015) examined the top 1 million websites' third-party HTTP requests demonstrated that Google, through its many third-party domains, is able to follow visitors across approximately 80% of websites. In 2015, online tracking increased from 10% in 2005 to 73% (Wambach & Bräunlich, 2016). Researchers discovered an average of about 6 external queries per webpage in 2015 (Wambach & Bräunlich, 2016). This signifies that at least six more hosts were notified of each website visit. Web tracking is encouraged since it often leads to increased profits. Companies may offer tailored experiences and advertising by building profiles of consumers' interests, traits, and demographics (Simpkins et al., 2015). Industries frequently utilise

(tracking) cookies to monitor customer surfing habits (Boerman et al., 2017). The browsing patterns of online users are regarded as a valuable resource for creating complete profiles (Falahrastegar et al., 2016) and are very relevant to enhancing the commercial operations (Roesner et al., 2012). The top 100 websites collect more than 6,000 cookies overall, 83% of which are third-party cookies with some websites gathering upwards of 350 cookies (Altaweel et al., 2015). These cookies make it possible for companies to gather detailed information on millions of users, among which will be leveraged for online behavioural advertisement OBA (Boerman et al., 2017). Online profiling, also known as online behavioural targeting, is the practice of tracking users' online activity and displaying adverts that are specifically tailored to each particular user (Zuiderveen Borgesius, 2015). OBA exploits personal information to personalise advertisements and make ads look more appealing pertinent to the user (Boerman et al., 2017).

In addition to targeted advertising, online tracking could be employed for social network integration (Mayer & Mitchell, 2012), sophisticated web site analytics (Bujlow et al., 2015), and personalization (Mayer & Mitchell, 2012; Roesner et al., 2012). Also, increased hazards such as pricing discrimination, governmental surveillance (Bujlow et al., 2015), and identity theft (Malandrino et al., 2013), are implied by online tracking practises. User information is also desired after by large-scale organizations and cybercriminals that are capable of conducting cyberattacks. Numerous studies highlight the risks of cookies on websites, which can subject user data to serious cyberattacks (Hussain et al., 2021; Jussila, 2018; Kaur & Garg, 2021; Kavisankar et al., 2016; Rodríguez et al., 2020; Sheikh, 2021; Tariq et al., 2021; WatchGuard, 2021; Zhou & Wang, 2019). Cookies and online tracking further expose users to challenges and invasions of privacy in addition to the security problems that pose. The gathering, use, and processing of this data, as well as its maintenance and storage, are all legally consider as violation for data privacy (Hamed & Ayed, 2015; Jegatheesan, 2013; Simpkins et al., 2015; Wambach & Bräunlich, 2016). Users' security and privacy are seriously jeopardized by the widespread gathering and tracking of end-user data through the use of cookies and other tracking technologies. The levels of awareness of the threats presented by cookies and web trackers globally have not been the subject of studies evaluating information awareness in terms of security and privacy. Additionally, research on privacy concerns has not focused on users' privacy concerns in relation to cookies and web tracking. The personal motivation for doing this research is that, with an estimated 7.9 billion people around the world today, 5.25 billion of them have access to and utilise the Internet. This indicates that 66.2% of the world's population utilise the Internet (Broadband Search, 2022). The majority of people worldwide are impacted by cookies, which reflect the global Internet user population. HTTP-cookies use tracking and identifiers to build user profiles and make individual identification possible. Although web cookies are now basically obligatory when visiting websites (Jegatheesan, 2013), most websites do not correspond to the laws and regulations that guarantee the

protection of users' privacy (Hu & Sastry, 2019; Li et al., 2019; Politou et al., 2018). Therefore, one of the primary issues in the contemporary digital world might be HTTP-cookies. Information identifying end-users is valuable. The perspectives and beliefs of users towards the possible threats that cookies bring to their personal data need thus to be studied. Users' attitudes can be identified by assessing their level of awareness and concern about the security and privacy risks posed by cookies. This evidence can then be used to influence decision-makers, organisations, and legislative bodies to strengthen public policies that protect users' privacy. The significance of this study lies in its thorough investigation of the end-user's level of privacy and security awareness in the context of cookies and tracking. This study, also, investigates the level to which cookie data collection, use, and storage raise privacy concerns.

Therefore, the following are the main research questions that this study aims to address:

Question 1 (Q1). *To what extent are computer users aware of the security and privacy threats associated with web cookies?*

Question 2 (Q2). *How concerned is the computer user about the use of their data that is collected through web cookies and Internet use trackers?*

1.2 Research Approach and Findings

To answer the research questions, an appropriate quantitative approach has been employed which is explained in Chapter 3. An anonymous online survey was adopted as the approach to collect data on end-user awareness of and concern over the consequences of web tracking and cookies on user data privacy and security. Based on the study's research questions, the survey's sections were designed and developed. The survey's first and second sections are intended to gather information on the participants' demographics data and technical backgrounds. The third, fourth, and sixth sections are intended to measure users' awareness of the privacy and security threats presented by cookies and web trackers. In terms of basic understanding and privacy practises, the results indicated high levels of end-user awareness, although users' understanding of legal aspects was less prevalent. The sixth and seventh sections of the questionnaire then examined levels of concern about security and privacy threats that web trackers posed to end-user. However, the fifth section primarily assesses users' concerns regarding the collection, storing, and usage of personal data via cookies. The seventh section involved a website that is part of this study in order to evaluate user privacy behaviors, comprehend user motivations for reading or not reading the privacy policy for cookies on websites, and compare this to users' attitudes toward privacy concerns. The survey's results revealed that the participants were extremely concern regarding cookie-based data collection, use, and storage. The

results also revealed that the primary reasons for reading the cookie policy are interest in the sorts of cookies being used on the website, while the main obstacles for not reading are the long and time-consuming style of the privacy policies. Additionally, the results showed that individuals' attitudes and behaviours regarding privacy concerns in context of cookies were consistent.

1.3 Thesis Structure

There are six chapters in this thesis. The thesis subject was presented in this chapter, along with some contextual information on its background and the problem of relevant research into the subject. Also, the motivation for conducting this research was explained, and its importance was highlighted. The research design employed in this study and the key findings were then indicated in this chapter. A review of the literature is presented in Chapter 2. It covers four key themes: the development of tracking technologies, associated consequences of web tracking on user data security, the concept of digital privacy in context of web tracking and web tracking in the fields of information privacy awareness (IPA), information security awareness (ISA) and information privacy concern (IPC). Chapter 2 opens with a review of the evolution of web-tracking on the Web, followed by an explanation of its concepts and a review of the most prominent and well-known web-tracking techniques and tools. Next, based on fundamental security goals including confidentiality, integrity, and availability, the chapter presented the most potential security repercussions that web trackers have on users' personal data, as evidenced by a range of recent empirical studies. The review of online tracking in the realm of digital privacy followed, showing the obstacles that web trackers provide for users' privacy in terms of data collection, usage, storage, and protection. The implementation of several existing online tracking prevention and regulation mechanisms is also reviewed. Chapter 2 concludes with a body of literature on conceptions of user information privacy awareness, information security awareness, and information privacy concerns in cyberspace in various online settings, identifying their impact on users, and highlighting the lack of focused examination in the context of cookies and web tracking. In order to fill in the gaps in the existing body of literature, the study's main research questions are identified in Chapter 3. Next, this chapter mainly focuses on the approach selected for this study and how it was developed from previous relevant studies. The findings of the online questionnaire are presented in Chapter 4. Chapter 4 analyses the findings from the online survey, starting with the demographic findings and moving on to the results on participant level of privacy and security awareness, participant level of privacy concern for privacy and participant access to the study website. Along with the outcomes of statistical analyses to determine the associations between participant demographics and their technical backgrounds in connection to levels of awareness and concern regarding data privacy and security in the context of web tracking. In order to respond to the research questions presented by this research, Chapter 5 discusses the findings from Chapter 4 and correlates them to the key issue outlined in the literature review.

This chapter addresses the implications of the findings and makes suggestions for boosting the privacy and security of end-user personal information as a result of cookies and online tracking. Finally, Chapter 6 concludes with recommendations for strengthening the security and privacy of end-user data in the context of web tracking. It also identifies possible future study areas that, if addressed, could further current research of end-user digital privacy protection.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

This chapter reviewed the literature on web tracking, examining its techniques and potential threats to end-user data security and privacy. The goals of this review are to summarize the current state of knowledge on web-based user tracking and the threats it poses to user data privacy and security, as well as to identify any limitations of the existing body of literature that require further investigation. The important challenges encountered in delivering a secure and private surfing experience for online users, which led to the development of the thesis's primary research topics, are also discussed. Section 2.2 described the most important aspects and motivations for the emergence and development of online tracking technologies, as well as the most prominent web-based methods for collecting, tracking, and storing user data, one of which is web cookies. Section 2.3 looked at the security issues and threats that tracking techniques potentially pose to users. Section 2.4 discussed the concept of privacy on the Internet and the most important principles of user data privacy that research indicates that web trackers violate and do not comply with based on many literary reviews in the same field. Section 2.5 discussed users' awareness of the major risks to their data's privacy and security on the Internet, as well as their concerns about this. Finally, Section 2.6 brings the literature review to a conclusion.

2.2 The Emergence of Web Tracking Technology

2.2.1 Online Business Development

The Web has promoted the growth of technologies and commercial practices, which has resulted in the emergence of innovative approaches in e-commerce, marketing, and sales. The Internet delivers a wealth of user data that is handled for a variety of purposes. In 2009, the volume of data generated on a daily basis was projected to be less than 1 petabyte; however, this has since expanded dramatically to 2.5 quintillion bytes (Jamiy et al., 2015). According to Statista, there will be 74 zettabytes of data generated in 2021. This represents an increase from 59 zettabytes in 2020 and 41 zettabytes in 2019. The growth in the amount of data generated on the Internet is more likely to be proportional to the increase in the number of Internet users. According to Statista, global Internet user growth from 2018 to 2023 will reach 5.3 billion by 2023. For the whole period from 2018 to 2023, the annual growth is 6%. Hence, the data received from the websites are processed using a variety of models. As a user surveillance model, selling collected data to advertising companies, predicting the characteristics, preferences, and activities of users, and interacting on the web (Angwin & McGinty 2010). The modern trend emerging in industries like marketing is using user data to establish lasting connections with customers. Online advertisements have benefited from the growth

of web tracking technology. Marketers receive a plethora of information on customer attributes and interests via the website, which is then utilized to better market segmentation and target marketing in order to engage users on an individual level (Sipior et al., 2011). The most extensively deployed form of online advertising is known as online behavioural advertising (OBA), which is described as the surveillance of users' online behaviour and the usage of that data to advertise to consumers with more targeted advertisements (Boerman et al., 2017). Otherwise, this is referred to as behavioural targeting, and it entails creating a user profile using web-based tracking tools in order to improve the efficiency of advertising. It is becoming increasingly significant in the online advertising industry (Jegatheesan, 2013; Yan et al., 2009). Website owners collaborate with advertising partners, generally known as vendors, who utilise cookies, which are text files stored on users' devices that allow them to be identified and tracked online even after they leave the actual webpage (Ur et al., 2012). First-party cookies are used by website owners to track visitor traffic for detection purposes; however, third-party ad vendors repeatedly add extra cookies to follow user activity even after they leave the primary webpage. This allows them to track the many Webpages the user visits, and appropriate adverts will be displayed on partner websites depending on the user's preferences (Ur et al., 2012). According to empirical research assessing user behaviours for digital advertising using data gathered from third parties that track users; the efficacy of commercials is greater and has a beneficial influence (Hoban & Bucklin, 2015). Since web marketing agencies can track users' actions when an ad is clicked and which ads drove the greatest traffic back to the website (Miller, 2010). Moreover, websites are becoming increasingly made up of information from a variety of unconnected "third-party" e-commerce sites in advertising, analytics, social networking, and other industries (Mayer & Mitchell, 2012). When people browse the web, service providers for example YouTube, elPais and content providers such as Google, Facebook, Amazon and other third parties like DoubleClick acquire huge quantities of data. This data collected by web tracking technologies varies and includes personal data, engagement data, behavioural data, and attitude data of users. Identities, postal addresses, email addresses, and demographic and behavioural information, such as prior purchases, are among the information collected (Bujlow et al., 2015). The data is utilised to generate profiled audiences and assess advertising effectiveness. As a result, the first party will be able to sell the identity to a third-party (Wills & Zeljkovic, 2011). Datalogix, for instance, is a data supplier that acquires data from consumer marketing firms and data compilers (Bujlow et al., 2015). Datalogix, additionally, receives and collects information on website visitors, including browser kinds, IP addresses, ad types served, and ad delivery dates (Bujlow et al., 2015). Therefore, user-website interaction yields a range of data on how people engage and behave while on the site. Meanwhile, because this data is often unstructured, it may not be immediately relevant for decision-making until it has been examined and data retrieved. So, the importance of web data analytics in the growth of a website cannot be overstated.

2.2.2 Web analysis

Website analytics assist website owners interpret their users by tracking user behaviours so they can see what users doing and how they're doing it. Web analytic is the method of quantifying, gathering, interpreting, and reporting on website data in order to understand how a website's audience uses it and how to improve it (Web Analytics Association,2008). It keeps track of crucial data including unique visitors, page views, traffic sources, and exiting pages, as well as analyzing website visitor behavior and traffic flow (Kumar et al., 2012). Web analytics, on the other hand, is not only a technique for evaluating website traffic; it is also a commercial tool for marketing research, as well as determining and improving the efficacy of a webpage. Through detailed research of visitor engagement with a website, Booth and Jansen (2010), introduce a set of approaches for assessing websites in order to maximize income and customer satisfaction. Moreover, Ellonen et al. (2015) use a unique dataset of real-life clickstream data from 295 visitors to a magazine's website to examine customer behavior patterns on the website. Researchers discovered some fascinating behavioral trends, such as the fact that 86 % of all sessions only go to the magazine's blogs (Ellonen et al., 2015). Furthermore, this is exceptionally effective for measuring traffic and recognizing new updates, both of which may be quite beneficial for marketing purposes (Phippen et al., 2004). As a result, monitoring tracked navigational data is vital to any online business's performance. However, the data tracking and data analysis processes are also part of the traffic analysis process. Capture, store, analyse, and report are the four steps of the traffic analysis process (Ehikioya & Lu, 2020). Off-site web analytics and on-site web analytics are the two main types of web analytics approaches. Off-site web analytics is mostly used to assess a website's potential visitors or customers (opportunity), buzz (comments), visibility (share of voice), and general online events in general. To assess the success of the website, the data supplied by this type of analytics is frequently linked with the agreed-upon key performance indicators (KPI) (Ehikioya & Lu, 2020). Whereas on-site web analytics is carried out by the website owners directly to evaluate the performance of the website from a business generation standpoint (Ehikioya & Lu, 2020). The most extensively used on-site web analytics services are Google Analytics and Adobe Analytics. Web analytics can collect a large number of distinct data points or metrics. This data is collected by techniques called information gathering technologies. Web-tracking and information-gathering technologies such as cookies, Flash cookies, and beacons make it easy to get personal information from web users, frequently without their awareness.

2.2.3 Web tracking mechanisms concept

The mechanism through which websites track, preserve, and gather information about users' web surfing behavior is known as web tracking. This is usually done using a portion of the user’s online browsing record. Web-tracking methods can be stateful or stateless.

2.2.3.1 Stateless tracking

A fingerprint-based tracking approach is known as a stateless tracking (Ishtiaq et al., 2017). Stateless tracking does not need the storage of any data on the computer of the user. Advertisement and anti-fraud businesses are increasingly using the stateless device fingerprinting (Sanchez-Rola et al., 2017). A webpage is able to gather information about the browser's attributes that, when combined, constitute a unique or nearly unique identification (Ehikioya & Lu, 2020; Mayer & Mitchell, 2012). Otherwise, the user's identity is determined by this identifier (Eckersley, 2010). These practices help advertisers to expand the amount of previously collected user data and to share user identifications more easily between various detection providers (Sanchez-Rola et al., 2017), since these fingerprints make it possible to correlate browsing sessions as well as a user's identity. However, some features necessitate the use of a script or plug-in for active discovery. Other features can be inferred from network traffic in an unobtrusive way (Mayer & Mitchell, 2012). Commonly, there are various features that can be obtained from different components of the browser, such as JavaScript or plugins, to uniquely fingerprint a device. As for the JavaScript technique, even when traditional modes of system identification such as the user-agent header are modified or concealed, the browser's JavaScript engine allows the detection of the browser version, operating system, and microarchitecture (Mowery et al., 2011). Mayer (2009), fingerprinted 1,328 users via hashing the contents of the JavaScript available browser properties navigation, screen, navigator, plugins, and navigator, which allowed researcher to uniquely fingerprint more than 96% of users. The list of installed typefaces was the most accurate feature for device fingerprinting, according to Eckersley. What is more, by abusing JavaScript's visited-link colour feature, the browsing history may be obtained (Eckersley, 2010). Additionally, canvas fingerprinting is another stateless tracking technique. This method takes advantage of the Canvas API found in more contemporary web browsers, and it may be used to derive a unique fingerprint by exploiting tiny changes in displaying the same text or WebGL scenes (Sanchez-Rola et al., 2017).

2.2.3.2 Stateful tracking

Stateful tracking, on the other hand, makes advantage of existing techniques to store data on the user's computer. For example, Implementing an HTTP cookie to uniquely identify a user by exploiting identifying information stored locally on the user's machine (Amarasekara et al., 2020). Commonly, websites often use stateful web technologies to incorporate a unique identification. To accomplish this, third-party websites get access to various components of the websites in order to gather and keep user information (Sanchez-Rola et al., 2017).

2.2.4 Web tracking technologies

2.2.4.1 HTTP-Cookies

HTTP Cookies, often known as cookies, are frequently used for session management, storing site settings, client authentication, and identification. Cookies were created by Netscape in 1994 to remedy a fundamental flaw in the Hyper Text Transfer Protocol (HTTP). The protocol did not allow for the storage of state or memory. HTTP is a stateless protocol in which transactions are unconnected (Sipior et al., 2011). Consequently, it was unable to recall the contents of a shopping cart, save language choices, or manage login status. Cookies, on the other hand, allow the Set-Cookie HTTP header to transmit information between the server that produced the cookie and the user's browser, making HTTP a stateful (Sipior et al., 2011). Through a unique visitor ID issued by the Web site and other information stored in the cookie, the server is able to recall details of prior interactions with the website. Cookies can utilise this ID to link numerous site visits to the same person (Felten & Schneider, 2000). So, adding unique IDs to each browser was one of the solutions to alleviate this restriction. A cookie, as defined by Millett et al. (2001), is a tiny text string transmitted by a web server to a browser. The text is then saved by the browser, usually on the user's hard drive, and eventually delivered back to a web server. Otherwise, cookies allow a website to store data on users' computers, which is then retrieved when the user visits the site again. These cookies are primarily used to preserve stateful information, such as goods placed in a shopping cart and data submitted into form fields such as name, address, and password (Ishtiaq et al., 2017; Li et al., 2015). Therefore, the server can store and retrieve data on the client side with the aid of HTTP cookies. However, HTTP cookies have been used for cross-domain tracking in the past (Kristol, 2001). Recently, emerging technologies such as HTML5's local storage and ETags might improve the HTTP cookie-based tracking process's dependability (Ayenson et al., 2011). Cookies are placed on the user's computer in one of two ways: through a JavaScript API call or through HTTP replies with the Set-Cookie header (Bujlow et al., 2015). Cookies can also be read by services in two ways. Using Cookie headers, they are first immediately linked to HTTP requests sent to the domain to which the cookies belong. They can also be obtained explicitly using a JavaScript API and then provided to the server in any method (Roesner et al., 2012). However, there are two types of cookies which are session cookies and persistent cookies. Firstly, session cookies are short-term cookies that are commonly used to remember user preferences or navigational status. When a user logs in, a service sets them and then removes them when the user logs out. Session IDs are the unique random numbers that are used to identify a session and are often kept inside cookies. Whereas, persistent cookies expire after a certain period of time, whereas session cookies expire when the user shuts the web browser (Bujlow et al., 2015). Persistent cookies are frequently used to keep an authorized session with a server by storing identifiable details, user interactions, or authentication fingerprints. Until they are

deliberately erased or expire, these files remain in the user's browser. They are transmitted back to the website by the browser unmodified each time it views it and may thus be used by websites to monitor people across visits. However, Kristol (2001), argued that this form of information gathering is a technological need that is transparent to the visitor and the website, and it is not harmful to users' privacy because no personally identifiable data is used. On the contrary, further research indicates that this technology poses a risk to users' security and privacy. It's estimated that 39% of users erase cookies at least once a month (Marshall, 2005). While keeping state throughout browser sessions and future visits to the same websites improves the user experience, most users frown on tracking user behaviours that do not directly contribute to the enrichment of user experiences (Hoofnagle, Urban, & Li, 2012). As users navigate from site to site and webpage to webpage for the purposes of browsing, reading, downloading, or engaging to register at a website, submit comments or inquiries, or make an order for products supplied, someone may utilise cookies without the user's permission (Sipior et al., 2011) .To give the user extra privacy, the popular Private Browsing mode was created to allow users to browse the web without leaving a trail in their local storage (Sanchez-Rola et al., 2017). Since May 25, 2018, up to 15.7 % of websites in various countries have introduced new privacy policies, culminating in 84.5 % of websites having privacy policies (Degeling et al., 2018). In Europe, 62.1 per cent of websites now display cookie consent messages, up 16 % from January 2018. These alerts notify users about a website's cookie policies and tracking practices (Degeling et al., 2018). As cookies' capacity to follow users improves, so do the threats. Cookies have progressed beyond just logging visits to a website to store personal information provided by users in order forms, registration or payment pages, and other online forms (Wagner, 2020).

2.2.4.2 Flash cookies

Flash cookies, like HTTP cookies, were created to improve a user's navigation experience by making it a stateful (Soltani et al., 2010). The Adobe Flash plugin creates flash cookies which is an embedded Flash programmer on a webpage (Simpkins et al., 2015). According to Adobe, Flash technology is used to provide over 75% of web videos (Sipior et al., 2011). It is also used by media firms to distribute games and animation. As a result, disabling Flash cookies may have an impact on a user's web experience (Sipior et al., 2011). As it is meant for basic Web tasks like preserving a user's setting and language choices or recalling where a user left off playing a video game (Sipior et al., 2011). Flash cookies can be maintained or retrieved by default anytime a user visits a page that contains a Flash app, and they are set to not ask for permission to keep the user's data. As HTTP cookies were increasingly deleted by users, in response Flash cookies were created (Sipior et al., 2011). Flash cookies differ from traditional browser cookies in terms of the amount of data they may hold and how they can be deleted. Adobe Flash uses Local Shared Objects (LOSes) to store data on users' systems that may be up to 100 KB in size, whereas HTTP cookies are only

4 KB (Ayenson et al., 2011; Sipior et al., 2011). The 100KB of data by default, makes them more appropriate for tracking than the HTTP cookies (Bujlow et al., 2015). Additionally, Flash cookies are not likely to be alerted when they are installed on a computer, and they do not expire like regular cookies (Simpkins et al., 2015). If a user deletes HTTP cookies, a Flash cookie can revive the cookies in a process known as the "respawning" (Simpkins et al., 2015; Sipior et al., 2011; Soltani et al., 2010). So, the cookie's unique ID can be re-assigned to a new cookie using Flash cookie data as a backup. Because Flash cookies are stored in a different place than HTTP cookies, users attempting to remove them may be unable to discern which files to delete (Soltani et al., 2010). They are more difficult for the user to wipe than HTTP cookies, and they are available from all of the system's browsers since all instances of Adobe Flash plugins share the same storage directory (Bujlow et al., 2015). This means that customers who "toss" their HTTP cookies to avoid surveillance or to stay anonymous online are still individually recognized by advertising organizations (Soltani et al., 2010). Consequently, third-party companies are allowed to follow users across many browsers (Mittal, 2010). Flash cookies are not restricted by browser tools. As a result, browser privacy settings like deleting HTTP cookies, clearing history, or clearing the cache are worthless when it comes to the Flash cookies (Soltani et al., 2010). The majority of people are unaware of Flash cookies and have no idea how to erase them (Sipior et al., 2011), as they are stored in a separate place than HTTP cookies, and browsers do not presently support displaying or deleting them in the same way that HTTP cookies do (Simpkins et al., 2015). The user must manually erase Flash cookie files since they are not managed by the Internet browser and are stored outside of the browser's control. On the top 100 websites, researchers discovered that Flash cookies are a common technique for storing data (Soltani et al., 2010). The study implies that this is harmful from a privacy standpoint since many websites saved identical information in both HTTP and Flash cookies, generally with telling variables, in addition to keeping user preferences (globally unique identifiers) (Soltani et al., 2010).

2.2.4.3 Evercookies

Evercookies (also called zombie cookies or supercookies) consider taking advantage of the weaknesses of a user interface for managing certain in-browser storage mechanisms to make them more resistant to user removal. A supercookie is a string of code injected into the data user downloading known as a "unique identification header,", rather than a little file stored by the web browser. When a user accesses a website that uses the ever cookie API, the web service produces an identifier and stores it in the browser's different storage mechanisms (Acar et al., 2014). Evercookies is a solution to Internet users' desire to wipe cookie storage for continued data tracking. In order to exist, reconstruct after deleting, or even reproduce in different browsers on the same machine, it uses a variety of storages (Bujlow et al., 2015). The evercookie makes advantage of several of the browser's storage features, including the following HTTP Cookies, Local

Shared Objects (Flash Cookies), Silverlight Isolated Storage, Web history, ETags, Web cache, and window. name DOM property, Internet Explorer user Data storage, HTML5 Session Storage, HTML5 Local Storage and among others (as cited in Bujlow et al., 2015). This is more likely to indicate that Evercookie is exceptionally resistant to user deletion. Evercookie is an innovative data tracking technology that includes numerous current tracking methods, minimizing the duplication of data gathering methods used by many commercial websites (Nielsen, 2019). Ayenson et al. (2011), discovered the first usage of caching ETags and localStorage for erasing cookies in 2011. This was the first research to discover evercookies in browser cache and HTML5 localStorage methods. Whereas Sörensen (2013), investigated into the usage of cache as a long-term storage method and discovered multiple instances of HTTP cookies respawning from cached page content. The author believed that cache-based storage is being exploited as a repository for personally identifiable information. The study implied that the topic of isolating information flow between multiple parties is a common focus of research in the realm of privacy-enhancing technology. We have proven that considering isolation across temporal domains is equally useful. Self-destructing identifiers are a simple and helpful notion to implement. Likewise, Acar et al. (2014), discovered a novel evercookie pattern, Indexed DB, that has never been recognized previously; additionally, it is widely used on the Internet. They have been reported to be employed by many major websites to avoid purposeful user actions (Acar et al., 2014). Recently, Schmidt (2020), examined evercookie to modern desktop and mobile browsers and discovered that it can still use virtually all of its original storing methods. Therefore, Evercookies make use of several storage methods that are less transparent to users, to keep track of browsers' exceptionally well actions (Acar et al., 2014).

2.2.4.4 Browser fingerprinting

Browser fingerprinting has become a popular method of tracking people without their knowledge. Fingerprinting, unlike cookies, is a stateless method that does not save anything on the client-side information storage. Rather, it queries specific characteristics that are available through the online browser (Boda et al., 2011; Vastel et al., 2018). This may enable a variety of commercial fraud detection services to recognize fraudulent transactions using a browser-based device fingerprinting (Alaca & Van Oorschot, 2016). Furthermore, even when people try to avoid being tracked, these fingerprints help in accumulating long-term records of their browsing history offering personalized advertising or targeted attacks (Abgrall et al., 2012; Nikiforakis et al., 2015). Meanwhile, due to the lack of transparency and control in terms of removing or deleting cookies, browser fingerprinting is more aggressive than cookie-based tracking (Iqbal et al., 2021), presenting a big issue for Internet privacy activists (Fifield & Egelman, 2015; Fouad et al., 2021). Several studies have demonstrated that fingerprints gathered via browsers may reliably identify the user. Figure 2.1 shows the several types of data that can be gathered from the browser fingerprinting. In the

literature, the Panopticlick fingerprinting experiment is the first serious assault (Eckersley, 2010). Eckersley examined the possibility of fingerprinting in the absence of common tracking mechanisms such as cookies. This research has available on the public webpage where users can produce fingerprints and know about their software configuration's uniqueness. The User-Agent string, HTTP request headers, when cookies are enabled, time zone, screen size, browser plugins and their versions, whether certain long-term state storage ("evercookies") is banned, and the list of system typefaces are all used to create fingerprints. 84 % of participants were recognized by these limiting characteristics (Eckersley, 2010). Subsequently, researchers have discovered that the IP address, font set, time zone, and screen resolution are all adequate to clearly identify the majority of users of the top five Web browsers (Boda et al., 2011). Similarly, Boda et al. (2012) investigated cross-browser fingerprinting using only JavaScript. The authors used JavaScript to establish browser independence as a user ID. The user ID is created using the initial 8 bits of the IP address, the operating system version, the screen resolution, the time zone, and a list of standard typefaces present in the system that is standardized across all web browsers. JavaScript and Flash can determine the operating system's version and architecture, as well as the set of existing fonts, colour depth, and panel sizes (AdobeFlash, n.d; Boda et al., 2011; Fifield & Egelman, 2015; Mowery et al., 2011). Several studies proposed ML detecting approaches for browser fingerprinting. Researchers have employed a syntactic-semantic technique based on machine learning to reliably identify browser fingerprinting on the top-100K websites. They detect previously undisclosed applications of JavaScript APIs by fingerprinting scripts and determine that browser fingerprinting is now prevalent on more than 10% of the top-100K websites (Vastel et al., 2018). While Acar et al. (2014), motioned that Canvas fingerprinting was identified on 5% of the top 100,000 Alexa sites, primarily due to third-party advertisement code. This might imply that, as technology progresses, this type of stateless tracking will become more prevalent. Recently, Nikkhah Bahrami et al. (2021), proposed a machine learning approach, which is FP-Radar for detecting a specific fingerprinting method. This approach is capable of detecting abuse presented properties such as longitude (Geolocation), DeviceMotionEvent (Sensor), and plugins (Navigator)at an early stage. Finally, although fingerprinting can assist protect against session hijacking (Spooren et al., 2015), which occurs when an attacker captures a victim's authentication cookies, Web-based malware can also utilise it to detect weak browsers (Acar, 2017).

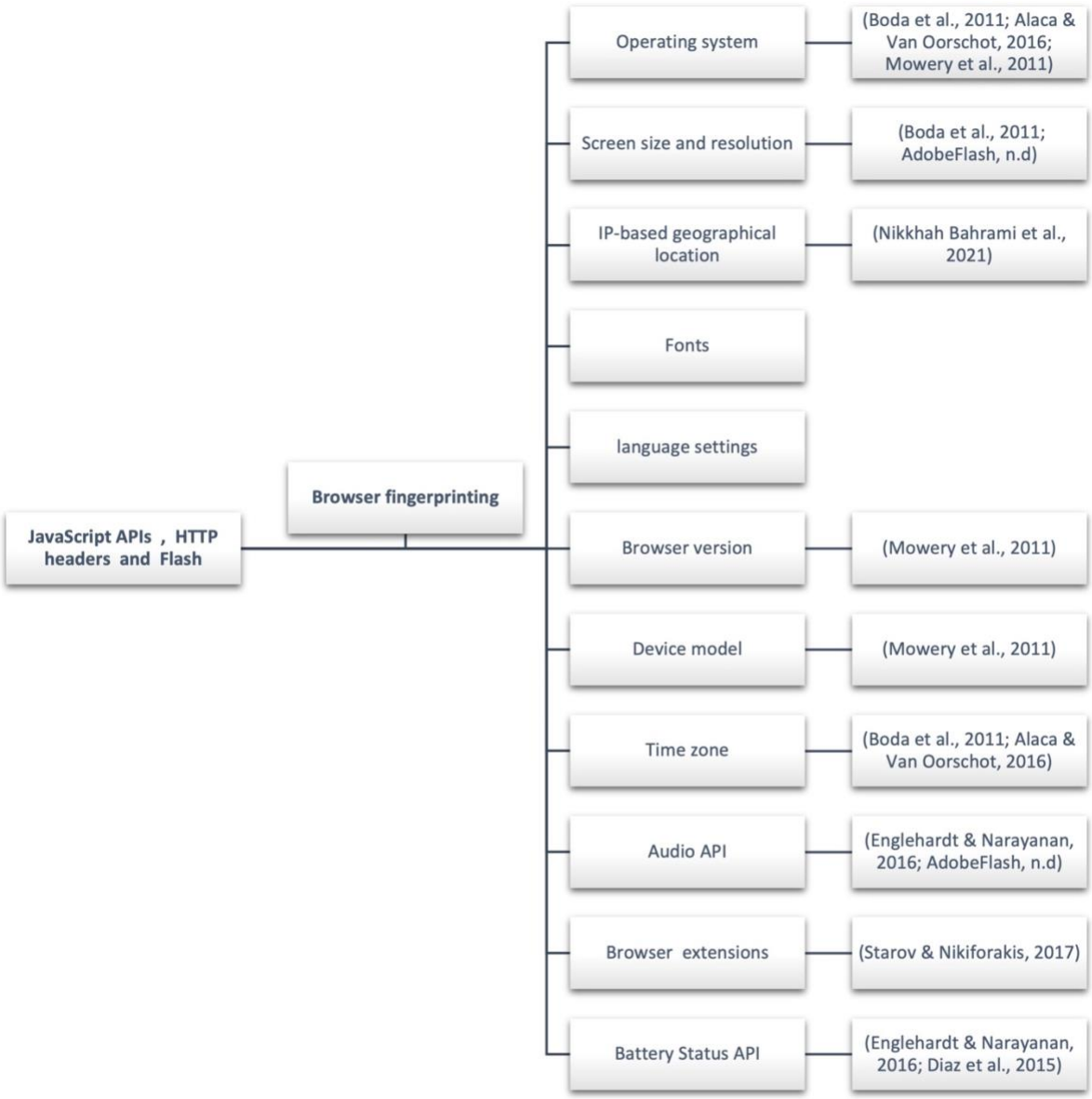


Figure 2.1 diagram illustrates information can be gathered from the browser fingerprinting

2.2.4.5 Mobile tracking

Cookies have a more restricted existence on mobile phones, where they are ineffective since they must be reset when a browser is quit, and they can't be transferred between applications or devices. To replace the mobile cookies gap, there are a few other tracking mechanisms that might be considered or used in conjunction to fill the gap left by cookies. Client/Device Generated Identifier, Statistical ID, HTML5 Cookie Tracking, and Universal Login Tracking are some of the alternative tracking technologies presently being applied (Frow, 2019). Several studies have shown how fingerprinting cell phones may be used to identify and track users. For example, Spooren et al. (2015), examine the variance and predictive capabilities of mobile device fingerprints such as time zone and geographic location that can be predicted and recognized 100%. While Bonneau (2012), showed how a range of onboard sensors may be used to fingerprint smartphones, including the device's accelerometer calibration error and the frequency response of the speakerphone-microphone combination. For multi-factor mobile authentication, Goethem et al. (2016), offered an accelerometer-based device fingerprinting approach. Various studies, on the other hand, investigated the use of sensing devices to uniquely identify a mobile device. Studies conducted by Bojinov et al. (2014), and Dey et al. (2013), looked at how a smartphone's sensors may be utilized to create a trustworthy hardware fingerprint. With the digital camera, similar device fingerprinting and entropy analysis work was done by (Li, 2010).

2.3 The Web's security consequences of cookies on users

Information security is built on the CIA triad of confidentiality, integrity, and availability. Generally, security vulnerabilities in online applications can lead to data integrity breaches, the theft of personal data, or a decrease in the availability of the Web application (Nagpure & Kurkure, 2017). Even when cookies are being used in conjunction with HTTPS, cookie confidentiality and integrity are not guaranteed (Jussila, 2018). As cookies retain user information, this information might be captured, allowing for impersonating or illegal access to a website (Wagner, 2020). The breach of a user's login is the most serious privacy offense as well as a significant security risk (Jegatheesan, 2013). So, since the core purpose of cookies is to establish a session between the webpage and the user. The user authenticates themself once, and then a token is provided back to the site in the cookie so that the site knows whose user is sending HTTP requests. Consequently, many websites hack attempt to steal or modify cookie data. The security flaw that is linked to Web tracking technology, specifically cookies, can be carried out in one of two ways: client-side cookie manipulation or cookie-based session exploitation. Technically, client-side cookie manipulation can execute an XSS attack to steal and simulate another user's cookie by leveraging a JavaScript injection vulnerability (Endler, 2002; Rahalkar, 2016). Another client-side cookie attack is CSRF (Kombade &

Meshram, 2012). The adversary can forge cookie data and deceive a website into believing they are a different user, resulting in privilege escalation flaws (cookies, n.d.). Privilege escalation occurs when a user attempts to get access to a service or software that they are not authorized to use (Rahalkar, 2016). While in the case of session attacks based on cookies the session identifier is the most important variable kept in user cookies since it facilitates for session hijacking and other threats. Hence, if an attacker exploits a man-in-the-middle attack to hijack HTTP traffic, they may be able to take another user's cookie and spoof them (Baitha & Vinod, 2018). Cookie-related security flaws occurred on major websites like Google and Bank of America and were discovered in popular Web browsers like Chrome, Firefox, and Safari (Zheng et al., 2015). Therefore, it appears that whenever data inputs are processed, there is a possibility of cyber-attacks. The tracking technologies incorporated in websites to give services to users may impact the user's security in some manner, more likely exposing them to hacker cyber-attacks. Accordingly, cookie security is a concern, as they are vulnerable to a variety of attacks including cross-site request forgery (XSRF), cross-site scripting (XSS) and Session hijacking.

2.3.1 Client-side cookie attacks

2.3.1.1 Cross-Site Scripting (XSS)

To begin with, XSS is one of the most dangerous and common online attack vulnerabilities, and it can impact both individual privacy and the economic (Zhou & Wang, 2019). In 2020, XSS vulnerabilities ranked first on the Common Weakness Enumeration's list of the 25 most dangerous software weaknesses (WatchGuard, 2021). XSS vulnerabilities don't directly affect the site or server, but rather the client side or the user, necessitating the use of social engineering to launch a powerful attack (Rodríguez et al., 2018). Cross-Site Scripting is a Java Script code injection technique (Martin & Lam, 2008), that allows an adversary to inject JavaScript into a victim's Web browser to get access to sensitive resources such as cookies, passwords, and credit card details (Zhenyu et al., 2007). In a study conducted by Rodríguez et al. (2018), the results indicated that 88.24% of cookies created have the potential to execute instructions, while only 11.76 % have the HTTP Only value, which is a browser feature that prohibits XSS attacks. Furthermore, 29.41% of malicious websites set cookies with a lifespan of more than two years (Rodríguez et al., 2018). XSS attacks, on the other hand, come in a variety of forms. Reflected XSS, stored XSS, and dome based XSS are the most prevalent forms (Kirda et al., 2006). These exploits are frequently used to steal cookies from a browser's database (Rodríguez et al., 2020). In a reflected XSS attack, the attacker controls the performance of a malicious payload to capture the victim's session cookies, for instance, to characterize himself as if it were his session. The attacker might use the stolen cookie to carry out operations

with the victim's privileges without the need for a password (Rodríguez et al., 2020). While, when a user's personal data is delivered to the targeted endpoint, an XSS stored attack is most likely to occur. Essentially, the attacker uses the website's flaw to deliver the exploit payload to the susceptible server. Afterwards, upon visiting the webpage with the attached XSS attack payload, a target user receives the attack (Prasetio et al., 2021; Rodríguez et al., 2020). Whereas, in the DOM XSS attack, the malicious script is injected by a link, but it is not incorporated into the site's source code as part of website. When a user visits an infected website, malicious code exploits a flaw to install itself in a Web browser file and execute without being verified (Rodríguez et al., 2020). Figure 2.2 is a workflow depicting the steps involved in launching an XSS attack. Research is continuously being undertaken to prevent this form of cyber-attack. Takahashi et al. (2013), suggested a method for determining whether or not a user is the true owner of a cookie that uses a one-time password and challenge-response authentication. Limiting the abuse of stolen cookies renders the XSS attack ineffective in capturing cookies. Recently, machine learning (ML) algorithms have demonstrated significant potential in addressing XSS threats, according to existing literature. Machine learning algorithms have recently been employed by several studies including (Abikoye et al., 2020; Fang et al., 2018; Prasetio et al., 2021; Zhou & Wang, 2019) in order to mitigate such forms of attacks. These strategies, on the other hand, are incapable of altering or adapting to new XSS assaults that are not similar to those on which the model was trained (Tariq et al., 2021).

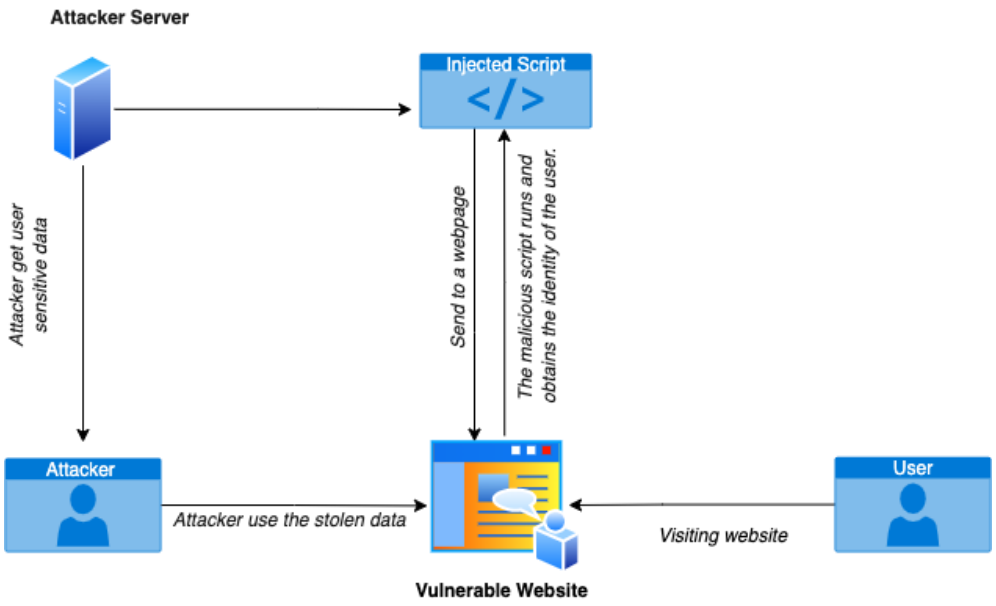


Figure 2.2 an illustration of an XSS attack workflow.

2.3.1.2 Cross-Site Request Forgery (XSRF)

Cross-site request forgery (CSRF or XSRF) is a sort of Web-based malicious vulnerability known as a one-click attack or session riding that takes advantage of cookies. In contrast to cross-site scripting (XSS), a CSRF attack is done without the user's awareness. CSRF is seen as a more serious threat than XSS (Jing et al., 2015). It is a new type of online application attack that takes advantage of authorized users' confidence (Alexenko et al., 2010). In this form of cyber-attack, cybercriminals constructed (fake) web pages with malicious code in order to deceive users into thinking they were accessing a legitimate website (Kour, 2020). Through social engineering, such as a phishing attack, the attacker convinces an authenticated or signed-in user to click a link to activate the malicious script (Lin et al., 2009; Yadav & Parekh, 2017), for example by delivering a link or an image to the victim via email (Rankothge & Randeniya, 2020). When a hacker is able to submit a constructed (malicious link) request to an audience user, the attacker modifies the required parameters in the script to perform a legitimate application request (Sentamilselvan et al., 2013; Wagner, 2020). Subsequently, an effective CSRF hack can access end user data and operations by circumventing the core authentication system and can compromise the entire website if the target user has administrator credentials (Sentamilselvan et al., 2013). Figure 2.3 shows a procedure for conducting a CSRF attack in a basic situation.

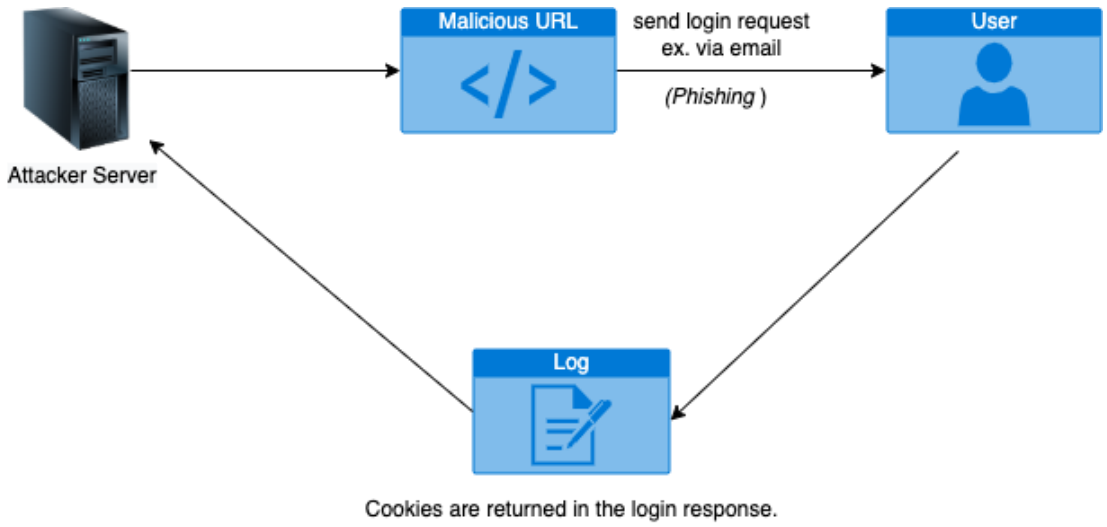


Figure 2.3An illustration CSRF attack technique.

CSRF attacks have been reported on a number of well-known websites. The first CSRF attack took place at ING Direct, a financial services company cyber attackers transfer funds from the user's bank account. The hacker is able to establish additional accounts on behalf of any user (Sentamilselvan et al., 2013). A

well-known viewed platform, Youtube, has been subjected to a CSRF attack; in addition to, Gmail, Netflix, Paypal and eBay (Coram, 2019; Sentamilselvan et al., 2013). Recently, in 2020 one of TikTok's subdomains, was determined to be vulnerable to XSS attacks, which allow a hacker to inject malware. The cybercriminals might utilise CSRF or cross-site scripting (XSS) attacks after deploying the virus, forcing other user accounts to send requests to the TikTok application on their behalf (Kaur & Garg, 2021).

2.3.2 Cookie-Based Session Attacks

2.3.2.1 Session Hijacking

Cookie-based sessions are vulnerable to major security risks since the unintended leak of a session cookie gives an attacker complete control over the browser identified by that cookie (Bugliesi et al., 2014). Cookie-based attacks targeting user sessions include session hijacking, session fixation (Kolsek, 2002), and session spoofing (Kavisankar et al., 2016). Session spoofing and session hijacking are closely linked (Kavisankar et al., 2016). The primary distinction is that in the first, the attack happens offline, while in the second, it occurs online. Session hijacking also referred to as cookie hijacking occurs when an attacker takes control of a user's established computer session across two devices (Sheikh, 2021). Session hijacking in a web-based application includes stealing a user's cookie (Jussila, 2018). The cookie is used to preserve private information like usernames and passwords. The cookie is used to preserve private information like usernames and passwords. The user is usually unaware of what is going on and receives a notification called, "session expired" or "login failed." An attacker might steal a session if session timeouts in the website server are wrongly set (Jussila, 2018). Research by Sivakorn et al. (2016), illustrated the severity of stealing users' cookies via a hijacking attack. For example, through HTTP cookie hijacking, private information such as browsing history, exploiting search optimization, location, and personal information can be obtained from a user's Google account (Sivakorn et al., 2016). This probably indicates more serious issues if accurate information, such as the user's location, is obtained. Since having access to location data exposes the user to physical dangers (Petsios & Keromytis, 2015). The user is usually unaware of what is going on and receives a notification that indicates "session expired" or "access failed." (Jussila, 2018). However, according to empirical research findings, HTTP cookie hijacking attacks may be used to obtain access to not just confidential and sensitive user data, but also to bypass verification and gain access to restricted account capabilities (Sivakorn et al., 2016). Likewise, in an empirical review of cookie attacks, including session hijacking, the authors of a (Zheng et al., 2015) research determined that cookies lack integrity. Privacy breaches, online abuse, and even money loss and account hijacking have all been accomplished attacks (Zheng et al., 2015). As a result, since cookies could be hijacking users' sessions and

compromising their confidentiality (Bortz et al., 2011), the consequences of these attacks on privacy are becoming ever more concerning (Sivakorn et al., 2016).

2.4 Web tracking and User Privacy

2.4.1 Digital privacy concept

As shown through, tracking technologies in a broad sense may compromise a user's security and expose data to a variety of cyber-attacks, which has a direct and noticeable impact on the privacy of users' data. The privacy issue over cookies, Flash cookies, and beacons is largely motivated by the risk of information being misused through secondary usage (Sipior et al., 2011). Generally, privacy on the Internet continues to be a major issue (Bouguettaya & Eltoweissy, 2003). Internet privacy refers to the right or obligation to protect an individual's personal information whether it is stored, repurposed, shared with third parties, or used on the Internet (Baumer et al., 2004). Otherwise, digital privacy refers to any personally identifiable information provided online when performing personal or corporate conversations over public networks (Hassan & Hijazi, 2017). In theory, data gathered from Internet interactions can be classified into two categories. The first is personally identifiable information (PII), also known as sensitive personal information (SPI). The anonymous information is the second type (Hassan & Hijazi, 2017). Personally identifiable information (PII) considered as one of the most important concepts in the privacy act (Schwartz & Solove, 2011). It includes a person's name, social security number, passport number, date/time, birthplace, gender, and any other information that may be used to identify or locate an individual (Hassan & Hijazi, 2017). On the other hand, anonymous information includes browser type, browser version, current location, school, country, and connected device type (Hassan & Hijazi, 2017). Although this metadata is not categorized as PII, Acar et al. (2013) claims that fingerprint tracking is linked to PII data, implying that the way users, organizations, and regulators deal with fingerprints needs to change. According to Mayer and Mitchell (2012), personal information is deeply linked to Web browsing history. A user's location, hobbies, purchases, work position, gender identity, financial state, medical concerns, and more can all be revealed by the page's user views. Many assumptions about a person may be drawn from individual page loads; studying patterns of behavior allows for even more inferences (Mayer & Mitchell, 2012). Owing to the extensive use of cookies, many websites show information hosted outside by third-party websites, which can follow users and become aware of their browsing behaviour. These third parties can uniquely identify the user (Li et al., 2015). At least one third-party tracker analyzes about 46% of the home pages of websites in the top 10,000 Alexa rankings. One-third of all queries to third-party websites were routed through a tracker. Google was responsible for monitoring 25% of the websites reviewed, Facebook for 13%, and

Twitter for 5%. (Li et al., 2015). So, third-party trackers (such as Doubleclick) pose a severe privacy risk since they may gather and aggregate browsing information across several websites (Bujlow et al., 2015).

2.4.2 Privacy challenges with cookies and other tracking technologies

2.4.2.1 Overview

The Internet cookie, often known as Web cookies, is the most well-known online PII collection method. Such technology is frequently used to acquire large quantities of data about users (Zuiderveen Borgesius et al., 2017). The gathered user data is used for a number of reasons, the most well-known of which is commercial tracking. Tracking data is widely used in the targeted advertising (Roesner et al., 2012). Tracking technology can be employed on websites and service providers, either in-house or by third-party organizations that offer tracking services to a variety of websites (Pugliese, 2015). Cookies, according to Pierson and Heyman (2011), are a commercial dataveillance technique. Cookies allow data surveillance strategies either through the deployment of zombie cookies to prevent user cookie deletion or through behavioral advertising via third-party tracking cookies (Pierson & Heyman, 2011). Third-party cookies are typically used to characterize users for behavioral targeting, a marketing approach that includes monitoring people's online behavior and utilizing the information gathered to send them individually focused advertisements (Pierson & Heyman, 2011; Zuiderveen Borgesius, 2017). This has resulted in the creation of hidden markets for targeted advertisements, such as real-time bidding for available ad spots on the sites that are presented to the user. Tracking and browsing data are also of value to large-scale data collectors and other data consumers that want to enrich data profiles on individual online users (Ermakova et al., 2019). When a user switches devices or moves to a different location, cross-device and mobile tracking can be employed. (Brookman et al., 2017). Although this may be upsetting to privacy-conscious users, no laws are broken. Because it hosts resources necessary by the website that the user wishes to access, the third-party tracker is properly accessed by the user's browser (Li et al., 2015). There are many web analytics services that collect and track user data. When web analytics services, such as Google Analytics, are used to track website visits, third-party tracking can occur (Mayer & Mitchell, 2012; O'Brien et al., 2018). These services give data on website usage and user behavior that utilize to improve their online services. The analytics services, on the other hand, utilize sophisticated algorithms and wide networks to follow users and their behavior across sites, collecting demographics and behavioral trends (O'Brien et al., 2018). However, O'Brien et al. (2018), argued that the analytics service providers' tracking has proven privacy standards useless. Browser cookies created by Google Analytics, for example, operate as leads for gathering and sharing user data over a broad network of commercial trackers (O'Brien et al., 2018). This might create

major challenges to the protection of personal data. To put it in other words, cookies and other tracking mechanisms have a major negative influence on user privacy. Online trackers do not appear to provide users with the data privacy that is stated in various Information Privacy Principles (IPP) laws, such as the New Zealand Privacy Act's data protection regulations. The IPPs are governed by various relevant ethical practice standards. In a broad set of governments throughout the world, similar principles underpin privacy and data protection policy. The principles relate to personal information collection, data use and disclosure, and data storage and security.

2.4.2.2 Web tracking and data gathering

The lack of transparency and ambiguity of the data gathering procedure are the concerns with data collecting using cookies and other related technologies. Although the Internet is widely thought to provide anonymity, online interaction does not (D'Ovidio & Doyle, 2003). Because of the hidden characteristics of data gathering tools, Web users may be unaware of the presence of cookies, Flash cookies, and Web beacons, as well as the sort of information, gathered about them (Sipior et al., 2011). These services try to create a user profile by gathering, aggregating, and correlating information about a person's browsing habits, demographics, interests, and temporal/spatial behaviour patterns (e.g., through smartphone localization, or location check-ins on Online Social Networks) (Falahrastegar et al., 2014). While cookie data is aggregated, continued data collection over time might result in entire profiles of individuals, including PII (Kirk, 2009). Also, some tools have the ability to directly acquire PII by collecting clickstream data (Sipior et al., 2011). Even if some users are taking the appropriate precautions to prevent it, the collecting of user data through trackers persists. Researchers discovered that users accessing the Web using a secure VPN face privacy-breaching concern (Papadopoulos et al., 2018). The hacks are made feasible via cookies (Papadopoulos et al., 2018). However, third parties gathering, and utilizing browsing activity have regularly been associated with aggression in user evaluations. In an interactive study conducted in mid-2011, 85 % of respondents stated they would not consent to monitoring for ad targeting, while 78 per cent said they would not accept tracking for Web analytics (as cited in Mayer & Mitchell, 2012). Users also decline to have their data collected for marketing purposes. 68 percent of participants said they were "not ok" with behavioral advertising (Purcell et al., 2012). Online tracking provides the ability to uniquely recognize website visitors without their awareness (Sipior et al., 2011). As a cookie contains a unique random identification, such as a numeric user code, as well as other processes of gathering (Sipior et al., 2011) So, the Web cookie can gather and retain online registration information, interests, and a record of the user's activity and purchases. All the data collected via Web tracking tools lie under PII. As a consequence, many users may be unaware of the extent to which Web trackers featured on most websites are transparent in identifying the sorts of

personal data gathered, either explicitly or implicitly. In addition to a lack of transparency for data gathering purposes, which creates a privacy issue from the user's perspective, there is a rising concern about privacy.

2.4.2.3 Web tracking data usage and disclosure

Users may not only be uninformed that cookies and Flash cookies may gather a wealth of information about them, but they may also be unaware of whether and to what extent their personal information is sold to other parties for illegitimate commercial purposes (Sipior et al., 2011). Some analytics services provide free analytics and sell the information they gather by using it for ad targeting such as Quantcast and market knowledge like Google Analytics (Mayer & Mitchell, 2012). Although the user may be unhappy or unaware of the data gathering practices, the usage of this data poses an additional threat to personal privacy. Trackers can leverage the respawning functionality to link a user's browsing records from before cookie clearing to after cookie clearing (Acar et al., 2014). For example, some tracking services are intended to be relevant for monitoring site traffic, tracking unique users, auditing and reporting on advertising, and customizing the online user's experience (Sipior et al., 2011). This surveillance technology, according to Pekala (2017), creates data profiling that "the data collector may sell to a data trader, where it will be integrated with profiles from other data brokers to construct an even more complete image of the user. Consequently, the capacity to link data obtained online with public documents, demographic data, and even statistical information presents a danger to individual privacy (Sipior et al., 2011). Further, targeting stakeholders may exploit sensitive information like a user's geographic location, which is automatically given owing to these technologies. For illustration, identifying the location of a website visitor is achievable if the website has a location-based server that can detect a user's location based on location information supplied by the application, allowing such a customer to be surveilled (Kumar & Ogunmola, 2020). However, several empirical research has demonstrated that Web trackers leak personal data, which may come under the use of personal data privacy standers. When personally identifiable information (PII) is leaked from a first party and given to a third party, this is known as PII leakage. Starov and Nikiforakis (2017), examined the data breaches caused by Google Chrome's 10,000 most popular browser extensions and discover that a sizable portion of them leak sensitive data about the user's surfing habits, such as browsing history and search-engine searches. PII leaking through contact forms, as well as data leakage from many online forms (Starov et al., 2016). A study by Ren et al. (2016), exemplified the detection of personal information (PII) leaks to third parties in smartphone apps. Englehardt et al. (2018), found that when emails are viewed, about 30% of them reveal the user's email address to one or more third parties. The bulk of these breaches is the result of email senders' purposeful actions, with additional leaks occurring when users click links in emails. Recently, Dao and Fukuda (2021), detect that 42.3 % of sites leak PII to third-party services by studying the authentication processes for 307 major shopping websites from the Tranco top 10,000 sites the most

popular is Facebook. Furthermore, they reveal a tracking technique for PII leaks that enables monitoring providers to continuously identify individuals across sites, browsers, and devices (5.1). The presence of PII leakage-based monitoring is next investigated, and it is shown that 20 monitoring providers employ PII to persistently follow user actions (5.2). Then, researchers stated that after reviewing the privacy policies of 130 first-party senders, they highlighted that many are unclear concerning PII shared with third parties (Dao & Fukuda, 2021). As a matter of fact, the websites that house contact forms must respect their users and not leak PII to third parties, either accidentally or maliciously (Starov et al., 2016). As it turns out, user data is being tracked for leaking and being used for a variety of purposes, the threat to this data's privacy is not limited to the non-transparent purposes of using and sharing personal user data obtained by trackers; it also poses a privacy threat in terms of how this data is protected and stored.

2.4.2.4 Data storage and security with Web tracking

As formerly indicated, the Web collects and uses a vast amount of personal data from users, which is clearly a violation of their privacy; however, there is another component to consider: how these parties protect and store such data. On the Internet, some sensitive information is being collected and stored. For example, date of birth, height and weight, marital status, medical issues, political party affiliation, and estimated property worth may be included in these records (Sipior et al., 2011). Users may be exposed to several threats if such data is not adequately protected. Cookies may be injected into extremely big websites and prominent open-source services, according to Zheng et al. (2015), including Google, Amazon, eBay, Apple, Bank of America, BitBucket, China Construction Bank, China Union-Pay, JD.com, phpMyAdmin, and MediaWiki. Injecting cookies has a number of implications for the privacy and security of users' data. XSS, privacy leakage, circumventing cross-site request forgery (CSRF) defences, money loss, and account hijacking are only some of the consequences of cyberattacks (Zheng et al., 2015). Some websites may not have sufficient encryption in place to secure data throughout its acquisition and storage. Users' personal information is purposely leaked to third parties when they are asked to fill out forms that contain third-party information and are poorly coded, resulting in thousands of third-party websites acquiring user PII without the user's knowledge or agreement (Starov et al., 2016). Many users' data were acquired from reputable sites and services on the Internet as a result of hacks meant to steal data from cookies. For example, Sivakorn et al. (2016), have access to user login, nickname, Gmail address, and profile image if the user utilize Google services. Also, the adversary can collect information that the victim uses to log in, such as the victim's login, email address, and/or mobile phone number, from Amazon. Furthermore, Amazon discloses the user's complete username and locality when you check out the things in your cart (used for shipping). Users may read user reviews to get a discount (which may include sensitive items) (Sivakorn et al., 2016). Along with several other research findings, there are many different types of attacks that target cookie data. Barth

et al. (2008), identified a kind of CSRF known as login CSRF. In this invasion, the attacker logs in to the victim's browser using his own account. If the victim is not aware, the attacker's account may be used to access the targeted website, resulting in security and privacy issues such as search history leaking, credit card theft, and XSS. According to Malandrino et al. (2013), identity theft, social engineering attacks, and online and physical stalking are all possible uses for such data. So, cookie injection is a big issue in the real world, according to the variety of susceptible online applications and exploitations, and it demands more attention from the Web security community (Zheng et al., 2015). This implies that some of the websites do not provide enough security for users' data and expose them to various types of cyber-attacks, which might be deemed a privacy violation. However, many providers are unclear about how long this information is retained. Iordanou et al. (2018), argue that there are enhanced laws that may have an influence on tracking in terms of security-related and data or server logs storage period but only have a national scope. In general, therefore, it seems that the information stored by cookies and other similar technologies may not be adequately protected and maintained independently and robustly, yet there is no defined duration for retaining this data. This, in turn, poses a significant threat to the privacy of users' data, exposing them to a variety of sophisticated cybersecurity threats that leverage data stored in cookies to gain more control over websites.

2.4.3 User's Countermeasures against Web Tracking Technologies

Cookies, as well as other tracking methods, can be disabled. Since maintaining online anonymity is a user's main priority, numerous protection strategies have been developed against web tracking techniques. Anti-tracking apps and services strive to reduce the amount of personal information that is exposed to tracking techniques and architecture (Ermakova et al., 2018). A well-known example of anti-tracking software is Tor. The Onion Router, also abbreviated as Tor, is a free application that enables users to connect anonymously. Tor's purpose is to safeguard its users' personal privacy, as well as their autonomy and capacity to communicate in confidence, by keeping their online actions untracked and blocking IP-based tracking but not non-IP-based threats such as fingerprinting and end-to-end timing (Acar et al., 2013). Tor is frequently used in conjunction with other applications, such as Privoxy, to increase privacy (Wills & Tatar, 2012). Privoxy is a free non-caching Web proxy with filtering features for boosting privacy, altering cookies, and changing Web page contents and HTTP headers before they are presented by the browser. Users may customise Privoxy to meet their own needs for both stand-alone and multi-user networks (Ishtiaq et al., 2017). In addition to the solutions mentioned above, users can use a private browsing mode to hide their identity when browsing. The browser's privacy mode is a security feature that protects users' online experience. The browser cache, search history, cookies, and local storage are all disabled in private browsing. This mode protects your privacy on the local system by hiding traces from anybody who has

physical access to the computer (Broenink, 2012). Despite this, some studies have discussed that some tracking mechanisms cannot easily be disrupted or prevented, due to the lack of efficiency or the spread of their own prevention tools. Flash cookies, for example, are a common tool for storing information on the top 100 websites. Users have restricted self-help options with this technique since robust anti-tracking solutions are few, and the use of Flash cookies is rarely acknowledged in the privacy rules (Soltani et al., 2010). Moreover, some sites also restrict users from viewing the contents of what are known as tracking walls if tracking avoidance technologies are employed or if the user rejects cookies. Organizations apply a variety of methods to obtain people's consent to Internet tracking. Offering individuals, the option to accept it or leave it strategy. Some webpages, for instance, deploy 'tracking walls,' also known as 'cookie walls,' which are barriers that visitors may only overcome if they consent to the site or its partners tracking them (Zuiderveen Borgesius et al., 2017); according to a survey conducted throughout the EU, 58 percent of respondents believe that providing personal data is the only way to obtain items or services (Zuiderveen Borgesius et al., 2017).

2.4.4 Related regulation

As a reflection of the privacy risks related to online tracking, some nations regulate the collection of personally identifiable information (PII) about their residents. Yet, privacy regulation varies greatly between nations, and many of them fail to handle the transnational characteristics of the Internet. In Europe for example, online tracking for personalized ads often needs the agreement of users. People's fundamental data protection rights should be respected when tracking (Iordanou et al., 2018). The European Union General Data Protection Regulation is one of the most significant reforms in how personal data is processed and stored (GDPR) (Hoofnagle et al., 2019). GDPR provides European citizens with protection from a wide range of privacy threats, including tracking on sensitive categories (Iordanou et al., 2018), including data about health, political beliefs, religion, or sexual orientation, as well as their desire to opt out. GDPR's ultimate purpose is to enable individuals to govern their personal data and is widely regarded as one of the tightest regulatory frameworks for data protection globally (Dabrowski et al., 2019). This legislation intends to provide users with a good level of control over their personal data in order to preserve their privacy (Strycharz et al., 2021). The GDPR impacts lots of Web services offered in Europe from all over the world. Companies must disclose how they handle personal data, the legal bases for their data processing, and offer frameworks for informed consent to their users, in addition to potentially modifying how they process personal data (Degeling et al., 2018). The GDPR offers customers a lot of power, including the ability to withdraw permission (Art. 7) and the right to be forgotten (Art.17). Simultaneously, stringent standards are set for data controllers and processors, including data protection by design and default (Art.25), as well as logging all processing operations (Art.30) (Li et al., 2019). GDPR appears to have had an influence on the

rate of third-party cookies immediately after the law went into force. Among the websites that empower individuals to make their own decisions (Hu & Sastry, 2019). 62.1 % of European websites now contain cookie consent policies, up 16 percent from before the GDPR went into effect (Degeling et al., 2018). However, some studies have found that users may have difficulty comprehending these policies based on their readings. The study by Becher and Benoliel (2021), measured the readability of cookie privacy policies on 300 of the world's most popular websites. According to their finding, despite the GDPR's obligation, users frequently face privacy regulations that are entirely unreadable. This is likely due to the privacy policies being written in a complex manner that is difficult for the average user to understand and may result in users being unaware of the personal data exchange in which they participate. Dorfleitner et al. (2021), illustrate that the readability of the privacy policies has deteriorated using textual analysis approaches. Users' comprehension of privacy statements has deteriorated as the texts have grown longer and more standardised (Dorfleitner et al., 2021). So, accepting the default selections on websites that keep giving users the option about whether and how they are tracked typically results in the storage of more cookies on average than accepting the default choices on websites that provide a notification of cookies stored but do not consider giving users the option about which cookies are stored, or those that do not provide a cookie notice at all (Hu & Sastry, 2019). Researchers note that although the majority of third parties reduced by more than 10% on average after GDPR, there is no content reduction in long-term statistics of third party cookies when we examine real users' browsing histories over a year, implying that users are not taking advantage of GDPR's enhanced privacy options (Hu & Sastry, 2019). Utz et al. (2019) found that people's consent behaviour is influenced by their status, presented options, prompting, and terminology. The findings also show that the GDPR's data protection by default and purposed-based consent standards would require websites to deploy permission notices that result in fewer than 0.1 % of users explicitly consenting to the usage of third-party cookies. However, it’s arguable whether users have real privacy rights if they have to accept tracking in order to utilize certain services or websites (Zuiderveen Borgesius, 2017). So, in practice, GDPR may fall short of the amount of protection that it promises (Hu & Sastry, 2019). This law still confronts several obstacles The researchers advocated for developing frameworks, methodologies, and structures that comply with GDPR standards for revocation of permission and permanent erasure of large amounts of personal data (Politou et al., 2018). Practically, Hu and Sastry (2019), evaluated how the top 100 Alexa.com sites in China and the United States handled cookie alerts when viewed from the United Kingdom. They found the great majority of leading sites (84 percent in China, 52 percent in the US) are now functioning without a cookie notice. Accordingly, Li et al. (2019), recommended identifying the factors that influence non-compliance and the application of the law in terms of several considerations, including the cost of achieving compliance with the General Data Protection

Regulation (GDPR)), investigating how cultural and national aspects affect the implementation and compliance of the GDPR.

2.5 End-users' online information awareness and concern

As previously stated, web trackers violate users' right to privacy and security of their personal information. Hence, the end-user's position and role in this context are to be aware of the possible consequences of violating online privacy. The user is typically unaware of the technology incorporated in websites that track online behaviour by collecting and analysing personal data in order to target them directly for commercial interests, such as advertising. The concept that cookies are implicit is a disempowering characteristic for all users who are unaware of the situation because it is not a consideration in their contextual appraisal (Pierson & Heyman, 2011). All additional activities to empower users in their privacy management are disabled due to the lack of cookies in the observed context (Pierson & Heyman, 2011). Which, in turn, is more likely to affect the level of awareness of the uses and understanding of this technology. Otherwise, many users are probably unaware of the purpose of cookies, let alone the security and privacy risk that technology poses. The public's awareness of privacy infractions is only slowly growing, despite the rapid growth and efficiency of information extraction to increase the effectiveness of the behavioural advertising (Krishnamurthy et al., 2011). Whereas the limited current studies have been considered that assess user's awareness in regard to their daily browsing activity cookies privacy and security threats, existing literature has expanded that examine information security awareness, information privacy awareness, online privacy concerns, and their relationships to user online behaviour, which can be used as a theoretical basis to assist in understanding this area and conduct further research.

2.5.1 Information Security Awareness (ISA)

By examining how the concept of awareness is viewed in the research on information system security. According to Jaeger (2018), awareness includes procedural aspects, such as the methods utilized to attain this state of mind and characteristics of an individual's cognitive state of mind, such as being conscious or having knowledge of something. Jaeger (2018) examined information security awareness from the standpoint of the cognitive state of mind to identify it from awareness-raising practices and subsequent outcome aspects like behavioural reactions. As per this viewpoint, awareness-raising processes are input factors, whereas later belief, attitude, and behavioural reactions are output variables. From this perspective, and in light of the security issues connected with cookies, it can be considered that a user's understanding of the security flaws in cookies in particular might assist them in implementing security measures and preventative approaches to combat cyber-attacks facilitated by this technology. Hadlington (2018), discovered that employees at large companies had a stronger awareness of cyber threats, which might be

due to more financial resources and organizational enforcement procedures. Individuals' information security awareness (ISA) is important in defining their security-related behaviour in both professional and personal scenarios (Jaeger, 2018). Recently, the study by Zwilling et al. (2022) looks at the connections between cyber security awareness, knowledge, and behaviour with protection tools among individuals in four nations. The findings demonstrate that while Internet users are aware of cyber threats, they only take modest precautions, which are often standard and simple. Higher levels of cyber awareness are likewise linked to higher levels of cyber knowledge. Nevertheless, studies have indicated that, despite the vital importance of cyber education/training in emphasizing suitable security procedures to enhance day-to-day online behaviour (McCrohan et al., 2010), a severe worldwide issue of cyber awareness still exists (Zwilling et al., 2022).

2.5.2 Information privacy awareness (IPA)

Notwithstanding its widely accepted significance, the meaning of information privacy awareness IPA is still unclear and is defined in various ways. Information privacy awareness (IPA) is defined by Correia and Compeau (2017), as the literacy of factors connected to information privacy, the knowledge that the elements exist in the current environment, and the prediction of their future implications. Protection of privacy is reliant on individual, organisational, and governmental decision-making. For instance, at the individual level, taking appropriate action is contingent on individuals being aware of the dangers of sharing personal information and setting appropriate sharing limits (Correia & Compeau, 2017). For people to take the necessary precautions to protect themselves, privacy awareness is crucial to support the current systems for ensuring that people have some understanding and acknowledgment of how information is gathered, used, and potentially abused in online contexts (Rotman, 2009). People appear to provide personal information to Internet-based services without realising the potential for privacy violations (Pitkänen & Tuunainen, 2012). Accordingly, users need to examine how they share information and comprehend how their data is utilised, disseminated, and loses its private nature (Givens, 2014). The technology, legislation, or common practices used by organisations or people to gather, handle, and distribute users' private information are elements in this case (Correia & Compeau, 2017). Further, the value of individuals being aware of their digital privacy contributes to the development of laws and policies to protect data; since the legislative activity is influenced by individual behaviour and awareness to some extent (Correia & Compeau, 2017). This implies that, in terms of law, user awareness may not be adequate. Researchers observed in the area of literacy on online privacy on the measuring of awareness of information privacy, it is crucial to take into account how familiar Internet users are with privacy-related rules and regulations (Prince et al., 2021). Users' knowledge of declarative privacy includes their comprehension of the laws or other legal elements of the online data protection (Prince et al., 2021).

2.5.3 Considering web tracking in the privacy (IPA) and security (ISA) of informational awareness.

In previous studies, examinations of users' privacy perceptions concerning the various Internet tracking modes were evaluated. Chanchary and Chiasson (2015), carried out a web-based user investigation to find out how well users perceived tracking prevention tools (TPT) and online behavioural advertising (OBA), in addition to whether users' desire to provide information to marketing firms varied according to the type of first-party website. They revealed that the majority of participants showed a minimum of some signs of protecting their online privacy, and that half of the participants were familiar with OBA. Recently, a survey conducted by Kashi and Zavou (2020), to determine whether users are aware of email tracking, showed that the majority of participants are aware that their online activity is being tracked. While Narayanan (2020), evaluated the perception of European-based users on how cookies work and discovered that 40% of respondents indicated they were extremely familiar with the functionality of cookies. In the same vein, a survey was conducted by Pinto et al. (2020), to find out more about how Portuguese Internet users perceive accepting cookies and the advantages they offer. Their study's results indicate that most participants do not inform enough about web cookies. Despite this, prior literature focused on a certain demographic and geographic region as (Pinto et al., 2020; Narayanan,2020). Further, earlier studies did not consider the end-user understanding of security and privacy in a manner that reflects both procedural and preventive aspects, including laws and privacy protection practices. Considering the existence of many research examining users' privacy-protective behaviour, such as those by Büchi et al. (2017), Chiasson et al. (2018), and Edith G Smit et al. (2014), these studies have not assessed users' privacy-protective behaviour with respect to cookies. Thus, given the significance of an individual's awareness of privacy and security hazards, end-user evaluation across several levels or contexts of awareness of online tracking and cookies is lacking.

2.5.4 End-user's information privacy concerns (IPC)

Sensitive personal information about people's regular activities and habits is increasingly likely to make up the majority of the data collected for the big data analysis (Janssen & Helbig, 2018). Users generally lose control over their data as a result of businesses collecting, using, and sharing information that is readily available online (Boerman et al., 2021). There have been increasing instances of private information leaks and data breaches in public agencies, financial institutions, and IT corporations' field (de Bruijn & Janssen, 2017). These events also make people more concerned about their personal information (Lee et al., 2019). The information privacy concern IPC is defined by Dinev and Hart (2006), as the assessments and decisions of whether or not to consider the disclosure of sensitive personal data to be a threat. Incidents involving

user personal data leakage also increase people's information privacy concerns (IPC) and, as a result, their interest in maintaining their private information (Lee et al., 2019). Lutz and Strathoff (2014), indicated a strong relationship between privacy concerns and protective behaviour. Likewise, according to Boyles et al. (2012), 30% of mobile application users uninstall software already installed on their smartphone after realising that it is gathering personal data without authorization. However, users' informational privacy concern is influenced by various factors. Informational awareness, regional characteristics, as well as culture, are all some considerations that influence users' privacy concerns that have been noted in the works of literature. The theoretical model conducted by Malhotra et al. (2004), introduces the concept of Internet users' information privacy concerns (IUPAC) model in terms of three factors: collection, control, and awareness of privacy practices. Correia and Compeau (2017), indicated that users' concerns about potential harm are typically linked to their awareness of possible threats. Schaub et al. (2016), performed qualitative lab research with 24 participants to see how three popular extensions including Ghostery, DoNotTrackMe, and Disconnect affect users' privacy awareness and concerns. Participants expressed higher privacy concerns when using a browser extension due to heightened awareness of tracking (Schaub et al., 2016). While Bellman et al. (2002) demonstrated that culture and national regulation considerably impact users' privacy concerns, implying that localised privacy regulations are necessary. likewise, under the impact of cross-cultural influences, the study by Wu et al. (2012), intends to evaluate trust and privacy concerns related to the desire to share personal information online. They identified a strong cross-cultural influence on the links between the substance of privacy policies and privacy concerns/trust. Additionally, Fujs et al. (2019), showed how privacy concerns are significantly impacted by information sensitivity, regulation, and surveillance concerns. So, it can be noted that some of the factors that influence users' concerns about their data, in general, are varied, and this, in turn, influences users' attitudes regarding the privacy of their data on the Internet.

However, studies in the area of information privacy concerns (IPCs) have observed that, even though online users value their privacy and are concerned of the hazards, they continue to give personal information. This case is recognized as the "Privacy Paradox" in the literature (Gerber et al., 2018; Kokolakis, 2017). The concept of "Privacy Paradox" was determined by considering this mismatch in attitudes and behaviours toward information privacy (Brown, 2001; Norberg et al., 2007). Otherwise, the privacy paradox refers to people who claim to be concerned about privacy yet behave differently when utilizing technology (Correia & Compeau, 2017) and the remedy, according to Barnes (2006), is consumers' understanding of how to preserve their privacy on the Internet. Results from existing literature showed that the privacy paradox phenomenon exists in a variety of settings, including e-commerce, social networking sites, and Internet use. An attitude vs. behaviour dichotomy was highlighted by (Lee et al., 2013). The research concluded that

users actively disclose personal information despite their concerns since they consider both the expected benefit and the threat of sharing. Hughes and Roberts (2013), concluded that a broad user concern is not a reliable indication of privacy behaviour within the network based on an analysis of a participant's Facebook accounts. Additionally, Taddicken (2014), showed that privacy concerns scarcely have an impact on self-disclosure. Despite this, Kokolakis (2017), indicated that the existence of the paradox of privacy as a phenomenon is still controversial. Nevertheless, the concern about privacy keeps evolving and has other implications that are of considerable relevance to the interest of the user; as stated by Correia and Compeau (2017), the interchange of citizens' concerns with politicians is a vital aspect of the evolution of public policy. According to a recent study by the Strycharz et al. (2021), it makes little difference whether a web shop or a news website requests cookies; what counts is how concerned one is about their privacy and how much one enjoys receiving tailored adverts and recommendations.

Studies that examine user awareness of tracking and cookies in various contexts, such as privacy practices, are lacking. Additionally, users' concerns about the transparency of the cookie-based data collection, processing, and protection were not a primary focus of the studies. Research on user behaviour regarding cookies is lacking related to the privacy paradox that has been addressed in several online settings. Existing digital studies discovered a link between a user's awareness of privacy countermeasures and their security and privacy practices (Dommeyer & Gross, 2003), stressing the need of examining the user's understanding and concern about the security and privacy of personal data as it relates to their everyday online browsing behaviours. As a result, a comprehensive assessment of user privacy and security knowledge, as well as concerns regarding web cookies and other associated web-tracking technologies, is important.

2.6 Conclusion

The literature review offered an outline of Web tracking studies and identified a growing body of information focusing on the security and privacy of Web tracking technologies in general. Nevertheless, it has uncovered unanswered questions and underlines the point that limited research has been done on the security and privacy risks presented by Web tracking technologies to online users. The focus of Section 2.2 was on a review of the main dimensions influencing the development of Web tracking technology, with online business being one of the most important of these considerations. Amongst the most essential marketing concepts mentioned: Online Behavioural advertising (OBA), which is the technique of tracking consumers and anticipating their online activity in order to target them in a personalized manner. This notion, which is closely related to online tracking, is explained in context. Web analysis is an essential strategy on which this strategy is based. Theoretically and practically, this approach has been described in depth. Web analytics, on the technical side, is built on a variety of methods and approaches that aid in the

development and improvement of user tracking. For gathering, storing, and tracking data, there are two main techniques. Cookies were described as one of the most essential techniques for tracking and collecting user data on the Internet. The subject of mobile tracking was also brought up. Section 2.3. has covered the theoretically and empirically studies into the security implications of Web tracking, particularly cookies, which are regarded as the most common kind of tracking. There have been several studies that have used cookie vulnerabilities to launch attacks and steal user data. Many papers were used to describe the most common attacks and their approaches. It explained how users' data is threatened and exposed to hacking and exploitation by hackers, highlighting users' online privacy issues. Section 2.4. the challenges and limitations of user data privacy via Web tracking are covered in this section. The concept of online privacy has been defined, and it is still one of the subjects that have been addressed and discussed to this moment. Web trackers collect data about users, which has been identified and labelled. In addition, the privacy challenges that Web trackers raise on user data from several dimensions such as collection, storage, processing, and usage are highlighted, based on the New Zealand Privacy Act 2020, which corresponds with many international standards. The existence of a variety of issues surrounding privacy from the aforementioned factors has been identified by studies, and they have been examined in depth. For more privacy, certain options are provided to avoid tracking to some extent on the Internet. The most well-known cookie regulation has also been discussed, which was recently enacted but has several flaws, prompting scholars to demand for more research and development of frameworks to comply with it. Section 2.5 Users' awareness and concerns regarding data privacy and security, along with Web trackers, are examined in this section. Based on theoretical models and prior studies, the awareness of users on both aspects of privacy and security was reviewed. The association between user awareness of privacy and security threats to their data and online behaviour was identified. Users' concerns about data privacy and how it relates to their online activity are also discussed. No particular literature was discovered in these evaluations that measured and examined users' awareness and concerns about the security and privacy of their data being tracked and gathered by Web trackers like cookies in general. As a result, this study will address this gap.

CHAPTER 3: RESEARCH DESIGN

3.1 Introduction

This chapter presents the research design for evaluating a user's awareness of and concerns about privacy and security issues with regard to web cookies. During the literature review, several technological, regulatory, and functional difficulties around the security and privacy of online tracking technologies such as cookies was discussed. Web-tracking poses privacy and security hazards to users' data in cyberspace, as identified by studies based on theoretical models and empirical evidence. Based on these findings, the research questions for this study are identified and developed in Section 3.2. To address the research questions given in the second section, a quantitative research design is defined in section 3.3. In Section 3.4, the survey's strategy and mode, sample, measures, scoring, validity and reliability, data collection and processing, and ethical considerations are all discussed.

3.2 Research Questions

Previous studies revealed that web trackers offer several security and privacy issues to user data. By design, web tracking is hidden from the user (Lutz et al., 2019). It is extensively used, with some form of online tracking being identified in 90% of popular websites and 60% of websites with extremely privacy-critical content (Schelter & Kunegis, 2018). According to statistics, the number of Internet users at the start of 2022 increased by 192 million (+4.0%) from the previous year (KEMP, 2022). Many studies have indicated the lack of user's awareness regarding web tracking (Kollnig et al., 2021; Larsson et al., 2021; Libert, 2018; Thode et al., 2015; Xu et al., 2018); this research will attempt to fill that void. The research will focus on evaluating the awareness of computer users around the world related to the privacy and security risks offered by online trackers, especially cookies, on their data.

This research will attempt to answer two research questions. The first question for this study is:
Question 1 (Q1). *To what extent are computer users aware of the security and privacy threats associated with web cookies?*

Users' concerns about the privacy and security of their data may be exacerbated by the prominent threats to users' data posed by trackers integrated into websites. Much of the research discussed in the literature review shows that users are concerned about the privacy of their data on the Internet, whether in the context of social media or the Internet in general. In addition, some studies have looked at users' concerns about the privacy of their data, which is enforced by third parties, from a geographical standpoint, such that the sample

is limited to a single geographic region, and the study of this has an influence on the amount of their concern. But users' concerns about the privacy and security of their data online have not been investigated in terms of other factors such as usage, storage, and collecting by web trackers such as cookies in a broad and complete context. As a result, this research will seek to fill this gap.

Therefore, the research also attempts to address the study's second question:
Question 2 (Q2). *How concerned is the computer user about the use of their data that is collected through web cookies and Internet use trackers?*

3.3 Research Approach

This study adopts a quantitative approach design. According to Babbie (2010), quantitative research is an approach that focuses on objective measurements and statistical, mathematical, or numerical analysis of data obtained through polls, questionnaires, and surveys, or by modifying pre-existing statistical data using computer tools. Statistics are a method of understanding numerical data in a manner that readers can understand, allowing the research findings to be used for evidence-based practice and thereby closing the gap between theory and practice (Marshall & Jonker, 2010). Quantitative research aims to gather numerical data and generalise the data across groups of individuals or to explain a specific event. This quantitative study follows a descriptive research design. Descriptive statistics, which can also be referred to as "explorative statistics" in some contexts, are ideal for gathering and summarising quantitative data (Marshall & Jonker, 2010). Descriptive research is used to characterise the current state of a variable that has been identified. These studies are intended to give detailed information on a phenomenon (Bloomfield & Fisher, 2019). The survey is a popular primary data collecting method for quantitative research. A survey for data collection is an effective technique to assess a user's level of awareness. Several research that looked at users' awareness of the security and privacy of their data on the Internet relied on surveys (Ali et al., 2019; Nyoni & Velempini, 2018; Parker et al., 2015). Thus, the data will be collected in this research by using the survey to contribute to answering the research questions for this study, Question 1 and Question 2. The survey method was chosen as a suitable strategy since this study aims to define users' concerns and awareness regarding web trackers. Figure 3.1.shows a visual representation of the research's selected approach design.

Figure 3.1The sequence displays the adopted research approach.

3.4 Survey Design

To answer Q1 and Q2, a cross-sectional survey was conducted. The survey's goal was to gather quantitative data for analysis in order to determine users' rates of concern and awareness about the privacy and security threats posed by cookies and other related technologies.

3.4.1 Approach and Mode

The survey was carried out using an Internet questionnaire. This questionnaire was designed according to Lumsden's (2007), guidelines as a practical reference guide for the design of online questionnaires illustrated in Figure 3.2. Closed-ended, dichotomic, or multiple-choice items with five-level Likert scales ranging from strongly disagree to strongly agree were used in this study. The closed questions were chosen so that a quantitative study could be conducted. A total of 34 closed-ended questions were included in the questionnaire. Qualtrics survey was chosen to provide the online questionnaire because of its simple navigation and structure. The convenience sample was gathered from users after the questionnaire was

published and distributed on social media sites such as Twitter and Facebook. All possible participants were directed to a Qualtrics online survey, which they were encouraged to complete. Appendix A contains a copy of the invitation's information and notifications. These messages were left up until the survey was completed. No personally identifiable information (PII) is gathered, and monitoring of the IP address and email address is deactivated to guarantee anonymity.

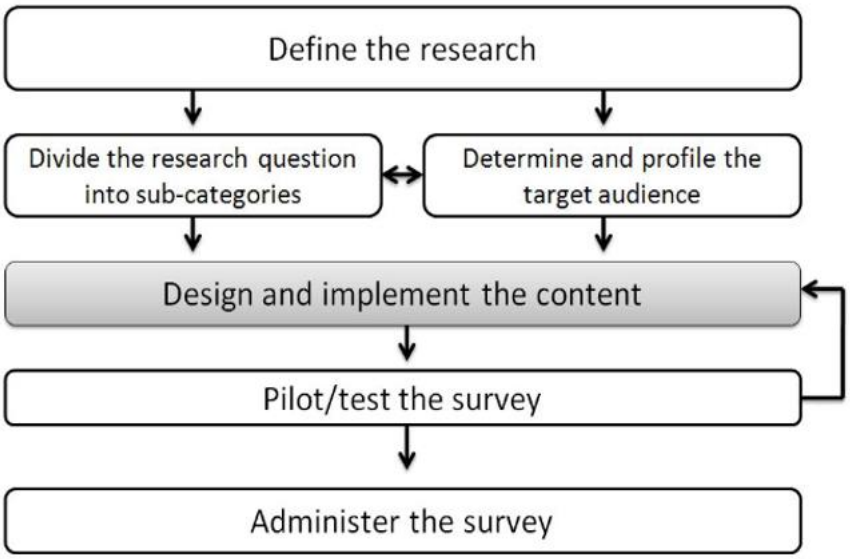


Figure 3.2 Web-based survey design process (Adapted from Lumsden, 2007, p. 46)

3.4.2 Sample

The non-probability sampling strategy employed in this study which satisfied the following criteria:

• Internet users on an international level.

• Adults over the age of 18

The worldwide Internet users' target audience was identified with the latest data, which represents nearly four billion global active Internet users (Johnson, 2021). As an outcome, a statistically significant sample size of 385 respondents was computed using the following formula to achieve an industry-standard confidence level of 95 per cent and a margin of error of 5%:

$$\frac{\frac{z^2 \times p(1-p)}{e^2}}{1+\left(\frac{z^2 \times p(1-p)}{e^2 N}\right)}$$

Where z (confidence level represented as a z-score) = 1.96, N (population size) = 4950000000, e (margin of error represented as a decimal) = 5%.

3.4.3 Questionnaire Design

Based on a comprehensive literature review and consultation with an experienced professional researcher, the survey was developed to gauge perceptions of concern and awareness about the privacy and security of web trackers leveraging cookies embedded in websites. The questionnaire's first two sections are designed to gather demographic information about the participants and determine their technical skills. The questionnaire was divided on the basis of the research questions into two parts, under each part subgroups that contribute to reaching results to answer the research questions see Figure 3.3. A copy of the questionnaire can be found in Appendix B.

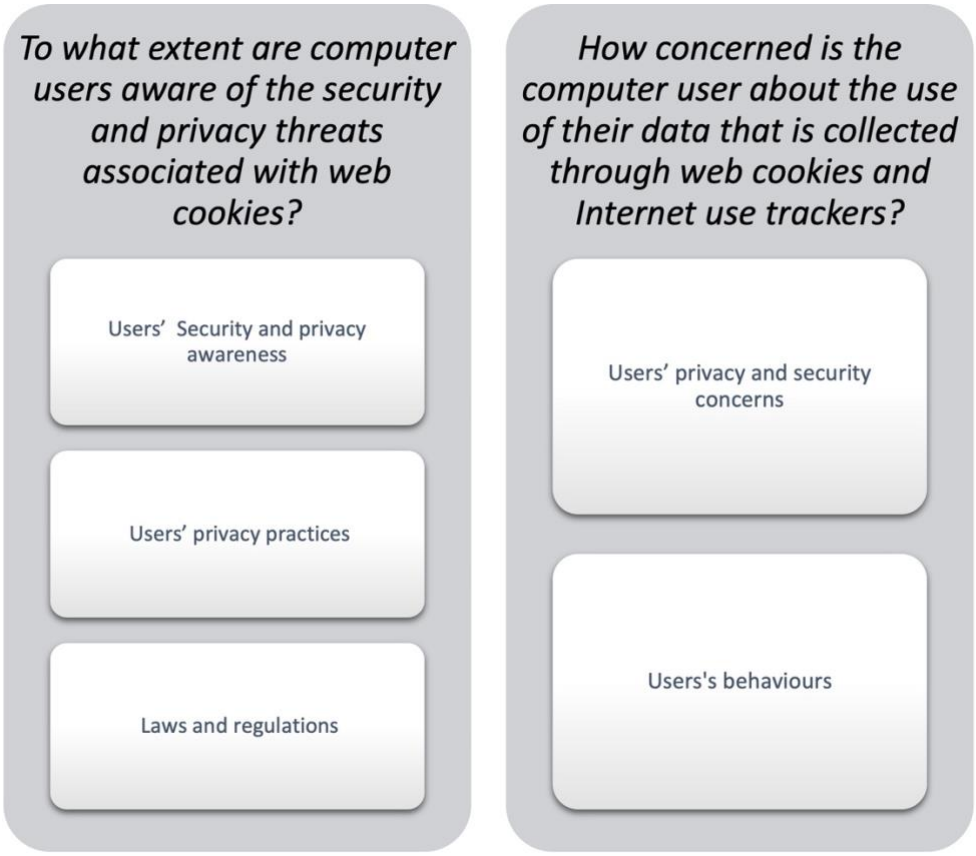


Figure 3.3 The questionnaire's structural design

3.4.3.1 Socio-demographic variables

The influence of demographic characteristics on the levels of information privacy awareness, information security awareness, and information privacy concerns have been the subject of several research, with mixed

results. Bergström (2015), indicated that IPC is a subjective term that, depending on demographic variables, can be expressed in several ways, even when the conditions are the same. Studies have examined the influence of gender, education and age on the level of concerns and awareness towards data privacy including (Baruh & Popescu, 2017; Baruh et al., 2017; Boerman et al., 2021; Chai et al., 2009; Hoy & Milne, 2010; Pinto et al., 2020; Edith G. Smit et al., 2014). The researchers also raised a point indicating that levels of awareness and concern about data privacy are influenced by an individual's cultural background, which varies between geographic regions (Bellman (Bellman et al., 2004; Milberg et al., 2000; Miltgen & Peyrat-Guillard, 2014; Taddicken, 2014). As a result, the first section of the survey was developed with the goal of gathering demographic information about the participants in order to identify the characteristics of the participants in this research that aid in selecting the right sample for the study. The four key variables in this section are age, education levels, geographic location, and gender. As part of the analysis, it was also determined to examine for associations and relationships between the demographic information of the participants and their levels of awareness as well as their concern of privacy and data security in relation to cookies.

3.4.3.2 Users Technical background

User privacy practices have been linked to their Internet skills. The body of research on digital inequality demonstrates how people's online sharing behaviours (Correa, 2010) and usage of privacy-enhancing technologies (Hargittai, 2010; Hargittai & Litt, 2013; Tomoya et al., 2012) are influenced, at least partially by their Web-using skills. Büchi et al. (2017), concluded that the factor most strongly influencing online privacy behavior is Internet skills. The presence of digital skills was proposed by Litt (2013), as a sign that users' privacy was protected. In light of this, the second component of the questionnaire was created to gauge the participants' technical skills with regard to their Internet experience, including their foundational control of computers and smartphones as well as online behaviors like website registration and purchases.

3.4.3.3 Level of Information Privacy (IPA) and Security (ISA) Awareness in the context of web tracking

In light of the objectives of this research, this section of the survey contributes to answer the first research question through three subsections. A prior literature review on the same topic served as the basis for the selection of this section. This section seeks to determine the level of user awareness of the privacy and security risks posed by web tracking, as well as protective and regulatory countermeasures in the context of tracking. This part includes three sub-sections, namely, the third, fourth and sixth sections. The third section consisted of a series of questions aimed at determining the participants' degree of understanding

regarding whether they are aware of the security and privacy risks of their data collected through web cookies. To accurately quantify these insights, there is a foundation of study literature that reflects online privacy awareness and understanding including studies (Chanchary & Chiasson, 2015; Narayanan, 2020; Pinto et al., 2020). Whereas the questions in the fourth section are designed to identify information on users' privacy practices, such as reading privacy policies, accepting or rejecting cookies, adjusting cookie settings on websites, and utilising web anti-trackers based on the following studies (Chanchary & Chiasson, 2015; Chiasson et al., 2018; Purcell et al., 2012). The questions in the sixth section are designed to assess participants' awareness and understanding of the usage of cookies and data privacy in local legislative implementations. Drawing on researchers' assertions that it's important to consider how well-versed Internet users are in privacy-related rules and regulations while evaluating online privacy literacy OPL (Prince et al., 2021).

3.4.3.4 Level of Information Privacy Concerns (IPC) in the context of web tracking

In line with the objectives of this study, this part contributes to answering the second research question. This section seeks to measure users' concerns and perceptions associated with web tracking violating their data. The fifth and seventh dimensions are where these questions are addressed. The questions in the fifth section are primarily based on the New Zealand Privacy Act 2020 information privacy principles. These principles have been used as a basis for assessing users' concerns in the context of web tracking about data collecting, data storage and security, as well as data usage and disclosure. To ascertain the extent to which individuals are concerned about data privacy, these principles have also been used since they are compatible with several theories and models which sought to conceptually define and measure IPC (Chang et al., 2018; Dinev & Hart, 2006; Malhotra et al., 2004). However, existing research has examined the connection between privacy concerns and behaviours in a range of academic disciplines (Acquisti et al., 2015; Baruh et al., 2017; Boyles et al., 2012; Fujs et al., 2019; Hughes-Roberts, 2013; Taddicken, 2014). According to Adjerid et al. (2018), there is a shortage of research on privacy that includes proof of actual behavioural decisions. In order to explore and analyse users' privacy and actual behaviours in the context of cookies, a website was created for this study. In the seventh section, as part of the survey, a separate link for a designed website for this research was included to evaluate users' behaviours. This section examines the key factors that influence users' propensity to read cookie privacy policies. In addition, it examines the most significant obstacles that prevent users from reading websites' cookie privacy policies. Whereas, much research has discovered that people's privacy concerns do not always correspond with the privacy decision they choose, a situation known as the "privacy paradox " (Brown, 2001; Norberg et al., 2007). So, the disparities between users' perspectives and attitudes toward privacy in the context of web cookies are also examined in this section.

3.4.4 Measures

3.4.4.1 Users' privacy and security awareness

Users' awareness of information security and the privacy of data associated with web tracking including cookies was evaluated via ten questions. The questions assess the extent to which users are aware that they are being tracked across the web, as well as their awareness of the data gathered from them via web browsers and the implications for data privacy and security. The questions are closed-ended, with yes/no, multiple-choice, and multiple-choice with five-level Likert scales that measure the level of agreement. Figure 3.4 is an exemplification of one of the questionnaire items used to assess users' awareness.

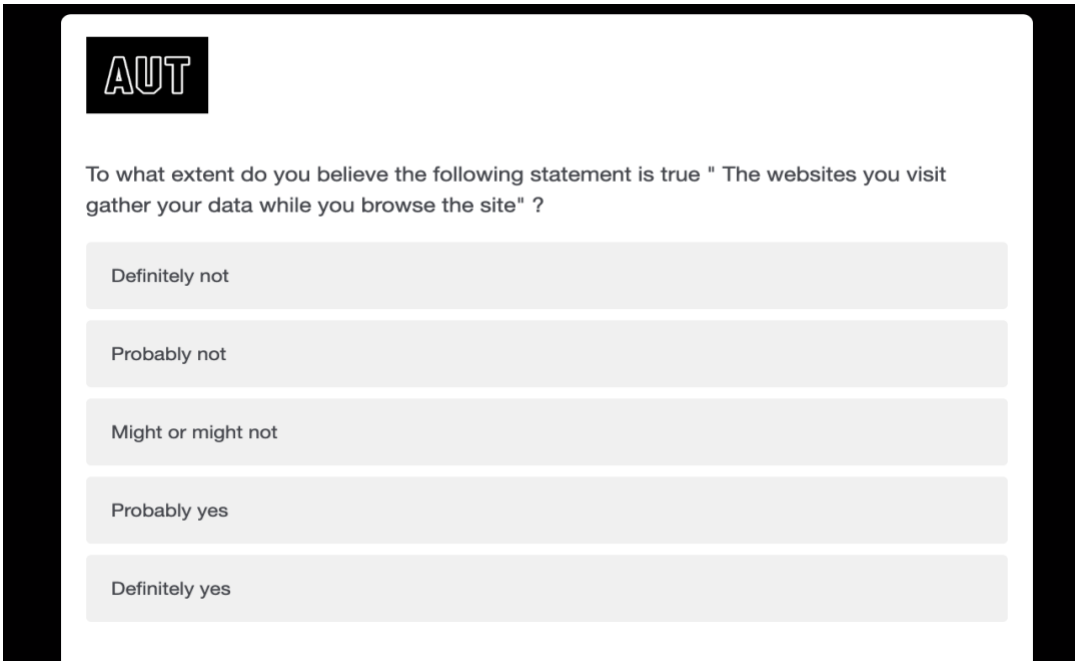


Figure 3.4 depicts one of the questionnaire items used to measure users' awareness.

3.4.4.2 Users' privacy practices

Users' comprehension and awareness of web trackers assessed via five questions regarding privacy practices based on how often they read cookie policies regulating PII collection, their control over acceptance and rejection, and the security techniques used to prevent tracking. One of the survey questions used to assess users' privacy practices is shown below in Figure 3.5.

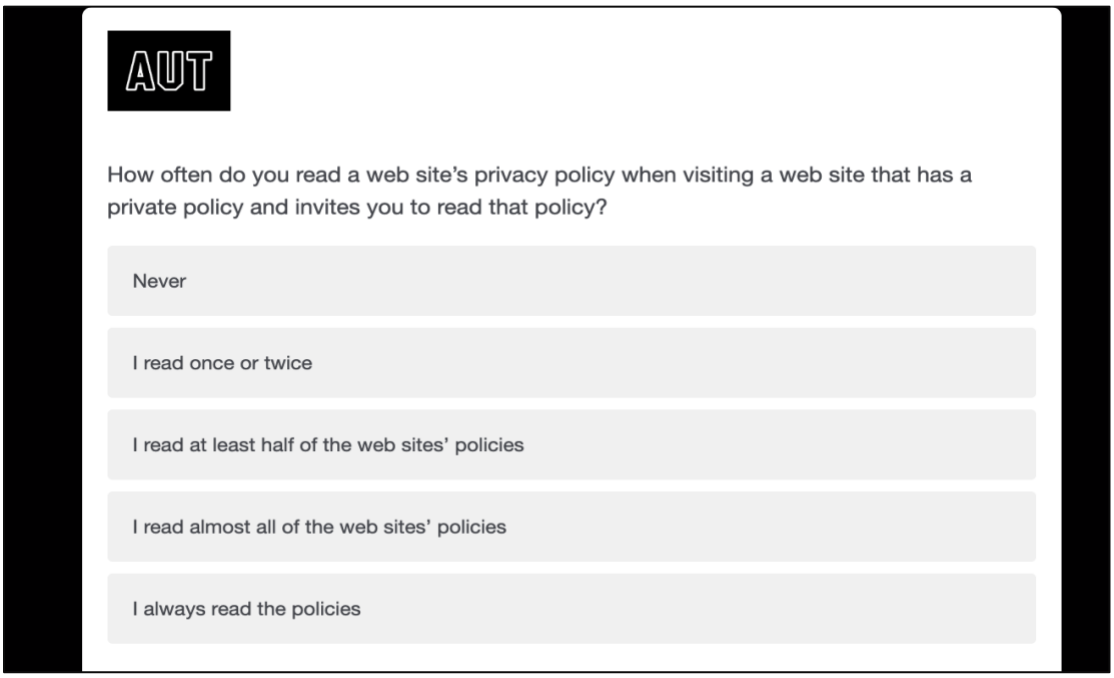


Figure 3.5 represents one of the questionnaire items used to examine users' privacy practices.

3.4.4.3 Users' privacy and security concerns

Users' concerns regarding the privacy and security of IPP data gathered by cookies and third parties were addressed in this measure, which contained eight questions; these questions were created in accordance with the New Zealand Privacy Act's data protection regulations, which are in line with international standards for legislative and regulatory regulations on information privacy (IPP). Users' concerns about personal information gathering, data usage and disclosure, and data storage and security are represented by the items.

3.4.4.4 Laws and regulations

This level of legal and regulatory awareness is related to other factors that have been brought up in several research with a focus on data privacy. Relating environmental influences, laws, rules, and social norms might have an impact on information privacy levels (Chen et al., 2008). Users' knowledge of declarative privacy is one primary factor for measuring users' online privacy literacy. The knowledge of the legislation or other legal aspects of online data protection falls under the umbrella of declarative privacy knowledge (Prince et al., 2021). The familiarity of Internet users with privacy-related laws and regulations must be considered (Prince et al., 2021). Accordingly, this section includes two questions designed to assess users' understanding of data privacy legislation and regulations, as well as their comprehension of local data privacy laws.

3.4.4.5 Users' behaviours

In this section, a website was designed for the study to analyse user behaviour in terms of cookie privacy policies awareness and the reasons for reading or not reading cookie privacy policies on websites. A sequence flow of the pilot site is shown in Figure 3.6; this part requires the user to visit this site. When the user clicks on the link, a pop-up cookies privacy policy will appear, giving the user the option of accepting or rejecting the site's cookies as shown in Figure 3.7. After then, the user was asked if had read the study site's cookie policy. In the case that the user read the privacy policy, the participants were questioned about the most significant reasons for doing so, and if they did not, the same question was asked.

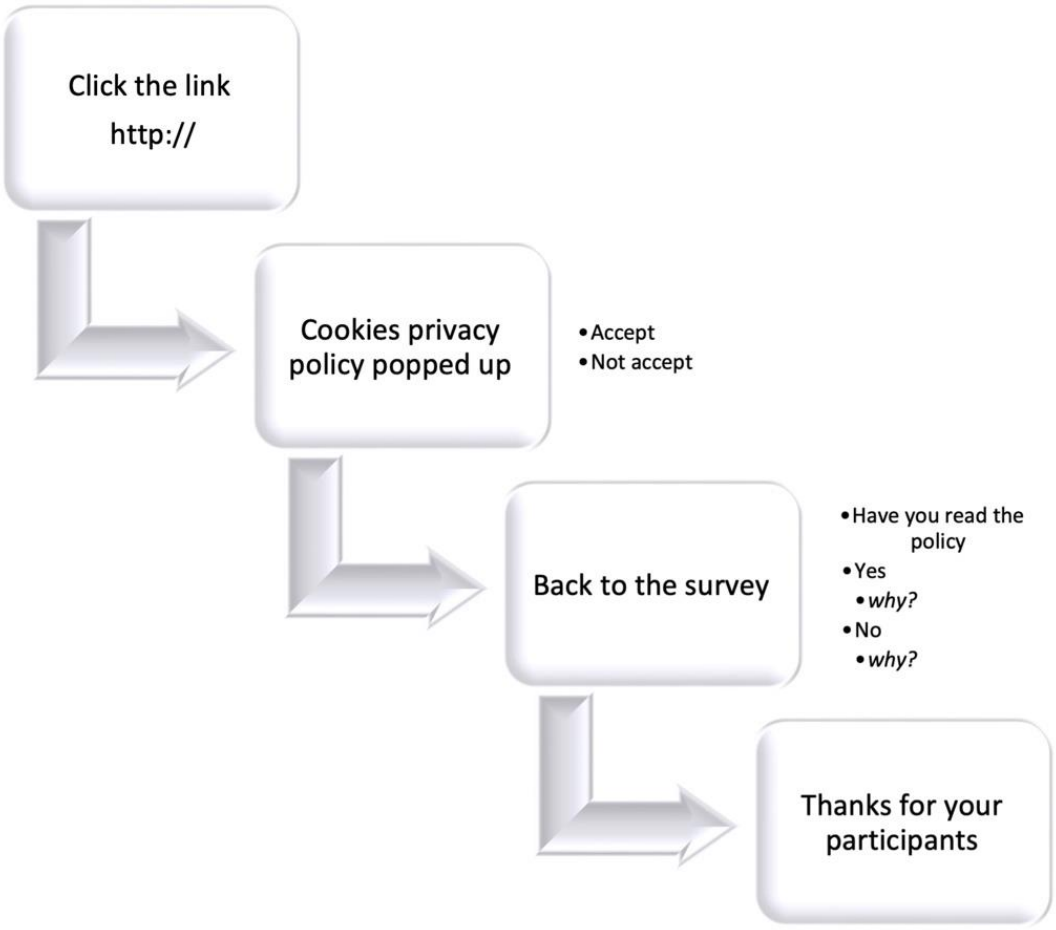


Figure 3.6 depicts the pilot website's process flow.

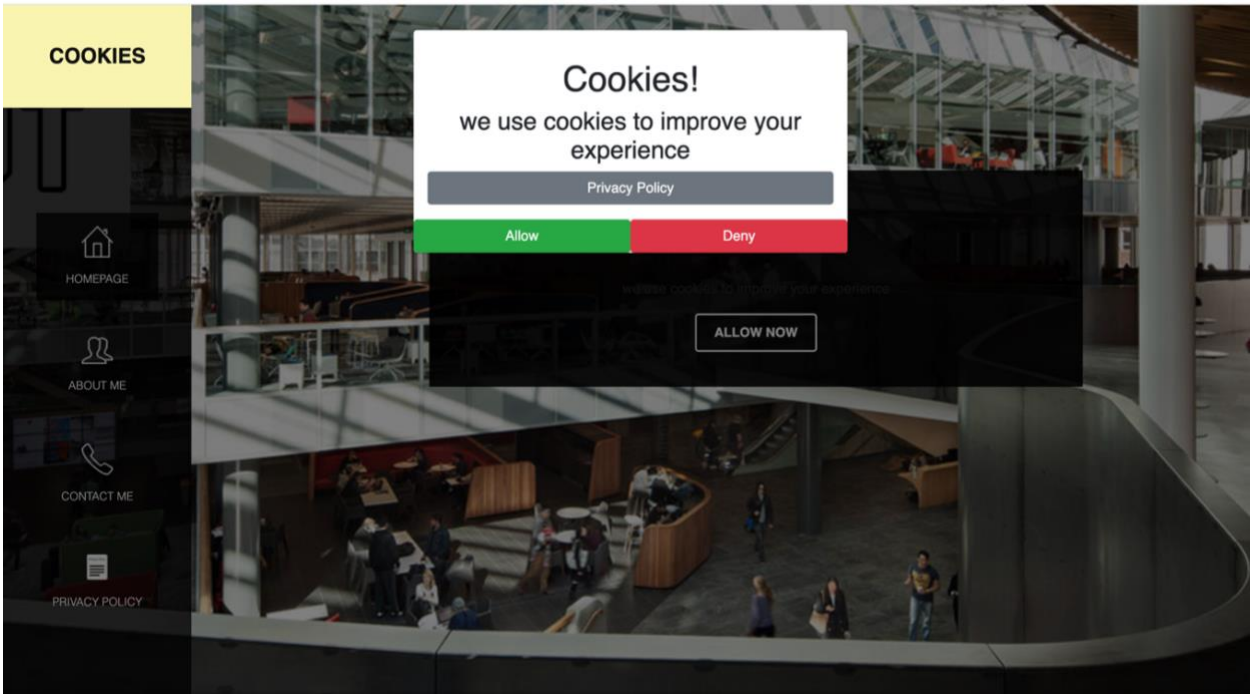


Figure 3.7 shows the popped-up cookies privacy policy for the experimental research website.

3.4.5 Scoring

To perform statistical analysis of data, items evaluating users' privacy and security knowledge, privacy practises, privacy and security concerns, laws and regulations, and user behaviours were scored and coded. Coding is the process of assigning numbers to the values or quantities of each variable (Morgan et al., 2019). A variable is defined by Morgan et al. (2019) as a parameter of the participants or circumstance in research that has a range of values. There are various scales of measurement for scoring variables in statistical terms (Marshall & Jonker, 2010). The variables analysed in this research are classified as categorical data. A categorical variable, also known as a qualitative variable, is a variable that can have one of a fixed number of potential values, typically fixed, and that classifies each individual or other unit of observations into a specific group or nominal category based on some qualitative feature (Starnes et al., 2010). In the social sciences, categorical scales are frequently used to assess attitudes and beliefs (Agresti, 2018). This level of data aids in measuring the participants' opinions and perceptions in order to determine their levels of concern and awareness regarding the threats that web tracking pose to data security and privacy. Nominal and ordinal scales are the two primary measuring types for categorical variables (Agresti, 2018; Marshall & Jonker, 2010). In this analysis, ordinal and nominal measures were utilised as variable-level measurements. Nominal variables are categorical variables with scales that are not ordered (Agresti, 2018). The inclusion of nominal variables in this study allowed for a closer examination of users' awareness

of and comprehension of various risks associated with web trackers, including risks to their privacy and data security. The nominal scored as the following example: 1 = No, 2 = Yes. Unlike nominal variables, ordinal variables are categorical variables with an ordered scale (Agresti, 2018). The use of ordinal variables in this study enabled it to quantify the level of agreement, significance, and frequency in a range of concerns related to the risks of web tracking, including threats to users' privacy and data security, which in turn aids in determining the levels of awareness and concern. The ordinal categorical variables measuring the level of agreement as a Likert Scale were scored as the following example: 1 = Strongly disagree, 2 = Disagree, 3 = Neither agree nor disagree, 4 = Agree, and 5 = Strongly agree. Similarly, 1=Definitely not,2 = probably not, 3 = might or might not, 4 = probably yes, and 5 = definitely yes. Likewise, the level of importance, 1=extremely important, 2 = very important, 3 = somewhat important, 4 = not so important, and 5 = not at all important.

3.4.6 Reliability &Validity

The validity of the survey measurement according to Etchegaray and Fischer (2010), is defined by whether or not the items are related to the domain of the measurement of concern. The studies on web tracking and cookies privacy and security threats were extensively researched in constructing the research questions in order to retain the validity of this study. A pilot test was undertaken to confirm the research's reliability by ensuring that questions were phrased appropriately and getting input from participants on if there was anything confusing that needed more explanation. Cronbach's alpha is another internal consistency-based measurement of reliability. The internal consistency test was used to assess the poll's item reliability, and the result was 0.685 after three items were excluded. Given that the survey components do not assess exact and narrow conceptions of the topic, but rather generic combinations; this result implies a fair degree of question consistency. High reliabilities (0.95 or greater) aren't always desirable because they suggest that the items are repetitive (Streiner, 2003). When creating a reliable instrument, the objective is for values on comparable items to be connected (internally consistent).

3.4.7 Data Collection

Social media platforms are most suited for gathering research data since the study intends to gauge individuals' broad knowledge of their online behaviours. According to Nurdin (2017), users can express their ideas and share their expertise and thoughts with others in the online realm through social media. This allows academics to broaden their study scope and collect data in a more diverse context (Nurdin, 2017). For data collection, to ensure that the intended participants were addressed, the questionnaire link was distributed to different social media channels. A survey link was shared on Twitter, for example, using relevant hashtags in order to increase wider acceptance. A summary of the survey was included in the

posting with the intention of motivating users to engage. The first page of the questionnaire submission was an information sheet, after which prospective participants were asked whether to begin the questionnaire in order to participate in the study. Consent to participate in the study is given by starting the questionnaire. Qualtrics hosted the survey online for around three months, from March 2022 to May 2022. This time range was established to guarantee that the sample size objective was met. The questionnaire took between 15 and 20 minutes to complete. Until the questionnaire was completed, data obtained from the questionnaire was gathered and kept online in compliance with Qualtrics Privacy Policy (Qualtrics, 2021). The data was exported to the researcher's personal computer for analysis once the survey was closed.

3.4.8 Data Preparation and Analysis

Initially, the goal of this study is to characterize a problem by establishing knowledge of the issue and a representation of probable interconnections. Generally, the descriptive statistics method allows forming conclusions about specific phenomena, and findings that, based on individual pieces of data, illustrate the validity or, in some circumstances, non-validity of existing expectations or theories (Cleff, 2014). Based on that, the data analysis procedure started with an examination of the data extracted from the questionnaire distributed to the sample. This is through the exploratory data analysis (EDA), according to Shreffler and Huecker (2021), conducting EDA approaches is critical for properly presenting findings to a target audience and creating appropriate graphs and figures. It can also detect whether outliers exist, whether data is missing, and how to interpret the data. The data was filtered and categorised after the process of exporting and altering the existing data in order to run statistical analyses and draw conclusions. SPSS.V28 was used to perform a descriptive statistical analysis on the data collected in order to identify possible implications and trends. Some of the procedures and techniques used in descriptive statistics include the calculation of numbers and parameters, as well as the creation of graphics and tables (Cleff, 2014). So, the researcher adopted theoretically three types of data analysis techniques: Univariate, Bivariate and Multivariate. The concept of univariate analysis is employed when there is just one item in the data and no cause-and-effect relationship (Babbie, 2007). In contrast, when there are two variables in a data set, bivariate analysis is used to compare the two parameters. Univariate analysis can be performed in a variety of methods, the most common of which are frequency distribution tables, frequency polygons, pie charts, and bar charts (Park, 2015). While bivariate analysis, such as the Mann-Whitney U test can be used to evaluate and compare two different the difference samples (Corder & Foreman, 2014; Zhang, 2016). Another example of performing bivariate statistical analyses between two independent samples is the chi-square test of independence (McHugh, 2013). When comparing several independent groups, a multivariate nonparametric Kruskal-Wallis test is used (Corder & Foreman, 2014; Katz & McSweeney, 1980). However, it is difficult to determine which central tendencies are substantially different from the others using the Kruskal-Wallis test.

Hence, post hoc tests like Dunn's pairwise comparisons test is required (Wallis, n.d.). Thus, to generate an overview of the univariate, bivariate and multivariate distribution of nominal- and ordinal-scaled variables, the researcher calculated and summarised graphical representations, frequency tables, and scale data means.

3.4.9 Ethics

The survey acquired ethical approval from the AUT Ethics Committee (Ethics Application Number 21/379). The survey was fully anonymous, did not ask the participant to provide any personal information, and was entirely optional. No data was exchanged with a third party, and none of the data contained any personally identifiable information (PII).

3.5 Conclusion

Chapter 3 discussed the research design. The primary research questions were identified and defined in this chapter. The research design, which is an online survey, was explained. The technique employed in the structure, the sample size, the design of the survey questions, the methods and classifications of the survey items, and the data collection and analysis process were all discussed. The results of the online survey are presented in Chapter 4.

CHAPTER 4: FINDINGS

4.1 Introduction

Chapter 3 outlines the research design for assessing user security awareness and concern about the security and privacy issues posed by web trackers, notably cookies. The research questions were developed based on the literature review conducted in Chapter 2, which identified potential security and privacy issues associated with web trackers. This chapter presents the findings that contribute to answering the study's research questions. The results of the online survey are discussed in the context of the study's questions and objectives, which were designed to determine the levels of concern and awareness among users about the privacy and security risks presented by web trackers. This section begins with an explanation of the completion rate and the refinement of the raw data before presenting the survey results. These findings include the conclusions of statistical analyses performed on the data in order to detect any patterns or trends within the collected data that will assist the researcher in better comprehending the findings and addressing the research questions.

4.2 Completion Rate and Exploratory Data Analysis

The questionnaire's completion rate was first determined, which Fincham (2008), defined as the number of acceptable responses received divided by the total number of eligible respondents in the sample chosen to calculate response rates. The goal of this study was to construct an Internet-based questionnaire written in English that could be directed at the international level, spanning all age groups for adults and all educational and technical skills levels. To obtain the required sample, social media channels were chosen as users on social media are diverse in terms of nationality, age, education, employment situation, and interests (Efthymiou & Antoniou, 2012). The approach used to gather the sample could result in a lower aggregate response rate. As Nayak and Narayan (2019) point out, the participation rate in online surveys is an essential issue because response rates are often low when compared to the offline survey approach. A study of multiple survey methods revealed a 33% average response rate for all approaches. In-person surveys had the greatest percentage of 50%. Email surveys received 30%, web surveys received 29%, and in-app surveys received 13% (Lindemann, 2021). Based on this data, the questionnaire of this study was accessible to all targeted respondents and was available online from March to May 2022. Throughout this time, 669 replies were received, with 470 of them completed, for a completion rate percentage of 70.25 %. Incomplete replies were eliminated during the exploratory data analysis stage to prevent utilizing a data set with missing values for analysis, which could misrepresent the results. Because a total of 470 completed replies was deemed sufficient to satisfy the previously determined aim of a representative sample of at least 385, the analysis was conducted using a representative sample of 470 completed responses.

4.3 Summary of Findings

This section provides a summary of the results of the survey questions starting with the demographic questions, qualifications, and technical background of the research sample. This is followed by a presentation of the results from the survey questions that contribute to answering the first research question of this study, which aims to measure the level of awareness among users regarding the risks to the privacy and security of user data collected by web trackers. It is based on three scales that included several items, which are as follows: Users' Security and privacy awareness, Users' privacy practices and Laws and regulations. Bivariate and multivariate analyses is then performed to determine whether there are any statistical discrepancies between the research sample's characteristics and the level of awareness. The results are then extracted from the two scopes of users' privacy and security concerns, along with users' behaviors, which involve a live experience of accepting cookies via a website created specifically for this research. These two sections contribute to addressing the study's second research question, which is to determine how concerned users are about web trackers, particularly cookies, collecting, using, and storing their data. Following this, bivariate and multivariate analyses are used to compare the level of concern to the characteristics of the research sample to determine whether any differences are statistically significant.

4.3.1 Preliminary data on the characteristics of the study population

The objective of this section is to provide some general information about the participants. The primary goal is to give an overview of the research participants' characteristics. In this case, descriptive statistical tools including frequency analysis, tables and charts are utilized to represent individuals based on age, gender, education, and geographic region.

4.3.1.1 Distribution of the respondents by gender, age, and education

The statistical makeup of the research sample is shown in Table 4.1. The answers to Questions 2,3 and 4 were diverse in terms of their characteristics. In respect to gender (51.3%, n =241) of respondents were male, while (47.7%, $n = 224$) were female. A small percentage of the participants were non-binary (.6%, n=3). This study adopted Johnson (2019), age ranges for Internet users. Six sequential patterns are used to compare the age groups of young people, middle age, and old age. The bulk of the research respondents (47.2 %, $n = 222$) were between the ages of 25 and 34, with (23.4 %, n =110) in the 18–24-year age group (4.5 %, n=21) in the 45–54-year age group, (2.1 %, $n = 10$) in the age group under 65 years, and only (1.1 %, n=5) in the age group over 65 years. In other words, 92.3 % of the survey participants were under 44 years old. The educational levels of the participants in this study varied, as the percentage of those with a bachelor's degree or diploma was the majority with (50.2%, $n = 263$) followed by the percentage of those

with a master's degree, (30.4%, n =143), participants with a high school make up a modest percentage of the total about (13.0%,n=61). While the percentage of participants with a PhD constituted the lowest, by approximately (4%, n =22).

Table 4.1
Sociodemographic Characteristics of the Participants

Demographic profile	Number of respondents (n)	Percentage (%)
Gender		
Male	241	51.3%
Female	224	47.7%
Non-binary	3	.6%
Prefer not to say	2	.4%
Age		
18-24	110	23.4%
25-34	222	47.2%
35-44	102	21.7%
45-54	21	4.5%
55-64	10	2.1%
over 65 years	5	1.1%
Level of education		
High School	61	13.0%
Bachelor Or Diploma	236	50.2%
Master's Degree (MS)	143	30.4%
Doctoral Degree (PhD)	22	4.7%
Other	8	1.7%

Note. n *= 470*

4.3.1.2 Distribution of the respondents by geographic region

The research focused on how users' cultural or geographic backgrounds influenced their awareness of data privacy and security posed by web-tracking, along with their levels of concern. This is owing to the existence of several studies in the same research context that demonstrated a link between geographical and cultural factors and privacy concerns (Bellman et al., 2004; Milberg et al., 2000; Taddicken, 2014). As a result, the participants in Question 5 were asked about the range of their geographical residency considering

the research's objectives, which are to evaluate levels of awareness and concern on a global scale. As shown in Figure 4.1, the participants in this study were widely dispersed geographically to approximately 13 geographical areas, the most prominent of which were located in USA by (36 %, $n = 169$) followed by the Middle East (23.2 %, $n = 109$), Asia (19.1%, $n = 90$), and New Zealand and Australia (7%, $n = 37$).

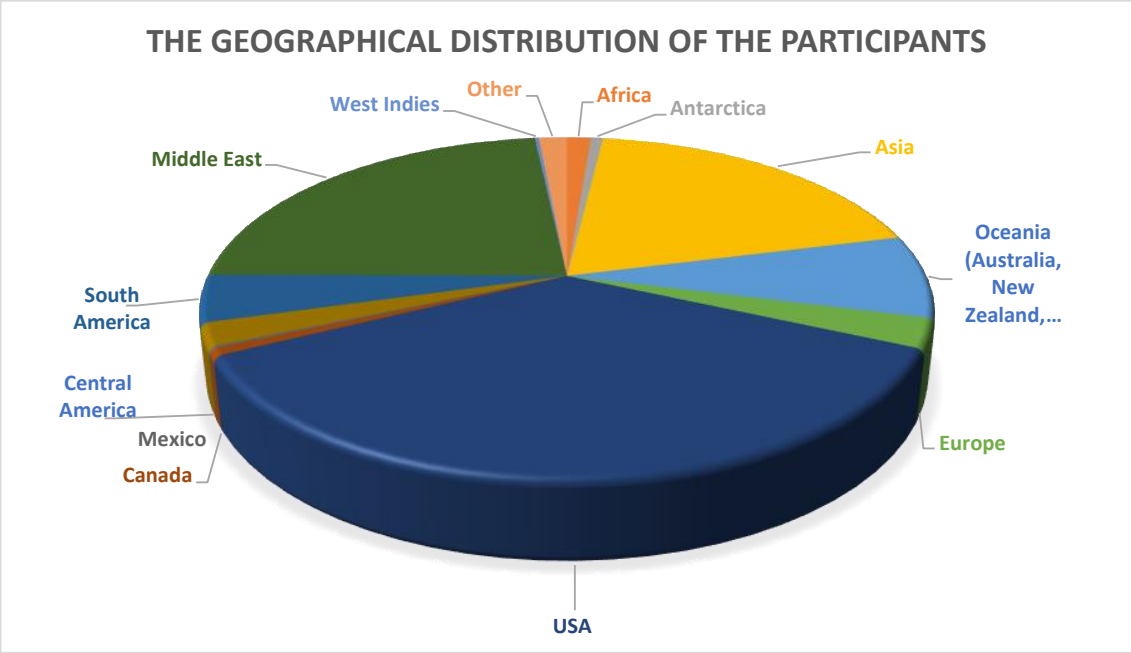


Figure 4.1 The geographic distribution of the participants.

4.3.2 Users' technical skills and background

The answers to Questions 6, 7, and 8 reveal the participants' level of technical skill in regard to their Internet experience, including their fundamental ability to utilize computers and cellophanes as well as online behaviors including web purchases and web registration. Participants were asked to rate their computer and smartphone abilities on a range of Unable to use, Normal User, and Advanced User. Results were roughly split evenly between normal users (51.3%, $n = 241$) and advanced users (43.0, $n = 202$). While the participants' online practices results demonstrated that nearly three-quarters of the participants found it is very easy to register on websites (73%, $n = 354$). Additionally, most of the participants (97%, $n = 456$) have had prior online purchasing experience.

4.3.3 User-level of information privacy awareness (IPA) and information security awareness (ISA) regarding web tracking cookies

This area of the analysis addresses the first research question, which aims to measure the current levels of awareness of web trackers and cookies. This is accomplished across three primary dimensions which assess user awareness of the security and privacy of their data, privacy practices, and web tracker regulations and legislation. There are 14 items distributed over the three dimensions. To describe and report the data, frequency rates, central tendency, tables, and graphs were generated.

4.3.3.1 Users' Security and privacy awareness

This section attempts to determine the general level of awareness of the sample regarding online tracking. Seven items from Questions 9 through 15 are covered in this section. Initially the participants were asked about the level of probability that their data would be collected during their visit via the Internet. As indicated by the results in Table 4.2, the participants support that they believe in a possible way that their data is collected approximately (34.0% $n = 160$) and almost likewise those who believe firmly about (33.6%, $n = 158$). While only (5.5 %, $n = 26$) of the participants were certain that their data was not gathered through websites. To put it another way, the findings indicate the average number of the respondents ($M = 3.83$) generally are likely to think that personal data is gathered when they are browsing the Internet.

Table 4.2
Participants' awareness of data gathering

Items	Number of respondents (*n*)	Percentage (%)
Definitely not	26	5.5%
Probably not	34	7.2%
Might or might not	92	19.6%
Probably yes	160	34.0%
Definitely yes	158	33.6%

Note. n = 470

The participants were asked to identify the types of personal information they believe websites gather about them when they browse. There are eight different types of data to select in Question 10 *"Which of your personal information do you think websites collect about you while you browse?".* Participants can choose

from a range of options. Almost all personal data that may be gathered by cookies and web trackers was included in the selections, including IP or MAC addresses, geographical location, browser and device information, and the pages you visit and click on. As indicated in Table 4.3 where n here refers to the number of cases that selected that response option. Approximately half of the participants identified and selected four distinct types of data, including IP or MAC address (48.7%, $n = 229$) Browser and device information (47.9%, $n =225$) the pages you visit and clicks you make (46.6%, n=219) and Geographical location (44.9%, $n = 211$). This is followed by those who have identified aggregated data on your visits to the site and E-mail addresses in the rate of (34.7% $n = 16$ and 30.0%, $n = 141$) respectively. While, just under a quarter of the participants have identified that their names can be collected via web trackers. Only (19.8%, $n = 93$) of those who specified a Telephone number.

Table 4.3
The types of data collected by web-tracking

Items	Number of respondents (*n*)	Percentage (%)
IP or MAC address	229	48.7%
Geographical location	211	44.9%
Browser and device information	225	47.9%
The pages you visit and clicks you make	219	46.6%
Aggregated data on your visits to the site	163	34.7%
Your name	109	23.2%
Telephone number	93	19.8%
E-mail addresses	141	30.0%

Note. n=470

Participants were asked their opinions and beliefs across several areas to measure their awareness of both privacy and security issues with cookies and web trackers. Table 4.4 demonstrated the summary statistics for Questions 11-15. According to Tirtea et al. (2011), users' privacy is often deemed violated when they are tracked without their agreement. As a result, the participants were asked if they considered that websites or third parties should inform them about the procedures of acquiring their data and how to use it for the aim of determining their level of awareness of their data's privacy. Table 4.4 shows that a large number of the participants, almost (77 %, $n = 365$) believe it is vital to be informed about how websites gather and use personal data. This indicates that the participants are generally aware of the importance of data privacy.

Users are concerned about the privacy and security of their activity records collected by online advertising, according to a survey performed in the United States (Madden & Rainie, 2015). In the same vein, participants in this study were questioned about their thoughts about tracking their activities, the goals for which their data will be used, and the security of their data. The participants were surveyed if they were aware that their online activities were being tracked. As shown in Table 4.4 most of the participants (79%, $n = 375$) are aware that their online activities are being tracked. Likewise, the majority of participants consciously expressed their opinions and beliefs regarding the websites selling their personal data and the security risks posed by cookies. As shown in Table 4.4 a significant percentage of participants (75.5%, n = 355) indicated that they are aware that websites sell their personal information. Additionally, a sizable portion of respondents are aware that hackers are attempting to steal private information saved in cookies (70.4 %, n= 331). Participants were also surveyed about their knowledge of managing cookies on their computer and erasing them. As indicated in Table 4.4 more than half of the respondents reported knowing how to manage and delete cookies on their computers (62%, n =292).

Table 4.4

Statistical summary of the percentages and frequency of participants' opinions and beliefs about their awareness of privacy and security issues with cookies and web trackers.

Items	Number of respondents *(n)*	Percentage (%)
Inform users about data usage and collection		
Yes	365	77.7%
No	105	22.3%
Tracking user online activity		
Yes	375	79.8%
No	95	20.2%
Sell personal information		
Yes	355	75.5%
No	115	24.5%
Exposing to cyber attacks		
Yes	331	70.4%
No	139	29.6%
Cookies management		
Yes	292	62.1%
No	178	37.9%

Note. n=470

4.3.3.2 Users' privacy practices

This section seeks to examine users' online privacy practices through five Questions 16-20 concerning reading, accepting, changing cookie privacy settings, refusing to use the website's service because of its policies, and utilizing anti-tracking tools. Participants were asked how often they read a site's privacy policy when they visit a site that contains a privacy policy and invite them to read that policy. As shown in Figure 4.2 , the chart demonstrates that about equal proportions of participants either do not read the policies at all (29%, $n = 136$), or read it only once or twice (30%, $n = 140$). While only (8%, $n = 38$) of respondents indicated they always read the cookie policies.

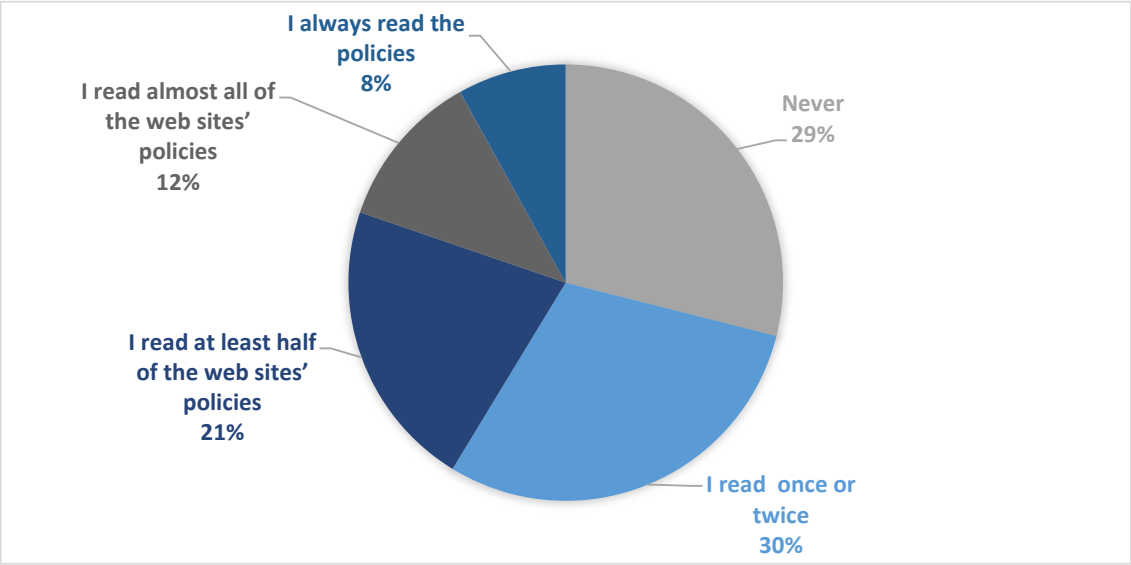


Figure 4.2 The frequency of reading the cookies privacy policy among participants.

Participants were asked how frequently they accepted the websites' cookie privacy policies. The frequency with which the respondents accepted cookies is depicted in the Figure 4.3. As shown in Figure 4.3, while a tiny minority said they never accept cookies (3%, $n = 15$), the majority (33%, $n = 156$) said they accept them sometimes. Also, another group accepts cookies policy most times (27%, $n = 128$) and some always accept it (23%, $n = 109$).

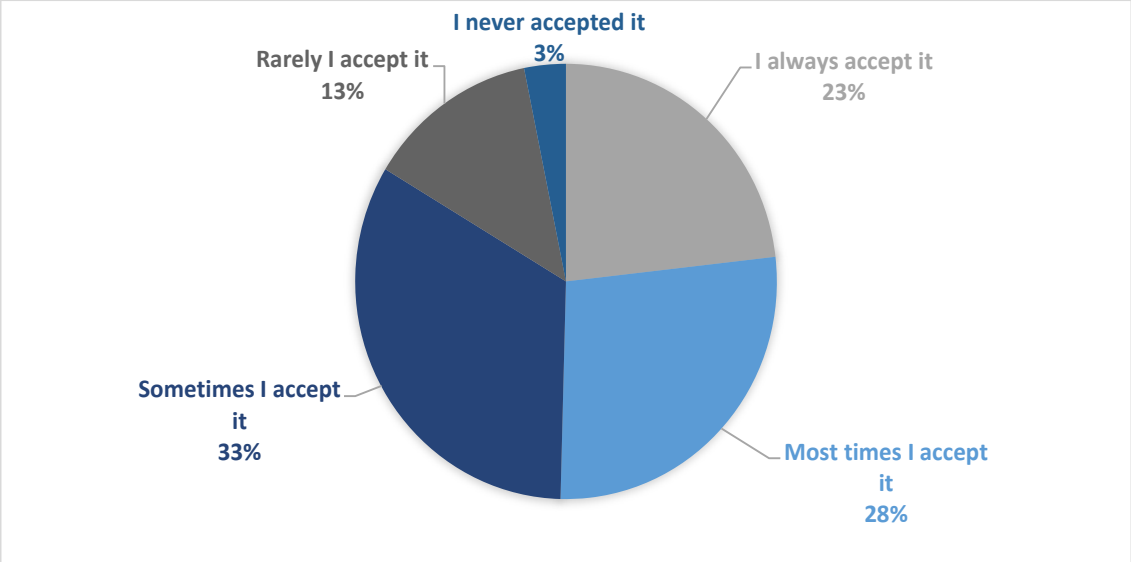


Figure 4.3 The frequency of acceptance the cookies privacy policy among participants

Participants were also surveyed about their habits while browsing the webpages in changing the cookie policies settings on the website. As Figure 4.4 shows, nearly half of the participants answered that they sometimes change the privacy policies of cookies on websites (47%, $n = 219$). Also, about a quarter of the participants always change their cookie settings (24%, $n = 115$). While (30%, $n = 136$) accept cookies as they are on their default settings on the site.

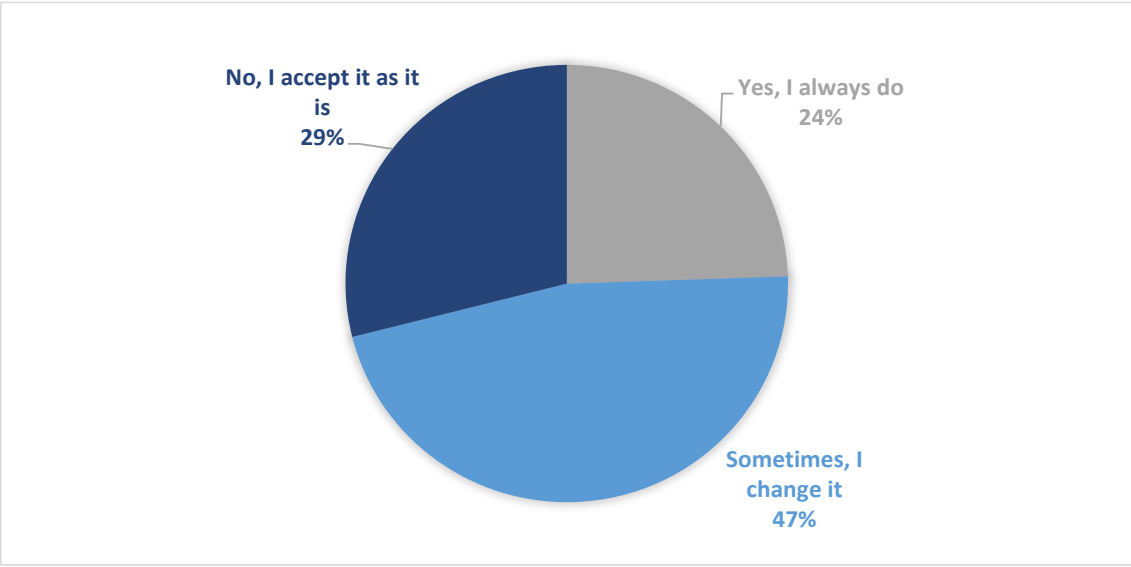


Figure 4.4 The frequency of changing the cookie policies settings on the website among participants.

Other related questions about respondents' privacy practices were addressed, including whether they avoided using websites because of their privacy policies. According to the findings depicted in the Figure 4.5 below the majority of participants had previously denied using the websites' services; of those who responded, (37%, $n = 176$) said they frequently decline, (45%, $n = 210$) said they sometimes decline, and only (18%, $n = 84$) said they have never declined because of the privacy policies.

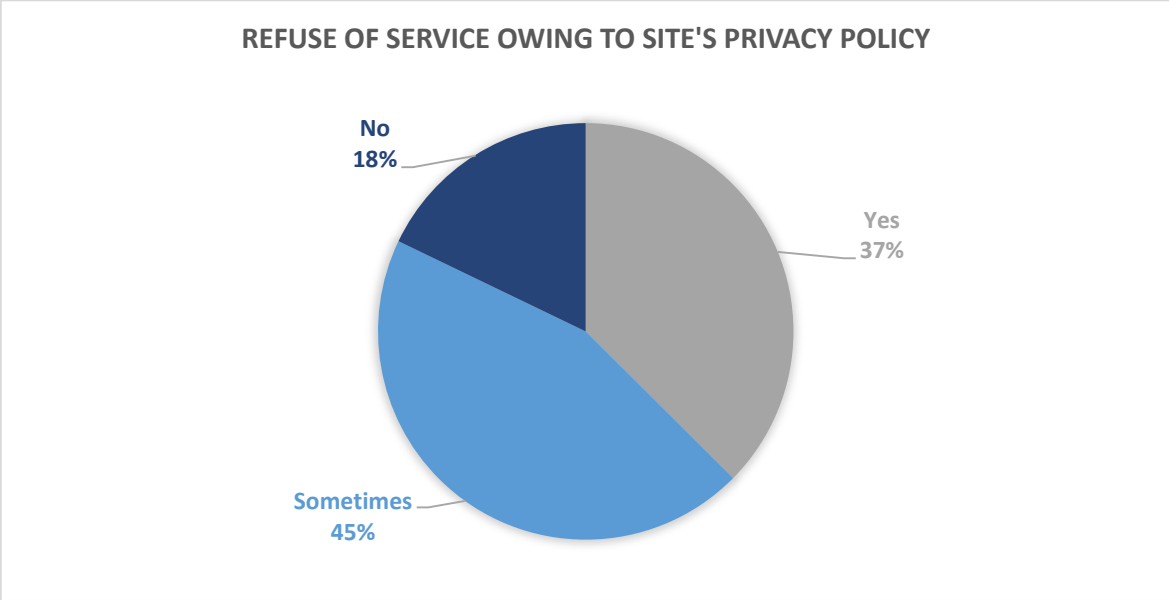


Figure 4.5 The rejection of website service due to privacy policies among participants

Participants were polled about their usage of privacy-protecting anti-tracking technologies, which restrict online activity from being recorded and help users maintain their privacy for example (Tor Browser, Privoxy). As shown in Figure 4.6, slightly above half of the participants (57 %, $n = 271$) said they use anti-tracking software, while (42%, $n = 199$) said they don't.

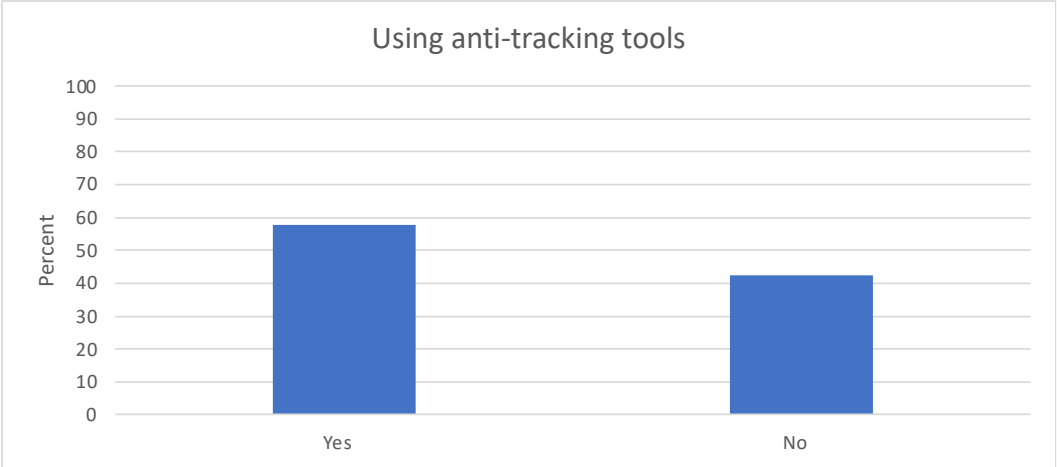


Figure 4.6 Use anti-tracking tools to protect user online privacy by preventing tracking between participants.

4.3.3.3 Laws and regulations

The awareness and understanding of international and local laws and regulations pertaining to the privacy and security of users' data on the Internet is a crucial component of online users' awareness of web trackers. As a result, the participants were asked in Question 29 and 30 about how much they understood about the local legal aspects of the legitimate systems for protecting data privacy and how much they believed these laws were committed to upholding the rights to data privacy. In terms of comprehension and understanding of the laws and legislation, participants were asked to rank their understanding from a scale of 1 extremely low to 5 very high of how the legislators in their nation consider privacy rights. As demonstrated on Figure 4.7, on average ($M=3$), the majority estimated that legislative bodies consider privacy rights by (40%, $n=189$). While some have evaluated a poor evaluation of the knowledge in the local legislation of data privacy rights from low to extremely low at (20 %, $n=94$) and (13.6 %, $n=64$%) respectively.

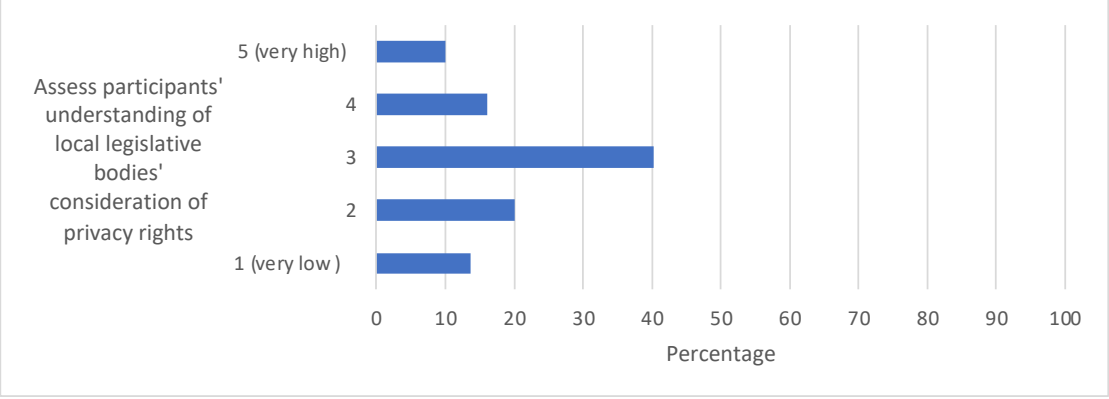


Figure 4.7 Participants' comprehension of local legislative bodies' observance of privacy rights

Participants were also surveyed about their views and level of agreement on the extent to which online service providers comply with privacy protection regulations and requirements imposed by regulatory bodies such as the General Data Protection Regulation (GDPR/CCPA). As illustrated in Figure 4.8, the participants' stance on whether they agree that online service providers adhere to international laws and procedures protecting personal information is neutral (M = 3), representing a (43.0%, n =202) response rate. While a small portion of people disagreed with the websites' adherence to the rules and regulations governing data privacy approximately (6.6%, n = 31).

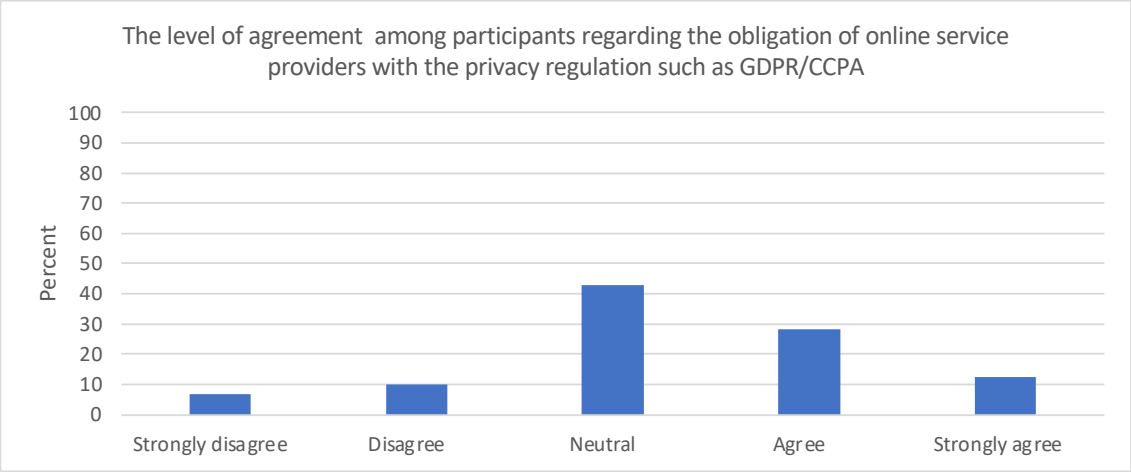


Figure 4.8 The level of participant agreement over Internet service providers' obligations to adhere to privacy regulations.

4.3.4 The statistically significant differences among participants' characteristics in regard to Web-tracking Awareness

4.3.4.1 Web-tracking Awareness in relation to Gender

Table 4.5 reveal that chi-square test of independence shows there is no significant association between the gender and the level of web tracking awareness with $x^2(3, N = 470) = 2.205$, $p = .531$,$\varphi = .068$. The value of phi-coefficient was .068 (< .50) which indicated a small effect size. The results demonstrated that there is no relationship between the awareness of online tracking and the gender variable.

Table 4.5
Frequencies and Chi-Square Results for The Level of Awareness and Gender Factor (N=470)

Awareness of web tracking	Male		Female		Non-binary / third gender		Prefer not to say		$x^2(1)$
	n	%	*n*	%	*n*	%	*n*	%	
Yes	189	78.4%	183	81.7%	2	66.7%	1	50%	2.205***
No	52	21%	41	18%	1	33.3%	1	50%	

$***\rho < .001$

4.3.4.2 Web-tracking Awareness in relation to the geographic region

The chi-square test of independence shows no significant association between the geographic region and the level of web tracking awareness among participants with x^2((12, N = 470) = 16.650, p = .163, φ=.188. The value of the phi-coefficient was .188 (< .50) which indicated a small effect size. Figure 4.9 illustrates the frequency of participant dispersion based on geographic range and awareness with online tracking. The results demonstrated no relationship between the awareness of online tracking and geographic region among participants.

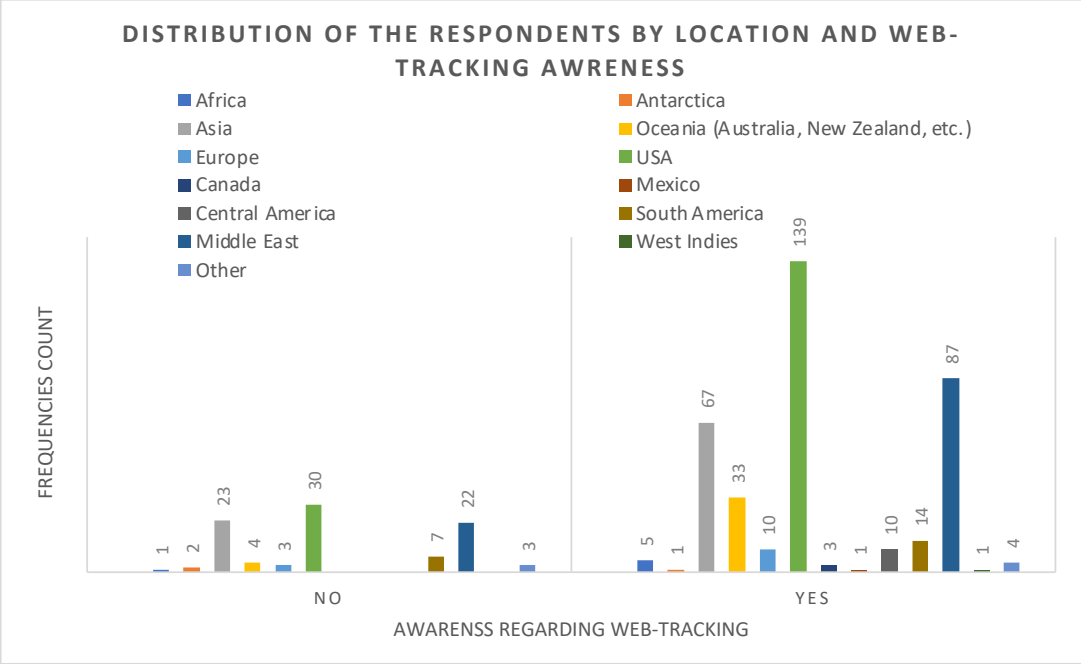


Figure 4.9 Distribution of responders based on geography and web-tracking awareness.

4.3.4.3 Web-tracking Awareness in relation to the level of education

A Kruskal-Wallis test showed that there are no statistically significant differences between the level of education and web-tracking awareness H (4) = 7.466, p=.113. Post-hoc Dunn's pairwise comparisons test, as illustrated in Figure 4.10, carried out for the five pairs of groups. There were no significant differences between the level of education and the awareness of web-tracking among participants.

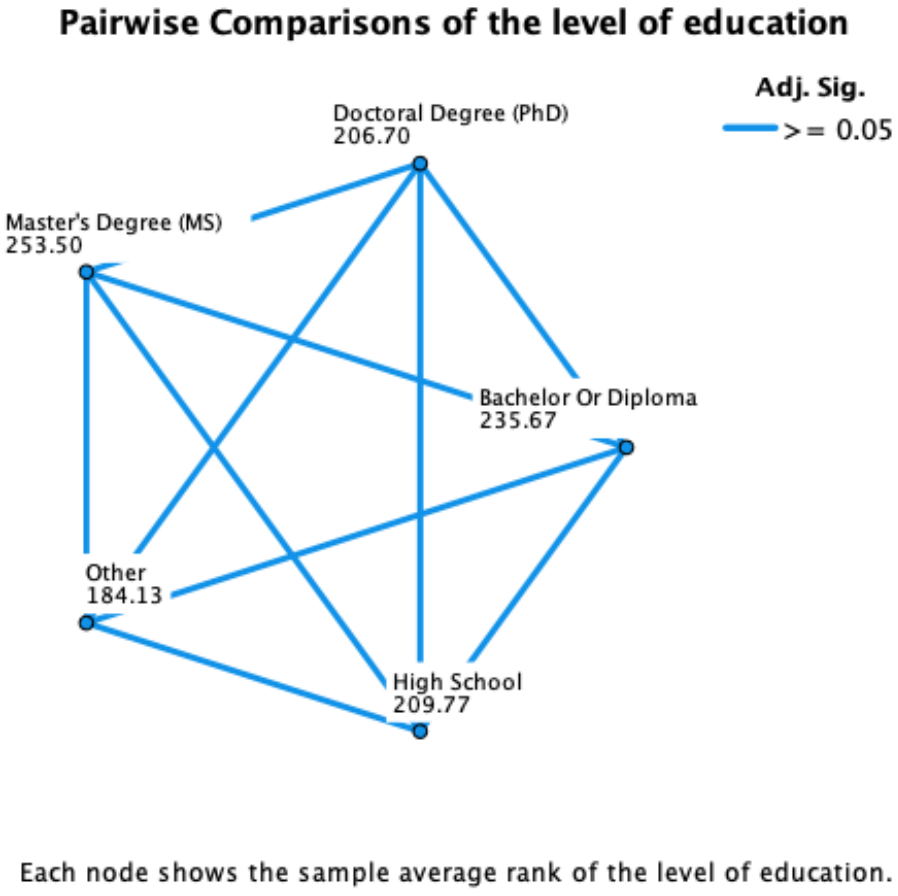


Figure 4.10 Post-hoc Dunn's pairwise comparisons test for the five pairs of groups and online tracking awareness.

4.3.4.4 Web-tracking Awareness in relation to the technical background

A Kruskal-Wallis test showed that there is a statistically significant difference between the level of technical skills and web-tracking awareness H (2) = 28.165, p=.001. Post-hoc Dunn's pairwise comparisons test, as illustrated in Figure 4.11, was carried out for the three pairs of groups. There were significant differences between the level of technical skills and the awareness of web-tracking among participants.

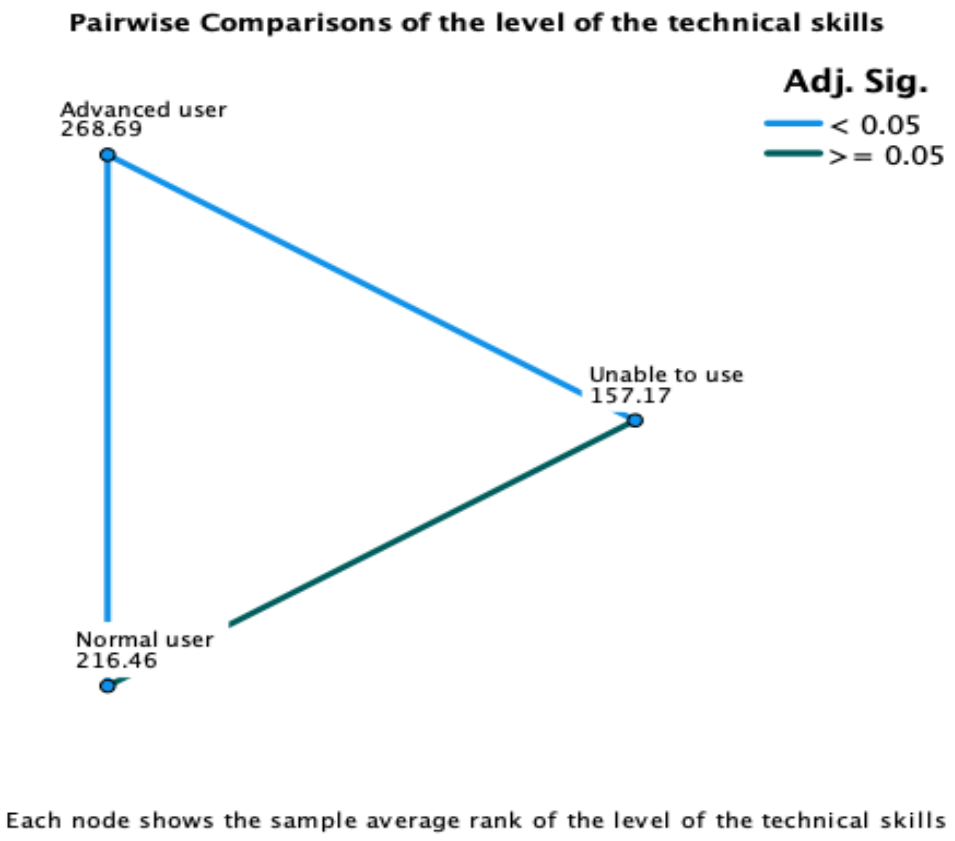


Figure 4.11 Post-hoc Dunn's pairwise comparisons test for the three pairs of groups and online tracking awareness.

4.3.5 Principal findings regarding participants' levels of awareness

The findings and related data in the three aforementioned sections help to address the study's first research question. The first research question aims to assess users' awareness of security and privacy issues with cookies and web trackers. Overall, the findings in the section on users' privacy and security awareness indicate that there is awareness among the participants, as evidenced by the average number of the sample and the perception that they are being tracked through the websites. Additionally, less than half of the participants were able to identify some of the types of personal information gathered from websites. Four of the eight most popular data types of IP visited webpages and clicks, geographic location, browser, and device information were chosen by the majority of participants that account for 40%. It's interesting that so few respondents have acknowledged that their names and phone number are being tracked. Likewise, a

large chunk of the participants nearly three-quarters expressed awareness and knowledge of the privacy and security issues associated with web trackers. Where they believe that they must be informed about the purposes and uses of their data, and the majority are aware that their online activities are tracked, selling their data, and the possibility of being exposed to cyber-attacks from their data stored in cookies. In like, more than 50% of the sample were aware of how to manage cookies on their devices. The findings on the participants' privacy practices revealed that the majority did not read the cookie policies and frequently accepted cookies. Nearly half of the participants sometimes alter cookie privacy settings on websites. Additionally, many of the participants make use of anti-tracking techniques. In terms of awareness of regulations and legislation, the results highlighted a neutral position by the participants regarding understanding local or global regulations and legislation in terms of privacy rights and the commitment of electronic service providers in these regulations and laws. In respect of awareness levels and the characteristics of the sample population in relation to gender, education, and geographical area, there are no statistical variances or differences. On the other hand, the findings showed that there was a significant difference between the participants' levels of technical background and awareness.

4.3.6 User-level of information privacy concern (IPCs)

This part of the analysis aims to address the second research question, which measure current levels of concern among users regarding web trackers and cookies. This is achieved through two sections, one of which is to directly measure participants about their concerns regarding the privacy of their data in the context of cookies, according to the approach proposed by Preibusch (2013) to measure the extent of concern about data privacy. These items are based on the data privacy principles of *New Zealand* law that are compatible with international regulations and legislation. However, a large and growing body of research in the area of information privacy concern (IPCs), which was discussed in chapter two, revealed a mismatch between people's self-reported perceptions of privacy concern and behaviors that fail to meet it. To describe this disparity in attitudes and behaviors towards information privacy, the concept "privacy paradox" was developed (Brown, 2001; Norberg et al., 2007). Accordingly, the analyses in this section are intended to look into any inconsistencies between participants' perceptions of privacy concerns related to cookies and their behaviors by the website designed for this research to examine the privacy paradox. This was implemented in order to determine how concerned users were about their privacy. To explain and present the data frequency rates, central tendency, and a Mann-Whitney U test were calculated.

4.3.6.1 Users' privacy and security concerns

In relation to web trackers, participants were questioned about their concerns regarding the privacy and security of their data from three primary angles: concerns about data collection, concerns about data usage

and disclosure, and concerns data storage and security. Eight variables, which are questions from 21 to 28, are used to accomplish this.

4.3.6.1.1 Data collection Concerns

The first item Question 21 aims to assess participants' concerns regarding the transparency of goals of data gathering their data. The majority of participants (54.7 %, n =257) believe that the goals of data collection by cookies and web trackers are at least partly transparent. By contrast, about (29.1%, n = 137) think that the goals of data collection are totally invisible. The second item Question 22 aims to assess participants' concerns by assessing the credibility of web trackers' use for data collecting. The majority of participants (47.7%, n =224) expressed concern about this and thought it was desirable that the information obtained through online cookies is directly tied to the purpose of the gathering. In contrast, only a very tiny percentage of participants (6.0 %, n = 28) express no interest at all in the authenticity of the purposes of the data collection.

4.3.6.1.2 Data use and disclosure Concerns

Concern about data usage and disclosure is a different aspect of measuring users' concerns about the privacy of their data in the context of cookies and web trackers. The results reported from Question 23 indicated that more than three-quarters of respondents are concerned about how third parties will access data and what they can do with it (77.4%, n = 364) where they selected ‘yes', concerned about third parties accessing the data and what they can do with it. While just a tiny proportion (22 %, n =106) are unconcerned with how their data may be accessed and used by third parties. Additionally, the significance of disclosing and supporting the legal justifications for data sharing with third parties was also brought up in Question 24 with the participants. The participant was required to determine how crucial it is to follow the legal justification for disclosing data to third parties. The results are similar among the participants who see that it is important to some extent and very important that the legal bases for sharing personal data be disclosed by (32%, n =149 and 33%, n = 155) respectively. Likewise, for those who believe that it is very important to justify disclosure and legally share data with third parties, at a rate of (27%, n =129). On the other hand, the minority of the participants expresses a lack of interest in the legal justification of disclosing data, with (5,5 %, n =26) understanding of them who do not care to some extent, and only (2%, n = 11) of them do not care at all. Otherwise, participants generally *(M* = 2) agreed that they were highly interested in knowing the legal justification for disclosing personal information to third parties.

4.3.6.1.3 Data storge and security Concerns

How personal data is stored and secured is another integral part of users' information privacy. Participants were asked about their concerns about the storage and security of their personal information gathered through web analytics and cookies. Regarding users' concerns about the storage of their data, participants

were asked in Question 25 about their concerns about where and how their personal data obtained through cookies is stored. The majority of participants (77 %, $n = 365$) admitted that they are concerned about the manner in which and the location of the storage of their personal information received through cookies. Concerns in Question 26 were also expressed over how long the information gathered should be kept on record, which is connected to concerns about data storage. How long web analysts should store their data is a concern for the vast majority of respondents (77%, $n = 365$). As for concerns about the security of users' personal data, participants were questioned through Question 27 about securing access to data and its consequences, such as cyber-attacks. Concerns about securing access to your data, such as being tracked for suspicious activity (e.g., browsing by employees) were prevalent among respondents (78.3% $n = 368$). Participants were also surveyed by Question 28 about the extent of concern about the security risks associated with the collection and storage of information by online cookies. Results showed that the average number of participants (M=3) are somewhat concerned about the security threats related to the online collecting of cookies and data storage.

4.3.6.2 Users' behavior

The aforementioned findings show that participants' concerns about the privacy of personal data in the context of online monitoring and cookies are rising. On the other side, the privacy paradox is the conclusion of several studies looking at users' privacy practices that have shown a disparity between what individuals claim about their privacy and how they actually behave it (Brown, 2001; Norberg et al., 2007). So, in light of the study's goals, the findings in this section evaluate how concerned users are with their online behavior and reading of cookie-related privacy policies. This is also designed to check and investigate any differences between what participants state and what they actually practice. Using a website run for the study to evaluate user behaviors during the opening of the screen to accept or reject the cookie policy, the participants were asked about their actions during the display of the cookie policy popup and whether they agreed or not.

4.3.6.2.1 Participants' privacy-related behaviors

The findings from Question 32 reveal that the respondents were nearly evenly distributed among those who read the site's cookie policy and those who did not (53.0 % and 47.0 %) respectively. The primary motivation for reading the study site's cookie policy was then elicited from the participants through Question 33. The major justifications given by participants for reading the cookie privacy policies are shown in Figure 4.12. According to the percentages in Figure 4.11, the primary variables that motivated the participants to read were their desire to know what type of cookies the website is using about (38%, n=94) and their concern about whether a third-party service was being used about (30%, $n = 74$). Approximately

(20% n =51) of people also want to know how the website utilizes their personal data. While just (12%, n =30) of users wanted to know what type of personal information the website acquired from them.

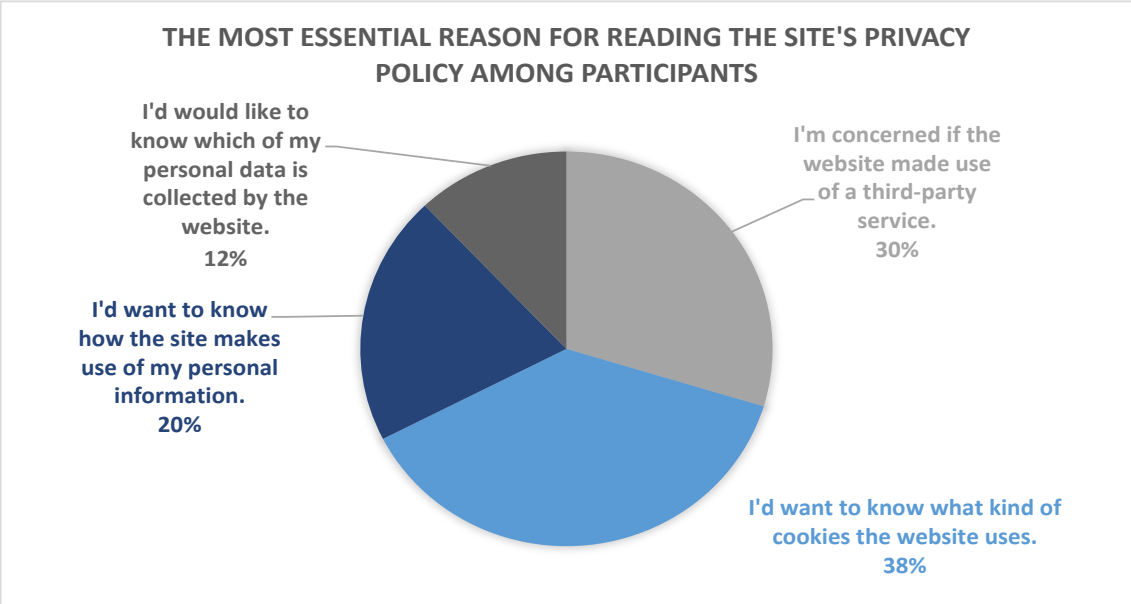


Figure 4.12 The participant's main reason for reading the website's cookie privacy policies.

A similar question applied to people who neglect to read the cookie policy through Question 34. The major justifications given by participants for not reading the cookie policy are shown in Figure 4.13. According to the percentages in Figure 4.13, participants' lack of reading the cookie policy is most frequently attributed to the fact that it takes a lot of time, with approximately (56 %, n=124) giving this as their reason. Following this are the (18%, n =39) of people who cannot grasp legalese. Some respondents also perceive that, by roughly (10%, n = 22) website surfing is more essential to them than privacy. Similarly, (10%, n=23) of those who consider that their privacy will not be protected by approval or rejection. While (6%, n=13) of people indicated that they don't care about content or privacy policies.

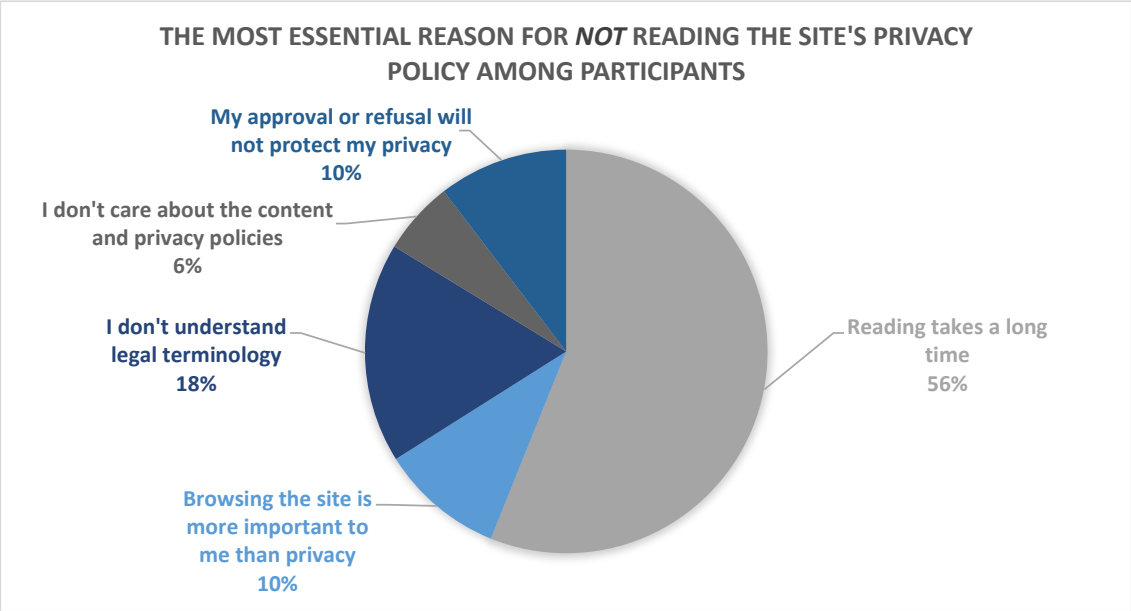


Figure 4.13 The participant's main reason for not reading the website's cookie privacy policies.

4.3.6.2.2 Participants privacy-related attitudes and behaviors

There are differences in users' beliefs and behaviors about privacy, according to a growing body of research on information privacy concerns (IPCs). The apparent dichotomy between attitudes and behavior concerning online privacy has been the subject of a number of research studies in various Internet contexts, analyzing and highlighting the phenomenon known as the privacy paradox (Hughes-Roberts, 2013; Kokolakis, 2017; Lee et al., 2013; Taddicken, 2014). As a result, the participants' experiences browsing the study website to show their behaviors while viewing the privacy statement for popped cookies sheds insight on any discrepancies between the users' attitudes and behaviors towards privacy. In order to identify any inconsistencies between users' actual behaviors and their attitudes regarding data privacy, this section sets out to perform an examination. Assuming three links exist between reading the cookie privacy policy and the level of concern users have about how personal data is collected, used, and secured. In the beginning, it was assessed whether there were any variations in the users' behavior toward reading the privacy policy and the level of concern about the collection of personal data. This is predicated on the assumption that user behavior and level of privacy concerns regarding data collection are mismatched. A Mann-Whitney U test, as illustrated in Figure 4.14, revealed that the level of concern regarding that data collection were distributed equally between behavior of the participants in both groups—those who read the cookie policy (M=23173, n=249) and those who did not read (M=239.74, n=221), U= 26576.500, p = .491, with a weak effect size r =26576.500. As a result, the results refute the existence of any privacy paradox among participates.

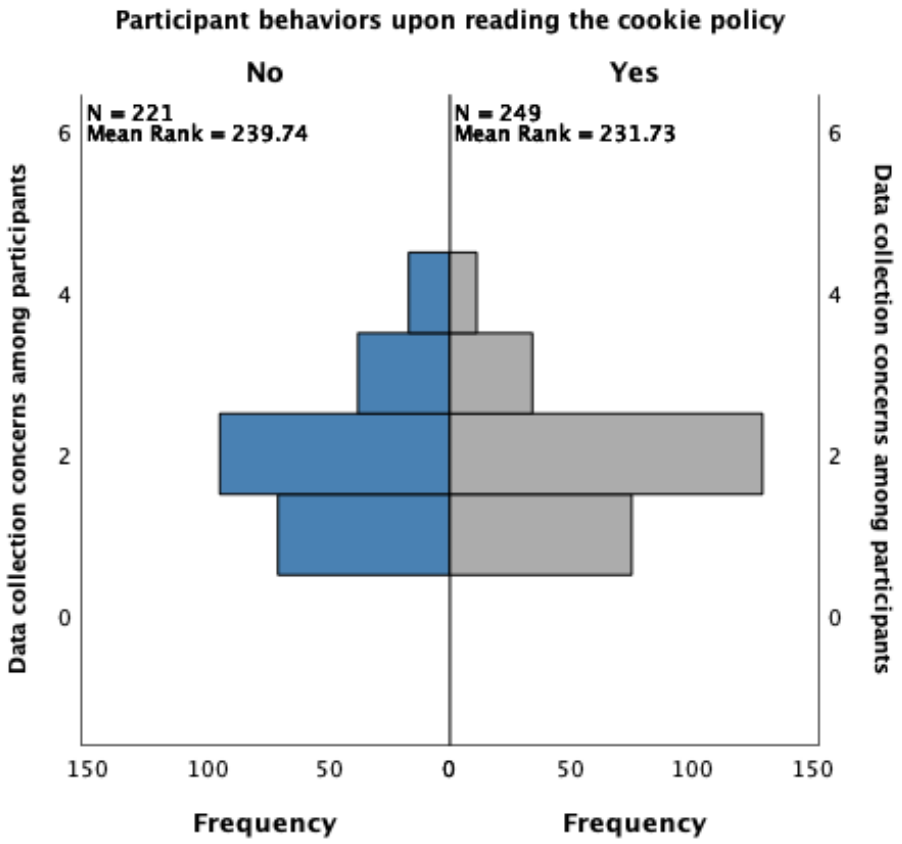


Figure 4.14 Distribution of Independent-Samples Mann-Whitney U Test of Participates who read the cookies policy and Participants who did not privacy read the cookies policy. Numbers on the X-axis indicate the frequency of the number of participants. Numbers on the Y-axis indicate the combined score of (Data collection concerns among participants).

The users' reading of the privacy policy and level of concern over the use of personal data were then examined to determine if there were any differences in actual behavior. This is predicated on the assumption that user behavior and level of privacy concerns regarding data usage are mismatched. A Mann-Whitney U test, as illustrated in Figure 4.15, revealed that the level of concern regarding that data usage was distributed equally between the behavior of the participants in both groups—those who read the cookie policy (M=241.70, n=249) and those who did not read (M=228,51 n=221), U= 29058.500, p = .271, with a weak effect size r =29058.500. The results, therefore, do not support the existence of any privacy paradox among participates.

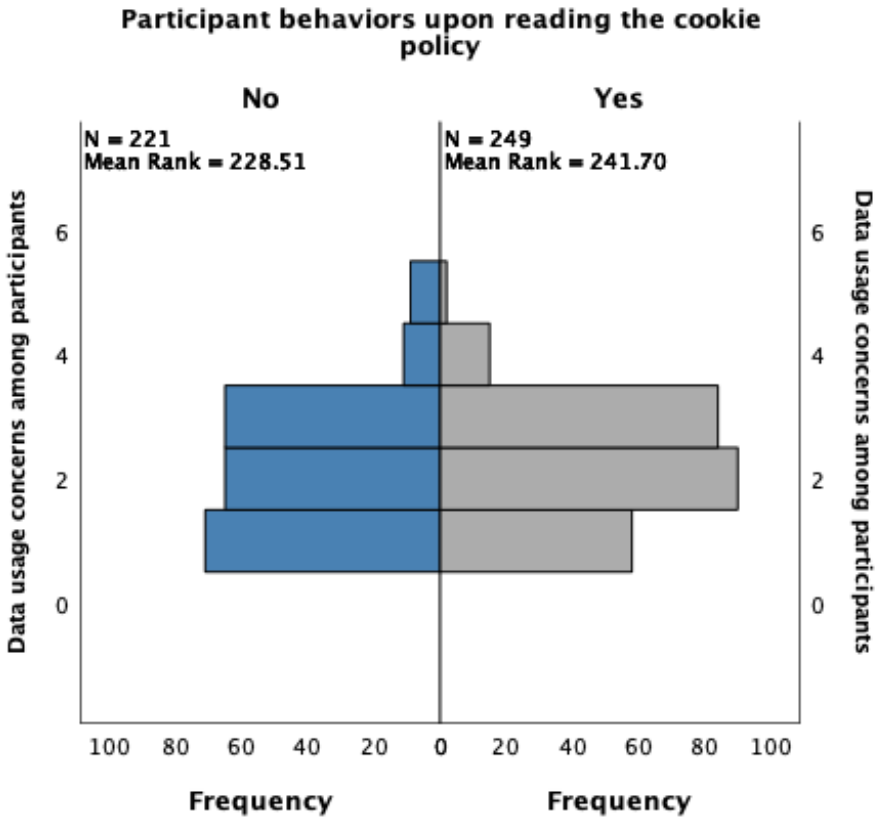


Figure 4.15 Distribution of Independent-Samples Mann-Whitney U Test of Participates who read the cookies policy and Participants who did not privacy read the cookies policy. Numbers on the X-axis indicate the frequency of the number of participants. Numbers on the Y-axis indicate the combined score of (Data usage concerns among participants).

Finally, the users' reading of the privacy policy and level of concern about the security threats of personal data weas assessed to determine whether there were any differences in users' behaviors. This is predicated on the assumption that user behaviors and level of privacy concerns regarding data security are mismatched. A Mann-Whitney U test, as illustrated in Figure 4.16, revealed that the level of concern regarding that data security was distributed equally between the behaviors of the participants in both groups—those who read the cookie policy (M=239.94, n=249) and those who did not read (M=230.50 n=221), U= 28619.000, p = .430, with a weak effect size r =228619.000. Thus, the findings thus disprove the existence of any privacy paradox among participates.

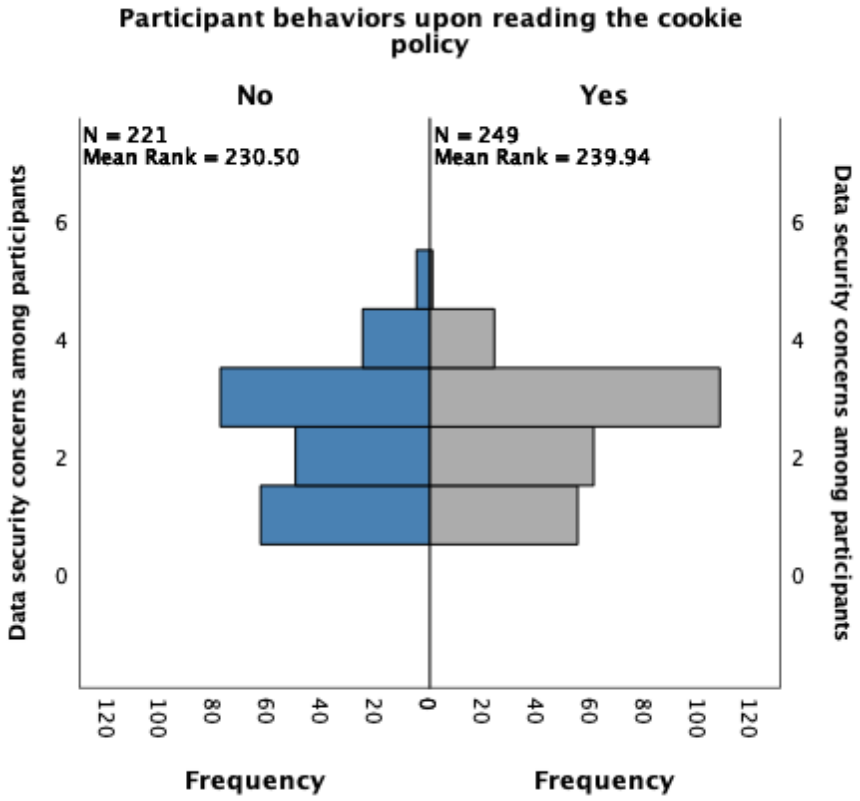


Figure 4.16 Distribution of Independent-Samples Mann-Whitney U Test of Participates who read the cookies policy and Participants who did not privacy read the cookies policy. Numbers on the X-axis indicate the frequency of the number of participants. Numbers on the Y-axis indicate the combined score of (Data security concerns among participants).

4.3.7 The statistically significant differences among participants' characteristics In regards to Web-Tracking Privacy Concern

4.3.7.1 Web-Tracking Privacy Concern In relation to Gender

Table 4.6 reveal that the chi-square test of independence shows there is a slightly significant difference between the gender particularly among the groups of male and female and the level of web tracking concern with $x^2(3, N = 470) = 7.910$, $p = .0481$, φ=.130. The value of the phi-coefficient was .130 (< .50) which indicated a small effect size. The results demonstrated that there is a slightly significant difference between the level of online tracking concern and the gender variable.

Table 4.6
Frequencies and Chi-Square Results for The Level of Concern and Gender Factor (N=470)

Privacy Concern of web tracking	Male		Female		Non-binary / third gender		Prefer not to say		$x^2(1)$
	n	%	*n*	%	*n*	%	*n*	%	
Yes	179	74.3%	183	81.7%	1	33.3%	1	50%	7.910***
No	62	25.7%	41	18.3%	2	66.7%	1	50%	

$***\rho < .001$

4.3.7.2 Web-Tracking Privacy Concern In relation to geographic region

The chi-square test of independence shows no significant association between the geographic region and the level of web tracking concern among participants with x^2 (12, N = 470) = 19.515, p = .077, φ=.204. The value of the phi-coefficient was .204 (< .50) which indicated a small effect size. Figure 4.17 illustrates the frequency of participant dispersion based on geographic range and privacy concerns with online tracking. The results demonstrated no relationship between the privacy concern of online tracking and geographic region among participants.

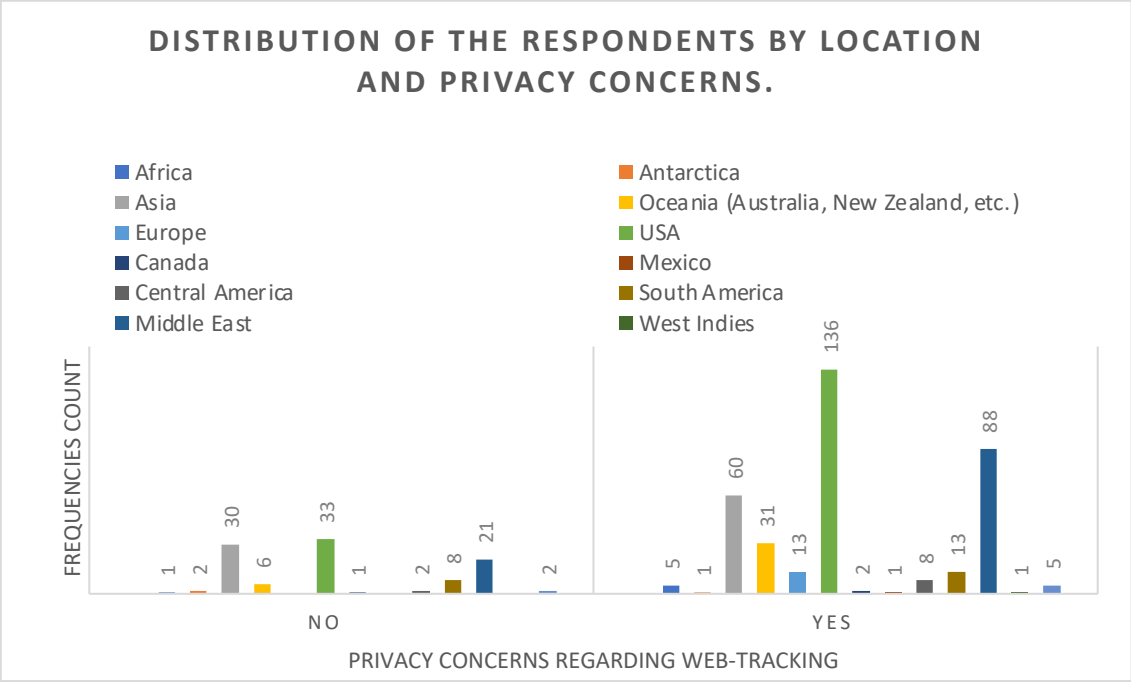


Figure 4.17 Distribution of responders based on geography and privacy concerns.

4.3.7.3 Web-Tracking Privacy Concern In relation to the level of education

A Kruskal-Wallis test showed that there are no statistically significant differences between the level of education and web-tracking privacy concerns, $H(4) = 1.557$, $p=.816$. Post-hoc Dunn's pairwise comparisons test, as illustrated in Figure 4.18, was carried out for the five pairs of groups. There were no significant differences between the level of education and the privacy concerns of web-tracking among participants.

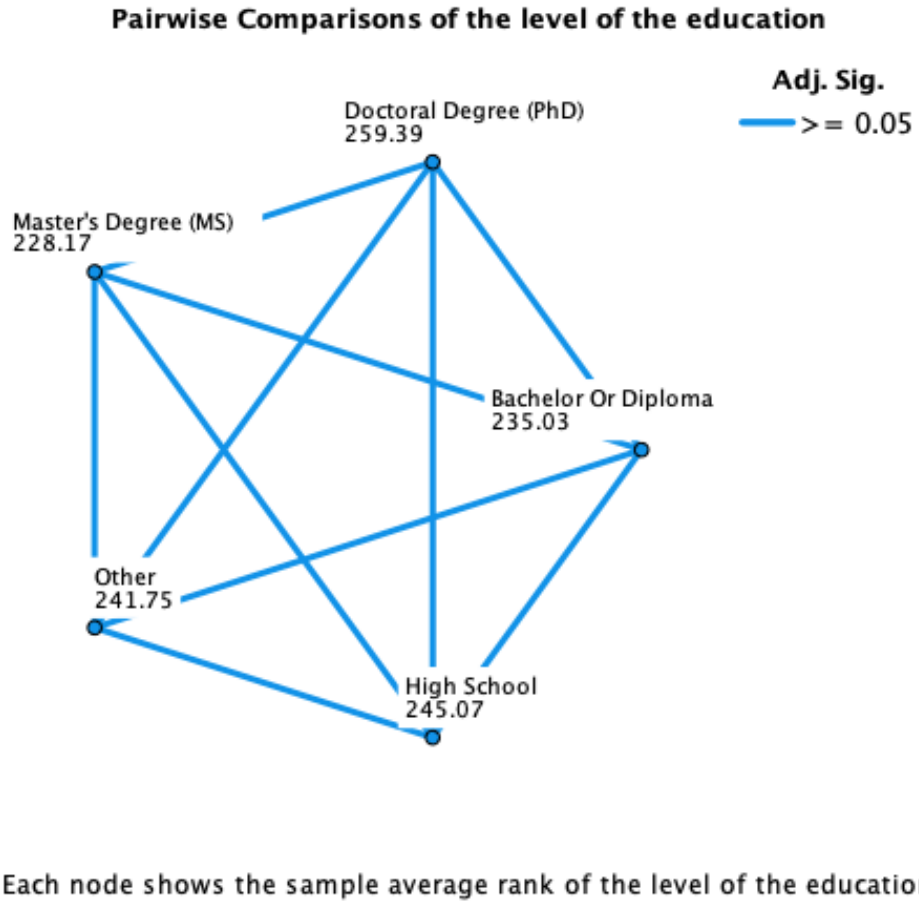


Figure 4.18 Post-hoc Dunn's pairwise comparisons test for the five pairs of groups and online tracking privacy concerns.

4.3.7.4 Web-Tracking Privacy Concern In relation to the technical background

A Kruskal-Wallis test showed that there is a statistically significant difference between the level of technical skills and web-tracking privacy concerns H (2) = 8.013, p=.018. Post-hoc Dunn's pairwise comparisons test, as illustrated in Figure 4.19, was carried out for the three pairs of groups. There were significant differences between the level of technical skills and the privacy concerns of web-tracking among participants.

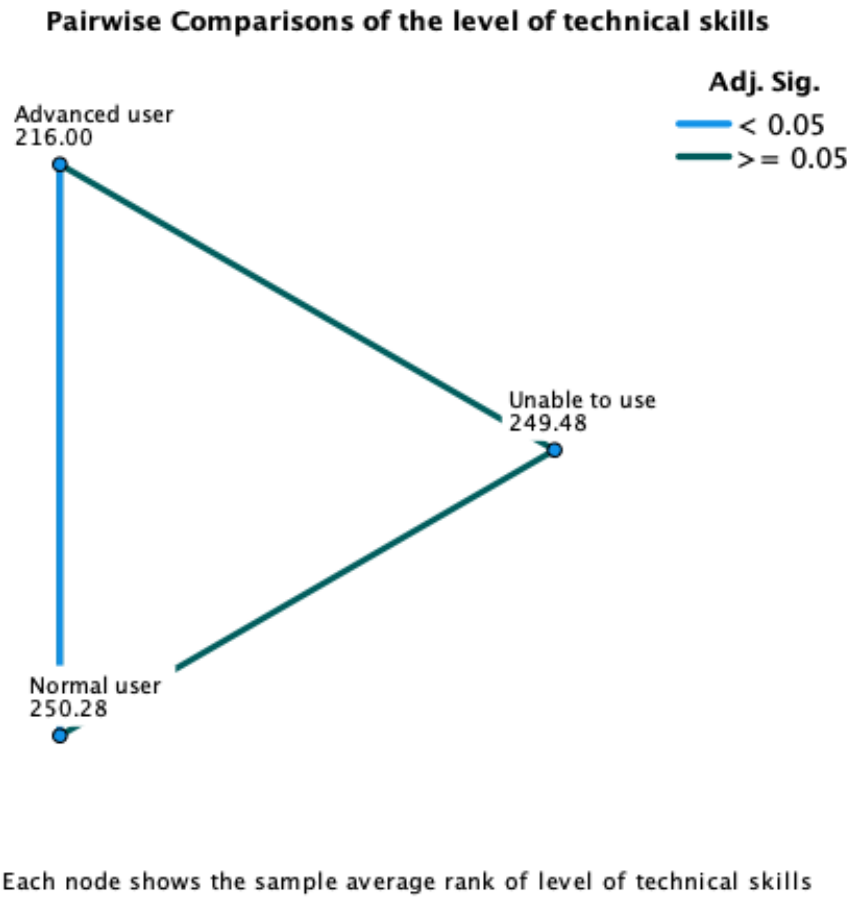


Figure 4.19 Post-hoc Dunn's pairwise comparisons test for the three pairs of groups and online tracking privacy concerns.

4.3.8 Principal findings regarding participants' levels of Privacy Concern

The findings and associated data throughout the last two sections aid in addressing the study's second research question. The second question aims to discover the level of concern people have regarding

security and privacy issues with cookies and web tracking. The findings from the questionnaire asking participants about their perceptions and opinions of privacy violations and security dangers brought on by cookies revealed that there is a significant level of concern among the respondents about the collecting, using, and storing of their data. Through the participants' experience of accessing website and viewing the privacy policy of cookies belonging to the study website, the results indicated that the level of concerns among the participants was similar, as the percentage of those who read the website's policy (53.0 %) was almost similar to those who did not (47.0 %). Recognizing what kind of cookies this website uses is one of the top reasons to read the privacy policy for cookies. However, the most common barrier for participants not reading the cookie privacy policy is the time required to read it. The analysis indicated that there were no clear statistical inconsistencies between the participants' attitudes regarding privacy concerns in terms of collecting, using, and storing their data, and their behaviors towards privacy. Regarding the levels of privacy concerns and the characteristics of the population sample in terms of educational level and geographic area, there were no statistically significant differences. On the other hand, the findings showed that there was a significant statistical difference between the gender and the level of privacy concern as well as participants' levels of technical background and privacy concern.

4.4 Conclusion

The findings of this an international Internet user survey are presented in Chapter 4 and show the extent to which users are concerned and aware of the privacy and security threats posed by web tracking and cookies. The survey confirmed that there is a comparable level of awareness and concern about privacy and security concerns with web tracking. The findings are covered in more detail in Chapter 5, which will also link results to the literature review and use the findings to address the study's core research questions.

CHAPTER 5: DISCUSSION

5.1 Introduction

The survey's findings on levels of awareness and concern about the consequences of cookies on data security and privacy are presented in chapter 4. The purpose of Chapter 5 is to analyse the results obtained in Chapter 4 and examine how they correspond to the security and privacy concerns raised by cookies and web trackers. These conclusions provide answers to research questions presented in Chapter 3. The primary research questions presented in Chapter 3 are addressed in Section 5.2, along with a discussion of whether the results support, contradict, or add to the pertinent prior research described in the literature review. With regard to the participants' levels of awareness and concern, Section 5.3 provides an interpretation of the statistical differences for a number of variables. In Section 5.4, the key conclusions of this study are further discussed. Finally, Section 5.5 discusses suggestions for improving user privacy and security within online tracking.

5.2 Research Questions

5.2.1 Question 1

Question 1 (Q1). *To what extent are computer users aware of the security and privacy threats associated with web cookies?*

Answer:

The findings show that while online users have a neutral position on the legal aspect, the majority were aware of the privacy and security risks posed by web trackers via cookies in terms of knowledge of online web tracking functions and privacy practices.

Discussion:

Three aspects were used to measure end-user awareness:

5.2.1.1 Users' privacy and security awareness

Section 4.3.3.1 found out that the majority of participants indicated awareness of their data being gathered and tracked using web trackers, which is compatible with existing literature on analysing awareness of information in the context of cookies. The findings of research to gauge users' comprehension of online tracking by Chanchary and Chiasson (2015), showed that about half of the participants were aware of the different tracking capabilities. Likewise, Narayanan (2020), assessed users' level of understanding of how cookies function and found that 40% of respondents said they were very familiar with cookie functionalities, while 30% said they were just somewhat aware. The survey by Kashi and Zavou (2020), also revealed that the majority of participants are aware that their online behaviour is being tracked.

Meanwhile, other research findings tend to support the opposite. The results of the survey according to Pinto et al., (2020), indicate that most people do not have enough information about online cookies as about 51.9%, are either wholly or partly unaware of web cookies. Information gathered by web tracking is referred to as personally identifiable information (PII) in the context of digital privacy (Hassan & Hijazi, 2017), as discussed in Chapter Two. According to the findings in section 4.3.3.1, many participants were unable to recognise all forms of personally identifiable information (PII) that had been obtained about them online, and very few were aware that their names and phone numbers were being collected. Despite being aware that their data is being gathered and tracked, participants were unable to accurately identify the various types of data collected. This is referred to as user disempowerment (Pierson & Heyman, 2011). The user in context of web tracking considered disempowered over their PII data if they are unable to recognise the data that has been acquired from them such as telephone number, name and E-mail addresses. There is less overlap between the apparent context and the whole context when utilising cookies as a method of collecting personal data since they usually hide the context for users (Pierson & Heyman, 2011). The findings in section 4.3.3.1 indicate that most participants were aware of the privacy and security risks associated with cookies and web tracking. The user understands that their online activities are being tracked, that their data being sold, and that data contained in cookies may be used in cyber-attacks. This level of end-user awareness may be attributed to the extensive usage of tracking across the Web, which arguably made the concept of tracking popular among users. For example, Binns et al. (2018) used 959,000 applications from Google Play shops in the US and UK to empirically highlight the proliferation of third-party trackers (Binns et al., 2018). Furthermore, tracking was not restricted to for-profit websites and services. It has recently expanded to include tracking information on hospital websites, notably with the COVID-19. According to McCoy et al. (2020), 89% of web sites linked to COVID-19 had a third-party cookie, and 99% of those pages contained a third-party data request. Comparatively, a prior investigation of 1 million popular web pages discovered that 91% of them featured a third-party data request and 70% of them contained a third-party cookie (McCoy et al., 2020). All of the hospital websites included in the analysis by Niforatos et al. (2021), employed ad trackers, and the majority of third-party cookies utilised and supplied user data to Facebook and Google. Compared to other websites in the healthcare industry, top USNWR Hospital websites employed more third-party ad tracking techniques and cookies (Niforatos et al., 2021).

5.2.1.2 Users' privacy practices

The findings in section 4.3.3.2 were corroborated by a number of cross-sectional studies looking at users' protective behaviour including (Büchi et al., 2017; Chiasson et al., 2018; Edith G. Smit et al., 2014) and showed that individuals typically make an effort to preserve their privacy online. There is, however, no

comprehensive explanation of the frequency of the potential ways individuals employ to preserve their online privacy since the studies evaluate diverse behaviours, and some do not offer descriptive statistics of the protective activities (Boerman et al., 2021). By giving descriptive statistics on user practice frequencies involving the readability, acceptance, modification and rejection of cookies privacy policies, this work adds to the body of knowledge addressing people's privacy practices in the context of web tracking and cookies. The findings showed that participants were interested in reading, approving, and changing websites' cookie rules from its default settings. End-user interest in altering and adjusting the cookie policies to suit their preferences suggests that users care about and are aware of the privacy of their personal information regarding web cookies. Also, findings showed participants' rejecting to utilize the services offered by the websites because of the sites' cookie privacy policies, which indicates that users are involved in and conscious of the privacy of their personal data on the webpages. Statistics on participants' usage of privacy-protecting anti-tracking technologies, such as Tor Browser and Privoxy, which enable users to retain their privacy and prevent online activity from being tracked, show that more than half of the respondents utilized anti-tracking tools. This result is consistent with that of Kashi and Zavou (2020), who indicated that 62.2% of users employ at least one of the common techniques to minimize tracking for example ad-blocking software or privacy-focused browsers. Accordingly, users' practises are considered to be solutions to improve their online privacy. Self-help tools that the user may employ to protect their privacy as indicated by Büchi et al. (2017) include the adoption of privacy improvement tools, cookie management solutions, and do-not-track technologies. Therefore, this demonstrates that individuals exercise their responsibility to protect their online privacy, indicating that users are aware of and interested in the privacy of their personal information on websites.

5.2.1.3 Laws and regulations

In evaluating the participants' understanding in terms of local or international privacy rights legislation, the results in section 4.3.3.3 showed that the average number of participants had an average level of understanding the regulations and laws in privacy rights on the Internet. This implies that, in terms of law, user awareness may not be adequate. Researchers observe in the area of literacy on online privacy on the measuring of awareness of information privacy It is crucial to take into account how familiar Internet users are with privacy-related rules and regulations (Prince et al., 2021). Users' knowledge of declarative privacy includes their comprehension of the laws or other legal elements of online data protection (Prince et al., 2021). This level of awareness in terms of the law and regulation can be linked to other elements that have been mentioned in several research with a focus on data protection. Becher and Benoliel (2021), study assessed the readability of cookie privacy policies on 300 of the most popular websites on the Internet. According to their results, despite compliance with the General Data Protection Regulation (GDPR), users

often confront privacy laws that are not totally understandable. As statements get longer and more standardized, users' comprehension of privacy declarations deteriorates (Dorfleitner et al., 2021). Due to the complexity of legal language and the difficulty in interpreting them, this may restrict the users' comprehension and awareness of the rules and regulations that protect their privacy. This calls into question why policymakers appear to assume that educated, empowered users are capable of making decisions that are in their best interests and why laws priorities user consent (Boerman et al., 2021). However, it is not apparent if people actually have the power and capacity to choose whether to provide their consent and then protect their online privacy once they have (Boerman et al., 2021). However, regarding knowledge of laws and regulations, the findings showed that participants had a neutral viewpoint on whether electronic service providers complied with national or international laws and regulations (GDPR/CCPA). Also, the cultural differences between the participants may be to blame for the participants' seeming skepticism regarding the service providers' adherence to rules and regulations (Wu et al., 2012). As in many nations, the legal framework does not adequately address contemporary data processing techniques (Baruh & Popescu, 2017; Zuiderveen Borgesius, 2014). Consequently, the lack of understanding of laws and regulations pertaining to online trackers may be linked to factors like the complexity of privacy rules and the cultural significance that laws and regulations play in different locations, where certain regions do not have laws governing cookies and web trackers.

5.2.2 Question 2

Question 2 (Q2). *How concerned is the computer user about the use of their data that is collected through web cookies and Internet use trackers?*

Answer:

The results showed that the respondents' level of concern about the gathering, utilising and storing of personal data via web tracking and cookies is significant.

Discussion:

User’s privacy concern was measured in two ways, as follows:

5.2.2.1 Users’ privacy and security concerns

5.2.2.1.1 Data collection Concerns

The results in section 4.3.6.1.1 show that in terms of concern about the data collection aspect, the majority of participants believed that the purposes of data collection by means of cookies are not clear. Also, the majority of participants believe that it is important for the purposes of sharing their data collected via cookies to be related to the purposes of the collection. This can indicate that users have concerns about the websites' transparency and credibility with regard to cookie-based data collection. It is also feasible to

suppose that concerns about the credibility of the purposes for which this data is shared may be connected to concerns about the transparency of the data collecting goals achieved via the use of cookies. Where the results converged among the participants regarding the concern about the transparency of the purposes of collection and the credibility of the purposes of sharing data with third parties. The empowerment of users over the privacy of their data is significantly aided by transparency in the context of cookies. Empowerment and freedom of choice are key to transparency (Laoutaris, 2018). Under the current technical paradigm, it is difficult to limit online leakage, but via transparency, such a goal may be attained on top of already-existing web technologies and business models without requiring some radical overhaul (Laoutaris, 2018). Transparency may be interpreted in the context of data protection and the privacy of online services as the capability to truthfully respond to questions like what information is gathered (stored and processed) about persons online? Who is the collector? The method of collection? What does it do? Has it leaked to any unauthorised parties? What effects may such Internet disclosure of private information have? (Laoutaris, 2018). Such questions comply to regional laws protecting information privacy, such as the New Zealand regulations. Thus, web tracking data collection may be considered a violation of privacy when required by local legislation, such as the New Zealand Privacy Act 2020, which governs data collecting. As, to comply with the principles in the New Zealand Privacy Act 2020 regarding data governance, under principle 1 purposes for collection of the Act, agencies are only allowed to collect personal information if the collection is lawful, the collection is connected with a function or activity of the agency, collecting that information is necessary to fulfil that function or activity (PrivacyCommissioner, n.d.-a). This is in accordance with the (GDPR), Art.6 which requires that websites gather data for lawful purposes (consulting, n.d.-b). According to the GDPR's Art.4 description of data processing, this includes activities such data collection, recording, structuring, storage, use, and disclosure. However, GDPR aims to protect only EU residents' personal information concerning cookies, while the NZ Privacy Act does not extend to the collecting of cookies data. Consequently, from a legal point of view, this may indicate that local privacy regulators must respond to users' concerns about cookies for the purposes of collecting personal data.

5.2.2.1.2 Data use and disclosure Concerns

The findings in section 4.3.6.1.2 show that with regard to concerns about data use and disclosure, the majority of respondents expressed concern about this aspect. Where a large number of participants expressed concern in terms of access to this data collected via cookies and its use by third parties. Also, it was important for the participants to know the legal justification for disclosing personal information to third parties. These results seem to be consistent with other research which found 79 % of American users are concerned about how businesses utilise personal data, in part because they are unaware of the data that businesses gather (Auxier et al., 2019). This implies that users' concerns about prevalent unauthorised access to their data may grow as a result of users' inability to manage who has access to their personal

information. According to Brooke et al. (2019), nearly half of Americans (48%) claim to feel powerless over who may access the search keywords they use. Users' personal information, including search preferences, names, and emails, are accessible to third parties. The privacy of Internet users may be violated by this access to and disclosure of data to third parties without clear legal explanations or basis. Thus, when mandated by regional law, such as the New Zealand Privacy Act 2020, which regulates data use and share, web tracking data usage and disclosure may be regarded as a violation of privacy. As, to comply with the principles in the New Zealand Privacy Act 2020 regarding data governance, under principle 10 Use of personal information, there are restrictions on the use of personal information for purposes other than those for which it was originally gathered by organisations (PrivacyCommissioner, n.d.-c). Also, in accordance to the data governance principles in the New Zealand Privacy Act 2020, under principle 11 of the Act Disclosing personal information, an organisation is typically only authorized to use or disclose personal information for the primary reason it was collected or gathered (PrivacyCommissioner, n.d.-d). This complies with the (GDPR), Art.6 requirement that websites process data only for legitimate reasons (consulting, n.d.-b). However, the GDPR only aims to regulate personal information belonging to EU individuals, whilst the NZ Privacy Act does not include the usage and disclosure of data obtained via cookies. Legal standpoint, this might imply that local privacy bodies need address users' concerns about cookies in terms of data usage and disclosure.

5.2.2.1.3 Data storge and security Concerns

The results presented in section 4.3.6.1.3 show that with regard to concerns about data storage and security, the majority of respondents expressed concern about this aspect. Most participants reported having concerns about how and where their personal information obtained through cookies is stored. Also, how long web tracker have to store their data is a concern of the vast majority of respondents. Concerns about securing access to the personal data, such as tracking suspicious activity (for example, browsing by employees) were prevalent among respondents. Additionally, there was a prevalent concern among participants about security threats related to the collection of cookies and online data storage. Participants generally tend to be concerned about the storage and security of personal data, which highlights the need of reacting to and contributing to address these concerns. As mentioned in Chapter Two, users may encounter security and privacy threats as a result of the data that is gathered and processed by cookies (Alexenko et al., 2010; Lin et al., 2009; Rodríguez et al., 2018; Rodríguez et al., 2020; Yadav & Parekh, 2017; Zheng et al., 2015). The cookies in extremely popular open-source services and websites may be injected including Google, Amazon, eBay, Apple, Bank of America, BitBucket, China Construction Bank, China Union-Pay, JD.com, phpMyAdmin and MediaWiki (Zheng et al., 2015). Many users may be exposed to many cyber-attacks associated with cookies, as it has been experimentally proven by a study conducted by Sivakorn et al. (2016), who demonstrated that the adversary can collect information that the victim uses to log in, such as

the victim's login, email address, and/or mobile phone number. The failure to provide adequate and suitable protection for user personal data obtained by cookies may have a number of negative effects, including usage for illegal and unethical purposes. Identity theft, social engineering attacks, and online physical surveillance are some these potential applications of the cookie-collected data (Malandrino et al., 2013). The information gathered via cookies is not regulated, subjected to regional or local regulations, and is not constrained by stringent international legal rules to ensure the adequate security for user's personal data. Even currently valid rules governing the data obtained by cookies, such as the European GDPR, nonetheless have issues with their practical and effective implementation. The issue with web tracking is that there is a lack of transparency from the stage of data collection through to its usage, storage, and security. With the involvement of the legislative bodies, the issue of cookies' lack of transparency may be addressed. Without regional or international regulatory limits that help to lessen the issues created by web trackers and help to preserve and provide more privacy for users online, the difficulties with web trackers that cause users to be concerned about their data privacy still persist. A fundamental need that complies among most data privacy laws is the implementation of adequate security protection on data gathered and stored via organizations. As, to comply with the principles in the New Zealand Privacy Act 2020 regarding data governance, under principle 5 Storage and security of information, organizations are required to put precautions in place that are appropriate given the circumstances to avoid the loss, misuse, or leakage of personal information (PrivacyCommissioner, n.d.-b). This complies with Art.5 of the General Data Protection Regulation (GDPR), which demands that the processing of personal data be done in a way that ensures appropriate security of the data, including protection against unauthorised or unlawful processing and against accidental loss, destruction, or damage using appropriate technical or organisational measures ('integrity and confidentiality') (consulting, n.d.-a). However, the NZ Privacy Act does not cover the security of data gathered via cookies, whilst GDPR only seeks to secure personal information pertaining to EU citizens. Accordingly, from a legal perspective, this may suggest that local privacy authorities must respond to users' concerns about cookies in terms of data storage and security.

5.2.2.2 Users' behaviours

5.2.2.2.1 Users 'privacy-related behaviours

The findings in section 4.3.6.2.1 demonstrate that there is no trend among participants in terms of who read the privacy policy and who did not, as the percentages were comparable in both groups when it came to user behaviour linked to the reading of cookies privacy policies. Recognizing what kinds of cookies, the website uses was one of the key motivations for reading the privacy policy. Additionally, reading the privacy statement was motivated by concerns regarding the usage of the website's third-party service. This

may imply that users' concerns about the usage of their data are a major consideration in the motivation for recognizing the kind of cookies and the use of third-party services. As the functionalities for utilising the data in which the information is kept are the key variations between the various sorts of cookies. Thus, in the trial, user behaviour in reading cookie privacy policies to know which type of cookies are being used and whether the website is used by third parties is more likely to be associated with privacy concerns in the context of web cookies regarding the usage of data. However, for those who did not read the cookie policy, there was a prevalent obstacle. Due to the lengthy reading duration, more than half of the participants did not select to read the website's cookie policy. Also, another notable obstacle to reading was the participants' inability to grasp legal terminology. Both of these considerations show that people disregard reading cookie privacy policies on websites because they are lengthy and difficult to comprehend. These results reflect those of Becher and Benoliel (2021), who discovered that while measuring the readability of cookies privacy policies on 300 of the most prominent websites in the world, users frequently encounter totally incomprehensible privacy policies. In a similar vein, Dorfleitner et al. (2021), who using a textual analysis approaches found that as privacy statement wording grew longer and more standardised, users' comprehension of them declined. So, it is evident from this that users' behaviour in ignoring the privacy policies of cookies is justified by their inability to read and comprehend them. Complicated reading and comprehension challenges may have further consequences, such as resulting in the acceptance of websites' cookie policies in their default settings. As indicated in chapter two Utz et al. (2019), found that the situation, options that are provided, prompting, and terminology all have an impact on people's consent behaviour. This is corroborated by early research by Acquisti and Grossklags (2005), that gathered survey data and validated the notion that psychological biases, limited rationality, and inadequate information all impact privacy decision-making. Accordingly, the length and terminology of websites' privacy policies may affect a user's ability to make rational, informed, self-serving decisions about providing consent and preserving privacy. This implies that the webpages must work to reduce the complexity of formulating policies and establish them in a simple and understandable manner in line with the knowledge background of the average individual in order to achieve the benefit of the privacy policies of cookies and to enable the user in their privacy on the web to make an informed and reasonable decisions.

5.2.2.2.2 Users 'privacy-related attitudes and behaviours

Contrary to expectations, this study did not find a significant difference between users' attitudes toward privacy concerns and users' behaviour in the context of web cookies. As the results indicated in section 4.3.6.2.2 when the Mann-Whitney U test was conducted to investigate the existence of inconsistencies between participants' attitudes regarding privacy concerns toward the collection, use, and security of their data and participants' behaviour toward privacy, the results refute any privacy paradox among participants.

These findings are unlike those of previous studies which showed that the existence of a privacy paradox phenomena in different contexts, such as e-commerce, SNSs, and Internet use. An attitude vs. behaviour dichotomy was confirmed by (Lee et al., 2013). Research concluded that users actively disclose personal information despite their concerns since they consider both the expected benefit and the threat of sharing. Hughes and Roberts (2013), came to the conclusion that a broad user concern is not a reliable indication of privacy behaviour within the network based on an analysis of a participant's Facebook accounts. Taddicken (2014), also demonstrated that self-disclosure is hardly affected by privacy concerns. In contrast to the results of this study, evidence from previous research shows that there is a privacy paradox present in many online contexts. However, the existence of the privacy paradox as a phenomenon is questioned by the present research. Supported by findings from multiple studies, this assertion that users' privacy attitudes and behaviours are different is brought into question. According to Boyles et al. (2012), found that 54 % application users have decided not to install a mobile phone application after discovering how much personal data they would have to reveal in order to use it and 30% of mobile application users have uninstalled an application that was already on their cell phone after realising it was gathering personal information they did not consent to. A favourable association between privacy concerns and protective behaviour is demonstrated by (Lutz & Strathoff, 2014). The study's findings, which were reported in section 4.3.3.2, which indicated that the majority of participants employ anti-tracking software is consistent with the findings in section 4.3.6.2.2 that disprove any disparities between users' opinions and behaviours regarding their privacy concerns. This demonstrates that users frequently adopt self-protective behaviour to secure their online privacy because they are concerned about the privacy risks provided by web trackers. Accordingly, this may be established on the argument that, compared to other contexts like SNSs and e-commerce, surveillance concerns caused by cookies may be more widespread among users and provide a stronger incentive for self-preservation measures to maintain privacy. That is evidenced by Fujs et al. (2019) who demonstrated that information sensitivity, regulation, and surveillance concerns all have a major impact on privacy concerns. On the basis of this, it is evident that privacy concerns are highly influenced by concerns about surveillance, which may contribute to minimize the inconsistencies between users' attitudes and behaviours. The lack of inconsistencies between users' privacy attitudes and behaviours in this study may be attributable to considerations related to web tracking, including a lack of transparency in web trackers and the absence of effective legislation that protects users' privacy rights and regulates data collection, use, and security via cookies. Thus, to develop a complete picture of the existence of the privacy paradox in the context of cookies and monitoring concerns, additional studies examining users' privacy attitudes and self-protective privacy behaviour will be required.

5.3 Statistical differences regarding web tracking Awareness and Concern

The findings showed that there were no statistically significant differences between the levels of awareness in section 4.3.4 and the level of concern in section 4.3.7 and the characteristics of the sample population in regard education, and geographical region. Results for the gender variable and awareness levels in section 4.3.4 showed no statistically significant differences, whereas findings for the gender factor and level of concern in section 4.3.7 showed significant statistical differences. On the other hand, the results revealed significant difference in the participants' levels of technical background, as well as the awareness and concern levels in sections 4.3.4 and 4.3.7 respectively.

5.3.1 Awareness and Concerns of Web-tracking in relation to Geographic Region, and Level of Education

The results in section 4.3.4.2 and in section 4.3.7.2 indicated that there were no statistically significant differences between geographic region and the levels of awareness and concerns among the participants. Although the research sample varied in terms of the geographic distribution of participants, this did not demonstrate a difference between participants from regions that regulate data collected through cookies, like the United States and Europe, and participants from regions that do not include their local laws to protect information privacy, like the Middle East. The findings of this study did not find any evidence of a regional effect on users' awareness or concern levels, based on the assumption that several studies in the area of privacy demonstrate that level of privacy awareness and concern vary between cultures (Bellman et al., 2004; Milberg et al., 2000; Miltgen & Peyrat-Guillard, 2014; Taddicken, 2014). This may imply that web tracking might become a prevalent problem and concern among users in many communities and regions; it could raise a lot of questions about the privacy and security threats that web tracking represents. The data that is monitored and gathered via cookies is also likely to cause greater thoughts and concerns among users in communities lacking local regulations or policies than among users in regions with such laws.

The results in section 4.3.4.3 and in section 4.3.7.3 indicated that there were no statistically significant differences between level of education and the levels of awareness and concerns among the participants. A factor that could have contributed to the lack of a statistical difference is the indication that the great majority of the participants in this research sample hold a university degree in various stages ranging from bachelor's to PhD holder. As indicated by Boerman et al. (2021), the effect of education related to the level of awareness that affects the levels of concern, thus making the user able to adopt self-protective methods to protect his privacy. Also, other studies in the field of information privacy confirmed that education has

an important role in the levels of knowledge and privacy behaviors of users (Baruh et al., 2017; Edith G Smit et al., 2014). Therefore, considering that the majority of participants in this research are educated, this may assist in understanding the high levels of awareness in this study. This in turn brought about a significant level of concern to be expressed by the study's findings on the security and privacy risks associated with cookies.

5.3.2 Awareness and Concerns of Web-tracking in relation to Gender

The results in section 4.3.4.1 showed that there were no statistically significant differences between gender and the levels of awareness among the participants. The data demonstrated that gender does not greatly influence one's level of awareness of the privacy and security issues posed by cookies. It is conceivable that this is because everyone is familiar with cookies because they're common on websites and because using the web is typically a routine that doesn't dominate a single gender. This result, however, is in contrast to other research that revealed a significant gender difference, with males consider to be more likely to be aware of cookies than women (Pinto et al., 2020). Considering that the sample size in this study is four times larger than the sample size in the study by Pinto et al. (2020), it is plausible that the difference is the result of statistical discrepancies arising from the disparities in sample sizes. Despite this, early research on the impact of gender and technology usage differences in real and perceived online skill levels supports the findings of this study, which found that, generally, online skill levels between men and women are not significantly different (Hargittai & Shafer, 2006). On the other hand, the findings in section 4.3.7.1 show that there are statistical differences in the correlation between concern levels and gender. It shows that women are likely to be more concerned than men about web trackers and third-party access to data. This result confirms the findings of other research on the context of privacy, such as those by Baruh et al. (2017), and Hoy and Milne (2010), which found that women tend to be more concerned about their online privacy. As indicated by Chai et al. (2009), Hoy and Milne (2010), women are more likely than men to practice privacy protection behaviours. This may imply that the adoption of privacy-protecting behaviours spurred by privacy concerns may also be significantly influenced by gender.

5.3.3 Awareness and Concerns of Web-tracking in relation to Technical Background

The results in section 4.3.4.4 and in section 4.3.7.4 showed that there were a statistically significant differences between the technical background and the levels of awareness and concerns among the participants. This difference may suggest that users with more advanced skills are more aware of cookie

tracking, which in turn suggests that users with advanced technical and normal technical skills have different levels of concern from users with weak technical skills. The user's technical experience may have a significant role in their level of comprehension the web architecture practically, which assists them better comprehend the technological framework in which web tracking and cookies handle user data and pose threats to data security and privacy. The importance of the user's technical knowledge and awareness of privacy threats has been supported by other studies in the same subject. Individuals who are more skilled online, more technically savvy, and knowledgeable of online behavioral advertisement are more likely to practice privacy protection (Baruh et al., 2017; Büchi et al., 2017; Ham & Nelson, 2016). This implies that the level of informational awareness of the security and privacy risks caused by web trackers reflects significantly on the user's technological background. Additionally, the user's technological background, which fosters knowledge and increases awareness of privacy and security on the web, imposes higher levels of concern about cookies and web trackers violating privacy.

5.4 Web-tracking in the context of online privacy literacy (OPL) and concerns (OPC)

The concept of privacy on the Internet is a broadening one that encompasses a bunch of concepts that influence people's levels of awareness and concern as well as the behaviours that follow within the restrictions of privacy. The term "online privacy literacy" (OPL) refers to a contemporary concept that concerns the privacy of personal data. OPL stands for online users' privacy-related knowledge and skills (Masur, 2020). According to Trepte et al. (2015), interpretation, "online privacy literacy" is "a combination of factual or declarative ("knowing that") and procedural ("knowing how") knowledge regarding online privacy." Users' privacy concerns, attitudes, and behaviours are significantly influenced by their level of online privacy literacy. The differences between users' privacy attitudes and behaviours have frequently been linked to privacy literacy as a proposed solution (Trepte et al., 2015).

One study indicated that individuals with significant privacy concerns and high trust in their ability to protect it have greater privacy literacy levels (Weinberger et al., 2017). Schubert et al. (2022), findings suggested that highly concerned individuals with high levels of privacy literacy appear to match their concerns with appropriate protective measures. Privacy literacy could provide individuals with the ability to challenge existing social norms and explore pathways for societal change toward more positive approaches of privacy, in addition to empowering them to defend themselves against unauthorised identity or access (Masur, 2020). Thus, users' concerns regarding threats presented by web trackers to the privacy

and security of personal data are minimized by strengthening privacy literacy. Additionally, privacy literacy encourages users to employ methods that increase online privacy.

5.5 Enhancing User Security and Privacy in Web Surveillance

5.5.1 End-user privacy control

One strategy to enhance the level of privacy in the context of online tracking is to adopt self-protection for the user in cyberspace via the use of the suggested tools and procedures. Self-help is a sufficient defence against privacy threats, based on a conceptual model of governance selection that takes into account contextual factors of governance such as incentives, conflicts of interest or intervention capacity (Büchi et al., 2017). End-user privacy protection behavior is varied and dependent on a variety of factors, including user concerns and threat propensity, digital literacy, and experience (Ebbers, 2019). Studies have found that using self-protection techniques to preserve individual's online privacy reduces the level of privacy concern. According to Chen and Chen (2015), managing privacy can overcome privacy concerns, especially for individuals who had previously had little privacy concerns. The protective privacy strategies include the use of cryptography and anonymization tools, VPNs, proxy and browser plugins that manage cookies or block tracking servers (Kashi & Zavou, 2020; Matzner et al., 2016; Prince et al., 2021; Rainie et al., 2013). However, the tools that are offered and proposed might offer a fraction of privacy protection from being tracked across the web. These technologies often work well when a stateful tracking is being used. However, the obstacles of evading tracking are becoming more difficult in the case of stateless tracking (Kashi & Zavou, 2020). Even with the adoption of the tools and means available to avoid tracking, these techniques provide a bigger difficulty, and individuals are unable to adequately defend themselves. Recently, monitoring through Bluetooth and GPS have gained popularity in light of the widespread popularity of mobile and IoT devices (Fawaz et al., 2016; Kirkpatrick, 2020). This demonstrates the tracking industry's rapid technological advancements, growing complexity, and commercial expansion, all of which might lead to a rise in privacy concerns. The third-party advertising technology ecosystem has expanded to incorporate increasingly complex partnerships between several third parties (Wambach & Bräunlich, 2017). Certain third parties are masters at connecting Internet users to the offline profiles that marketing firms hold of them such as LiveRamp (Binns & Bietti, 2020). Therefore, despite offering an acceptable level of protection, the concept of user self-protection will need to be developed more carefully due to the lack of standard and efficient rules and regulations governing cookies to protect users' privacy.

5.5.2 Legal restrictions

The success of the legislative efforts in mandating websites to regulate cookie-based data gathering and allowing end users more control over their data is observed. According to Degeling et al. (2018), 62.1% of websites in Europe now have cookie consent policies, up 16% from before the GDPR took force. Also, the average number of third parties decreased by more than 10% after GDPR Hu and Sastry (2019), it shows that the law had an impact on the rate of third-party cookies shortly following its implementation. It is important to highlight that, although the GDPR recently amended tighter guidelines for user privacy protection, personal information is still in risk (Bornschein et al., 2020). This observes that even if such regulations exist in some nations of the globe, they still have problems protecting users' privacy. According to the empirical findings of Prince et al. (2021), the large percentage of cookie notifications are either not visible to users or do not provide users with a choice regarding data collection practices, indicating the need for strengthened regulations regarding personal information privacy practices. The special concerns for privacy and basic rights that are raised by acquisitions and mergers between companies engaged in third-party tracking are frequently overlooked in governmental decisions and scholarly debates of data and market concentration (Binns & Bietti, 2020). The privacy regulations in various jurisdictions are mostly different (Centeno, 2016). The perspectives of the US versus with that of the EU, Argentina, and Canada on privacy are some illustrations of differences (Centeno, 2016). There is no universally recognized definition or standard for data privacy in the digital environment, and there are no comprehensive, legally enforceable multilateral agreements that address privacy and cross-border data flows (Fefer, 2019). Due of the various manners that information is disseminated, there is an expanding issue. Thus, on the local or international levels, it is possible to argue that more effective legal involvement and uniform standards will help regulate the personal data collected by cookies and cookie tracking, boost Internet privacy, and alleviate people's concerns about their privacy in terms of data collection, use, and storage.

5.6 Conclusion

The main research questions raised by this study were addressed in chapter 5, which also provided a discussion of the data that were initially presented in chapter 4. The findings indicated a high level of awareness and concern regarding online tracking. This chapter examined some of the variables that might affect an individual’s levels of concern and awareness. Further, the significance of personal privacy awareness and comprehension has been made apparent, which has an impact on establishing guidelines for online privacy. Finally, the concerns posed by online tracking to data privacy and security have been highlighted with some suggested measures that may be pertinent to both end users and policymakers.

CHAPTER 6: CONCLUSION

In Chapter 1, the research topic linked to the security and privacy threats that online trackers pose to the user was introduced; the thesis structure was explained, and the purpose of the study was described. The literature review in Chapter 2 outlined the characteristics that contributed to the evolution and growth of online trackers, its consequences on user data security and associated implications for users' privacy under the New Zealand Information Privacy Act 2020. The literature review, also, highlighted the user's information awareness of the privacy and security of their data, as well as the end-user's concerns associated with online privacy. The challenges and concerns presented in Chapter 2 served as the basis for the research questions for this study, which concentrate on user awareness and concern about the threats posed by online tracking to data security and privacy. Relevant studies were reviewed in order to develop an appropriate study design for addressing the research questions. In Chapter 3, a survey questionnaire was presented as the approach of research design employed in this study. In Chapter 4, the survey findings for this study were reported. These findings revealed that individuals all across the world had a high level of awareness and concern about web tracking. The findings were further evaluated and discussed in Chapter 5 in order to connect findings to the literature review and, most essential, to provide answers to the study's main questions. Additionally, suggestions were provided on how to help protect users' privacy online from the risks of tracking that may be carried out by regulatory entities and end users. The research is summarized in this chapter's conclusion, along with the research contributions to the fields of information privacy awareness (IPA), information security awareness (ISA), and information privacy concern (IPC) in the context of online tracking. This chapter also makes recommendations for future research.

6.1 Summary of Research

The goal of this study was to examine users' levels of awareness and concern regarding the threats that web tracking brings to the security and privacy of users' personal information. The focus of this study was on adult users among all age categories and countries around the world. Examining end-user awareness of the privacy and security risks posed by web tracking and cookies was the first objective of this thesis. The results of this study showed that, with regard to the level of end-user awareness of the privacy and security risks posed by cookies and web trackers, the majority of users were aware of the privacy and security risks posed by web trackers via cookies in terms of knowledge of online web tracking functions and privacy practices; while, from a legal aspect, most people have a neutral stance regarding privacy regulation understanding and functionality. The findings demonstrated statistical differences in the users' levels of awareness and technical skills. However, there was no statistically significant difference between the users'

levels of awareness and the factors of gender, education, and culture. The level of end-user privacy concerns relating to web tracking and cookies was the second focus of this thesis. The study's findings, which evaluate end-users' privacy and security concerns imposed by cookies and web tracking, revealed that end-user privacy concerns were high significantly. Users were concerned about the collection, usage, and storage of personal data via cookies and web tracking. Statistics revealed differences in the association between privacy concern levels and the gender factor, with women being more concerned about privacy. As with the association between privacy concern levels and user technical backgrounds, the data showed statistical differences.

However, there was no association between users' levels of privacy concern and either cultural or educational factors. Finally, this study demonstrated how users have serious privacy concerns when it comes to cookies and online tracking since users' behaviours are consistent with their concerned attitudes regarding privacy. The data that is gathered, processed, and stored via cookies is not governed by local laws that are typically concerned with protecting the privacy of information, such as the New Zealand Privacy Act 2020. Although, there are laws and regulations governing online tracking in various nations, including Europe and America, research has demonstrated that the privacy of cross-border data flows is not covered by any comprehensive international regulations (Fefer, 2019), calling for the development of an international framework of uniform standards.

6.2 Research Methods and Limitations

An appropriate method for gathering information on end-users' awareness and concerns about the risks of web tracking was the implementation of an anonymous online questionnaire involving the study's experimental website. This study used a non-probability sampling technique, namely a voluntary response sample. The approach followed to collect the data from the targeted sample, however, resulted in a low overall response rate. The participation rate in online surveys is a significant challenge, as Nayak and Narayan (2019) indicated, as response rates are frequently low when compared to the offline survey technique. Likewise, despite that the majority of respondents showed a high level of awareness that contributed to the development of novel insights, the sample technique made it harder to defend making generalisations about those outcomes for the whole population. Even though the survey was designed for adult end-users worldwide, due to time limitations, it was only published in English language. Non-native English speakers who responded may have had some difficulty understanding the question and providing appropriate answers (Wenz et al., 2020). Despite gathering user demographics, this study disregarded age as a component that may affect concern and awareness levels. Additionally, this study did not take into consideration other demographic parameters that may be evaluated to determine links and relationships

between these variables and the levels of awareness and concern, such as the participants' occupations, work sectors, and income levels. There are restrictions to information gathering through self-reports. Self-reported measures of behaviour can be inaccurate, as participant responses are subject to cognitive, social and communication biases (Parry et al., 2021). So, participants' self-reported data on privacy practices, such as changing cookie privacy settings, can be susceptible to various biases and limitations. Despite these limitations, the study evidently advances current understanding of the user's frequently privacy practices, such as setting cookie privacy policies as well as using protective privacy techniques to avoid web tracking. Although the current investigation into the paradox of privacy was based on a single user behaviour, the findings showed that there was no inconsistency between attitudes and behaviour. This cannot, however, be generalized and requires deeper research in the context of cookies and web tracking. In spite of a significant amount of research on the privacy paradox, one may counter that it is still an open question (Kokolakis, 2017). Finally, some factors, such as the importance and relevance of the website's content to the user, were not taken into account when designing the website for this study.

6.3 Recommendations and Contributions

This study shed light on online tracking and cookies from a range of perspectives, including information privacy awareness (IPA), information security awareness (ISA), and information privacy concerns (IPC) among adult Internet users in an international level. This research expands current knowledge of digital data privacy and illustrates the extent of users' awareness and privacy concerns associated with web tracking. The research found that awareness levels of the security and privacy issues of web trackers were comparable to levels of privacy concern related to privacy violations by web trackers. The analysis showed correlations between privacy awareness and concern levels and the technical background of the user. The study also revealed variations in privacy concern levels among users with the gender factor. These findings provide the following insights for future research on online privacy literacy (OPL) with respect to web tracking in order to better understand how it affects user privacy concerns (IPC). Likewise, further study on a privacy-enhancing technology (PET) in regard to online tracking might be carried out to investigate its effects on the information privacy concern (IPC) for users. However, the study's evidence-based findings revealed that the length of cookie policies, which take a lot of time to read, is the biggest barrier for users not to read the privacy policies of cookies on websites. Hence, the study recommends that website owners should empower users to make informed choices over personal privacy while approving cookie policies by minimizing and/or simplifying existing cookies policies so that users can comprehend it. Finally, from the perspective of legal consideration, the study's outcomes have a number of important implications for future practise. The local or national data privacy regulation must take into consideration the privacy and security

concerns with regard to the data gathered, used, and kept via cookies in order to address end-user concerns that have been raised internationally and have been highlighted in this research. The data collected through web tracking and cookies is not governed by local legislation, such as the New Zealand privacy Act of 2020, which represented as the study's foundation. As a result, in order to address end-user privacy concerns, a uniform framework that regulates the data collection, use, and storage via web surveillance tools, such as web cookies, must be developed.

6.4 Future Research

There are a number of possible related subjects for additional study that, if addressed, might further the comprehension of the level of privacy awareness and concern regarding web tracking and cookie challenges. The concept of privacy is a highly contextual phenomena should be taken into consideration in both survey and experimental research (Kokolakis, 2017). In further research, it will be important to focus closely on the impacts of demographic factors and privacy awareness levels (IPA) as well as the impacts of demographic factors on the level of privacy concern (OPC). A greater focus on a specific culture could produce interesting findings that account more for the impact of culture on privacy awareness and concerns regarding web tracking. Furthermore, to determine the effect of privacy level of awareness and concern in user behaviours regarding web tracking, more modelling work will need to be undertaken using advanced statistical analysis such as PLS-SEM. Participants' self-reported findings regarding privacy practices, such adjusting cookie privacy settings, might be biased. Therefore, rather than self-reports of more objective results, future studies should employ evidence of real behaviors in privacy practices in the context of cookies. Investigations might be carried out in actual environments that offer a rich and relevant background. Moreover, a deeper comprehension of the privacy paradox can open up a new viewpoint for looking at the ethical and legal foundations of information privacy (Kokolakis, 2017). For a better understanding, comparative future studies may conduct a deeper analysis into the privacy paradox phenomena linked to cookies on the web.

References

Abgrall, E., Traon, Y. L., Monperrus, M., Gombault, S., Heiderich, M., & Ribault, A. (2012). XSS-FP: Browser fingerprinting using HTML parser quirks. arXiv preprint arXiv:1211.481. https://doi.org/10.48550/arXiv.1211.4812

Abikoye, O. C., Abubakar, A., Dokoro, A. H., Akande, O. N., & Kayode, A. A. (2020). A novel technique to prevent SQL injection and cross-site scripting attacks using Knuth-Morris-Pratt string match algorithm. *EURASIP Journal on Information Security*, *2020*(1), 1-14. https://doi.org/10.1186/s13635-020-00113-y

Acar, G., Eubank, C., Englehardt, S., Juarez, M., Narayanan, A., & Diaz, C. (2014). The Web Never Forgets: Persistent Tracking Mechanisms in the Wild. *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, Scottsdale, Arizona, USA*(pp.697-689). Association for Computing Machinery. https://doi.org/10.1145/2660267.2660347

Acar, G., Juarez, M., Nikiforakis, N., Diaz, C., Gürses, S., Piessens, F., & Preneel, B. (2013). FPDetective: dusting the web for fingerprinters. *Proceedings of the 2013 ACM SIGSAC conference on Computer & communications security* (pp. 1129–1140). Association for Computing Machinery. https://doi.org/10.1145/2508859.2516674

Acar, M. G. C., Diaz, C., & Preneel, B. (2017). *Online Tracking Technologies and Web Privacy* [Doctoral thesis, Princeton University]. KU Leuven. https://lirias.kuleuven.be/1662331?limo=0

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. *Science, 347*(6221), 509-514. https://doi.org/doi:10.1126/science.aaa1465

Acquisti, A., & Grossklags, J. (2005). Privacy and rationality in individual decision making. *IEEE Security & Privacy, 3*(1), 26-33. https://doi.org/10.1109/MSP.2005.22

Adjerid, I., Peer, E., & Acquisti, A. (2018). Beyond the privacy paradox: Objective versus relative risk in privacy decision making. *MIS Quarterly, 42*(2), 465-488. http://doi.org/10.25300/MISQ/2018/14316

AdobeFlash. (n.d). *ActionScript® 3.0 Reference for the Adobe® Flash® Platform*. ActionScript. https://help.adobe.com/en_US/FlashPlatform/reference/actionscript/3/flash/system/Capabilities.html

Alaca, F., & Van Oorschot, P. C. (2016). Device fingerprinting for augmenting web authentication: classification and analysis of methods. *Proceedings of the 32nd annual conference on computer security applications* (pp. 289–301). Association for Computing Machinery. https://doi.org/10.1145/2991079.2991091

Alexenko, T., Jenne, M., Roy, S. D., & Zeng, W. (2010). Cross-Site Request Forgery: Attack and Defense. *2010 7th IEEE Consumer Communications and Networking Conference (pp. 1-2)*. https://doi.org/10.1109/CCNC.2010.5421782

Altaweel, I., Good, N., & Hoofnagle, C. J. (2015). Web privacy census. *Technology Science, 2015121502*. https://ssrn.com/abstract=2703814

Amarasekara, B. R., Mathrani, A., & Scogings, C. (2020). Crookies: Tampering With Cookies to Defraud E-Marketing. In *Encyclopedia of Criminal Activities and the Deep Web* (pp. 1062-1076). IGI Global. https://doi.org/10.4018/978-1-5225-9715-5.ch073

Auxier, B., Rainie, L., Anderson, M., Perrin, A., Kumar, M., & Turner, E. (2019). *Americans and privacy: Concerned, confused and feeling lack of control over their personal information*. https://policycommons.net/artifacts/616499/americans-and-privacy/

Ayenson, M. D., Wambach, D. J., Soltani, A., Good, N., & Hoofnagle, C. J. (2011). Flash cookies and privacy II: Now with HTML5 and ETag respawning. *SSRN 1898390*. http://dx.doi.org/10.2139/ssrn.1898390

Baitha, A. K., & Vinod, S. (2018). Session hijacking and prevention technique. *International Journal of Engineering & Technology, 7*(2.6), 193-198. https://doi.org/10.14419/ijet.v7i2.6.10566

Barnes, S. B. (2006). A privacy paradox: Social networking in the United States. *First Monday 11*(9). https://doi.org/10.5210/fm.v11i9.1394

Barth, A., Jackson, C., & Mitchell, J. C. (2008). Robust defenses for cross-site request forgery. *Proceedings of the 15th ACM conference on Computer and communications security*(pp. 75-88). Association for Computing Machinery. https://doi.org/10.1145/1455770.1455782

Baruh, L., & Popescu, M. (2017). Big data analytics and the limits of privacy self-management. *New media & society, 19*(4), 579-596. https://doi.org/10.1177/1461444815614001

Baruh, L., Secinti, E., & Cemalcilar, Z. (2017). Online privacy concerns and privacy management: A meta-analytical review. *Journal of Communication, 67*(1), 26-53. https://doi.org/10.1111/jcom.12276

Baumer, D. L., Earp, J. B., & Poindexter, J. (2004). Internet privacy law: a comparison between the United States and the European Union. *Computers & Security, 23*(5), 400-412. https://doi.org/10.1016/j.cose.2003.11.001

Becher, S. I., & Benoliel, U. (2021). Law in books and law in action: the readability of privacy policies and the gdpr. In *Consumer Law and Economics* (pp. 179-204). Springer. https://doi.org/10.1007/978-3-030-49028-7_9

Bellman, S., Johnson, E. J., Kobrin, S. J., & Lohse, G. L. (2002). Regional Differences in Privacy Preferences: Implications for the Globalization of Electronic Commerce. *NY: Columbia University, The Columbia Center for Excellence in E-Business.*

Bellman, S., Johnson, E. J., Kobrin, S. J., & Lohse, G. L. (2004). International Differences in Information Privacy Concerns: A Global Survey of Consumers. *The Information Society, 20*(5), 313-324. https://doi.org/10.1080/01972240490507956

Bergström, A. (2015). Online privacy concerns: A broad approach to understanding the concerns of different groups for different uses. *Computers in Human Behavior, 53*, 419-426. https://doi.org/10.1016/j.chb.2015.07.025

Binns, R., & Bietti, E. (2020). Dissolving privacy, one merger at a time: Competition, data and third party tracking. *Computer Law & Security Review, 36*, 105369. https://doi.org/https://doi.org/10.1016/j.clsr.2019.105369

Binns, R., Lyngs, U., Van Kleek, M., Zhao, J., Libert, T., & Shadbolt, N. (2018). Third party tracking in the mobile ecosystem. *Proceedings of the 10th ACM Conference on Web Science* (pp.23-31). Association for Computing Machinery. https://doi.org/10.1145/3201064.3201089

Boda, K., Földes, Á. M., Gulyás, G. G., & Imre, S. (2011). User tracking on the web via cross-browser fingerprinting. *Nordic conference on secure it systems* (pp. 31-46). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-29615-4_4

Boerman, S. C., Kruikemeier, S., & Zuiderveen Borgesius, F. J. (2017). Online behavioral advertising: A literature review and research agenda. *Journal of advertising*, *46*(3), 363-376. https://doi.org/10.1080/00913367.2017.1339368

Boerman, S. C., Kruikemeier, S., & Zuiderveen Borgesius, F. J. (2021). Exploring motivations for online privacy protection behavior: Insights from panel data. *Communication Research, 48*(7), 953-977. https://doi.org/10.1177/0093650218800915

Bojinov, H., Michalevsky, Y., Nakibly, G., & Boneh, D. (2014). Mobile device identification via sensor fingerprinting. *arXiv*. https://doi.org/10.48550/arXiv.1408.1416

Bonneau, J. (2012). The science of guessing: analyzing an anonymized corpus of 70 million passwords. *2012 IEEE symposium on security and privacy* (pp. 538-552). IEEEXplor. https://doi.org/10.1109/SP.2012.49

Booth, D., & Jansen, B. J. (2010). A review of methodologies for analyzing websites. In A. Tatnall (Ed.), *Web technologies: Concepts, methodologies, tools, and applications* (pp. 145-166). IGI Global. https://doi.org/10.4018/978-1-60566-982-3.ch009

Bornschein, R., Schmidt, L., & Maier, E. (2020). The effect of consumers' perceived power and risk in digital information privacy: The example of cookie notices. *Journal of Public Policy & Marketing, 39*(2), 135-154. https://doi.org/10.1177/0743915620902143

Bortz, A., Barth, A., & Czeskis, A. (2011). Origin cookies: Session integrity for web applications. *Web 2.0 Security and Privacy (W2SP)*. http://sharif.edu/~kharrazi/courses/40442-952/read/session-integrity.pdf

Bouguettaya, A. R. A., & Eltoweissy, M. Y. (2003). Privacy on the Web: facts, challenges, and solutions. *IEEE Security & Privacy, 1*(6), 40-49. https://doi.org/10.1109/MSECP.2003.1253567

Boyles, J. L., Smith, A., & Madden, M. (2012). *Privacy and data management on mobile devices*. Pew Internet & American Life Project. https://www.pewresearch.org/internet/2012/09/05/privacy-and-data-management-on-mobile-devices/

Broadband Search. (2022). *Key Internet Statistics to Know in 2022*. https://www.broadbandsearch.net/blog/internet-statistics

Broenink, R. (2012). Using browser properties for fingerprinting purposes. *16th biennial Twente Student Conference on IT, Enschede, Holanda* (p.85).

Brooke, A., Lee, R., Monica, A., Andrew, P., Madhu, K., & Erica, T. (2019). *Americans and Privacy: Concerned, Confused and Feeling Lack of Control Over Their Personal Information*. Pew Research Center: Internet, Science & Tech. https://policycommons.net/artifacts/616499/americans-and-privacy/

Brookman, J., Rouge, P., Alva, A., & Yeung, C. (2017). Cross-device tracking: Measurement and disclosures. *Proceedings on Privacy Enhancing Technologies, 2017*(2), 133-148. https://doi.org/10.1515/popets-2017-0020

Brown, B. (2001). Studying the internet experience. *HP laboratories technical report HPL, 49*.

Büchi, M., Just, N., & Latzer, M. (2017). Caring is not enough: the importance of Internet skills for online privacy protection. *Information, Communication & Society, 20*(8), 1261-1278. https://doi.org/10.1080/1369118X.2016.1229001

Bugliesi, M., Calzavara, S., Focardi, R., & Khan, W. (2014). Automatic and robust client-side protection for cookie-based sessions. *International Symposium on Engineering Secure Software and Systems(pp. 161–178)*. Springer,Cham.https://doi.org/10.1007/978-3-319-04897-0_11

Bujlow, T., Carela-Español, V., Solé-Pareta, J., & Barlet-Ros, P. (2015). Web tracking: Mechanisms, implications, and defenses. *arXiv*. https://doi.org/10.48550/arXiv.1507.07872

Centeno, G. L. D. (2016). *Public opinion on Internet privacy: A Puerto Rico* survey [Master's thesis, POLYTECHNIC UNIVERSITY OF PUERTO RICO].ProQuest. https://www.proquest.com/dissertations-theses/public-opinion-on-internet-privacy-puerto-rico/docview/1836098604/se-2

Chai, S., Bagchi-Sen, S., Morrell, C., Rao, H. R., & Upadhyaya, S. J. (2009). Internet and online information privacy: An exploratory study of preteens and early teens. *IEEE Transactions on Professional Communication, 52*(2), 167-182. https://doi.org/10.1109/TPC.2009.2017985

Chanchary, F., & Chiasson, S. (2015). User perceptions of sharing, advertising, and tracking. *Eleventh Symposium On Usable Privacy and Security* (pp.53-67). USENIX Association. https://www.usenix.org/conference/soups2015/proceedings/presentation/chanchary

Chang, Y., Wong, S. F., Libaque-Saenz, C. F., & Lee, H. (2018, 2018/09/01/). The role of privacy policy on consumers' perceived privacy. *Government Information Quarterly, 35*(3), 445-459. https://doi.org/https://doi.org/10.1016/j.giq.2018.04.002

Chen, H.-T., & Chen, W. (2015). Couldn't or wouldn't? The influence of privacy concerns and self-efficacy in privacy management on privacy protection. *Cyberpsychology, Behavior, and Social Networking, 18*(1), 13-19. https://doi.org/10.1089/cyber.2014.0456

Chiasson, S., Abdelaziz, Y., & Chanchary, F. (2018). Privacy concerns amidst OBA and the need for alternative models. *IEEE Internet Computing, 22*(2), 52-61. https://doi.org/10.1109/MIC.2017.3301625

Consulting, I. (n.d.-a). *Art. 5 GDPRPrinciples relating to processing of personal data*. https://gdpr-info.eu/art-5-gdpr/

Consulting, i. (n.d.-b). Art. 6 GDPRLawfulness of processing. https://gdpr-info.eu/art-6-gdpr/

Cookies. (n.d.). hacksplaining. https://www.hacksplaining.com/glossary/cookies

Coram, M. (2019). *Cross-Site Request Forgery Challenges and Solutions*. OSTI. https://www.osti.gov/biblio/1639993

Correa, T. (2010). The participation divide among “online experts”: Experience, skills and psychological factors as predictors of college students' web content creation. *Journal of computer-mediated communication, 16*(1), 71-92. https://doi.org/10.1111/j.1083-6101.2010.01532.x

Correia, J., & Compeau, D. (2017). Information privacy awareness (IPA): a review of the use, definition and measurement of IPA. *Proceedings of the 50th Hawaii International Conference on System Sciences* (pp. 4021-4030). http://hdl.handle.net/10125/41646

D'Ovidio, R., & Doyle, J. (2003). A study on cyberstalking: Understanding investigative hurdles. *FBI Law Enforcement Bulletin, 72(3), 10-17.*

Dabrowski, A., Merzdovnik, G., Ullrich, J., Sendera, G., & Weippl, E. (2019). Measuring cookies and web privacy in a post-gdpr world. *International Conference on Passive and Active Network Measurement* (pp.258- 270). Springer, Cham. https://doi.org/10.1007/978-3-030-15986-3_17

Dao, H., & Fukuda, K. (2021). Alternative to third-party cookies: investigating persistent PII leakage-based web tracking. *Proceedings of the 17th International Conference on emerging Networking EXperiments and Technologies (pp.223-229).* Association for Computing Machinery. https://doi.org/10.1145/3485983.3494860

De Bruijn, H., & Janssen, M. (2017). Building cybersecurity awareness: The need for evidence-based framing strategies. *Government Information Quarterly, 34*(1), 1-7. https://doi.org/10.1016/j.giq.2017.02.007

Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F., & Holz, T. (2018). We value your privacy... now take some cookies: Measuring the GDPR's impact on web privacy. *arXiv*. https://doi.org/10.14722/ndss.2019.23378

Dey, S., Roy, N., Xu, W., & Nelakuditi, S. (2013). Acm hotmobile 2013 poster: Leveraging imperfections of sensors for fingerprinting smartphones. *ACM SIGMOBILE Mobile Computing and Communications Review, 17*(3), 21-22. https://doi.org/10.1145/2542095.2542107

Dinev, T., & Hart, P. (2006). An extended privacy calculus model for e-commerce transactions. *Information systems research, 17*(1), 61-80. https://doi.org/10.1287/isre.1060.0080

Dommeyer, C. J., & Gross, B. L. (2003). What consumers know and what they do: An investigation of consumer knowledge, awareness, and use of privacy protection strategies. *Journal of Interactive Marketing, 17*(2), 34-51. https://doi.org/10.1002/dir.10053

Dorfleitner, G., Hornuf, L., & Kreppmeier, J. (2021). Promise not Fulfilled: FinTech Data Privacy, and the GDPR. CESifo. http://dx.doi.org/10.2139/ssrn.3950094

Ebbers, F. (2019). How to Protect My Privacy?-Classifying End-User Information Privacy Protection Behaviors. *IFIP International Summer School on Privacy and Identity Management (*pp. 327–342). Springer, Cham. https://doi.org/10.1007/978-3-030-42504-3_21

Eckersley, P. (2010). How Unique Is Your Web Browser? Privacy Enhancing Technologies. *International Symposium on Privacy Enhancing Technologies Symposium* (pp. 1–18). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-14527-8_1

Efthymiou, D., & Antoniou, C. (2012). Use of Social Media for Transport Data Collection. *Procedia - Social and Behavioral Sciences, 48,* 775-785. https://doi.org/https://doi.org/10.1016/j.sbspro.2012.06.1055

Ehikioya, S. A., & Lu, S. (2020). A Traffic Tracking Analysis Model for the Effective Management of E-commerce Transactions. *International Journal of Networked and Distributed Computing, 8*(3), 171-193. https://doi.org/10.2991/ijndc.k.200515.006

Ellonen, H.-K., Wikström, P., & Johansson, A. (2015). The role of the website in a magazine business – revisiting old truths. *Journal of Media Business Studies, 12*(4), 238-249. https://doi.org/10.1080/16522354.2015.1107334

Endler, D. (2002). *The evolution of cross site scripting attacks*. http://www.dankalia.com/tutor/01001/0100101079.pdf

Englehardt, S., Han, J., & Narayanan, A. (2018). I never signed up for this! Privacy implications of email tracking. *Proceedings on Privacy Enhancing Technologies, 2018*(1), 109-126. https://doi.org/10.1515/popets-2018-0006

Englehardt, S., & Narayanan, A. (2016). Online tracking: A 1-million-site measurement and analysis. *Proceedings of the 2016 ACM SIGSAC conference on computer and communications security* (pp. 1388-1401). Association for Computing Machinery. https://doi.org/10.1145/2976749.2978313

Ermakova, T., Fabian, B., Bender, B., & Klimek, K. (2018). Web tracking-A literature review on the state of research. *Proceedings of the 51st Hawaii International Conference on System Sciences*(pp.4732-4741). AIS eLibrary. http://hdl.handle.net/10125/50485

Ermakova, T., Hohensee, A., Orlamünde, I., & Fabian, B. (2019). Privacy-invading mechanisms in e-commerce-a case study on German tourism websites. *International Journal of Networking and Virtual Organisations, 20*(2), 105-126.

Falahrastegar, M., Haddadi, H., Uhlig, S., & Mortier, R. (2014). Anatomy of the third-party web tracking ecosystem. *arXiv*. https://doi.org/10.48550/arXiv.1409.1066

Falahrastegar, M., Haddadi, H., Uhlig, S., & Mortier, R. (2016). Tracking personal identifiers across the web. *International conference on passive and active network measurement* (pp.30-41). Springer, Cham. https://doi.org/10.1007/978-3-319-30505-9_3

Fang, Y., Li, Y., Liu, L., & Huang, C. (2018). DeepXSS: Cross site scripting detection based on deep learning. *Proceedings of the 2018 International Conference on Computing and Artificial Intelligence* (pp.47-51). https://doi.org/10.1145/3194452.3194469

Fawaz, K., Kim, K.-H., & Shin, K. G. (2016). Protecting privacy of {BLE} device users. *25th USENIX Security Symposium(pp.1205-1221)*. USENIX Association.

Fefer, R. F. (2019). Data flows, online privacy, and trade policy. *Congressional Research Service*. https://www.everycrsreport.com/files/20190311_R45584_caa16f89385a9b89430cfaac5d6633c392313c45.pdf

Felten, E. W., & Schneider, M. A. (2000). Timing attacks on web privacy. *Proceedings of the 7th ACM Conference on Computer and Communications Security* (pp.25-32). Association for Computing Machinery. https://doi.org/10.1145/352600.352606

Fifield, D., & Egelman, S. (2015). Fingerprinting web users through font metrics. *International Conference on Financial Cryptography and Data Security* (pp.107-124). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-47854-7_7

Fincham, J. E. (2008). Response rates and responsiveness for surveys, standards, and the Journal. *American journal of pharmaceutical education, 72*(2), 43-43. https://doi.org/10.5688/aj720243

Fouad, I., Santos, C., Legout, A., & Bielova, N. (2021). Did I delete my cookies? Cookies respawning with browser fingerprinting. *arXiv*. https://doi.org/10.48550/arXiv.2105.04381

Frow, J. (2019). Cookie. *Cultural Studies Review, 25*(2), 208-210. https://search.informit.org/doi/10.3316/informit.915946672490262

Fujs, D., Mihelic, A., & Vrhovec, S. (2019). Social Network Self-Protection Model: What Motivates Users to Self-Protect? *Journal of Cyber Security and Mobility*, 467–492. https://doi.org/10.13052/2245-1439.844

Gerber, N., Gerber, P., & Volkamer, M. (2018). Explaining the privacy paradox: A systematic review of literature investigating privacy attitude and behavior. *Computers & Security*, *77*, 226-261. https://doi.org/10.1016/j.cose.2018.04.002

Givens, C. L. (2014). *Information privacy fundamentals for librarians and information professionals*. Rowman & Littlefield.

Goethem, T. V., Scheepers, W., Preuveneers, D., & Joosen, W. (2016). Accelerometer-based device fingerprinting for multi-factor mobile authentication. *International Symposium on Engineering Secure Software and Systems* (106-121). Springer, Cham. https://doi.org/10.1007/978-3-319-30806-7_7

Hadlington, L. (2018). Employees attitudes towards cyber security and risky online behaviours: an empirical assessment in the United Kingdom. *International Journal of Cyber Criminology*, *12*(1), 269–281. https://doi.org/10.5281/zenodo.1467909

Ham, C.-D., & Nelson, M. R. (2016). The role of persuasion knowledge, assessment of benefit and harm, and third-person perception in coping with online behavioral advertising. *Computers in Human Behavior, 62*, 689-702. https://doi.org/10.1016/j.chb.2016.03.076

Hamed, A., & Ayed, H. K.-B. (2015). Privacy scoring and users' awareness for Web tracking. *2015 6th International Conference on Information and Communication Systems (ICICS)* (pp. 100-105*)*.IEEEXplore. https://doi.org/10.1109/IACS.2015.7103210

Hargittai, E. (2010). Facebook privacy settings: Who cares? *First Monday*. 15(8). https://doi.org/10.5210/fm.v15i8.3086

Hargittai, E., & Litt, E. (2013). New strategies for employment? internet skills and online privacy practices during people's job search. *IEEE Security & Privacy, 11*(3), 38-45. https://doi.org/10.1109/MSP.2013.64

Hargittai, E., & Shafer, S. (2006). Differences in actual and perceived online skills: The role of gender. *Social Science Quarterly, 87*(2), 432-448. https://doi.org/10.1111/j.1540-6237.2006.00389.x

Hassan, N., & Hijazi, R. (2017). *Digital Privacy and Security Using Windows: A Practical Guide*. Apress.

Hoban, P. R., & Bucklin, R. E. (2015). Effects of internet display advertising in the purchase funnel: Model-based insights from a randomized field experiment. *Journal of Marketing Research, 52*(3), 375-393. https://doi.org/10.1509/jmr.13.0277

Hoofnagle, C. J., van der Sloot, B., & Borgesius, F. Z. (2019). The European Union general data protection regulation: what it is and what it means. *Information & Communications Technology Law, 28*(1), 65-98. https://doi.org/10.1080/13600834.2019.1573501

Hoy, M. G., & Milne, G. (2010). Gender differences in privacy-related measures for young adult Facebook users. *Journal of interactive advertising, 10*(2), 28-45. https://doi.org/10.1080/15252019.2010.10722168

Hu, X., & Sastry, N. (2019). *Characterising Third Party Cookie Usage in the EU after GDPR* Proceedings of the 10th ACM Conference on Web Science, Boston, Massachusetts, USA. https://doi.org/10.1145/3292522.3326039

Hughes-Roberts, T. (2013). Privacy and social networks: Is concern a valid indicator of intention and behaviour?. *2013 International Conference on Social Computing* (pp. 909-912). IEEE Xplore. https://doi.org/10.1109/SocialCom.2013.140

Hussain, A., Tahir, A., Hussain, Z., Sheikh, Z., Gogate, M., Dashtipour, K., Ali, A., & Sheikh, A. (2021). Artificial Intelligence–Enabled Analysis of Public Attitudes on Facebook and Twitter Toward COVID-19 Vaccines in the United Kingdom and the United States: Observational Study. *Journal of medical Internet research, 23*(4), e26627. https://doi.org/10.2196/26627

Iordanou, C., Smaragdakis, G., Poese, I., & Laoutaris, N. (2018). Tracing cross border web tracking. *Proceedings of the Internet Measurement Conference 2018* (pp.329-342). Association for Computing Machinery. https://doi.org/10.1145/3278532.3278561

Iqbal, U., Englehardt, S., & Shafiq, Z. (2021). Fingerprinting the fingerprinters: Learning to detect browser fingerprinting behaviors. *2021 IEEE Symposium on Security and Privacy (SP)* (pp. 1143-1161).IEEE Xplore. https://doi.org/10.1109/SP40001.2021.00017

Ishtiaq, A., Abbasi, S. H., Aleem, M., & Islam, M. A. I. (2017). User tracking mechanisms and counter measures. *International Journal of Applied Mathematics Electronics and Computers, 5*(2), 33-40. https://doi.org/10.18100/ijamec.2017528829

Jaeger, L. (2018). Information security awareness: literature review and integrative framework. *Proceedings of the 51st Hawaii International Conference on System Sciences* (pp. 4703- 4712).AIS eLibrary. http://hdl.handle.net/10125/50482

Jamiy, F. E., Daif, A., Azouazi, M., & Marzak, A. (2015). The potential and challenges of Big data-Recommendation systems next level application. *arXiv*. https://doi.org/10.48550/arXiv.1501.03424

Janssen, M., & Helbig, N. (2018). Innovating and changing the policy-cycle: Policy-makers be prepared! *Government Information Quarterly, 35*(4), S99-S105. https://doi.org/10.1016/j.giq.2015.11.009

Jegatheesan, S. (2013). Cookies Invading Our Privacy for Marketing Advertising and Security Issues. *arXiv*. https://doi.org/10.48550/arXiv.1305.2306

Jing, F., Shenwen, L., Xiongjie, D., Yong, Y., & Shaohua, W. (2015). 2*015 5th International Conference on Information Science and Technology (ICIST)* (pp. 408-412).IEEE Xplore. https://doi.org/10.1109/ICIST.2015.7289006

Johnson, J. (2019). *Distribution of internet users worldwide as of 2019, by age group*. https://www.statista.com/statistics/272365/age-distribution-of-internet-users-worldwide/

Jussila, J. (2018). *HTTP cookie weaknesses, attack methods and defense mechanisms: a systematic literature review.* [Master's thesis, UNIVERSITY OF JYVÄSKYLÄ]. JYX. http://urn.fi/URN:NBN:fi:jyu-201808023720

Kashi, E., & Zavou, A. (2020). Did I Agree to This? Silent Tracking Through Beacons. *International Conference on Human-Computer Interaction* (pp 427–444). Springer, Cham. https://doi.org/10.1007/978-3-030-50309-3_28

Kaur, J., & Garg, U. (2021). A Detailed Survey on Recent XSS Web-Attacks Machine Learning Detection Techniques. *2021 2nd Global Conference for Advancement in Technology (GCAT) (*pp. 1-6).IEEE Xplore. https://doi.org/10.1109/GCAT52182.2021.9587569

Kavisankar, L., Chellappan, C., & Poovammal, E. (2016). Against spoofing attacks in network layer. In *Combating Security Breaches and Criminal Activity in the Digital Sphere* (pp. 41-56). IGI Global. https://doi.org/10.4018/978-1-5225-0193-0.ch003

Kirda, E., Kruegel, C., Vigna, G., & Jovanovic, N. (2006). Noxes: a client-side solution for mitigating cross-site scripting attacks. *Proceedings of the 2006 ACM symposium on Applied computing* (pp.330-337). Association for Computing Machinery. https://doi.org/10.1145/1141277.1141357

Kirkpatrick, K. (2020). Tracking shoppers. *Commun. ACM*, *63*(2), 19–21. https://doi.org/10.1145/3374876

Kokolakis, S. (2017). Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Computers & Security*, *64*, 122-134. https://doi.org/10.1016/j.cose.2015.07.002

Kolsek, M. (2002). Session fixation vulnerability in web-based applications. *ACROS Security*, http://www. acrossecurity. com/papers/session fixation. pdf.

Kombade, R. D., & Meshram, B. (2012). CSRF vulnerabilities and defensive techniques. *International Journal of Computer Network and Information Security*, *4*(1), 31-37. https://doi.org/10.5815/ijcnis.2012.01.04

Kour, P. (2020). A Study on Cross-Site Request Forgery Attack and its Prevention Measures. *International Journal of Advanced Networking and Applications*, *12*(2), 4561-4566. https://doi.org/10.35444/IJANA.2020.12204

Krishnamurthy, B., Naryshkin, K., & Wills, C. (2011). Privacy leakage vs. protection measures: the growing disconnect. *Proceedings of the Web* (pp. 1-10).

Krishnamurthy, B., & Wills, C. (2009). Privacy diffusion on the web: a longitudinal perspective. *Proceedings of the 18th international conference on World wide web* (pp.541-550). Association for Computing Machinery. https://doi.org/10.1145/1526709.1526782

Kristol, D. M. (2001). HTTP Cookies: Standards, privacy, and politics. *ACM Trans. Internet Technol.*, *1*(2), 151–198. https://doi.org/10.1145/502152.502153

Kumar, L., Singh, H., & Kaur, R. (2012). *Web analytics and metrics: a survey* Proceedings of the International Conference on Advances in Computing, Communications and Informatics, Chennai, India. https://doi.org/10.1145/2345396.2345552

Kumar, V., & Ogunmola, G. A. (2020). Web analytics for knowledge creation: a systematic review of tools, techniques, and practices. *International Journal of Cyber Behavior, Psychology and Learning (IJCBPL), 10*(1), 1-14. https://doi.org/10.4018/IJCBPL.2020010101

Laoutaris, N. (2018). Data Transparency: Concerns and Prospects [Point of View]. *Proceedings of the IEEE, 106*(11), 1867-1871. https://doi.org/10.1109/JPROC.2018.2872313

Lee, H., Park, H., & Kim, J. (2013). Why do people share their context information on Social Network Services? A qualitative study and an experimental study on users' behavior of balancing perceived benefit and risk. *International Journal of Human-Computer Studies, 71*(9), 862-877. https://doi.org/10.1016/j.ijhcs.2013.01.005

Lee, H., Wong, S. F., Oh, J., & Chang, Y. (2019). Information privacy concerns and demographic characteristics: Data from a Korean media panel survey. *Government Information Quarterly*, *36*(2), 294-303. https://doi.org/10.1016/j.giq.2019.01.002

Li, C.-T. (2010). Source camera identification using enhanced sensor pattern noise. *IEEE Transactions on Information Forensics and Security*, *5*(2), 280-287. https://doi.org/10.1109/TIFS.2010.2046268

Li, T.-C., Hang, H., Faloutsos, M., & Efstathopoulos, P. (2015). Trackadvisor: Taking back browsing privacy from third-party trackers. *International Conference on Passive and Active Network Measurement* (pp.277-289). Springer, Cham. https://doi.org/10.1007/978-3-319-15509-8_21

Libert, T. (2015). Exposing the hidden web: An analysis of third-party HTTP requests on 1 million websites. *arXiv*. https://doi.org/10.48550/arXiv.1511.00619

Lin, X., Zavarsky, P., Ruhl, R., & Lindskog, D. (2009). Threat Modeling for CSRF Attacks. 2009 *International Conference on Computational Science and Engineering* (pp. 486-491). IEEE Xplore. https://doi.org/10.1109/CSE.2009.372

Lindemann, N. (2021). *What's The Average Survey Response Rate? [2021 Benchmark]*. https://surveyanyplace.com/blog/average-survey-response-rate

Litt, E. (2013). Understanding social network site users' privacy tool use. *Computers in Human Behavior*, *29*(4), 1649-1656. https://doi.org/10.1016/j.chb.2013.01.049

Lutz, C., & Strathoff, P. (2014). Privacy concerns and online behavior–Not so paradoxical after all? Viewing the privacy paradox through different theoretical lenses. *Viewing the Privacy Paradox Through Different Theoretical Lenses (April 15, 2014.)* http://dx.doi.org/10.2139/ssrn.2425132

Madden, M., & Rainie, L. (2015). *Americans' attitudes about privacy, security and surveillance.* https://policycommons.net/artifacts/619110/americans-attitudes-about-privacy-security-and-surveillance/

Malandrino, D., Petta, A., Scarano, V., Serra, L., Spinelli, R., & Krishnamurthy, B. (2013). *Privacy awareness about information leakage: who knows what about me? Proceedings of the 12th ACM workshop on Workshop on privacy in the electronic society* (pp.279-284). Association for Computing Machinery. https://doi.org/10.1145/2517840.2517868

Malhotra, N. K., Kim, S. S., & Agarwal, J. (2004). Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. *Information systems research*, *15*(4), 336-355. https://doi.org/10.1287/isre.1040.0032

Marshall, M. (2005). New cookies much harder to crumble. *The Standard-Times, May, 15*.

Martin, M. C., & Lam, M. S. (2008). Automatic Generation of XSS and SQL Injection Attacks with Goal-Directed Model Checking. *17th USENIX Security Symposium* (pp.31-43). USENIX Association. http://usenix.org/event/sec08/tech/full_papers/martin/martin.pdf

Masur, P. K. (2020). How online privacy literacy supports self-data protection and self-determination in the age of information. *Media and Communication, 8*(2), 258-269. https://doi.org/10.17645/mac.v8i2.2855

Matzner, T., Masur, P. K., Ochs, C., & Pape, T. v. (2016). Do-It-yourself data protection—Empowerment or burden? In *Data protection on the move* (pp. 277-305). Springer. https://doi.org/10.1007/978-94-017-7376-8_11

Mayer, J. R. (2009). Any person... a pamphleteer": Internet Anonymity in the Age of Web 2.0. *Undergraduate Senior Thesis, Princeton University*, 85.

Mayer, J. R., & Mitchell, J. C. (2012). Third-party web tracking: Policy and technology. *2012 IEEE symposium on security and privacy (pp. 413-427).IEEE Explore. https://doi.org/10.1109/SP.2012.47*

McCoy, M. S., Libert, T., Buckler, D., Grande, D. T., & Friedman, A. B. (2020). Prevalence of third-party tracking on COVID-19–related web pages. *Jama, 324*(14), 1462-1464. https://doi.org/10.1001/jama.2020.16178

McCrohan, K. F., Engel, K., & Harvey, J. W. (2010). Influence of awareness and training on cyber security. *Journal of internet commerce, 9*(1), 23-41. https://doi.org/10.1080/15332861.2010.487415

Milberg, S. J., Smith, H. J., & Burke, S. J. (2000). Information privacy: Corporate management and national regulation. *Organization science, 11*(1), 35-57. http://dx.doi.org/10.1287/orsc.11.1.35.12567

Miller, M. (2010). *The ultimate web marketing guide*. Pearson Education.

Millett, L. I., Friedman, B., & Felten, E. (2001). *Cookies and Web browser design: toward realizing informed consent online* Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Seattle, Washington, USA. https://doi.org/10.1145/365024.365034

Miltgen, C. L., & Peyrat-Guillard, D. (2014). Cultural and generational influences on privacy concerns: a qualitative study in seven European countries. *European Journal of Information Systems, 23*(2), 103-125. https://doi.org/10.1057/ejis.2013.17

Mittal, S. (2010). User privacy and the evolution of third-party tracking mechanisms on the world wide web. *Available at SSRN 2005252*. http://dx.doi.org/10.2139/ssrn.2005252

Mowery, K., Bogenreif, D., Yilek, S., & Shacham, H. (2011). Fingerprinting information in JavaScript implementations. *Proceedings of W2SP*, *2*(11).

Nagpure, S., & Kurkure, S. (2017). Vulnerability Assessment and Penetration Testing of Web Application. *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA*) (pp. 1-6). IEEE Xplore. https://doi.org/10.1109/ICCUBEA.2017.8463920

Narayanan, L. (2020). *Cookies 'n'consent: An empirical study on the factors influencing customer attitudes towards cookie consent among internet users in EU Dublin Business School* [Master's thesis, Dublin Business School]. Dublin Business School. https://esource.dbs.ie/handle/10788/4115

Nayak, M. S. D. P., & Narayan, K. (2019). Strengths and weaknesses of online surveys. *Journal of Humanities and Social Sciences 24*(5),31-38.

Nielsen, J. (2019). Experimenting with computational methods for large-scale studies of tracking technologies in web archives. *Internet Histories, 3*(3-4), 293-315. https://doi.org/10.1080/24701475.2019.1671074

Niforatos, J. D., Zheutlin, A. R., & Sussman, J. B. (2021). Prevalence of third-party data tracking by US hospital websites. *JAMA network open, 4*(9). https://doi.org/10.1001/jamanetworkopen.2021.26121

Nikiforakis, N., Joosen, W., & Livshits, B. (2015). *PriVaricator: Deceiving Fingerprinters with Little White Lies* Proceedings of the 24th International Conference on World Wide Web, Florence, Italy. https://doi.org/10.1145/2736277.2741090

Nikkhah Bahrami, P., Iqbal, U., & Shafiq, Z. (2021). FP-Radar: Longitudinal Measurement and Early Detection of Browser Fingerprinting. *arXiv.* https://doi.org/10.48550/arXiv.2112.01662

Norberg, P. A., Horne, D. R., & Horne, D. A. (2007). The privacy paradox: Personal information disclosure intentions versus behaviors. *Journal of consumer affairs, 41*(1), 100-126. https://doi.org/10.1111/j.1745-6606.2006.00070.x

O'Brien, P., Young, S. W., Arlitsch, K., & Benedict, K. (2018). Protecting privacy on the web: A study of HTTPS and Google Analytics implementation in academic library websites. *Online Information Review 42* (6),734-751. https://doi.org/10.1108/OIR-02-2018-0056

Papadopoulos, P., Kourtellis, N., & Markatos, E. P. (2018). Exclusive: How the (synced) cookie monster breached my encrypted vpn session. *Proceedings of the 11th European Workshop on Systems*

Security (pp.1-6). Association for Computing Machinery. https://doi.org/10.1145/3193111.3193117

Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour, 5*(11), 1535-1547. https://doi.org/10.1038/s41562-021-01117-5

Pekala, S. (2017). Privacy and user experience in 21st century library discovery. *Information technology and libraries, 36*(2), 48-58. https://doi.org/10.6017/ital.v36i2.9817

Petsios, I. P. G. A. T., & Keromytis, S. S. A. D. (2015).Where's Wally? Precise User Discovery Attacks in Location Proximity Services. *In Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security* (pp. 817-828). Association for Computing Machinery. https://doi.org/10.1145/2810103.2813605

Phippen, A., Sheppard, L., & Furnell, S. (2004). A practical evaluation of Web analytics. *Internet Research. 14*(4),284-293. https://doi.org/10.1108/10662240410555306

Pierson, J., & Heyman, R. (2011). Social media and cookies: challenges for online privacy. *Info 13*(6),30-42. https://doi.org/10.1108/14636691111174243

Pinto, P., Lages, R., & Au-Yong-Oliveira, M. (2020). Web cookies: Is there a trade-off between website efficiency and user privacy? *World Conference on Information Systems and Technologies* (pp.713-722). Springer, Cham. https://doi.org/10.1007/978-3-030-45691-7_67

Pitkänen, O., & Tuunainen, V. K. (2012). Disclosing personal data socially—An empirical study on Facebook users' privacy awareness. *Journal of Information Privacy and Security, 8*(1), 3-29. https://doi.org/10.1080/15536548.2012.11082759

Politou, E., Alepis, E., & Patsakis, C. (2018). Forgetting personal data and revoking consent under the GDPR: Challenges and proposed solutions. *Journal of cybersecurity, 4*(1), tyy001. https://doi.org/10.1093/cybsec/tyy001

Prasetio, D. A., Kusrini, K., & Arief, M. R. (2021). Cross-site Scripting Attack Detection Using Machine Learning with Hybrid Features. *JURNAL INFOTEL, 13*(1), 1-6. https://doi.org/10.20895/infotel.v13i1.606

Preibusch, S. (2013). Guide to measuring privacy concern: Review of survey and observational instruments. *International Journal of Human-Computer Studies, 71*(12), 1133-1143. https://doi.org/https://doi.org/10.1016/j.ijhcs.2013.09.002

Prince, C., Omrani, N., Maalaoui, A., Dabic, M., & Kraus, S. (2021). Are We Living in Surveillance Societies and Is Privacy an Illusion? An Empirical Study on Privacy Literacy and Privacy Concerns. *IEEE Transactions on Engineering Management*. https://doi.org/10.1109/TEM.2021.3092702

PrivacyCommissioner. (n.d.-a). *Principle 1 - Purpose for collection of personal information*. https://www.privacy.org.nz/privacy-act-2020/privacy-principles/1/

PrivacyCommissioner. (n.d.-b). *Principle 5 - Storage and security of information*. https://www.privacy.org.nz/privacy-act-2020/privacy-principles/5/

PrivacyCommissioner. (n.d.-c). *Principle 10 - Limits on use of personal information*. https://www.privacy.org.nz/privacy-act-2020/privacy-principles/10/

PrivacyCommissioner. (n.d.-d). *Principle 11 - Disclosure of personal information*. https://www.privacy.org.nz/privacy-act-2020/privacy-principles/limits-on-disclosure-of-personal-information-principle-11/

Pugliese, G. (2015). Web Tracking: Overview and applicability in digital investigations. *it-Information Technology, 57*(6), 366-375. https://doi.org/10.1515/itit-2015-0015

Purcell, K., Rainie, L., & Brenner, J. (2012). *Search engine use 2012*. Pew Research Center. https://www.pewresearch.org/internet/2012/03/09/search-engine-use-2012/

Rahalkar, S. A. (2016). Web Application Hacking. In *Certified Ethical Hacker (CEH) Foundation Guide* (pp. 131-141). Springer.

Rainie, L., Kiesler, S., Kang, R., Madden, M., Duggan, M., Brown, S., & Dabbish, L. (2013). *Anonymity, privacy, and security online*. Pew Research Center. https://www.pewresearch.org/internet/2013/09/05/anonymity-privacy-and-security-online/

Rankothge, W. H., & Randeniya, S. M. N. (2020). Identification and Mitigation Tool For Cross-Site Request Forgery (CSRF). *2020 IEEE 8th R10 Humanitarian Technology Conference (R10-HTC)*(pp.1-5). IEEE Xplore. https://doi.org/10.1109/R10-HTC49770.2020.9357029

Ren, J., Rao, A., Lindorfer, M., Legout, A., & Choffnes, D. (2016). Recon: Revealing and controlling pii leaks in mobile network traffic. *Proceedings of the 14th Annual International Conference on Mobile Systems, Applications, and Services* (pp.361-374). Association for Computing Machinery. https://doi.org/10.1145/2906388.2906392

Rodríguez, G. E., Benavides, D. E., Torres, J., Flores, P., & Fuertes, W. (2018). Cookie scout: An analytic model for prevention of cross-site scripting (XSS) using a cookie classifier. International Conference on Information Technology & Systems (pp.497-507). Springer, Cham. https://doi.org/10.1007/978-3-319-73450-7_47

Rodríguez, G. E., Torres, J. G., Flores, P., & Benavides, D. E. (2020). Cross-site scripting (XSS) attacks and mitigation: A survey. *Computer Networks, 166*, 106960. https://doi.org/https://doi.org/10.1016/j.comnet.2019.106960

Roesner, F., Kohno, T., & Wetherall, D. (2012). Detecting and Defending Against {Third-Party} Tracking on the Web. *9th USENIX Symposium on Networked Systems Design and Implementation (NSDI 12)*. USENIX Association. https://www.usenix.org/conference/nsdi12/technical-sessions/presentation/roesner

Rotman, D. (2009). Are You Looking At Me?-Social Media and Privacy Literacy. http://hdl.handle.net/2142/15339

Sanchez-Rola, I., Ugarte-Pedrero, X., Santos, I., & Bringas, P. G. (2017). The web is watching you: A comprehensive review of web-tracking techniques and countermeasures. *Logic Journal of the IGPL, 25*(1), 18-29. https://doi.org/10.1093/jigpal/jzw041

Schaub, F., Marella, A., Kalvani, P., Ur, B., Pan, C., Forney, E., & Cranor, L. F. (2016). Watching them watching me: Browser extensions' impact on user privacy awareness and concern. *NDSS workshop on usable security*. https://doi.org/10.14722/USEC.2016.23017

Schubert, R., Marinica, I., Mosetti, L., & Bajka, S. (2022). *Mitigating the Privacy Paradox Through Higher Privacy Literacy? Insights from a Lab Experiment Based on Facebook Data* (Collegium Helveticum Working Papers NO. 3). https://doi.org/10.3929/ethz-b-000574848

Schwartz, P. M., & Solove, D. J. (2011). The PII problem: Privacy and a new concept of personally identifiable information. *NYUL rev.*, *86*, 1814.

Sentamilselvan, K., Pandian, D. S. L., & Sathiyamurthy, D. K. (2013). Survey on Cross Site Request Forgery. *International Conference on Research and Development Prospects on Engineering and Technology* (pp.159-164*).*

Sheikh, A. (2021). Session Hijacking. In *Certified Ethical Hacker (CEH) Preparation Guide* (pp. 93-102). Springer.

Simpkins, L., Yuan, X., Modi, J., Zhan, J., & Yang, L. (2015). *A course module on web tracking and privacy* Proceedings of the 2015 Information Security Curriculum Development Conference, Kennesaw, Georgia. https://doi-org.ezproxy.aut.ac.nz/10.1145/2885990.2886000

Sipior, J. C., Ward, B. T., & Mendoza, R. A. (2011). Online privacy concerns associated with cookies, flash cookies, and web beacons. *Journal of internet commerce*, *10*(1), 1-16. https://doi.org/10.1080/15332861.2011.558454

Sivakorn, S., Polakis, J., & Keromytis, A. D. (2016). Http cookie hijacking in the wild: Security and privacy implications. *Black Hat USA*. https://www.blackhat.com/docs/us-16/materials/us-16-Sivakorn-HTTP-Cookie-Hijacking-In-The-Wild-Security-And-Privacy-Implications-wp.pdf

Smit, E. G., Van Noort, G., & Voorveld, H. A. (2014). Understanding online behavioural advertising: User knowledge, privacy concerns and online coping behaviour in Europe. *Computers in Human Behavior, 32*, 15-22. https://doi.org/10.1016/j.chb.2013.11.008

Soltani, A., Canty, S., Mayo, Q., Thomas, L., & Hoofnagle, C. J. (2010). Flash Cookies and Privacy. *AAAI Spring Symposium: Intelligent Information Privacy Management* (pp.158-163). http://dx.doi.org/10.2139/ssrn.1446862

Sörensen, O. (2013). Zombie-cookies: Case studies and mitigation. *8th International Conference for Internet Technology and Secured Transactions (ICITST-2013)*(pp. 321-326). IEEE Xplore. https://doi.org/10.1109/ICITST.2013.6750214

Spooren, J., Preuveneers, D., & Joosen, W. (2015). Mobile device fingerprinting considered harmful for risk-based authentication. *Proceedings of the Eighth European Workshop on System Security* (pp.1-6). Association for Computing Machinery. https://doi.org/10.1145/2751323.2751329

Sprankel, S. (2011). Online tracking targeted advertising and user privacy-the technical part. *Privacy and Web 2.0 Seminar Summer Term 2011*(pp.1-6).

Starov, O., Gill, P., & Nikiforakis, N. (2016). Are you sure you want to contact us? quantifying the leakage of pii via website contact forms. *Proceedings on Privacy Enhancing Technologies, 2016*(1), 20-33.

Starov, O., & Nikiforakis, N. (2017). Extended tracking powers: Measuring the privacy diffusion enabled by browser extensions. *Proceedings of the 26th International Conference on World Wide Web* (pp.1481-1490). Association for Computing Machinery. https://doi.org/10.1145/3038912.3052596

Strycharz, J., Smit, E., Helberger, N., & van Noort, G. (2021). No to cookies: Empowering impact of technical and legal knowledge on rejecting tracking cookies. *Computers in Human Behavior, 120*, 106750. https://doi.org/10.1016/j.chb.2021.106750

Taddicken, M. (2014). The 'privacy paradox'in the social web: The impact of privacy concerns, individual characteristics, and the perceived social relevance on different forms of self-disclosure. *Journal of computer-mediated communication, 19*(2), 248-273. https://doi.org/10.1111/jcc4.12052

Takahashi, H., Yasunaga, K., Mambo, M., Kim, K., & Youm, H. Y. (2013). Preventing Abuse of Cookies Stolen by XSS. *2013 Eighth Asia Joint Conference on Information Security* (pp. 85-89). IEEE Xplore. https://doi.org/10.1109/ASIAJCIS.2013.20

Tariq, I., Sindhu, M. A., Abbasi, R. A., Khattak, A. S., Maqbool, O., & Siddiqui, G. F. (2021). Resolving cross-site scripting attacks through genetic algorithm and reinforcement learning. *Expert Systems with Applications, 168*, 114386. https://doi.org/https://doi.org/10.1016/j.eswa.2020.114386

Tirtea, R., Castelluccia, C., & Ikonomou, D. (2011). Bittersweet cookies. Some security and privacy considerations. *European Union Agency for Network and Information Security-ENISA*.

Tomoya, H., Tomoki, N., Akio, M., & Kazumasa, H. (2012). Regional differences in survey response rates and their individual and geographic determinants: A multilevel analysis. *Geographical review of Japan series B, 85*(5), 447-467.

Trepte, S., Teutsch, D., Masur, P. K., Eichler, C., Fischer, M., Hennhöfer, A., & Lind, F. (2015). Do People Know About Privacy and Data Protection Strategies? Towards the "Online Privacy Literacy Scale" (OPLIS). *Reforming European Data Protection Law* (pp. 333-365). Springer, Dordrecht. https://doi.org/10.1007/978-94-017-9385-8

Ur, B., Leon, P. G., Cranor, L. F., Shay, R., & Wang, Y. (2012). Smart, useful, scary, creepy: perceptions of online behavioral advertising. *proceedings of the eighth symposium on usable privacy and security(pp.1-6).* Association for Computing Machinery. https://doi.org/10.1145/2335356.2335362

Utz, C., Degeling, M., Fahl, S., Schaub, F., & Holz, T. (2019). *(Un)informed Consent: Studying GDPR Consent Notices in the Field* Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security, London, United Kingdom. https://doi.org/10.1145/3319535.3354212

Vastel, A., Laperdrix, P., Rudametkin, W., & Rouvoy, R. (2018). FP-STALKER: Tracking Browser Fingerprint Evolutions. *2018 IEEE Symposium on Security and Privacy (SP)* (pp.728-741). IEEE Xplore. https://doi.org/10.1109/SP.2018.00008

Wagner, P. (2020). *Cookies: Privacy Risks, Attacks, and Recommendations. Attacks, and Recommendations.* SSRN. http://dx.doi.org/10.2139/ssrn.3761967

Wambach, T., & Bräunlich, K. (2016). The evolution of third-party web tracking. *International Conference on Information Systems Security and Privacy* (pp.130-147). Springer, Cham. https://doi.org/10.1007/978-3-319-54433-5_8

WatchGuard. (2021). *Understanding the Impact of Cross-Site Scripting in 2021*. https://www.watchguard.com/wgrd-news/blog/understanding-impact-cross-site-scripting-2021

Weinberger, M., Zhitomirsky-Geffet, M., & Bouhnik, D. (2017). Factors affecting users' online privacy literacy among students in Israel. *Online Information Review 41(5),655-671*. https://doi.org/10.1108/OIR-05-2016-0127

Wenz, A., Al Baghal, T., & Gaia, A. (2020). Language Proficiency Among Respondents: Implications for Data Quality in a Longitudinal Face-To-Face Survey. *Journal of Survey Statistics and Methodology*, *9*(1), 73-93. https://doi.org/10.1093/jssam/smz045

Wills, C. E., & Tatar, C. (2012). Understanding what they do with what they know. *Proceedings of the 2012 ACM Workshop on Privacy in the Electronic Society (pp.13-18)* . Association for Computing Machinery. https://doi.org/10.1145/2381966.2381969

Wills, C. E., & Zeljkovic, M. (2011). A personalized approach to web privacy: awareness, attitudes and actions. *Information Management & Computer Security 19(1)*,53-73. https://doi.org/10.1108/09685221111115863

Wu, K.-W., Huang, S. Y., Yen, D. C., & Popova, I. (2012). The effect of online privacy policy on consumer privacy concern and trust. *Computers in Human Behavior, 28*(3), 889-897. https://doi.org/https://doi.org/10.1016/j.chb.2011.12.008

Yadav, P., & Parekh, C. D. (2017). A report on CSRF security challenges & prevention techniques. *2017 international conference on innovations in information, embedded and communication systems (ICIIECS)* (pp.1-4). IEEE Xpolre. https://doi.org/10.1109/ICIIECS.2017.8275852

Yan, J., Liu, N., Wang, G., Zhang, W., Jiang, Y., & Chen, Z. (2009). How much can behavioral targeting help online advertising? *Proceedings of the 18th international conference on World wide web,* Madrid, Spain. https://doi.org/10.1145/1526709.1526745

Zheng, X., Jiang, J., Liang, J., Duan, H., Chen, S., Wan, T., & Weaver, N. C. (2015). Cookies Lack Integrity: Real-World Implications. *24th USENIX Security Symposium 7* (pp.707-721).USENIX Association.https://www.usenix.org/conference/usenixsecurity15/technical-sessions/presentation/zheng

Zhenyu, Q., Jing, X., Baoguo, L., & Fang, T. (2007). MBDS: model-based detection system for cross site scripting. *2007 IET Conference on Wireless, Mobile and Sensor Networks (CCWMSN07)(pp. 849-852)*. https://doi.org/10.1049/cp:20070282

Zhou, Y., & Wang, P. (2019). An ensemble learning approach for XSS attack detection with domain knowledge and threat intelligence. *Computers & Security*, *82*, 261-269. https://doi.org/10.1016/j.cose.2018.12.016

Zuiderveen Borgesius, F. (2014). Behavioural Sciences and the Regulation of Privacy on the Internet. *Draft chapter for the book 'Nudging and the Law-What can EU Law learn from Behavioural Sciences*. SSRN.https://ssrn.com/abstract=2513771

Zuiderveen Borgesius, F. (2015). Improving privacy protection in the area of behavioural targeting. *Available at SSRN 2654213*. http://dx.doi.org/10.2139/ssrn.2654213

Zuiderveen Borgesius, F. J., Kruikemeier, S., Boerman, S. C., & Helberger, N. (2017). Tracking walls, take-it-or-leave-it choices, the GDPR, and the ePrivacy regulation. *Eur. Data Prot. L. Rev., 3,* 353.

Zwilling, M., Klien, G., Lesjak, D., Wiechetek, Ł., Cetin, F., & Basim, H. N. (2022). Cyber security awareness, knowledge and behavior: A comparative study. *Journal of Computer Information Systems, 62*(1), 82-97. https://doi.org/10.1080/08874417.2020.1712269

Appendix A: Survey Participant Invitation and Information Notices

Project Title *Security & Privacy Awareness & Concerns of Computer Users Posed by Web Cookies and Trackers*

An Invitation

Hello, my name is Smah Almotiri, and I am currently a master's student at the design and creative technology at Auckland University of Technology. You are invited to take part in a study of the Security & Privacy Awareness & Concerns of Computer Users Posed by Web Cookies and Trackers. This survey is part of research conducted at Auckland University of Technology in Auckland, New Zealand and will contribute to my obtaining a master's degree qualification.
Thank you for reading this.

What is the purpose of this research?

As online Cookie's prevalence on web websites offers many advantages for the growth of websites but also endanger users' security and privacy, this research aims to assess computer users' awareness of the security and privacy risks of cookies on websites. Besides measuring the extent to which a user of the computer is concerned in using the data acquired over the Internet and web cookies. The findings of this research may be used for academic publications and presentations.

How was I identified and why am I being invited to participate in this research?

You are invited to participate in this research since you are a general Internet user, therefore your activity in Internet browsing and visiting websites will greatly contribute to the improvement of search results. You must be 18 years old or older to participate in this study.

How do I agree to participate in this research?

Your participation in this research is voluntary (it is your choice) and whether or not you choose to participate will neither advantage nor disadvantage you. You are able to withdraw from the study at any time and you do not have to answer any question you don't wish to. You agree to participate by visiting the survey link and completing the questionnaire. Completion of the questionnaire will be taken as your consent to participate.

What will happen in this research?

This research will include an online survey that covers most concerns connected to web tracking or cookies relating to privacy and security risks. You will be asked to complete a web-based questionnaire which almost will take you 9 to 15 minutes. Including visiting a website belonging to this study at the end of the survey. It does not contain any personal inquiries; all questions are relating in general to your activities over the Internet.

What are the discomforts and risks?

Participation will be anonymous and there will be no personally identifying information collected such as names, addresses and emails. If you do not want to answer any of the questions, you are able to withdraw from the survey or simply move on to the next question, at any time.

What are the benefits?

Whilst there will be no immediate advantages to individuals engaged in the study, this effort hopes to have a positive influence on how the privacy and security of online users could be considered and

maintained from web cookies and trackers. Results will be shared with participants in order to inform their professional work.

How will my privacy be protected? All the information that we collect about you during the course of the research will be kept strictly confidential. You will not be able to be identified or identifiable in any reports or publications. Any data collected about you in the online questionnaire will be stored online in a form protected by passwords and other relevant security processes and technologies. In addition, all data will be destroyed 6 years after completion of the of study findings.

What are the costs of participating in this research?
Participation is free of charge and completion time is expected to take about 15 minutes at a maximum.

What opportunity do I have to consider this invitation?
Taking part in this survey will be available for three weeks from the date of survey publication.

Will I receive feedback on the results of this research?
There will be a summary of the findings of this research provided on this link http://01code.net/samah/finding.php

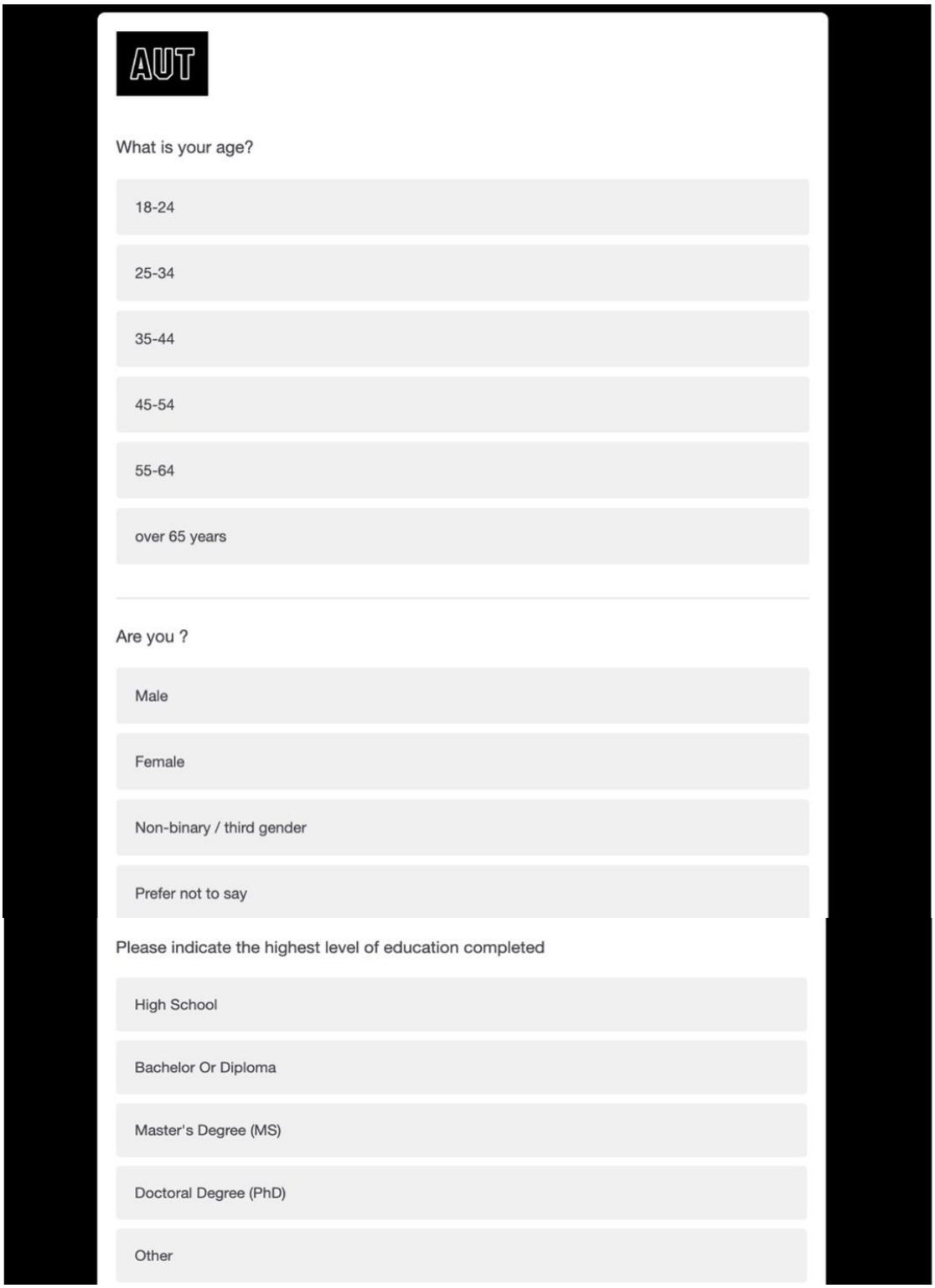
What do I do if I have concerns about this research?
Any concerns regarding the nature of this project should be notified in the first instance to the Project Supervisor, Dr. Alastair Nisbet, alastair.nisbet@aut.ac.nz, (+649) 921 9999 ext. 5879.
Concerns regarding the conduct of the research should be notified to the Executive Secretary of AUTEC, *ethics@aut.ac.nz*, (+649) 921 9999 ext 6038.

Whom do I contact for further information about this research?
Please keep this Information Sheet and a copy of the Consent Form for your future reference. You are also able to contact the research team as follows:

Researcher Contact Details:
Smah Almotiri, Auckland University of Technology, email: vdg1321@autuni.ac.nz
Project Supervisor Contact Details:
Dr. Alastair Nisbet, Auckland University of Technology, email: alastair.nisbet@aut.ac.nz

Approved by the Auckland University of Technology Ethics Committee on 8 November ,2021, AUTEC Reference number 21/379.

Appendix B: Survey Questionnaire



AUT

What is your age?

- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- over 65 years

Are you ?

- Male
- Female
- Non-binary / third gender
- Prefer not to say

Please indicate the highest level of education completed

- High School
- Bachelor Or Diploma
- Master's Degree (MS)
- Doctoral Degree (PhD)
- Other

Where are you located?

- Africa
- Antarctica
- Asia
- Oceania (Australia, New Zealand, etc.)
- Europe
- USA
- Canada
- Mexico
- Central America
- South America
- Middle East
- West Indies
- Other

AUT

How would you evaluate your ability to use a computer and smartphone?

- Unable to use
- Normal user
- Advanced user

How easy do you find it to register with a web site?

- I don't know, someone else is signing me up
- not easy to some extent
- very easy

Have you ever purchased online?

- Yes
- No

AUT

To what extent do you believe the following statement is true " The websites you visit gather your data while you browse the site" ?

- Definitely not
- Probably not
- Might or might not
- Probably yes
- Definitely yes

Which of your personal information do you think websites collect about you while you browse?

- Your IP or MAC address
- Geographical location
- Browser and device information
- The pages you visit and clicks you make
- Aggregated data on your visits to the site
- Your name
- Telephone number
- E-mail addresses

Do you think it is essential for websites to inform users about their data collection and usage practice ?

No

Yes

Are you aware that websites track your online activities?

No

Yes

Do you believe that websites sell personal information about their users?

No

Yes

Do you realize that hackers are attempting to access your sensitive personal data stored in cookies ?

No

Yes

Do you know how to manage cookies on your computer and erase them?

No

Yes

AUT

How often do you read a web site's privacy policy when visiting a web site that has a private policy and invites you to read that policy?

- Never
- I read once or twice
- I read at least half of the web sites' policies
- I read almost all of the web sites' policies
- I always read the policies

How frequently do you accept cookie policies on websites?

- I always accept it
- Most times I accept it
- Sometimes I accept it
- Rarely I accept it
- I never accepted it

Have you ever changed the cookie settings on the website?

- Yes, I always do
- Sometimes, I change it
- No, I accept it as it is

Have you ever refused to use a website's service due to its privacy policies?

Yes

Sometimes

No

Do you use anti-tracking tools that help you protect your privacy online by preventing your activities from being tracked? Example (Tor Browser, Privoxy)

Yes

No

Do you think that the purpose of collecting your personal information through cookies is clear?

No, not at all clear

Somewhat clear

Yes, it is clear

Is it important to you that the sharing of your web cookie-collected data is directly connected to the collection purpose?

- very important to me
- It's good to tell the user that
- I don't care too much
- never interested

Do you have any concerns about third parties having access to the data and what they could do with it?

- No
- Yes

Some cookie data is shared with third parties; is it important to you as to what legal basis the sharing is based on to justify this disclosure?

- Extremely important
- Very important
- Somewhat important
- Not so important
- Not at all important

Do you have concerns about where and how your personal data acquired by cookies will be stored?

No

Yes

Do you have any concerns regarding how long your gathered data should be retained?

No

Yes

Are you concerned about securing access to your data, such as being tracked for suspicious activity (e.g., surfing by employees)?

No

Yes

How concerned are you about the security risks connected with online cookies collecting and storing your information?

Extremely worried

Very worried

Somewhat worried

Not so worried

Not at all worried

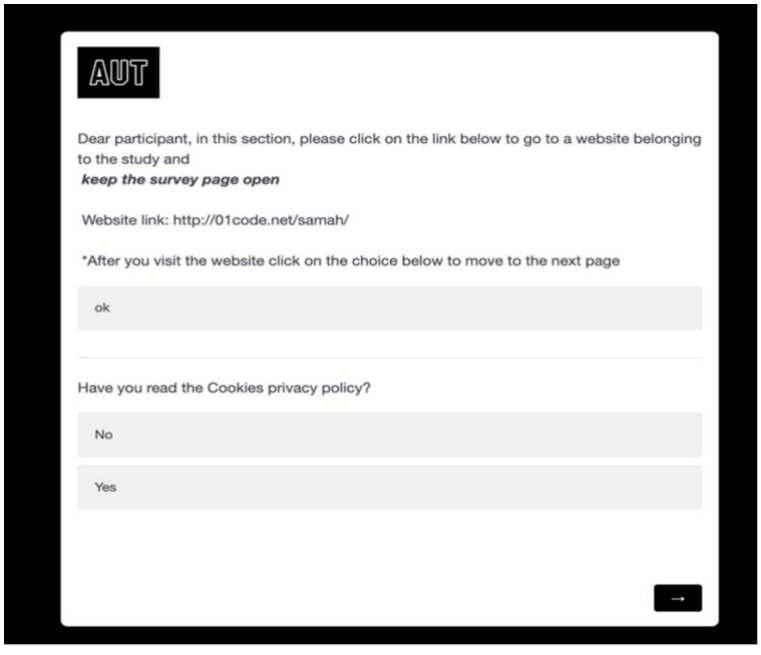
AUT

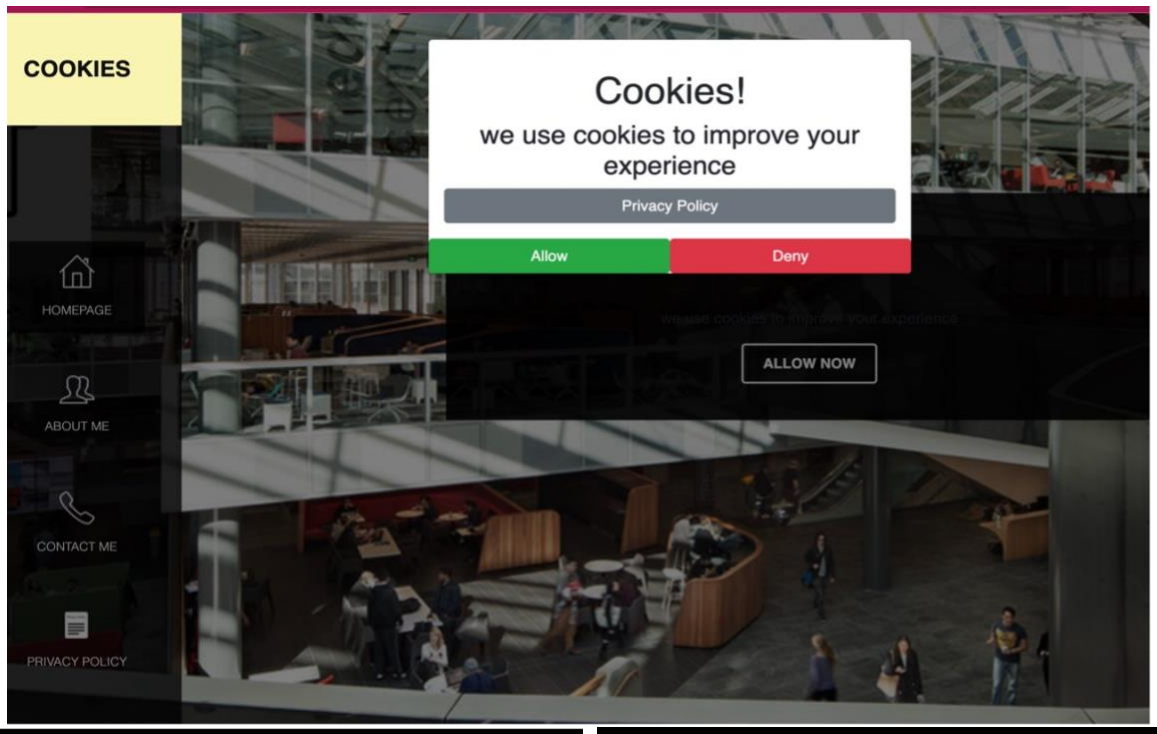
How would you rate your understanding of how your country's legislative bodies consider privacy rights?

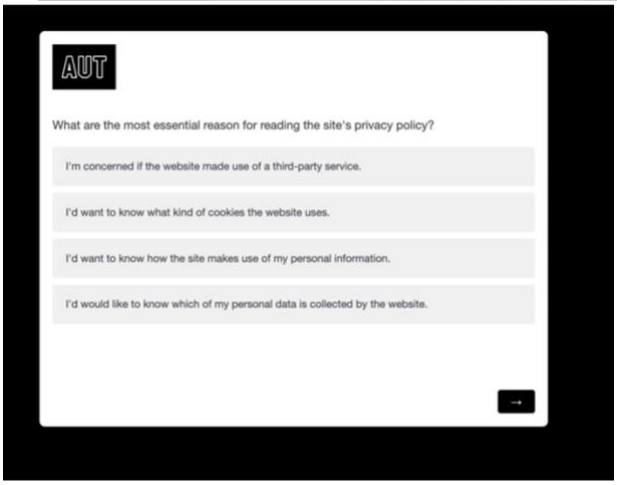
- 1 (very low)
- 2
- 3
- 4
- 5 (very high)

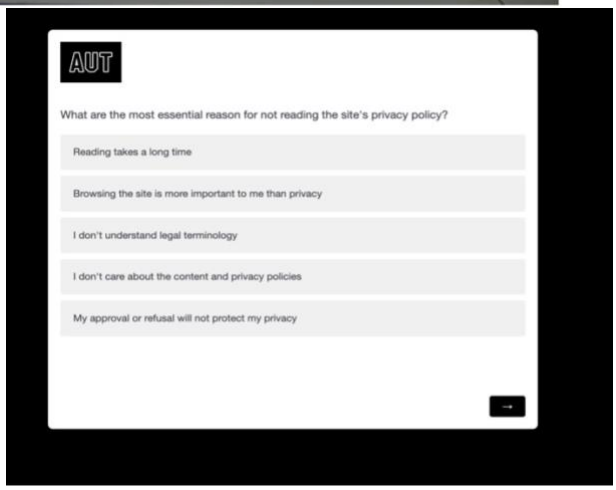
To what extant do you agree with this statement "Online service providers are obligated by the privacy protection regulations and requirements imposed by regulators such as GDPR/CCPA

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly agree




COOKIES
HOMEPAGE
ABOUT ME
CONTACT ME
PRIVACY POLICY
Cookies!
we use cookies to improve your experience
Privacy Policy
Allow
Deny
ALLOW NOW


AUT
What are the most essential reason for reading the site's privacy policy?
I'm concerned if the website made use of a third-party service.
I'd want to know what kind of cookies the website uses.
I'd want to know how the site makes use of my personal information.
I'd would like to know which of my personal data is collected by the website.


AUT
What are the most essential reason for not reading the site's privacy policy?
Reading takes a long time
Browsing the site is more important to me than privacy
I don't understand legal terminology
I don't care about the content and privacy policies
My approval or refusal will not protect my privacy