TALLINN UNIVERSITY OF TECHNOLOGY

School of Information Technologies

Chinmay Khandekar 177232IVCM

Cookie Security and its Implementation in the Light of GD-PR and E-Privacy Regulation

Master's thesis

Supervisor: Eneken Tikk, Dr. Iur.

Tallinn 2019

TALLINNA TEHNIKAÜLIKOOL

Infotehnoloogia teaduskond

Chinmay Khandekar 177232IVCM

Küpsiste turvalisus ja selle rakendamine isikuandmete kaitse üldmääruse (GDPR) ja E-Privaatsuse määruse valguses

Magistritöö

Juhendaja: Eneken Tikk, Dr. Iur.

Tallinn 2019

Author's declaration of originality

I hereby certify that I am the sole author of this thesis. All the used materials, references to the literature and the work of others have been referred to. This thesis has not been presented for examination anywhere else.

Author: Chinmay Vidyadhar Khandekar

13.05.2019

Abstract

The online identifiers or cookies are used for storing user information for personalization, authentication and advertisement. The usage of such web trackers drives the web economy where these trackers can be used to target users based on their online surfing behaviour. They are crucial in web user experience and their compromise leads to critical security problems. The European Union's General Data Protection Regulation (EU GDPR) enforced important aspects of data protection for web activity and consent for the usage of cookies. The proposed ePrivacy Regulation is set to address certain privacy and vague cookie-consent notices on websites and offer better privacy by design than that of the GDPR.

This research stands out as online identifier provisions in the legal framework, such that of GDPR-cookies, have seen recent research but none specific for the ePR and cookies. This study analyses the specifics of both GDPR and ePR in the context of the cookie, user consent and browser privacy settings. This study identifies key gaps in the cookie provisions within the GDPR such as the failure to clarify the explicit usage of the tracking cookies, lack of clarity in protecting session ID as personal data, limitations in the privacy by design to revoke the cookie consent and limited choices on acceptable cookies. The ePR would address the gaps in the GDPR such as bringing transparency in the cookie consent notices and choices, implementing privacy choices in the browser, and protecting the transmission of data.

The findings of this research increase user awareness encourage further discussion on privacy by design in the proposed ePR and web publishers and provide a comparative study for the web publishers for the development of applications for the existing GDPR and upcoming ePR.

This thesis, written in English, is 59 pages long, including 7 chapters, 33 figures and 6 tables.

List of abbreviations and terms

DPI	Dots per inch
TUT	Tallinn University of Technology
HTTP	Hyper Text Transfer Protocol
URL	Uniform Resource Location
ENISA	European Union Agency for Network and Information Security
GDPR	General Data Protection Regulation
RFC	Request for Comments
OWASP	Open Web Application Security Project
ePD	ePrivacy Directive
ePR	ePrivacy Regulation
DNT	Do Not Track
P3P	Platform for Privacy Preferences Project
DVWA	Damm Vulnerable Web App

Table of Contents

1 Introduction 11
1.1 Background 12
1.2 Problem Statement and Purpose 13
1.3 Significance of this Research 14
2 Literature Review 15
3 Methodology 17
4 Technical Background 19
4.1 About Cookies 19
4.1.1 HTTP or the (HyperText Transfer Protocol) 20
4.1.2 Session 22
4.2 Cookie Implementation 23
4.3 Cookie Components 25
4.4 Cookie Classification 26
4.4.1 Persistent Cookies 26
4.4.2 Session Cookies 27
4.4.3 Tracking cookie 27
4.4.4 Local Storage Objects 27
4.5 HTTP Cookie Vulnerabilities 28
4.6 Summary 29
5 Legal Framework for Cookies 30
5.1 General Data Protection Regulation (GDPR) 30
5.1.1 Transparency 31
5.1.2 Data protection by design and by default 33
5.1.3 Consent 33
5.2 Proposed e-Privacy Regulation 34
5.2.1 Transparency 35
5.2.2 Consent to the use of cookies 35
5.2.3 Protection for Communications data in transit 38
5.2.4 Tracking walls 38
5.2.5 Privacy and settings by Design 39
5.3 Analysis of regulatory frameworks based on cookie classification 39
5.3.1 Strictly Necessary Cookies 41
5.3.2 Performance Cookies 42
5.3.3 Targeting Cookies 43
5.3.4 Functionality Cookies 43
6 Cookie-related Issues 46
6.1 Cookies Consent and Browser 46
6.1.1 Do Not Track 54
6.1.2 Platform for Privacy Preferences (P3P) 54
6.2 Privacy evasive behaviour 55
6.2.1 Cookie related cybersecurity crimes 55

6.2.2 Susceptibility of Session IDs to Attack 56
6.2.3 Session Hijacking attack 57
7 Conclusions and Recommendations 63
7.1 Future Work 66
References 67

List of figures

Figure 1 Online Identifier creation process 20

Figure 2 HTTP Evolution 22

Figure 3 Cookie view in Browser 23

Figure 4 Third Party Cookies 24

Figure 5 Cookies stored in the browser from book.lufthansa.com 25

Figure 6 Cookies + Consent 40

Figure 7 Flow of Cookie protection through Regulations 40

Figure 8 Essential Cookies in use on https://nytimes.com from [98] 41

Figure 9 Performance Cookie used on https://nationalgeographic.com from [100] 43

Figure 10 Targeting Cookie from https://nytimes.com [98] 43

Figure 11 Functional Cookies from https://clym.io from [102] 44

Figure 12 Cookie consent usage disclaimer as seen on [106] and [107] 47

Figure 13 Cookie acceptance valued stored in the Cookie 47

Figure 14 Cookie consent request from [72] 48

Figure 15 Cookie in use [72] 48

Figure 16 Cookie usage notice from [110] 49

Figure 17 Cookies used on [110] 49

Figure 18 Cookies Consent from [114] 50

Figure 19 Cookie usage overview from [114] 50

Figure 20 Cookie consent from [115] 50

Figure 21 Group websites cookie comparison from [115] and [116] 51

Figure 22 Cookie consent on [117] 51

Figure 23 Cookies in use on [117] 52

Figure 24 Group cookie consent on [119] 52

Figure 25 Cookie used on [119] 53

Figure 26 Do Not Track setting in Google Chrome from [122] 54

Figure 27 XSS Workflow 59

Figure 28 Reflected XSS 59

Figure 29 BeEF Network Map 60

Figure 30 BeEF Control Panel 60

Figure 31 XSS Apache Cookie Disclosure Results (Session ID) 61
Figure 32 DVWA User Smithy 61
Figure 33 Smithy's Stolen Pablo Session 62

List of tables

Table 1 GDPR Cookie and Data Protection requirement 31
Table 2 Cookie provisions in the ePrivacy Regulations 34
Table 3 Difference between First Party and Third-Party cookies from [91] 37
Table 4 The Cookie Table 44
Table 5 Cookies which collect personal information 45
Table 6 Online Identifier with their Susceptibility 65

1 Introduction

The European Union's General Data Protection Regulation (GDPR), enforced on May 25, 2018, introduced significant changes to user privacy and user rights over their personal data. The regulation requires organizations to detail their processing of personal data, including the use of online identifiers also known as cookies. To complement the General Data Protection Regulation (GDPR), the European Union has proposed the e-Privacy regulation which will introduce further requirements of personal data protection and user privacy. The upcoming ePrivacy Regulation (ePR) is also known as the cookie law, since it covers the aspects of user privacy and cookies to compliment the GDPR.

This thesis analyses how the GDPR and the draft ePR address the use of online identifiers and whether the current and envisaged regulations adequately respond to privacy and security concerns regarding the use of cookies. Accordingly, this thesis first introduces the history and the concept of cookies as well as the different types of cookies used on websites. It then proceeds to discussing the privacy and security threats associated with the use of cookies. The author then analyses the requirements related to the use of cookies from the GDPR and the e-Privacy regulations, including the requirement of consent, transparency of their use etc. Finally, the author offers recommendations on potential steps to how to enhance existing user privacy and security regulation in relation to the use of cookies.

Cookies generated by the web server communicate through the HTTP protocol to keep track of all browser-server interactions. The evolution of the internet has raised both security and privacy concerns, where the different types of cookies cater to the needs of industry (i.e., advertising and e-commerce companies) [1]. The objective of this thesis is to determine the most exploited type of the cookies against the safety provisions of GDPR and the upcoming E-privacy regulation.

1.1 Background

One of the earliest web browsers was the Lynx, which in 1991 merged with the Mosaic Communications Corporation, and later become Netscape, in 1994. Cookies were developed by Lou Montulli to support web browsers [2]. A cookie is a text string that is placed on a client browser when it accesses a given server. The cookie is then transmitted back to that server in the header of subsequent requests. Initial support for such online identifiers was provided in 1994 in pre-1.0 versions of the Netscape browser [3], and the first standard for web cookies was published in 1997 [4] [5].

The term "cookie" is derived from a popular concept UNIX computing "magic cookie," which inspired both the idea and the name of it [6]. Online identifiers are created by the web server and shared between the web browser and the server via the HTTP Header, Cookie [7]. The most common way to track activities of a user via a web browser is through HTTP cookies, often set by third party analytics and advertising domains [8].

A security aware web user recognises HTTP cookies are a severe threat to privacy. Such users block, limit or periodically delete them; however, a majority remain ignorant. The cookies outside the browser also known as super cookies [9]. Awareness of supercookies is even lower, but political and public relations (PR) pressures [10] may eventually force firms to make their super cookies comply with the browser's standard HTTP cookie privacy settings [11].

The processing of personal data through websites is regulated by the General Data Protection Regulation (EU) 679/2016 (GDPR) [12]. The GDPR, is the primary data protection legal framework in the European Union, directly applicable to all Member States, repealing the older Data Protection Directive 95/46/EC [13]. The enforcement of the data protection principles, its obligations and rights enshrined in the Directive 679/2016 [12], the GDPR includes additional protection mechanisms to allow individuals to better control over their personal data, which is primarily a challenge in the dynamic online and mobile environment.

In addition to the GDPR provisions, cookie privacy and data protection requirements are also derived from the Directive on privacy and electronic communications 2002/58/EC [14] (ePrivacy Directive) currently in force. The latest amendment by the Directive

2009/136/EC [15] has added universal service and users' rights relating to electronic communications networks and services.

In January 2017, the proposal for a new ePrivacy Regulation [16] was made by the European Commission and is currently being debated in the European Parliament and the Council. The ePrivacy Regulation Proposal 15333/2017 [16] is concerned with the processing of personal data and the protection of privacy in the electronic communications which is currently under review to be updated and aligned with GDPR.

This proposal sets out essential provisions for the privacy of cookies, the confidentiality of communications channels and related metadata, the placement of software and data (e.g., cookies) on user devices and the regulation of privacy settings by levels with informed consent for tracking.

1.2 Problem Statement and Purpose

This study explores the phenomenon of cookies in the context of harmful and malicious uses of Information and communications technology (ICT). It examines the role of cookies in cybercrime and other malicious activities online and analyses the anticipated effect of enforcing the GDPR on such cases. The effect of GDPR can be observed by the reduction of third-party tracking cookies by 22% [15]. However, there are still large numbers of websites which were observed to be using vague and unclear cookie consent requests [17]. Another effect of GDPR was seen with massive reduction in botnet based spamming and phishing attacks by e-mails to Domain Name [18] registrants as WHOIS data [19] for Domain Names is no longer public [20].

The thesis first explores how cookies reveal the user's preferences and the role of cookies in cybercrime and other harmful online activities. It then goes on to study the provisions for cookie and its data in the GDPR, the existing ePrivacy Directive, and the upcoming ePrivacy Regulation. Using that, a gap analysis is done to understand the effectiveness of the provisions under these laws against cookie-based cybercrime. Finally, the author offers recommendations on reducing the risks related to cookies and the scope for future work.

To reduce the opportunity of using cookies for malicious purposes, this thesis explores the regulations provided in the GDPR and the draft e-Privacy Regulation. The research question and theme of this thesis is to determine "Does the GDPR and the upcoming ePrivacy Regulation adequately and sufficiently protect users from cookie-based cyber-attacks?". This research question is examined further by answering the following sub-questions:

- What are cookies and how do they affect user privacy?
- What malicious and harmful uses of cookies have emerged?
- Which legal requirements are applicable to the use of cookies under
    - The GDPR
    - The ePrivacy Regulation Proposal 15333/2017
- Which practical problems related to malicious and harmful uses of cookies remain after the enforcement of the GDPR and the proposed e-Privacy Regulation?

1.3 Significance of this Research

The research stands out as online identifier provisions in the legal framework, such that of GDPR-cookies, have seen recent research but none specific for the ePR and cookies. This study analyses the specifics of both GDPR and ePR in the context of the cookie, user consent and browser privacy settings.

The findings of this research increase user awareness encourage further discussion on privacy by design in the proposed ePR and web publishers and provide a comparative study for the web publishers for the development of applications for the existing GDPR and upcoming ePR.

2 Literature Review

To understand cookies, their background needs to be explored. They have been studied in detail through research since 1994 with the Netscape browser functioning [21]. The Internet Engineering Task Force (IETF) in 1995 initiated a standardisation process for cookies [6]. In 2000, IETF published RFC 2965 "HTTP State Management Mechanism", which specifies methods to create a stateful session with HTTP requests and responses [22]. During the standardization process, privacy concerns have been raised, especially for third-party cookies [1]. The online identifiers have an important property to send the cookie automatically, which makes it very unobtrusive. The cookie enables positive and empowering uses to improve browsing experience by adding a social or personalized layer [21]. There is existing literature review on cookies, their vulnerability and what remedies could be offered to enhance their protection by Jussila [23].

There have been many studies on internet user privacy and the potential for information leakage via web cookies. Privacy is a central concern of internet users and this work examines one privacy issue which is the potential to track and correlate knowledge about a user's actions across seemingly unrelated websites in the work by Krishnamurthy *et al.* [8] .

An Empirical study on web cookies conducted revealed that 3rd party cookies outnumbered first party cookies by Cahn *et al.* [5]. When web applications deploy more than one type of cookies, 52.1% use both first- and third-party cookies and 74.1% use a combination of session and persistent cookies as defined in the study by Tappenden and Miller [24]. A similar study was undertaken to analyse the usage of cookies for targeting marketing purposes, where online user behaviour within the advertisement network was used to develop a more responsive and effective communication campaign by Zarouali *et al.* [25]. A similar web tracking technique research by Iskander *et al.* [26] highlighted privacy issues from the third-party cookies loading and sending content to the website. The initial analysis by Krishnamurthy and Wills [27] on existing privacy protection techniques showed their limitations in preventing 3rd party cookie tracking. The research by Roesner *et al.* [28] examines 3rd party tracking of cookies using a small set of web crawlers. Another study on web tracking and policy by Mayer and Mitchell [29] analysed third-party cookies with Do Not Track enabled and other privacy protective browser tools.

There has also been research done on HTTP session management in the context to cookie security by Ayandi *et al.* [30] in terms of reverse proxy components but the session management architecture would help us understand session creation in relevant context of cookies. A study was conducted in Finland by Ruohonen and Leppänen [31] to analyse persistent third-party cookie presence on Finnish News portals. Which showed ambiguous cookie consent in compliance with GDPR and show no success on the Do Not Track initiative or opt-out.

The Cookie has been studied before and after GDPR enforcement. A pre-GDPR enforcement study by Bader *et al.* [32] presents how web users have misconceptions about websites using cookies, when in fact the reason for using cookies was to track their paying customers for measuring the performance of their services and products, mainly for marketing purposes. Research done on online profiling post GDPR by student number 8027 of Oslo University [33] which proved GDPR enhances protection in privacy and data protection in the digital context but, yet regulations provide flexibility to agents engaged in online behaviour tracking. The user transparency of cookies syncing with third-party under GDPR was researched by Urban *et al.* [34] where third-party data sharing was studied before the GDPR implementation and post GDPR to analyse how data is shared in the advertising providers network. A similar analysis was undertaken to measure GDPR's impact on web privacy by Degeling *et al.* [35], with the research establishing the fact that while GDPR has bought in transparency, it lacks effective mechanisms for users to deny processing of personal data.

The upcoming ePrivacy Regulation, which is an amendment to the existing ePrivacy Directive to support the GDPR. There was no literature found on the ePrivacy Directive in the context of cookies until the time this thesis was written. Cookies have long been studied as privacy evasive behavior as mentioned previously, but there is no literature available on first-party cookie analysis specific to session cookies focussing on their privacy evasiveness. Our study identifies provisions for cookies that forthcoming directives could use to regulate cookies that are excluded from the current framework.

3 Methodology

The research followed the guidelines of systematic literature review research methods. Previous literature was studied to develop knowledge of the essence of cookies, their vulnerabilities, legal framework provision in GDPR and cookie based cyber-attacks. The cookies study helps identify their susceptibility and implementation structure which makes them an easy target for the hackers. The existing research gaps were identified to come up with specific research questions. The research problem was defined based on the study of previous literature, to understand whether the existing and proposed regulations adequately address the challenge of cookie security.

There has been extensive literature already in context of cookies. The evidence was gathered from available resources addressing cookie development, classification and vulnerabilities that exist. Cookie based cyber-attacks is a vague terminology. The objective of this research is to study cookies which were excluded under the GDPR [12] and ePrivacy Directive [14] [15]. The study of frameworks is undertaken to understand each article and recital which would be applicable to the online identifiers and the type of cookies that are protected.

The research addresses practical problems which exist post GDPR enforcement with respect to privacy, browser-cookies and consent. Specifically, certain portals were analysed to understand how essential about cookie consent is displayed to the user. An experiment was setup where consent for cookies used on a news portal specifically was analysed as existing research showed changes in the third-party tracking methods post GDPR. The findings will address cookies which are excluded by the existing legal framework but continue to be used for cyber-security attacks.

The flow and structure of the research is as follows:

Cookie: The background and technical implementation of cookies is analysed. The background leads the research into classification of online identifiers used on websites. OWASP vulnerabilities regarding cookies are then analysed.

Legal Framework: The European Union has several legal legislatures, GDPR and ePR were specifically chosen for this research as GDPR was implemented last year and the much anticipated ePR currently debating in the European Parliament is likely to be enforced later in 2019 or early 2020. The research analyses cookie specific provisions and compares them on how they will complement each other post enforcement. The cookie classification was used to understand which cookies require no consent for usage on websites. The two frameworks are compared to understand which cookies are excluded from consent which can be an attack vector for cyber-attacks.

Practical issues which exist for excluded cookies: The result is derived from legal framework analysis post which cookies are categorized to fit in the post GDPR requirement. Then a research is undertaken to understand what practical issues exist in cookie from the cybersecurity perspective. To understand practical issues, an experiment was setup to study the cookie consent notice with the cookie used on the website:

Site Selection**:** The Alexa ranking of top News portal list in Estonia was used to analyse compliance with both frameworks in context of cookies. The following list of websites was chosen for two primary reasons. First, the Article 6 of GDPR establishes display of cookie usage information with option to deny consent. The second reason being the provision in Recital 20 [36] of the proposed ePR [16] which prohibit ambiguous cookie consent through which web publishers monetize on third-party cookies.

Tools for Analysis**:** To analyse the cookie used, two browser extension Edit this Cookie[1] and Cookie Checker[2] by Cookiebot were used. These extensions analyse consent with cookie usage that detects privacy issues to comply with ePR and GDPR.

Session: Session cookies have been identified as being excluded from consent in both the legal frameworks that have been analysed. The vulnerability and cyber-crime instances are quite high for session-based attacks. Susceptibility of session ID is discussed. Where a link is established of session ID and cookie to the personal data through the experimentation setup. In the last stages of research another experimentation is conducted to prove session cookies and session ID theft is possible which offer no legal provisions.

[1] http://www.editthiscookie.com

[2] https://www.cookie-checker.com

4 Technical Background

This chapter builds the background on cookies, their implementations and provisions in RFC. The browser footprint generation of cookies is discussed along with cookie components, values and classification. The cookies like any other technology is subject to vulnerabilities.

4.1 About Cookies

Cookies are a text files or HTTP (Hypertext Transfer Protocol), as are also known as online identifiers. When visiting a website, a browser asks the website in question for a page from the server. The server sends back the page and attaches a piece of data. The page the browser receives from the server is the one shown to the user. The data that is sent along with the page is called a Cookie. This cookie will be stored on the user's computer and will be used every time they revisit the same webpage, unless they delete the cookie. The data stored in the text files may differ for each website with the specific information stored depending on what the owner of the website wants to track from its visitors. The cookie helps a website remember its visitors using identifiers such as "user=1337".

Cookies are a common mechanism for the websites to maintain its state during an e-commerce transaction or to maintain personalized content to the user based on settings (i.e: language or region) [8].

Cookies are not very data intensive, varying in size from 4 kilobytes to 100 kilobytes and are stored in the web browser or in a special folder on the device. They store all sorts of information about the user including their surfing behavior, browser agents, and the websites visited.

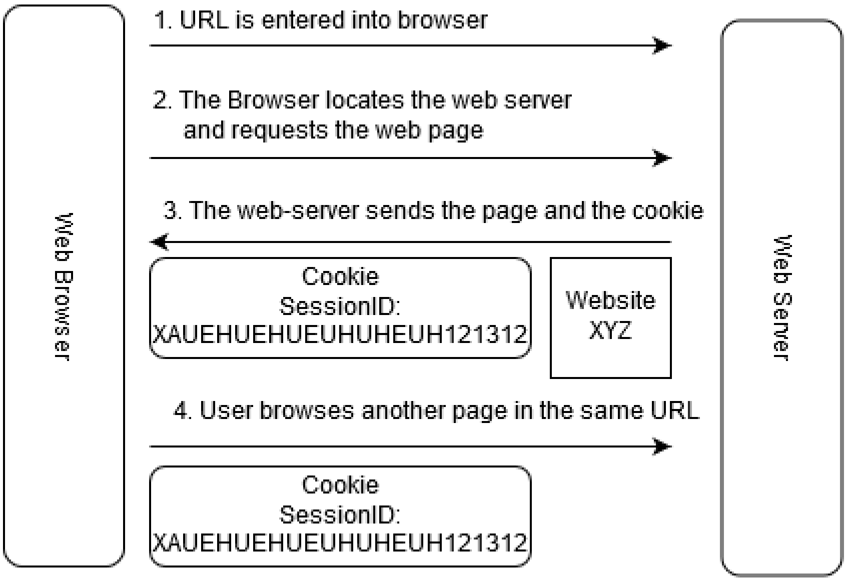

1. URL is entered into browser
2. The Browser locates the web server and requests the web page
3. The web-server sends the page and the cookie
Cookie SessionID: XAUEHUEHUEUHUHEUH121312
Website XYZ
4. User browses another page in the same URL
Cookie SessionID: XAUEHUEHUEUHUHEUH121312
Web Browser
Web Server

Figure 1 Online Identifier creation process

The web server distinguishes each user session and communication with a unique identifer and the process is visually explained in Figure 1 above. Each online identifier has its own file and sets a session ID in the cookie header for each web server where the Uniform Resource Locater (URL)is located. Hence a new cookie file is created by the web server on every visit to the URL. The page is received from the server side by the browser, data is retrieved, and the existing cookie file is updated if a user authenticates or enables any function or preference on the webpage. The cookie file remains on the browser/ computer until deleted [32].

The advent of session identification cookies resolved several the issue of keeping track of user sessions and connection timeouts. Session identification cookies are used for establishing and keeping track of the browser-server relationship throughout the session [37]. The flexibility which cookies offered due to their size helps to keep connections open and close them once the browser closes or moves away from the website [38].

4.1.1 HTTP or the (HyperText Transfer Protocol)

HTTP is a request-reply communication protocol, with each version of HTTP different in their interaction models. The request-reply communication model was the most basic communication model designed [39]. The message exchange pattern is one of common client-server architectures. The architecture allows a client to request information from the web server that receives and returns a responses message [40]. The Figure 2 shows example of HTTP evolution through its three versions (v1.0, v1.1 and v2.0) and their request-reply interaction model.

The HTTP version 1.0, maintained and supported a content ratio of 1:1 which means TCP connection was open and then closed after single HTTP request/reply pair. So only one request-response pair was supported per connection [41] [39].

In HTTP 1.1, a keep-alive function was introduced. It provided the ability for a reusing the TCP connection for sending multiple requests to the web server without waiting for a response [42]. HTTP 1.1 specifies that the server must send its responses to the requests in the same order in which it was received, the browser starts listening once all requests are sent [39]. This modification was done to support the growing complexity of the webpages that included web browser component and web elements that needed to be transferred from the server to the client [43].

The HTTP evolution was required with the internet growing as it became more than just a simple function from transferring text and images for the web server to a browser; it became a platform for web applications [7]. The browser-server interactions were more elaborate than the simple request-response interactions facilitated by the client/server model of the web [40] [39].

The new HTTP 2.0 introduced multiplexing mechanism, where the browser and server could send and receive responses asynchronously through a single TCP connection [44]. Of many changes which the HTTP 2.0 bought over web the noticeable ones were a many-request-per-connection model and the exchange of headers and the transition from a text-based transfer to binary [43].

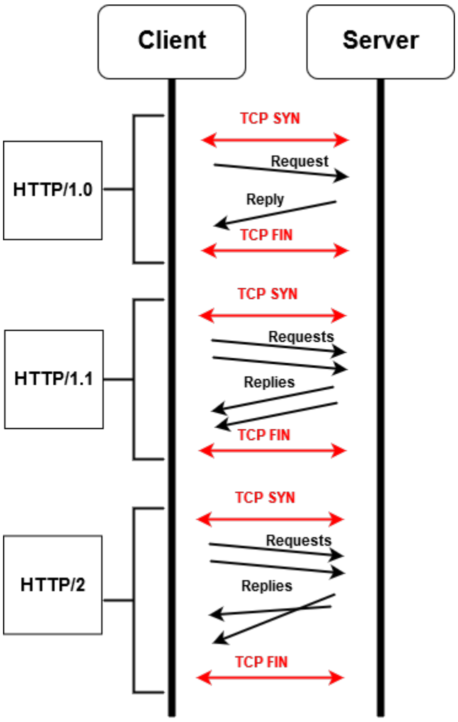


Figure 2 HTTP Evolution

The HTTP is a stateless protocol but is one of the primary application transport protocol of the web. The HTTP had the capability to track through the use of cookies and sessions come into play for the application [7].

4.1.2 Session

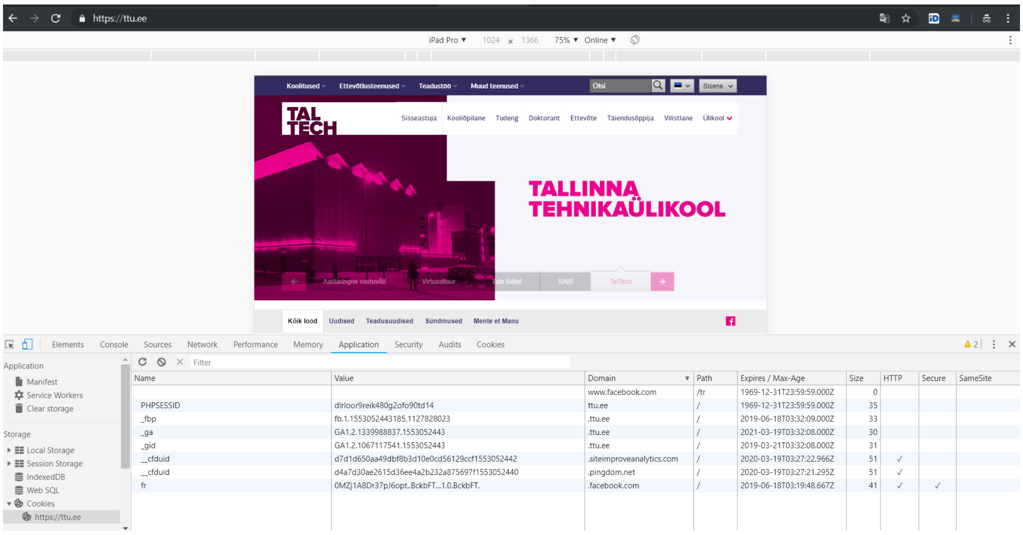
Session is one the many methods used to maintain the web and application communication state. These are simple chunks of data which is stored in the memory that is associated with every TCP connection made to a web or application server and serve as in-memory storage for information in on the HTTP-based applications [7].

The process of managing the state of the web-based client is performed by session. The session is unique to every TCP connection or user browser. The session ID is established when the user first connects to the web or application server. The Session IDs are used by the application to uniquely identify a client browser, while background (server-side) processes are used to associate the session ID with a level of access once they authenticate [45].

An example of session usefulness is thatwhile shopping on a e-commerce website, items added to the shopping cart remain throughout a "session" because every item added to the shopping cart is updated in the session and also updated on the web server. This is possible even without logging into the e-commerce website as the server has identified the browser uniquely with the session ID [7].

4.2 Cookie Implementation

A web server uses "cookies" to offer an enhanced user experience when a user visits a website. By default, there are some cookies which are optional and some which are required for the smooth functioning of the website as seen in Figure 3. Optional cookies allow users to choose whether to share data or not which will limit the personalization of the website.



Name	Value	Domain	Path	Expires / Max-Age	Size	HTTP	Secure	SameSite
		www.facebook.com	/tr	1969-12-31T23:59:59.000Z	0			
PHPSESSID	dlrloor9reik480g2ofo90td14	ttu.ee	/	1969-12-31T23:59:59.000Z	35			
_fbp	fb.1.1553052443185.1127828023	.ttu.ee	/	2019-06-18T03:32:09.000Z	33			
_ga	GA1.2.1339988837.1553052443	.ttu.ee	/	2021-03-19T03:32:08.000Z	30			
_gid	GA1.2.1067117541.1553052443	.ttu.ee	/	2019-03-21T03:32:08.000Z	31			
__cfduid	d7d1d650aa49dbf8b3d10e0cd56129ccf1553052442	.siteimproveanalytics.com	/	2020-03-19T03:27:22.966Z	51	✓		
__cfduid	d4a7d30ae2615d36ee4a2b232a875697f1553052440	.pingdom.net	/	2020-03-19T03:27:21.295Z	51	✓		
fr	0MZj1A8Dr37pJ6opt..BckbFT...1.0.BckbFT.	.facebook.com	/	2019-06-18T03:19:48.667Z	41	✓	✓	

Figure 3 Cookie view in Browser

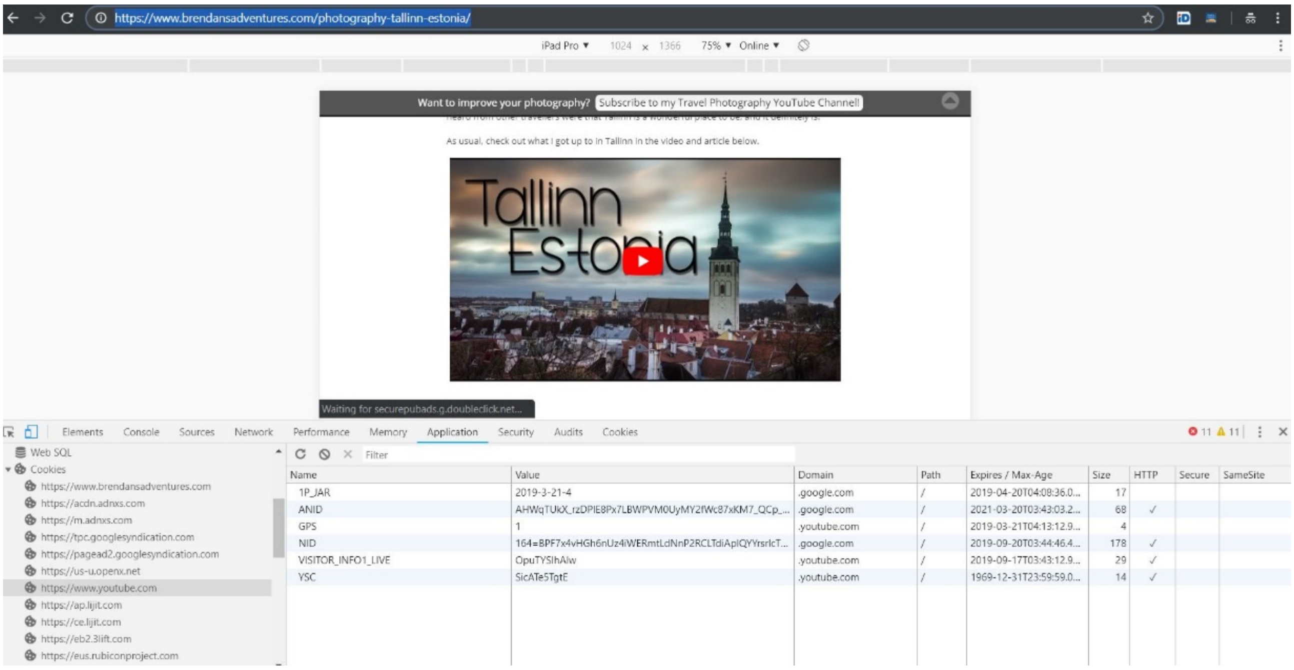
Various RFC's have been published over time that focus on cookie functioning in the browser. RFC 7231 [46] defines the value which is commonly generated by the generic web server for cookies. RFC 6265 [47] obsoletes RFC 2109 [4] and RFC 2965 [22] which prevents the browser from sharing cookies with the web server. The development of RFC 6265 was led by to the advancement of modern browsers which performed stateful session over the stateless HTTP Protocol. Even though the usage of cookies degrades security and privacy, cookie and set-cookie header fields are widely used over the internet.

Cookie implementations offer the following flexibility to web developers:

- A cookie stores the current state of a webpage based on the selections of every individual user thereby helping users navigate between the different pages of a website efficiently.
- A cookie can help identify user behavior and keep track of user clicks on most frequently visited pages.
- Cookies are unique identifiers to distinguish each user and browser request.

In today's competitive market, online advertising serves as a critical source of income for websites, applications, and hosted platforms. Third-party cookies enabled advertisers to maintain a common platform for exchange of cookies for displaying relevant advertisements based on their online behavioral advertising (OBA) which gets data from a user’s surfing behavior and is uniquely identified by the IP Address. For this form of targeted advertising, data is collected by installing either "personally identifiable cookies or non-personally identifiable information [34].

The third party and tracking online identifiers in addition to the one established by the website visited are also created as shown in Figure 4. These are mostly tracking cookies, a YouTube player is embedded in the webpage which calls for storing of YouTube cookies, tracking cookies from advertisements and other web analytic scripts.

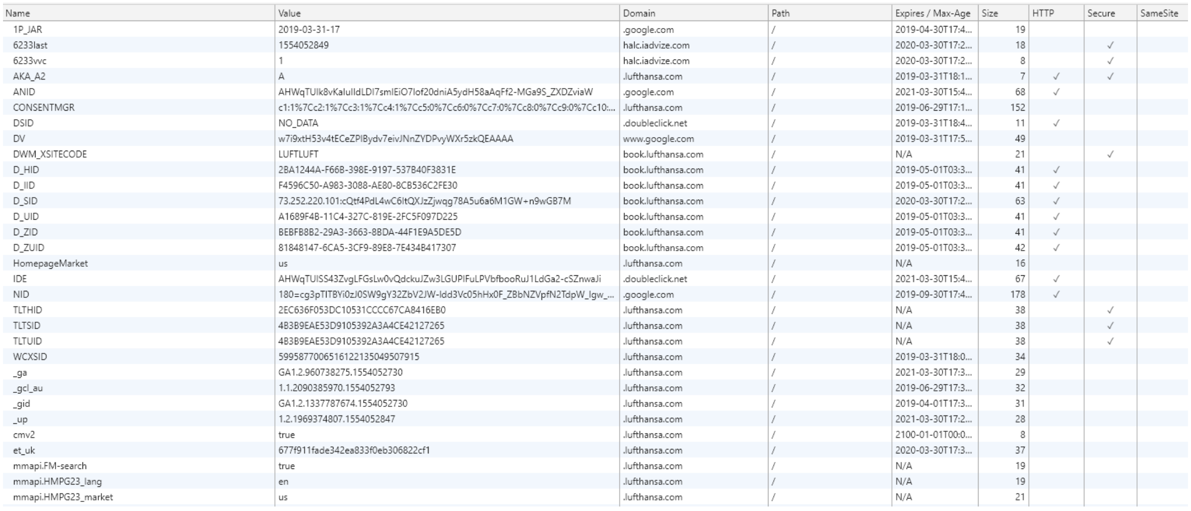

https://www.brendansadventures.com/photography-tallinn-estonia/
iPad Pro ▾ 1024 × 1366 75% ▾ Online ▾
Want to improve your photography? Subscribe to my Travel Photography YouTube Channel!
As usual, check out what I got up to in Tallinn in the video and article below.
Tallinn Estonia
Waiting for securepubads.g.doubleclick.net...
Elements
Console
Sources
Network
Performance
Memory
Application
Security
Audits
Cookies
11
11
Web SQL
Cookies
https://www.brendansadventures.com
https://acdn.adnxs.com
https://m.adnxs.com
https://tpc.googlesyndication.com
https://pagead2.googlesyndication.com
https://us-u.openx.net
https://www.youtube.com
https://ap.lijit.com
https://ce.lijit.com
https://eb2.3lift.com
https://eus.rubiconproject.com
Filter

Name	Value	Domain	Path	Expires / Max-Age	Size	HTTP	Secure	SameSite
1P_JAR	2019-3-21-4	.google.com	/	2019-04-20T04:08:36.0...	17			
ANID	AHWqTUkX_rzDPIE8Px7LBWPVM0UyMY2fWc87xKM7_QCp_...	.google.com	/	2021-03-20T03:43:03.2...	68	✓		
GPS	1	.youtube.com	/	2019-03-21T04:13:12.9...	4			
NID	164=BPF7x4vHGh6nUz4iWERmtLdNnP2RCLTdiApIQYYrsrlcT...	.google.com	/	2019-09-20T03:44:46.4...	178	✓		
VISITOR_INFO1_LIVE	OpuTYSIhAlw	.youtube.com	/	2019-09-17T03:43:12.9...	29	✓		
YSC	SicATe5TgtE	.youtube.com	/	1969-12-31T23:59:59.0...	14	✓		

Figure 4 Third Party Cookies

4.3 Cookie Components

A web cookie is a formatted string consisting of semi-colon which separates the value pairs. Figure 5 gives an insight on how cookies are appearing in the browser on visiting a website. The Cookie would appear as follows:



Name	Value	Domain	Path	Expires / Max-Age	Size	HTTP	Secure	SameSite
1P_JAR	2019-03-31-17	.google.com	/	2019-04-30T17:4...	19			
6233last	1554052849	halc.iadvize.com	/	2020-03-30T17:2...	18		✓	
6233vvc	1	halc.iadvize.com	/	2020-03-30T17:2...	8		✓	
AKA_A2	A	.lufthansa.com	/	2019-03-31T18:1...	7	✓	✓	
ANID	AHWqTUIk8vKalulldLDI7smlEiO7lof20dniA5ydH58aAqFf2-MGa9S_ZXDZviaW	.google.com	/	2021-03-30T15:4...	68	✓		
CONSENTMGR	c1:1%7Cc2:1%7Cc3:1%7Cc4:1%7Cc5:0%7Cc6:0%7Cc7:0%7Cc8:0%7Cc9:0%7Cc10:...	.lufthansa.com	/	2019-06-29T17:1...	152			
DSID	NO_DATA	.doubleclick.net	/	2019-03-31T18:4...	11	✓		
DV	w7i9xtH53v4tECeZPIBydv7eivJNnZYDPvyWXr5zkQEAAAA	www.google.com	/	2019-03-31T17:5...	49			
DWM_XSITECODE	LUFTLUFT	book.lufthansa.com	/	N/A	21		✓	
D_HID	2BA1244A-F66B-398E-9197-537B40F3831E	book.lufthansa.com	/	2019-05-01T03:3...	41	✓		
D_IID	F4596C50-A983-3088-AE80-8CB536C2FE30	book.lufthansa.com	/	2019-05-01T03:3...	41	✓		
D_SID	73.252.220.101:cQtf4PdL4wC6ltQXJzZjwqg78A5u6a6M1GW+n9wGB7M	book.lufthansa.com	/	2020-03-30T17:2...	63	✓		
D_UID	A1689F4B-11C4-327C-819E-2FC5F097D225	book.lufthansa.com	/	2019-05-01T03:3...	41	✓		
D_ZID	BEBFB8B2-29A3-3663-8BDA-44F1E9A5DE5D	book.lufthansa.com	/	2019-05-01T03:3...	41	✓		
D_ZUID	81848147-6CA5-3CF9-89E8-7E434B417307	book.lufthansa.com	/	2019-05-01T03:3...	42	✓		
HomepageMarket	us	.lufthansa.com	/	N/A	16			
IDE	AHWqTUISS43ZvgLFGsLw0vQdckuJZw3LGUPIFuLPVbfbooRuJ1LdGa2-cSZnwaJi	.doubleclick.net	/	2021-03-30T15:4...	67	✓		
NID	180=cg3pTITBYi0zJ0SW9gY32ZbV2JW-Idd3Vc05hHx0F_ZBbNZVpfN2TdpW_Igw_...	.google.com	/	2019-09-30T17:4...	178	✓		
TLTHID	2EC636F053DC10531CCCC67CA8416EB0	.lufthansa.com	/	N/A	38		✓	
TLTSID	4B3B9EAE53D9105392A3A4CE42127265	.lufthansa.com	/	N/A	38		✓	
TLTUID	4B3B9EAE53D9105392A3A4CE42127265	.lufthansa.com	/	N/A	38		✓	
WCXSID	59958770065161221350495O7915	.lufthansa.com	/	2019-03-31T18:0...	34			
_ga	GA1.2.960738275.1554052730	.lufthansa.com	/	2021-03-30T17:3...	29			
_gcl_au	1.1.2090385970.1554052793	.lufthansa.com	/	2019-06-29T17:3...	32			
_gid	GA1.2.1337787674.1554052730	.lufthansa.com	/	2019-04-01T17:3...	31			
_up	1.2.1969374807.1554052847	.lufthansa.com	/	2021-03-30T17:2...	28			
cmv2	true	.lufthansa.com	/	2100-01-01T00:0...	8			
et_uk	677f911fade342ea833f0eb306822cf1	.lufthansa.com	/	2020-03-30T17:3...	37			
mmapi.FM-search	true	.lufthansa.com	/	N/A	19			
mmapi.HMPG23_lang	en	.lufthansa.com	/	N/A	19			
mmapi.HMPG23_market	us	.lufthansa.com	/	N/A	21			

Figure 5 Cookies stored in the browser from book.lufthansa.com

Name: The attribute name given to a cookie sent by the web server. This is a unique keyword to identify cookies to a web server [37] [47] [48].

Value: The value attributes to the data the cookie is transmitting between the web server and the browser. The value is either in plain text, encrypted or obfuscated for security and privacy reasons. The cookie value is stored in the client's computer or web browser [37] [47] [48].

Domain: The domain is the web server from where the cookies originated. This would allow the browser to send cookies back to the appropriate web server from where it originated. The domain column from Figure 4 also distinguishes First and Third-party cookies. The first party cookie is the cookies which the URL generates that in this example is book.lufthansa.com where the third-party cookies are distinguished by a "." or dot before the URL of origination in this column [37] [47] [48] [49].

Path: The path attribute restricts the browser from sending a cookie back to the web server. The path is the location of the cookie location on the URL of the web server. The path is set based on the web application which created the cookie. i.e.: "/php/" the cookie will only be available within the php directory and all sub-directories of php [37] [47] [48] [49].

Expires / Max-Age: This is the validity of the cookies in date/time till the cookie lives or is valid till. If the cookie posts an expiry or maximum age, then it is termed as a permanent cookie, while N/A specifies a Session Cookie [37] [47] [48] [49].

HTTP: The Http or HTTP Only is a unique cookie flag to tell the browser not to display the cookie through other communication channels except web server (HTTP/HTTPS) as it only supposed to be communicating with the web server. Setting cookies to HTTP only can reduce make common XSS attack harder to be successful [37] [47] [48] [49] [50].

Secure: The secure is another cookie flag for the browser limiting communication of the cookies only over secure and encrypted transmission, directing transmission only over an SSL or HTTPS [37] [47] [48] [49] [51].

Same Site: The Same-Site flag are relatively new and is supported by all browsers. The cookie flag prevents the browser from sending the cookie along with other cross-site requests [52]. The flag are used for protection against cross-site request forgery attacks [51] [53].

4.4 Cookie Classification

Cookies can be classified into three types as described [54], but for research purpose the classification was expanded with an additional category of local storage cookies as it has some unique characteristics unlike others:

4.4.1 Persistent Cookies

These are anonymous tracking cookies which are used for analytical purposes for tracking the number of websites, page visit specifics, duration of the visit, the language of the website, authentication information and user settings for the website. These cookies are typically not deleted after closing the browser, hence called persistent. The cookies data is available even after user's session termination or browser closed. Web applications with the login credential "remember me" functionality use these cookies [55] [23] [32].

4.4.2 Session Cookies

These are per visit cookies which are stored in-browser and are per session cookies. These are deleted when the user closes the browser. These cookies are also shared in case you would want online chat support or customer support. These cookies contain the information of session, credential and authentication data for the web server [23] [32] [51].

4.4.3 Tracking cookie

These are cookies which are permanently stored on a user's computer. These are used for tracking technologies that do not rely on HTTP cookies. These cookies have similar functionality to a regular cookie. They are used to store information like browsing history, authentication data, and ad-related data [23] [32] [51].

4.4.4 Local Storage Objects

Local Shared Objects (LSO), also called Flash Cookies, are pieces of data that are saved on your computer by websites that use Adobe Flash software [56]. Adobe Flash Player uses its own proprietory cookies, which are not manageable through your browser settings but are used by the Flash Player for similar purposes, such as storing preferences or tracking users [57]. These works like persistent cookies and continue to store in user's PC even after session termination [55].

The Local storage cookies work differ from web browser cookies as they are createaed and set by the browser . But these function specific cookies, as the website is restricted from storing all data in one cookie. The flash cookies since are proprietary cookies which are not stored in the browser [58]. There are some other function specific cookies in the browser

4.4.4.1 Zombie cookies

Zombie cookies or ever cookies are those cookies that "respawn". They recreate themselves automatically even after being deleted [59]. This is possible because this cookie is stored in multiple locations: Flash Local shared object, HTML5 Web storage and other client-side and even on the web server locations [60]. The cookies are a combination of various tracking mechanisms, each reinforcing the others, and can identify a client even when standard cookies and Flash cookies have been removed [61].

4.4.4.2 Super cookies

Super cookies are more persistent, more difficult to delete and are more effective at tracking user data. The super cookies are flexible and can grow as large as 100 KB each, compared with 4 KB for a regular tracking cookie. Super cookies are also browser-independent, so they're still able to track user activity even if the user switches browsers [62]. The browsers can't access the super cookies as they are stored in folders not accessible to the browser [61] [33].

4.5 HTTP Cookie Vulnerabilities

The cookies contain personal and public information about the user browsing the website, which makes them susceptible to vulnerabilities and the lucrative proposition to hack. The vulnerabilities are highlighted below:

Cookie manipulation is one of the known cookie vulnerabilities which can alter data stored in the cookie. The cookie manipulation ranges from session tokens to arrays which make product or pricing authorization decisions. Cookie poisoning can leave severe vulnerabilities such as SQL Injection and cross-site scripting. A shopping cart example shows the manipulation in clear text [63].

Part of a Shopping Cart Application's Cookie:

```
item1_ID=12369&item1_pr=27,95&item2_ID=10334&item2_pr=19,95
                   > Total Amount: $47,90
```

Manipulated Cookie:

```
item1_ID=12369&item1_pr=0,95&item2_ID=10334&item2_pr=1,95
                  > Total Amount: $2,90
```

HTTP Cookie hijacking is another known cookie-based vulnerability when the user browses and requests cookies over an unsecured wireless network. The attacker is also on the same wireless network as the users who are monitoring the network traffic. Since the Wireless connection is unencrypted, the data can be extracted by the attacker, who can extract the cookie from the network traffic and can use the cookie to log in to any secured portal which the victim accessed. The attacker can use the cookie until it expires [64].

Cross-site scripting (XSS) is one of the top 10 OWASP vulnerability. This type of attack involves stealing sensitive information, hijacking user sessions, and compromising web browsers and system integrity. The cookie is exploited for attack initiation with persistent Cross-site scripting requests. These cookies are stored in the user's session with the web server in various locations such as the browser and the website. The cookie content and information can be accessed maliciously via XSS [23].

There are three forms of XSS that usually targets the user's browsers: reflected, stored and DOM. In Reflected XSS, the application (service or API) allows unvalidated and un-sanitized input as part of the output. The user will have to interact with some malicious link or sent via a phishing email [65].

Cookie Poisoning Cookies store varied information in the browser which sends back to the web server unaltered. An Attacker may change the value in the cookie and send the cookie back to the web server. Hence the original cookie is poisoned with altered values, and it is sent back to the web server. The web server identifies the header as its values and processes the cookie. This may allow the attackers to gain access to the web server sensitive information. The attack may also impersonate the session of the user [23].

4.6 Summary

Cookies form the basis of this research. The background covers HTTP evolution which paved development of a communication protocol to align itself with the growth of internet applications. The web servers require various cookies to keep up the management of session for the internet applications. The cookie implementation and development highlights how its components and their flags can be vulnerable to attack. Cookie classification defines the different types of cookies used on a website with evolution of cookies outside the browser for tracking purposes. Cookie vulnerabilities are critical for our next analysis of cookie-based cybercrime. The cookie classification would be later split into cookie categories for aligning with the legal provisions.

5 Legal Framework for Cookies

The usage of cookies has significantly evolved from being simple personalization tools for individual users to becoming one of the key tools for targeted marketing. Due to this, the provisions of ePrivacy Directive 2002/58/EC [14] as it pertains to cookies are inadequate for the protection of user privacy and security. Two regulations that are aimed at addressing this regulatory gap are the GDPR and the ePR.

The following sections provide an overview of the GDPR and the proposed ePrivacy Regulation and an overview of regulations relevant to cookie implementation within them. The next section classifies cookies that need consent according to the ICC UK Cookie Guide [66]. The concluding section analyses the two frameworks by highlighting their inclusions and exclusions.

5.1 General Data Protection Regulation (GDPR)

The GDPR or Regulation 2016/679) [12] is a regulatory initiative by the European Union (EU) to harmonize data protection laws across the EU. After a transition period of two years, the regulation came into effect on May 25, 2018. The GDPR specifies under what circumstances personal data may be processed and includes several rights of data subjects and obligations for those processing personal data of EU-residents. It makes it mandatory for every website, accessible to users within the European Union, to specify their policy on privacy and the purpose of data collection and sharing. The latest transparency requirements and the user rights (wherein the right to retrieve a copy was recently introduced) will shed light on some of the practices of online advertising services.

The Table 1 summarizes the above discussion in a format where each article relevant to cookies in the GDPR text is included:

Article 4 [67]	Free will or wishes of an individual
Article 6 [68]	Consent for Processing Personal Data
Article 9 [69]	Higher Data Protection standards for sensitive personal information
Article 12 [70]	Privacy Policy Requirement
Article 20 [71]	Right to receive a copy of Processed Data

Article 22 [72]	User Profiling
Article 25 [73]	Data Protection by Design (Default settings of Cookies “accept all”)

Table 1 GDPR Cookie and Data Protection requirement

5.1.1 Transparency

Article 12 [70] of the GDPR requires the personal data processor to inform the user about how their information will be processed through their privacy policy which needs to be presented in *“a concise, transparent, intelligible, and easily accessible form, using the clear and plain language.”* [70] According to the GDPR the following are considered personal data as detailed in Article 13 [74] and on the European Union data protection website [75]:

- “a name and surname;
- a home address;
- an email address such as name.surname@company.com;
- an identification card number;
- location data (for example the location data function on a mobile phone);
- an Internet Protocol (IP) address.” [75]

The privacy policy also incorporates contact information of the data officer, the purposes for the data processing and the data subject’s rights regarding their personal data, e. g., the right to access, rectification, or forgotten (deleted). To fulfill these requirements it is mandatory for every website in the European Union to have a privacy policy and modify existing privacy policies to comply with the new transparency requirements [35].

The IP addresses version 4 or version 6 are unique identifiers and are considered personal data in the European Union. This means, every website and the underlying web server which processes these addresses is required to provide privacy policy information [75].

Article 20 [71] requires that the user receive a copy of all of their personal data that a website has processed. The GDPR specifies a timeline for response from the data processor to queries from a user to be within one month (Art. 12, No. 2), it can be extended up to two months [34].

The data subject rights to retrieve a copy of data processed and stored–is described in Article 20 [71]. According to recital 68 [76] of the GDPR (recitals describe the reasoning behind regulations), and their right to data portability which is meant to support them gain control over their personal data by allowing access to the data stored "in a structured, commonly used, machine-readable and inter-operable format.

Article 22 [72] also focuses on automated user profiling aspect:

> *"Any form of automated processing [..] to analyze or predict aspects concerning that natural person's performance at work, economic situation, health, personal preferences, interests, reliability, behavior, location or movements;"* [12]

Any company that infers information about individuals for advertising purposes and needs to disclose this. However, they are not bound to the additional requirements for profiling mentioned particularly in Article 22 [72], [12] i.e. "to enable human intervention, as profiling for advertising and tracking purposes most likely does not have any legal or significant effects" [77].

Tracking companies claim the data they use is not personal information since it is pseudonymized. If that were true, it would free them from any data protection related obligations. But Article 29 [77] of the Working Group, a committee of European data protection officials, had clarified in 2010 that storing and accessing a cookie on a user's device or browser is indeed processing of personal data since it

> "*enables data subjects to be 'singled out', even if their real names are not known,"* [78]

And therefore requires consent. The Opinion 2/2010 on online behavioural advertising document [78] was written in respect to the previous directives, but the assessment has been confirmed by court rulings that, e. g., found IP addresses, sometimes also considered pseudonyms, to be personal data [79]. Our study/this thesis focuses on GDPR and the cookie aspect including the third-party scripts on their websites that are responsible for the data processing. It was argued, since advertisers rent the space on publisher websites which set cookies linked to their hosts they are liable for the data processing.

The organizations representing the online advertising industry interpret the data subject rights differently. The Interactive Advertising Bureau Europe has published a working paper [80] for data subject requests in April 2018. The focus of the paper was on two provisions that limit the data processor's obligations to answer user requests from individuals. The first being, Article 12 which states that "a data controller does not need to act on requests if they *demonstrate that [they are] not in a position to identify the data subject"* [12].

The argument is that if the information were processed in a pseudonymous fashion, i.e., with a cookie ID, data subjects would need to prove that the information connected to that ID is actually about them.

5.1.2 Data protection by design and by default

Article 25 [73] states that entities processing personal data should

> *"[..] implement appropriate technical and organizational measures, such as pseudonymization, which are designed to implement data-protection principles, such as data minimization, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects. [..] In particular, such measures shall ensure that by default personal data are not made accessible without the individual's intervention to an indefinite number of natural persons" [12].*

There are higher protection standards that are required for sensitive categories of personal information like health care data (Article 9 [69]).

5.1.3 Consent

Article 6 [68] mentions that the processing of personal data is only lawful if either of the scenarios applies.First,it includes a situation when the processing is necessary *"for the legitimate interests [of] the controller or [...] a third party"* (Article 6(1)(f)) [68] and the Second, to comply with a legal obligation (Article 6(1)(c)) [68]. Most importantly, the processing of personal data is lawful only when*"the data subject has given consent"* (Article 6(1)(a)) [68]. Consent is defined in Article 4(11) [67] as *"any freely given,* *specific, informed and unambiguous indication of the data subject's wishes [...]"*. Here, *"freely given"* means the data subject has to be offered real choice and control; if they feel compelled to agree to the processing of their personal data, this does not constitute valid consent. For children under the age of 16 consent can only be given by the holder of parental responsibi/lity (Article 8) [12] [34].

5.2 Proposed e-Privacy Regulation

The objective of the Digital Single Market Strategy (DSM Strategy) of the European Union is to increase trust in the security of digital services. The reform of the data protection framework, and the adoption of Regulation (EU) 2016/679 also called the GDPR was a key action to harmonize data protection laws as part of the DSM. The Strategy also announced to review the Directive 2002/58/EC (ePrivacy Directive) and provide a higher level of privacy protection for users of electronic communications services of all market players. The proposal reviews the ePrivacy Regulation (ePR) in the DSM strategy objectives and ensures consistency with the GDPR [81].

The fundamental basis of the ePrivacy directive is to ensure [81]:

- Protection to respect individual private life
- Confidentiality and protection of personal data communication
- Free movement of electronic communication data

The overview of each article has been summarized in Table 2 below, which will be discussed detailing its relationship with cookies.

Article 5 [82]	Protection of Communication Data
Article 6 [83]	Processing of Personal Data
Article 8 [84]	Protecting Users Terminal Equipment (Browser)
Article 10 [85]	Privacy by Design
Recital 20 [36]	User Profiling
Recital 21 [86]	Consent to use Cookies
Recital 23 [87]	Cookies settings in Browser

Table 2 Cookie provisions in the ePrivacy Regulations

5.2.1 Transparency

Article 6 [83] of the ePR mentions, “the processing of personal data in connection with the provision of publicly available electronic communications services in public communications networks in the Community, including public communications networks supporting data collection and identification device.” i.e.: These can be demonstrated through cookies or the browser and server communication which support the collection of data of the browser footprint [88].

5.2.2 Consent to the use of cookies

The Recital 21 of the ePR requires consent to set cookies [86]. In the harmonization effort, Directive 2009/136/EC had changed Article 5(3) of the ePrivacy Directive (2002/58/EC) to state that *“the storing of information [...] in the terminal equipment of a [...] user”* is only allowed if the user *“has given his or her consent, having been provided with [...] information [...] about the purposes of the processing”* [14]. The consent requirement does not affect all the types of cookies; those that are deemed to be "*strictly necessary for the delivery of a service requested by the user*", for example, cookies that track the contents of a user's shopping cart in an online shopping portal or the user logged in to retrieve the shopping cart, are exempted. Since the consent of user is deemed necessary, this has enforced all the EU based websites to display *cookie consent notices* often referred to as *cookie banners*. These are boxes or pop-up’s informing users about the use of different cookies by the website and associated third parties.

The notices may either explicitly ask users their preferences of the cookies for their consent or interpret a user’s continued website browsing as implied consent. However, according to the EU guidelines, a valid consent needs to be freely given with an active choice based on the processing of specific information about the purpose of the data before the processing starts [89]. Article 5(3) is applicable to any information stored on the user’s browser or system though but it does not contain personal information or credentials. However, when it does contain personal information, the consent in accordance to GDPR rules are required, though the two types may be merged into practice [34].

While the focus of the ePR was consent for storing information including personal data, the study by the European Union concluded that the cookie popup or the cookie consent post GDPR implementations were either over-inclusive or under-inclusive without clear information on the choices selected by the users or the consent ignored i.e.: the users were given choices of the Necessary Cookies, Marketing Cookies, Statistical cookies but without information on how they processed and used it. The consent rule is not only over-inclusive because it also covers non-privacy intrusive practices, but also under-inclusive since it does not clearly cover some tracking techniques (e.g., device fingerprinting) which may not entail access/storage in the device [81]. The Article 5(3) of the ePD

> *"requires prior informed consent for storage or for access to information stored on a user's terminal equipment. In other words, you must ask users if they agree to most cookies and similar technologies (e.g., web beacons, Flash cookies, etc.) before the site starts to use them" [90].*

The proposed ePR focuses on simplifying consent of cookie rules to protect citizens against unsolicited marketing it proposes to add an exception to opt-out in the cookie consent choices. Article 8 focuses on the protection of end-user's terminal equipment information [84].

The regulation makes a clear distinction between the third-party persistent tracking cookies and the non-privacy intrusive cookies (i.e., first-party cookie). To understand the difference between the two types of cookies in Table 3 is a brief comparison:

	First-Party Cookies	**Third-Party Cookies**
Setting and Reading the Cookie	Can be set by the web server or any JavaScript loaded on the website	Can be set by a third-party server (e.g. an AdTech platform) via Iframe or code loaded on the website
Availability	A first-party cookie is only accessible via the domain that created it.	A third-party cookie is accessible on any website which loads the third-party code or iframe.
Browser Support, Blocking and Deletion	Supported by all browsers and can be blocked and deleted by the users; doing so resets all user preferences	Supported by all browsers, features such as Do Not Track block certain abilities to create third-party cookies. These cookies can be deleted by users too.

Table 3 Difference between First Party and Third-Party cookies from [91]

5.2.2.1 First party cookie

These are associated with the web server where the URL (Uniform Resource Locator) resides. As per the RFC 2965 which describes the three headers Cookie, Cookie2, and set-Cookie2 are originating cookies from the web server and browser [92]. The cookies have various attributes such as Name, Origination, Storage Path, Time live (Age) and Size of the cookie. The first party cookie establishes a session ID for differentiating the communication with the server and the browser [93].

5.2.2.2 Third party cookie

Third party cookies are created by a webpage to load content from another domain or webpages such as advertisements. These cookies are used for tracking user behavior and sharing that data with advertising companies.

The proposed ePrivacy regulation referring to Article 25 of the Regulation 2016/679 [12] (GDPR) included in Recital 23 [87], specifically addresses the default settings for cookies which are set in most current browsers to accept all cookies. It requires browser or software providers to offer the users a choice of privacy level settings such as:

> *"higher (for example, 'never accept cookies') to lower (for example, 'always accept cookies') and intermediate (for example, 'reject third-party cookies' or 'only accept first-party cookies')" [81].*

The effects of accepting 'all cookies' by various third-party scripts that can harvest data can

> *"reveal details of an individual's emotional, political, social complexities, including the content of communications, pictures, the location of individuals by accessing the device's GPS capabilities, contact lists, and other information already stored in the device, the information related to such equipment requires enhanced privacy protection. Furthermore, the so-called spyware, web bugs, hidden identifiers, tracking cookies and other similar unwanted tracking tools can enter end user's terminal equipment without their knowledge*

> *in order to gain access to information, to store hidden information and to trace the activities" [81].*

Cookies not only track user actions resulting in a direct threat to their privacy and create an indirect threat by the sharing of data within network connected equipment. This can cause user data to be harvested without user knowledge by surfing website and this is considered as accepting all cookies.

The ePR mentions that the consent should be in a user-friendly and transparent manner where the cookie notices should detail out every possible

> *"use of tracking cookies and other tracking techniques, end-users are increasingly requested to provide consent to store such tracking cookies in their browsers" [81].*

5.2.3 Protection for Communications data in transit

Article 5 ensures that electronic communications data is protected under the ePrivacy Regulation and data remains confidential [82]. The consent and choices through the communication channel by the end-user are under the ePR, whereas the processing of electronic data and data protection aspects fall under the GDPR. While the cookies are originating from the server to the user equipment software (the browser), the ultimate storage of information is on the web server as part of service-related processing. The modern-day secure communication example would be the usage of tokenized authentication system where the user must generate One Time Password (OTP) on the device which is used to securely login onto the browser. The authentication data is then stored in cookies as a unique identifier of the session and user browser [94].

5.2.4 Tracking walls

The ePR and GDPR both prohibit user profiling and usage of cookie tracking. It has been a practice of websites to provide content without a monetary payment if the user accepts third party cookies. The ePR specifically denies tracking walls in Recital 20 which reads

> *"to provide for additional benefits of the website operator. In some cases, making access to website content conditional to consent to the use of such cookies may be considered to be disproportionate. This is, for example, the case for websites providing certain services, such as*

> *those provided by public authorities, where the user could be seen as having few or no other options but to use the service, and thus having no real choice as to the usage of cookies" [36].*

An example would be when the user is visiting the news portal. Where the content (news) is provided to the users without any monetary benefit as the revenue is through advertisements which are third-party cookies which the user agrees too without having any choice. The net effect would be that the user is trading personal data (through the tracking cookies recording user behaviour and specifics over the internet). The revised draft posted in March 2019 alters the conditional context of consent for using cookies on such websites [95].

5.2.5 Privacy and settings by Design

The Article 10 requires Web Browsers to offer options of preventing third parties from storing information such as the cookies in the browser (Terminal Equipment) or from processing information already stored in the browser in the form of cookies. The effect would enable users to make choices on their cookies and privacy setting within the browser to set those options by default on all the websites the user visits. Hence, the browser must provide users with the option to revise settings on every update. The provision states as:

> *"inform the end-user about the privacy settings options and the way the end-user may use them. The software shall offer the end-user the choice to be reminded about the privacy settings options" [85].*

As per the last changes suggested by the Austrian presidency of the council in November 2018 [96] Article 10 has been removed from the Draft text of the proposal [95].

5.3 Analysis of regulatory frameworks based on cookie classification

While both the regulations support each other in the enforcement of protection and management of cookie data. For a better understanding, Figure 6 illustrates the relationship of GDPR and ePrivacy in relation to consent and cookies.

GDPR + ePrivacy Regulation = Cookies + Consent

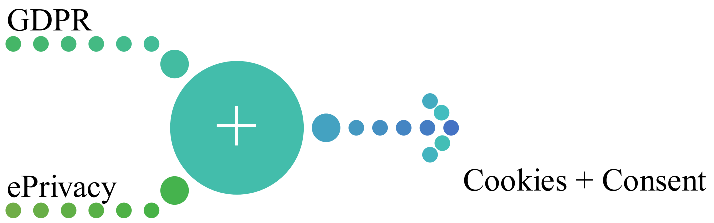


Figure 6 Cookies + Consent

Though the GDPR does not directly address cookies, it re-defines consent which must be given “freely.” The ePR in turn references the definition of cookie consent from the GDPR. Consent to use cookies is covered under the ePR, but it relies on user choices to determine which cookies are acceptable for which user [97].

Figure 7 highlights how the flow of information through cookies is covered under which regulation. The communication data shown in the Figure is entirely covered by ePR.

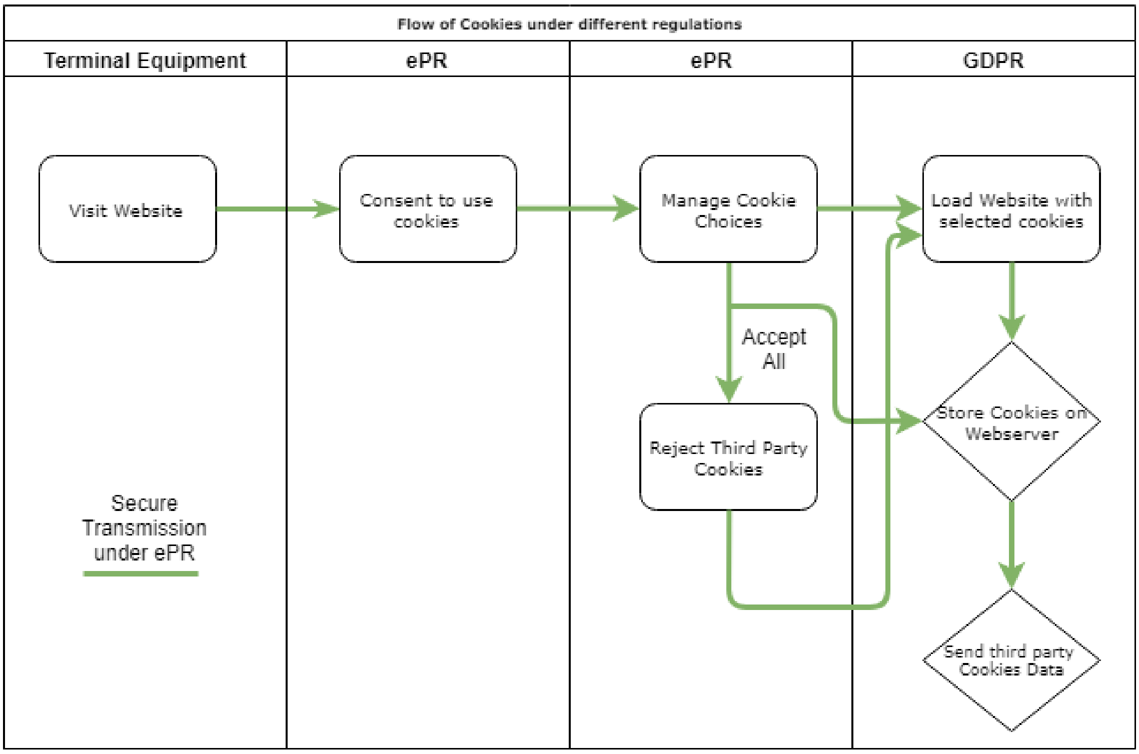

Flow of Cookies under different regulations
Terminal Equipment
ePR
ePR
GDPR
Visit Website
Consent to use cookies
Manage Cookie Choices
Load Website with selected cookies
Accept All
Reject Third Party Cookies
Store Cookies on Webserver
Send third party Cookies Data
Secure Transmission under ePR

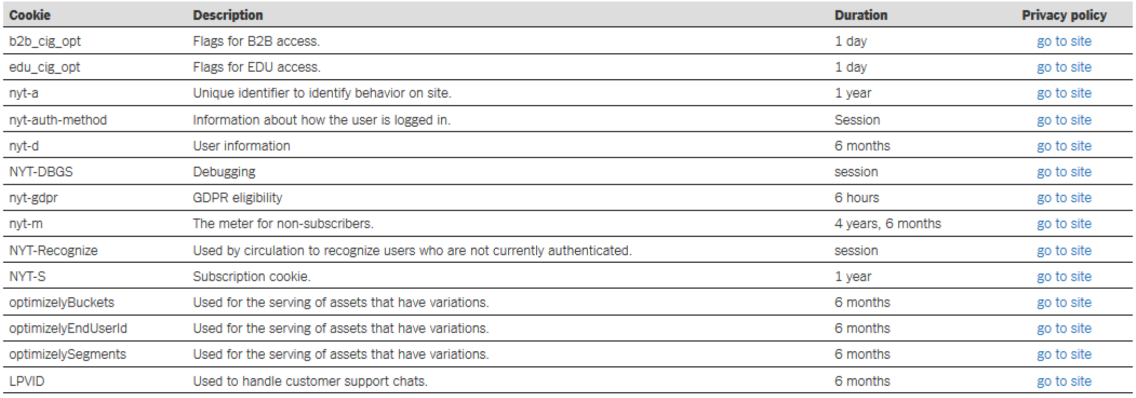
Figure 7 Flow of Cookie protection through Regulations

In order to better understand how cookies have changed due to the enforcement of GDPR, the next section describes another way of classifying cookies as compared to Chapter 2 by using the categories recommended by the International Chamber of Commerce (ICC) [66]. According to the ICC, there are four different types of cookies. These are:

- Strictly Necessary Cookies
- Performance Cookies
- Functionality Cookies
- Targeting Cookies or Advertising Cookies

5.3.1 Strictly Necessary Cookies

These are cookies which are essential for a user to visit and access the website and use its features, such as logging into their account. These cookies require no consent, and by surfing or browsing the website, the user acknowledges this fact. Figure 8 shows an extract from the Cookie policy which is enforced by The New York Times [98].



Cookie	Description	Duration	Privacy policy
b2b_cig_opt	Flags for B2B access.	1 day	go to site
edu_cig_opt	Flags for EDU access.	1 day	go to site
nyt-a	Unique identifier to identify behavior on site.	1 year	go to site
nyt-auth-method	Information about how the user is logged in.	Session	go to site
nyt-d	User information	6 months	go to site
NYT-DBGS	Debugging	session	go to site
nyt-gdpr	GDPR eligibility	6 hours	go to site
nyt-m	The meter for non-subscribers.	4 years, 6 months	go to site
NYT-Recognize	Used by circulation to recognize users who are not currently authenticated.	session	go to site
NYT-S	Subscription cookie.	1 year	go to site
optimizelyBuckets	Used for the serving of assets that have variations.	6 months	go to site
optimizelyEndUserId	Used for the serving of assets that have variations.	6 months	go to site
optimizelySegments	Used for the serving of assets that have variations.	6 months	go to site
LPVID	Used to handle customer support chats.	6 months	go to site

Figure 8 Essential Cookies in use on https://nytimes.com from [98]

These cookies are clearly exempt from consent according to the EU advisory body on data protection – WP29 [81] [99] which describes the exceptions from asking for user consent as follows:

- User-input cookies (session-id) such as first-party cookies: keep track of the user's input when filling online forms, shopping carts, etc., for the duration of a session or persistent cookies limited to a few hours in some cases.

- Authentication cookies: to identify the user once he has logged in, for the duration of a session.

- User-centric security cookies: used to detect authentication abuses, for a limited persistent duration.

- Multimedia content player cookies: used to store technical data to play video or audio content, for the duration of a session.

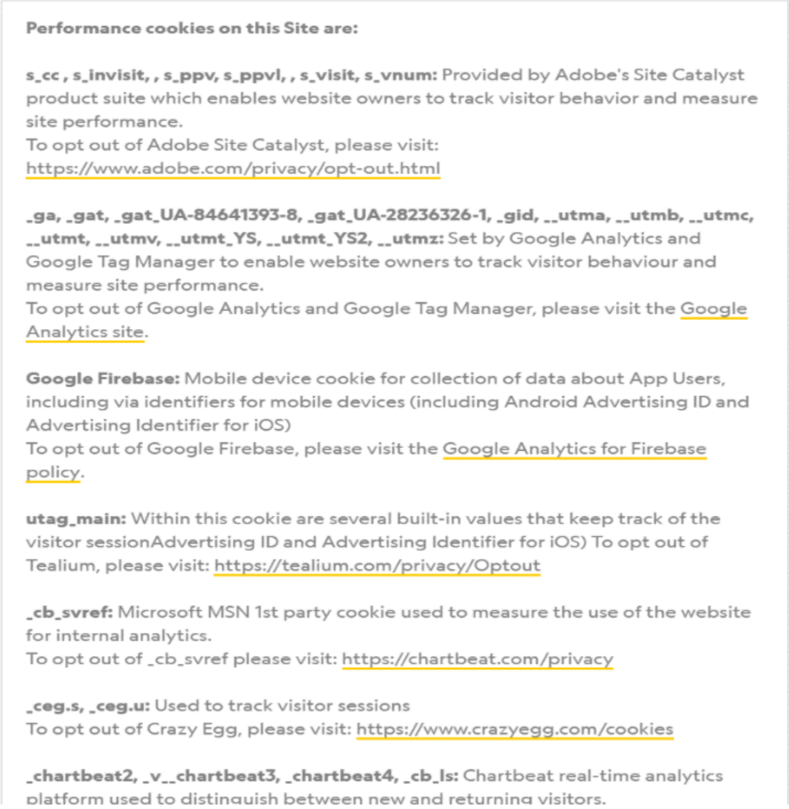
- Load-balancing cookies: for the duration of session.

- User-interface customisation cookies such as language or font preferences: for the duration of a session (or slightly longer).

- Third-party social plug-in content-sharing cookies: for logged-in members of a social network.

5.3.2 Performance Cookies

These are cookies that collect information about user activity on the website. These can include third-party analytical and performance measuring cookies from Google Analytics, Clicky Analytics or Adobe Analytics and many such web analytical platforms. These cookies are subject to user consent, and the potential user is listed in the terms and conditions. The GDPR focuses on the processing of personally identifiable information, but since analytical data has undergone pseudonymization it does not fit the GDPR context. However, the ePrivacy Regulation under Recital 23 [87] explicitly states that the consent for each cookie type is necessary. These are also known as statistical cookies. Figure 9 shows an example of performance cookies in action.



Performance cookies on this Site are:

s_cc , s_invisit, , s_ppv, s_ppvl, , s_visit, s_vnum: Provided by Adobe's Site Catalyst product suite which enables website owners to track visitor behavior and measure site performance.
To opt out of Adobe Site Catalyst, please visit:
https://www.adobe.com/privacy/opt-out.html

_ga, _gat, _gat_UA-84641393-8, _gat_UA-28236326-1, _gid, __utma, __utmb, __utmc, __utmt, __utmv, __utmt_YS, __utmt_YS2, __utmz: Set by Google Analytics and Google Tag Manager to enable website owners to track visitor behaviour and measure site performance.
To opt out of Google Analytics and Google Tag Manager, please visit the Google Analytics site.

Google Firebase: Mobile device cookie for collection of data about App Users, including via identifiers for mobile devices (including Android Advertising ID and Advertising Identifier for iOS)
To opt out of Google Firebase, please visit the Google Analytics for Firebase policy.

utag_main: Within this cookie are several built-in values that keep track of the visitor sessionAdvertising ID and Advertising Identifier for iOS) To opt out of Tealium, please visit: https://tealium.com/privacy/Optout

_cb_svref: Microsoft MSN 1st party cookie used to measure the use of the website for internal analytics.
To opt out of _cb_svref please visit: https://chartbeat.com/privacy

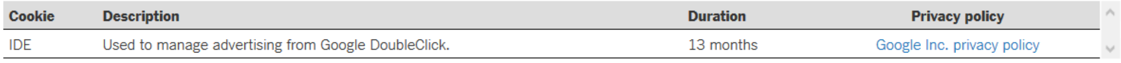
_ceg.s, _ceg.u: Used to track visitor sessions
To opt out of Crazy Egg, please visit: https://www.crazyegg.com/cookies

_chartbeat2, _v__chartbeat3, _chartbeat4, _cb_ls: Chartbeat real-time analytics platform used to distinguish between new and returning visitors.

Figure 9 Performance Cookie used on https://nationalgeographic.com from [100]

5.3.3 Targeting Cookies

These cookies focus on the personalized content matching of interest based on all the cookies stored in the browser. These cookies are used for displaying targeted advertising campaigns based on personal interest and website visit preferences. These are also termed as 3rd Party or Advertisement cookies. These cookies require explicit consent for usage and storage. These cookies have the most extended life span, as shown in Figure 10 below:

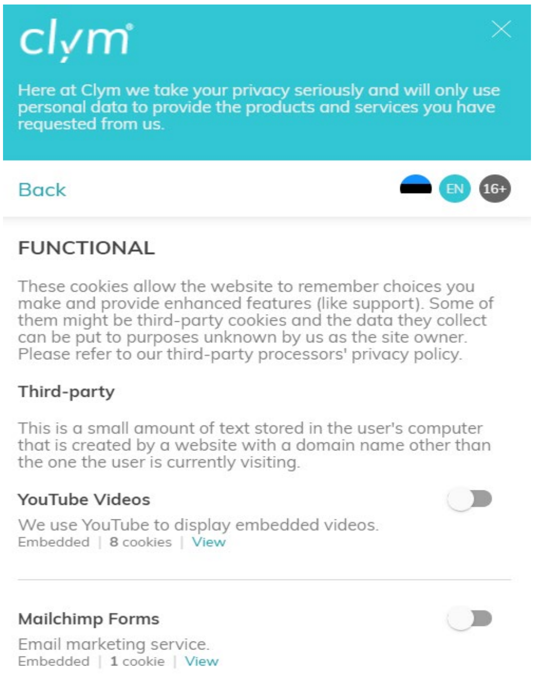


Cookie	Description	Duration	Privacy policy
IDE	Used to manage advertising from Google DoubleClick.	13 months	Google Inc. privacy policy

Figure 10 Targeting Cookie from https://nytimes.com [98]

5.3.4 Functionality Cookies

These cookies are website specific and are linked to choices a user makes during the browsing session on the website. The cookie contains data from all previous website visits including previously searched terms, pages visited or similar interest content display for the searched term [101]. A cookie consent form for the functionality cookie is shown in Figure 11 below.



Here at Clym we take your privacy seriously and will only use personal data to provide the products and services you have requested from us.

FUNCTIONAL

These cookies allow the website to remember choices you make and provide enhanced features (like support). Some of them might be third-party cookies and the data they collect can be put to purposes unknown by us as the site owner. Please refer to our third-party processors' privacy policy.

Third-party

This is a small amount of text stored in the user's computer that is created by a website with a domain name other than the one the user is currently visiting.

YouTube Videos

We use YouTube to display embedded videos.
Embedded | 8 cookies | View

Mailchimp Forms

Email marketing service.
Embedded | 1 cookie | View

Figure 11 Functional Cookies from https://clym.io from [102]

The Functional cookies are also at times referred to as preference or comfort cookies. i.e.: These cookies are used to display personalized content matching the user’s interests. These are for targeted special offers in line with each user’s search history or upcoming trips. These contain recommendation and history of the most frequently searched terms or your recent flight history [103].

The cookie category and cookie types are summarized in the Table 4 below:

Cookie Categories	**Cookie Types**	
Strictly Necessary Cookies	Session Cookie	First Party Cookies
Performance Cookies	Personalization Cookies	
Functionality Cookies	Persistent Cookies	Third Party Cookie
Targeting Cookies	Super and Zombie Cookies	

Table 4 The Cookie Table

The GDPR excludes the strictly necessary cookies for website functioning from being listed in a Cookie Policy. The ePR would enable individual consent of the user on each cookie category rather than the existing “accept all” GDPR cookie consent popup. The Session cookie contains most data of interest to a cybercriminal as it contains login information of a user and the session identifier contains all user activity and interaction with the server. Table 5 summarizes each cookie that collect personal data.

Types of Cookies	**Do these cookies collect my personal data/identify me?**
Strictly Necessary Cookies	These cookies do not identify any user as an individual. Unless the cookies are not accepted, the web page would not be displayed affecting the web page performance [57]. These cookies can identify the individual once a user authentication logon to the website or web application is performed and it can possibly connect future visits to your profile through these cookies [45].
Performance Cookies	These cookies do not identify the individual particularly. The cookie are performance and analytical cookies. All data is collected and aggregated anonymously [104].
Functionality Cookies	The information these cookies collect can include personally identifiable information that was disclosed or entered, such as your user name or contact information. There is always a transparency in the information collected by the websites. Denying or blocking these cookies may affect the performance and functionality of the website and may restrict access to content on the website [57].
Targeting Cookies	Most types of these cookies track consumers via their IP address so that it may collect some personally identifiable information. These cookies data are used to deliver advertising that is relevant to your interests. These cookies can remember that your device has visited a site or service and may also be able to track your device’s browsing activity on other sites [105]. These online identifiers share information with the advertisers and advertising networks to deliver targeted advertising, and to help measure the effectiveness of an advertising campaign or other business partners to provide aggregate Service usage statistics and aggregate service testing.

Table 5 Cookies which collect personal information

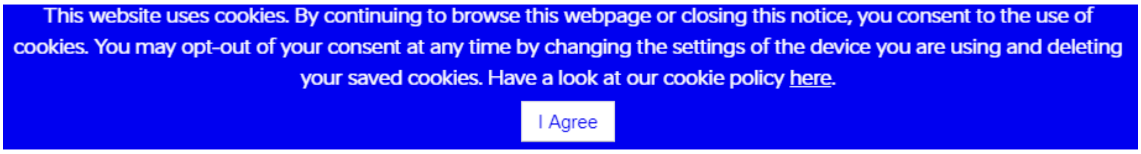
6 Cookie-related Issues

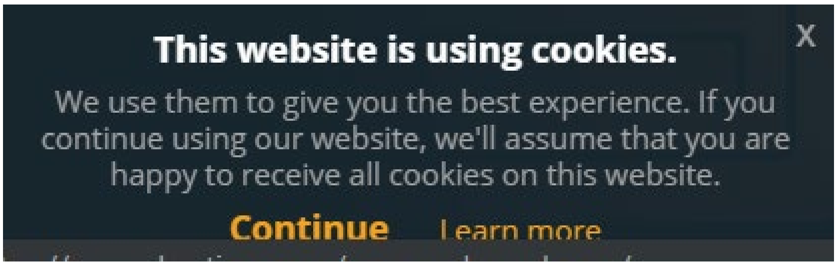
The problematic aspect of information security is that attack methods evolve rapidly, and the defence mechanism evolve slowly. Cookies are an important component in enhancing user experience for the delivery of web services. The cookies are known vulnerable in nature [23]. Setting insecure parameters in the web application and their communication protocol, can cause provocative attack parameters which can have serious arbitrary impacts or a security breach with vulnerability.

In this section the results reached in the research process are addressed. First, the weakness of cookie implementation process with respect to consent and browser settings is presented. This gap was identified in the GDPR which the ePR would fill in. The analyses by the browser consortium and internet researchers' standard suggestions and settings which can outrightly reject third-party cookies similar regulation provision in Article 10 [85] of ePR. Next, the strictly necessary cookie is studied against privacy evasive components with focus on the Yahoo! cookie forging incident from the past. Further on analysis of susceptibility of session cookie or session ID to understand why they remain an issue even after enforcement of GDPR and the upcoming ePR. An experimental setup is proved by session ID theft which is equivalent to personal identity theft. The experimental setup is to perform an XSS attack in a simple network by stealing the session ID. The setup then uses the stolen session ID to gain unauthorized access to the users' account. The access can give me details about the new stolen identity. Name is also a considered a personal data in the GDPR.

6.1 Cookies Consent and Browser

The legal framework discussion presented earlier concluded that consent is the basis of how cookies are covered under the GDPR and the proposed ePR. The consent pop-up post GDPR was ambiguous as shown in Figure 12 shown:


This website uses cookies. By continuing to browse this webpage or closing this notice, you consent to the use of cookies. You may opt-out of your consent at any time by changing the settings of the device you are using and deleting your saved cookies. Have a look at our cookie policy here.

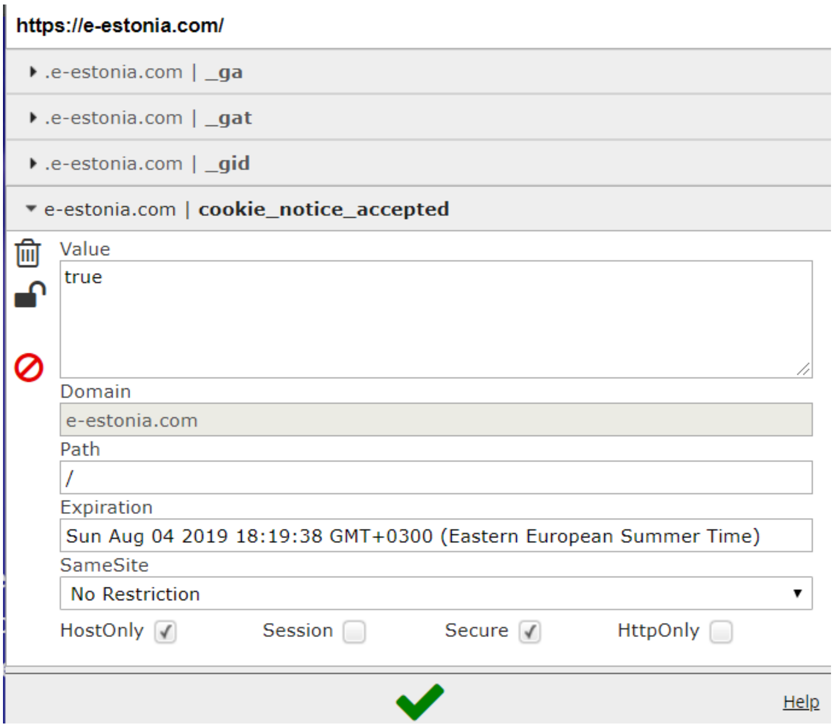


This website is using cookies.

We use them to give you the best experience. If you continue using our website, we'll assume that you are happy to receive all cookies on this website.

Figure 12 Cookie consent usage disclaimer as seen on [106] and [107]

The browser stores the user acceptance in the cookie and would not display the consent again to the users as shown in the Figure 13 below



https://e-estonia.com/

- .e-estonia.com | _ga
- .e-estonia.com | _gat
- .e-estonia.com | _gid
- e-estonia.com | cookie_notice_accepted

Value: true

Domain: e-estonia.com

Path: /

Expiration: Sun Aug 04 2019 18:19:38 GMT+0300 (Eastern European Summer Time)

SameSite: No Restriction

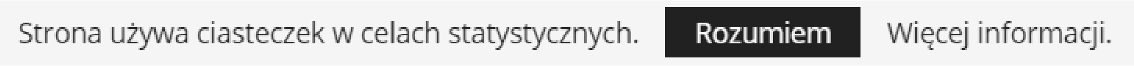
HostOnly ☑ Session □ Secure ☑ HttpOnly □

Figure 13 Cookie acceptance valued stored in the Cookie

The cookie consent is stored in the browser, and it would not be triggered again unless the cookies are cleared from the system.

The experimental setup analysed cookie consent of top news and media portals according to their Alexa ranking [108] in Estonia. This category was chosen for two reasons. Firstly, existing research [17] done to observe the changes in third-party trackers across all website categories identified News Portals as having the highest number of tracking cookies. In the proposed ePR, Recital 20 [36] notes that websites are often forcing users indirectly to accept tracking cookies to deliver its content free of charges.

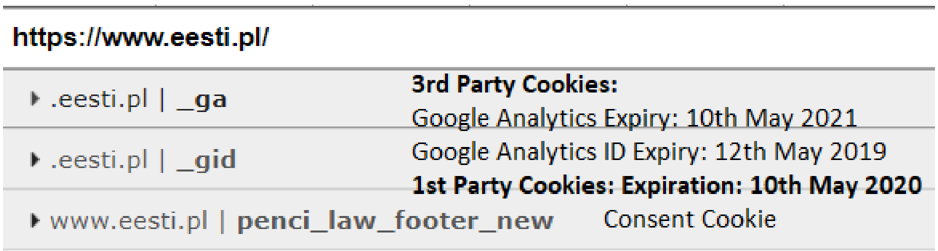
The first website analysed is eesti.pl which is in Polish display a consent of cookie usage as seen in Figure 14. The website has no advertisements throughout and doesn’t add external tracking cookies like most of the news portals. The website uses Google Analytics for measuring performance and the browsing style of users for their website. The analytics would be used to deliver and track most read content as per their privacy policy as shown in Figure 15. The consent cookie has an expiration of 1 year from being accepted. However, there is no cookie management on the website where users can have a choice of cookies [109].



Strona używa ciasteczek w celach statystycznych. Rozumiem Więcej informacji.

Figure 14 Cookie consent request from [72]

The website is using Google Analytics (3rd Party Cookies) similar to those used by all the other websites that were tested for consent and browser behaviour. Google Analytics is a simple, easy-to-use tool that helps website owners measure how the visitors interact and use the website content [109]. There is a JavaScript inserted in each webpage to measure performance when the user navigates to another page which use HTTP Cookies to ‘remember’ what the user has viewed or what content was viewed including other interactions and visits from the past.



https://www.eesti.pl/

▸ .eesti.pl | _ga

3rd Party Cookies:

Google Analytics Expiry: 10th May 2021

▸ .eesti.pl | _gid

Google Analytics ID Expiry: 12th May 2019

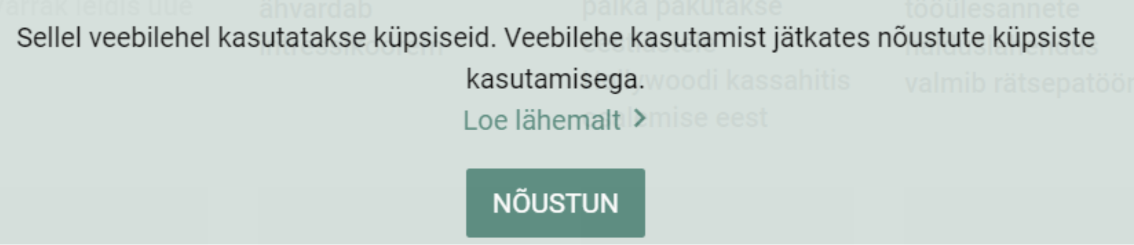
1st Party Cookies: Expiration: 10th May 2020

▸ www.eesti.pl | penci_law_footer_new

Consent Cookie

Figure 15 Cookie in use [72]

On the portal of aripaev.ee [110] the ambiguous cookie usage consent message displayed is shown in Figure 15. The notice doesn’t explicitly mention any usage of 3rd Party cookies or tracking cookies used on the portal. Figure 17 shows the usage of different types of online identifiers including 1st and 3rd party identifiers. The 3rd party identifiers have a longer life span than first party identifiers. There are some application specific cookies which are being used on the website for weather, stock exchange and currency values displayed on the website. The website uses Facebook Pixel [111], an analytical tool to enhance re-targeting of advertisements to measure conversion to visits through web advertisements.

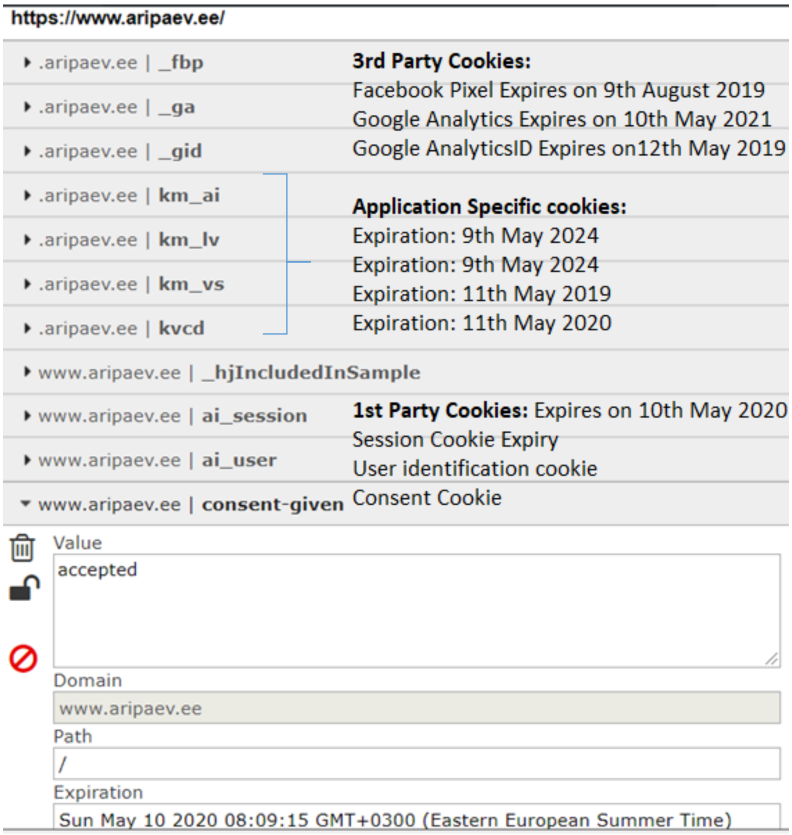


Sellel veebilehel kasutatakse küpsiseid. Veebilehe kasutamist jätkates nõustute küpsiste kasutamisega.

Loe lähemalt ›

NÕUSTUN

Figure 16 Cookie usage notice from [110]



https://www.aripaev.ee/

- .aripaev.ee | _fbp
- .aripaev.ee | _ga
- .aripaev.ee | _gid
- .aripaev.ee | km_ai
- .aripaev.ee | km_lv
- .aripaev.ee | km_vs
- .aripaev.ee | kvcd
- www.aripaev.ee | _hjIncludedInSample
- www.aripaev.ee | ai_session
- www.aripaev.ee | ai_user
- www.aripaev.ee | consent-given

3rd Party Cookies:
Facebook Pixel Expires on 9th August 2019
Google Analytics Expires on 10th May 2021
Google AnalyticsID Expires on12th May 2019

Application Specific cookies:
Expiration: 9th May 2024
Expiration: 9th May 2024
Expiration: 11th May 2019
Expiration: 11th May 2020

1st Party Cookies: Expires on 10th May 2020
Session Cookie Expiry
User identification cookie
Consent Cookie

Value
accepted

Domain
www.aripaev.ee

Path
/

Expiration
Sun May 10 2020 08:09:15 GMT+0300 (Eastern European Summer Time)

Figure 17 Cookies used on [110].

The next website analysed is postimees.ee website one of the most popular websites [112] in Estonia. The website's consent pop-up is shown in Figure 18 for the usage of online identifiers than the ones analysed before. The cookies are used for offering pleasant reading and personally appealing offers for the readers visiting the website. The Figure 18 shows the use of 3rd party cookies such as that of Google analytics, AdWords and Facebook pixel. The Google AdWords cookie is used to contact their backend network to display relevant 3rd Party advertisements from the advertisers targeting specific users in a location [113].

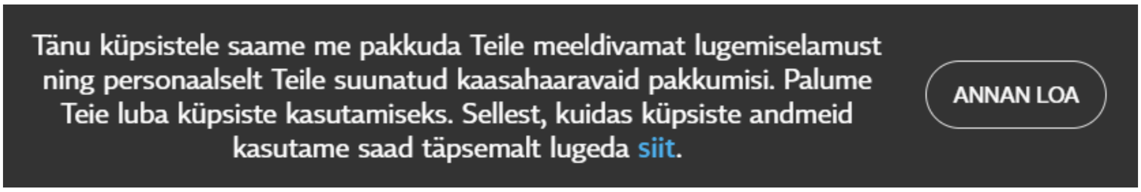
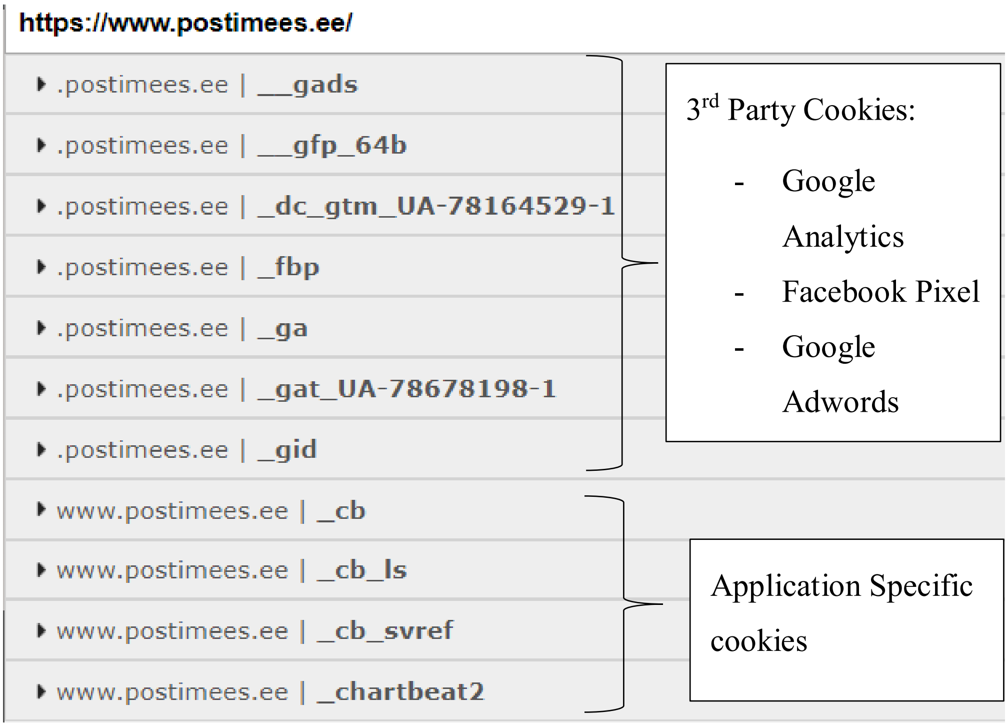

Tänu küpsistele saame me pakkuda Teile meeldivamat lugemiselamust ning personaalselt Teile suunatud kaasahaaravaid pakkumisi. Palume Teie luba küpsiste kasutamiseks. Sellest, kuidas küpsiste andmeid kasutame saad täpsemalt lugeda siit.

Figure 18 Cookies Consent from [114]


https://www.postimees.ee/

- .postimees.ee | __gads
- .postimees.ee | __gfp_64b
- .postimees.ee | _dc_gtm_UA-78164529-1
- .postimees.ee | _fbp
- .postimees.ee | _ga
- .postimees.ee | _gat_UA-78678198-1
- .postimees.ee | _gid
- www.postimees.ee | _cb
- www.postimees.ee | _cb_ls
- www.postimees.ee | _cb_svref
- www.postimees.ee | _chartbeat2

3rd Party Cookies:

- Google Analytics
- Facebook Pixel
- Google Adwords

Application Specific cookies

Figure 19 Cookie usage overview from [114]

The next website analysed is of delfi.ee, another popular news portal in Estonia. This website is part of a group AS Ekspress Meedia unlike the previous websites that were analysed. The cookie consent banner is common to all the group websites as seen in Figure 20 and consent is synced to all websites as the cookie consent is stored in the online identifiers.

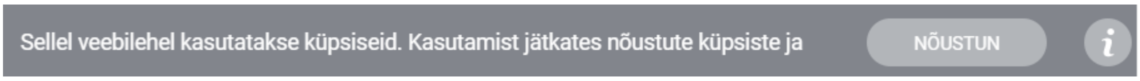
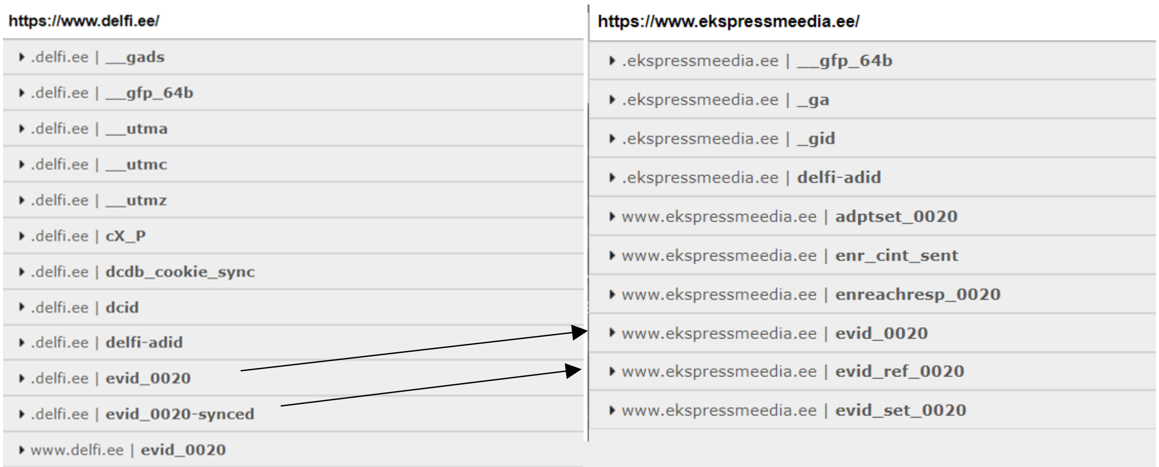

Sellel veebilehel kasutatakse küpsiseid. Kasutamist jätkates nõustute küpsiste ja

Figure 20 Cookie consent from [115]

In Figure 20 two group websites are compared to understand cookie syncing, a transparent mechanism to sync cookie consent shared in one of group websites to another website within the same group. It should be noted that both group websites have separate domain names and URL’s. The cookie consent syncing offers flexibility to not display the consent again on either of the group websites and would only request cookie consent on expiration of the cookie. The website also deploys third-party analytical snippets and independent advertisement performance trackers.



https://www.delfi.ee/

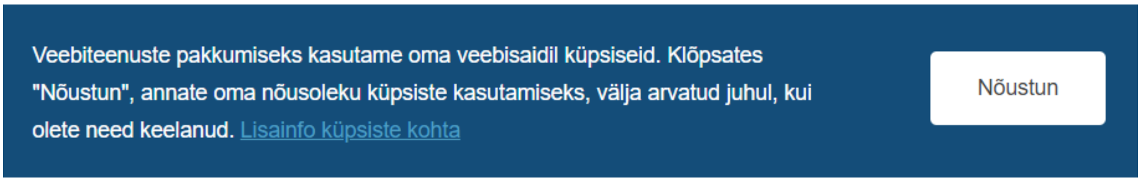
- .delfi.ee | __gads
- .delfi.ee | __gfp_64b
- .delfi.ee | __utma
- .delfi.ee | __utmc
- .delfi.ee | __utmz
- .delfi.ee | cX_P
- .delfi.ee | dcdb_cookie_sync
- .delfi.ee | dcid
- .delfi.ee | delfi-adid
- .delfi.ee | evid_0020
- .delfi.ee | evid_0020-synced
- www.delfi.ee | evid_0020

https://www.ekspressmeedia.ee/

- .ekspressmeedia.ee | __gfp_64b
- .ekspressmeedia.ee | _ga
- .ekspressmeedia.ee | _gid
- .ekspressmeedia.ee | delfi-adid
- www.ekspressmeedia.ee | adptset_0020
- www.ekspressmeedia.ee | enr_cint_sent
- www.ekspressmeedia.ee | enreachresp_0020
- www.ekspressmeedia.ee | evid_0020
- www.ekspressmeedia.ee | evid_ref_0020
- www.ekspressmeedia.ee | evid_set_0020

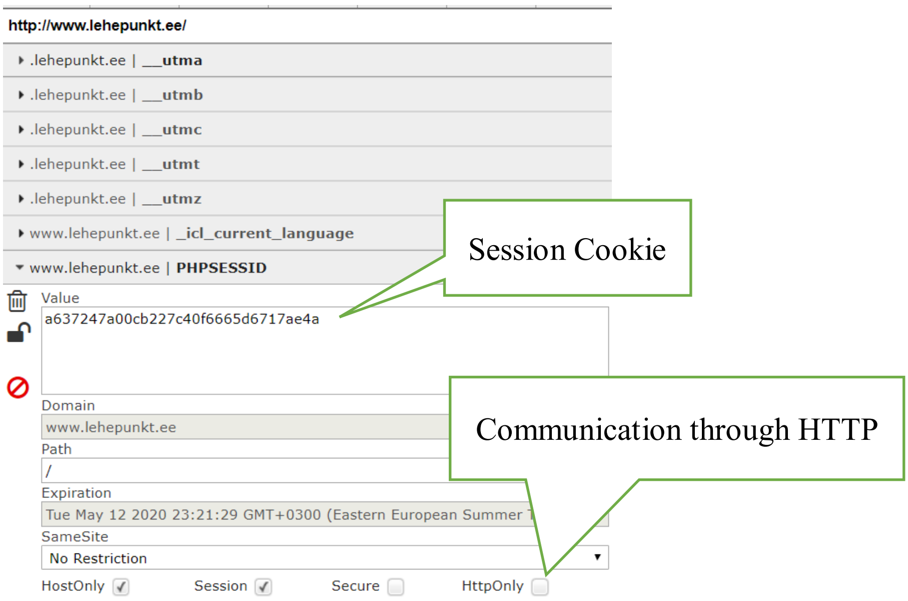
Figure 21 Group websites cookie comparison from [115] and [116]

The next website analysed is lehepunkt.ee which also displays (in Figure 22) the cookie consent notice in a detailed manner in comparison to other websites. Figure 23 has similar Google Analytics cookies with a unique PHP session ID [48] for the session established which has a 1-year expiration date and can be misused for a Cross-Site Forgery request as it is not set to HTTP only cookie flag.



Veebiteenuste pakkumiseks kasutame oma veebisaidil küpsiseid. Klõpsates "Nõustun", annate oma nõusoleku küpsiste kasutamiseks, välja arvatud juhul, kui olete need keelanud. Lisainfo küpsiste kohta

Figure 22 Cookie consent on [117]



http://www.lehepunkt.ee/

- .lehepunkt.ee | __utma
- .lehepunkt.ee | __utmb
- .lehepunkt.ee | __utmc
- .lehepunkt.ee | __utmt
- .lehepunkt.ee | __utmz
- www.lehepunkt.ee | _icl_current_language
- www.lehepunkt.ee | PHPSESSID

Value: a637247a00cb227c40f6665d6717ae4a

Domain: www.lehepunkt.ee

Path: /

Expiration: Tue May 12 2020 23:21:29 GMT+0300 (Eastern European Summer T

SameSite: No Restriction

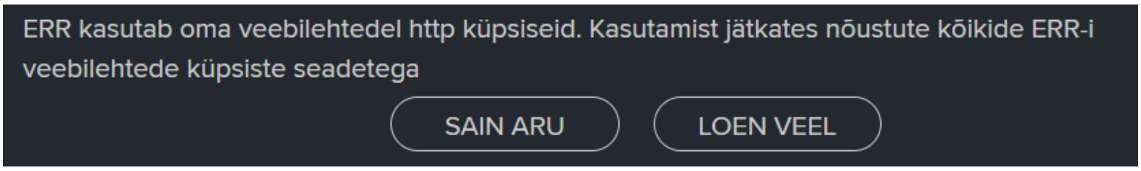
HostOnly ☑ Session ☑ Secure □ HttpOnly □

Session Cookie

Communication through HTTP

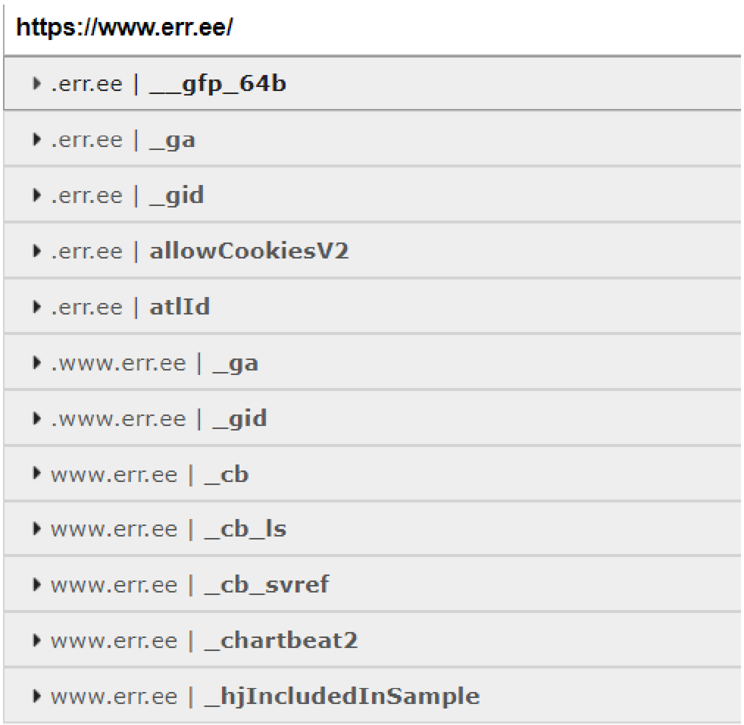
Figure 23 Cookies in use on [117]

The last website analysed is err.ee website, one of the most popular and visited website with the highest readership [118]. The website is a group website like that of delfi.ee and they have a similar cookie usage structure and consent. The consent (shown in Figure 24) is applicable to all the group websites of ERR. The website like all others, deploy Google analytics and an independent advertisement network to display advertisements on website illustrated in Figure 25.



ERR kasutab oma veebilehtedel http küpsiseid. Kasutamist jätkates nõustute kõikide ERR-i veebilehtede küpsiste seadetega

Figure 24 Group cookie consent on [119]



https://www.err.ee/

- .err.ee | __gfp_64b
- .err.ee | _ga
- .err.ee | _gid
- .err.ee | allowCookiesV2
- .err.ee | atlId
- .www.err.ee | _ga
- .www.err.ee | _gid
- www.err.ee | _cb
- www.err.ee | _cb_ls
- www.err.ee | _cb_svref
- www.err.ee | _chartbeat2
- www.err.ee | _hjIncludedInSample

Figure 25 Cookie used on [119]

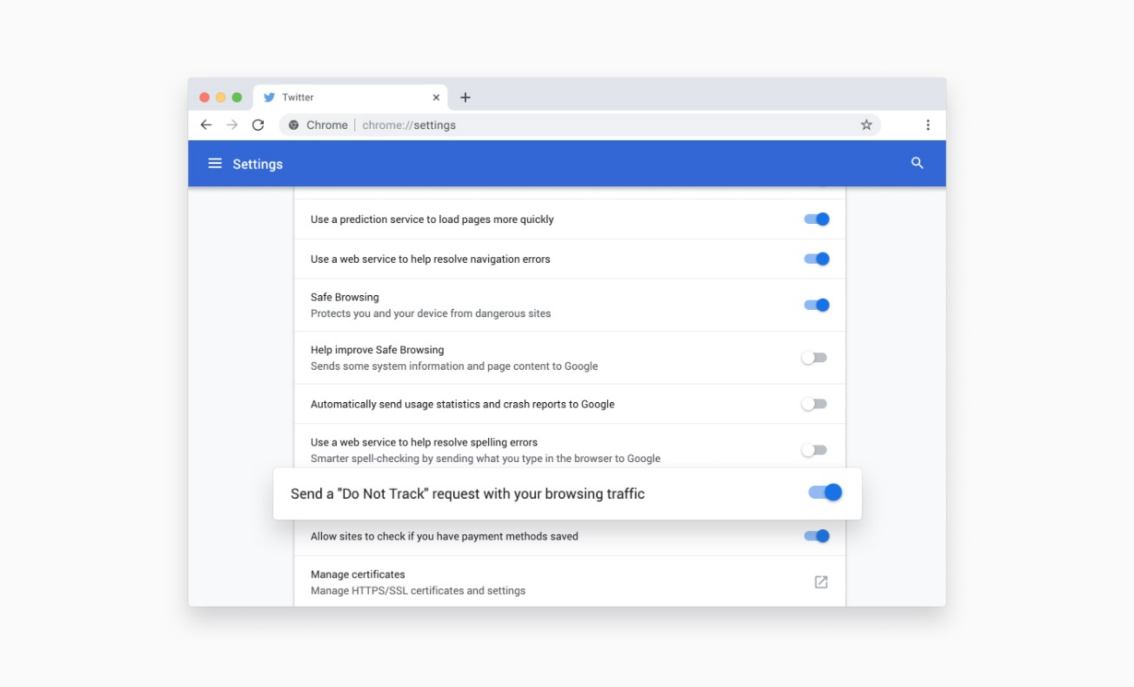
In summary, it can be concluded that, all news portals are using Google Analytics for website performance measurement. While some of the portals also use Google AdWords to monetize its web content, the cookie consent being ambiguous in nature doesn't seek a well-informed consent. Though this would change and ePR would offer protection from the 3rd party data sharing cookie network as required by Recital 20. Other recommendations from the analysis of consents would be to:

1. Inform the visitors in plain language about the purpose of cookies and trackers (ePR)
2. Provide options for visitors to change or withdraw a consent (GDPR and ePR)
3. The cookie choices and revocation setting are not provided on the website. The only way a user can revoke his/her consent is through the browser by deleting the cookies. (GDPR)
4. The users lack the ability to choose online identifier acceptable to them on these websites (GDPR)
5. The users visiting these websites are not aware that their visit is personalized based on automatic decisions including profiling to display relevant advertisements and news articles based on topic of interest.

6.1.1 Do Not Track

An important proposal to limit web tracking is the initiative Do Not Track [120] promoted by Mayer et al. that allows users to express their willingness to avoid being tracked [26]. The Do Not Track (DNT) is a web browser setting that requests that a web application disable its tracking of an individual user. When you choose to turn on the DNT setting in your browser, your browser sends a special signal to websites, analytics companies, ad networks, plug in providers, and other web services you encounter force browsing to stop tracking your activity [121].

Most browsers have a "Do Not Track" (DNT) setting (such as Figure 26) that sends "a special flag to websites, analytics companies, ad networks, plug in providers, and other web services while browsing, to stop tracking your activity [122]. The Do Not Track represent users' choices for the third-party cookies as per the ePR. But the latest revision of the draft removes the browser's obligation to process users' cookie preferences [95].



Use a prediction service to load pages more quickly ●

Use a web service to help resolve navigation errors ●

Safe Browsing
Protects you and your device from dangerous sites ●

Help improve Safe Browsing
Sends some system information and page content to Google ○

Automatically send usage statistics and crash reports to Google ○

Use a web service to help resolve spelling errors
Smarter spell-checking by sending what you type in the browser to Google ○

Send a "Do Not Track" request with your browsing traffic ●

Allow sites to check if you have payment methods saved ●

Manage certificates
Manage HTTPS/SSL certificates and settings

Figure 26 Do Not Track setting in Google Chrome from [122]

6.1.2 Platform for Privacy Preferences (P3P)

A similar concept to Do Not Track is a part of Internet Explorer which poses a unique set of challenges to web applications. The feature is set to reject any third-party cookies set by a third-party host that does not contain a compact Platform for Privacy Preferences

(P3P) policy. P3P is an Internet Explorer standard that exists to allow automatic retrieval and processing of privacy policy by a user agent (or browser) [123].

This function is more standardized for websites specific to run on Internet Explorer and is not compatible with other browsers.

6.2 Privacy evasive behaviour

One of the main sources of information used for profiling comes from web tracking, i.e. tracking users across different visits or across different sites. Data collected includes the sequence of visited sites and viewed pages, and the time spent on each page. However, behavioural tracking is not allowed in cases of data that contain any personally identifiable information, such as name, address, phone number and so forth. Web tracking is mainly performed by monitoring IP addresses, and using techniques such as cookies, or the more powerful super cookies [1].

6.2.1 Cookie related cybersecurity crimes

The Yahoo! data breach which revealed that 32 million user accounts were accessed in 2015 or 2016 by hackers who used forged cookies to log in without a password [124]. The November 2016 SEC filing noted that the company believed the data breach had been conducted through a cookie-based attack that allowed hackers to authenticate as any other user without their password [125] [126]. Yahoo! and its outside security analysts confirmed this was the method of intrusion in their December 2016 disclosure announcement in regards to the data breach which happened in August 2013 and had invalidated all previous cookies to eliminate this route. [127] [128] The evidence suggested that the hackers had created forged cookies that let them bypass the need to enter password to access users' accounts. [125] [129] The company's proprietary code was studied to understand how to forge certain cookies [128] [129].

The leaked data included names, email addresses, phone numbers, birthdays, hashed passwords, and a mix of encrypted and unencrypted security questions and answers. Yahoo claimed that the breach did not include unencrypted passwords, credit card numbers, or bank account information [127] [126].

Due to the involvement of state-sponsored attack elements, details on the leak and how attackers explicitly "forged cookies" is not understood and clearly published. But available data, it can be concluded that session cookies can be forged to impersonate another users' access.

The session cookies, which identifies a session is generated based on the authentication by a user. One possible attack vector could be session prediction. "The session prediction attack focuses on predicting session ID values that permit an attacker to bypass the authentication schema of an application. By analysing and understanding the session ID generation process, an attacker can predict a valid session ID value and get access to the application. In addition, the attacker can implement a brute force technique to generate and test different values of session ID until they have successfully gained access to the application" [130]. Another attack vector could involve stealing the session ID and using the persistent cookies and enhancing the settings instead of temporary session cookie, thereby fixing the user's session for a longer period [131]. The discussion of susceptibility of attack on Session ID is discussed in the next subsection.

6.2.2 Susceptibility of Session IDs to Attack

Session IDs are long random generated alphanumeric strings which are attached to the header and transmitted between a client and server through the cookies. The modern web applications store authentication data in these session IDs, which prevents the users from re-entering authentication data again. "The most common flaw in session ID usage has always been predictability" [45].

"Many websites are designed to authenticate and track a user when communication is first established. To do this, users must prove their identity to the website, typically by supplying a username/password (credentials) combination. Rather than passing these confidential credentials back and forth with each transaction, websites will generate a unique "session ID" to identify the user session as authenticated" [132].

The security problems behind session ID-based authentication can be summarized into categories as described below [133]:

Predictable Algorithm: Most websites are currently using linear algorithm based on easily predictable variables, such as time or IP address. It leaves the attackers with limited combination to produce a valid session ID if sequential pattern of cookie ID is followed by generating random requests in a sequential pattern [133] [45].

Limited Account Lockout Systems: Websites with custom web platforms have poor prohibition and security mechanism in place for fingerprinting and active scanning of brute force attempts. The attacker can try random combinations of session ID embedded in a legitimate URL without a single complaint from the web server [133]. A brute force attack can conduct as many as 1000 session ID attacks per second [45].

Transmission over insecure protocol: Transmission of web identifiers between the web browser and web server over a non-secure protocol or in an unencrypted manner [55]. In a "flat network topology the network can be sniffed" [133] to capture the packet and steal the session ID. Network traffic sniffing can reveal packet with login credentials if they are transmitted in clear text [131]. In Figure 22 the communication of cookies over insecure protocol (i.e.: missing secure flag)

Short Length: "The most cryptographically strong algorithm still allows an active session ID to be easily determined if the length of the string is not sufficiently long" [133].

Longer expiration of cookies on the server: The cookies have a set expiration although it is noticed in many cases, such as in Figure 16 where the expiration is over a year. This enables the attackers allow over a year to guess a valid session ID for access. If the cookie file is intercepted the static session ID which are valid over a year can allow repetitive access for misuse till expiration [133].

Insecure cookie retrieval: The cookies can be retrieved from session fixation, where the attacker maintains a trap session until the user logs into [131]. The user is tricked into visiting another website or a trap page, from where the attacker can retrieve stored session ID from the browser [133].

6.2.3 Session Hijacking attack

As session ID and session cookies are susceptible to attacks. The session hijacking attack consists of the exploitation of the web session control mechanism, which is normally managed for a session token.

The http communication uses multiple TCP connections, the web server identifies every user's connections. The most useful method depends on a token or a session ID that the Web Server sends to the client browser after a successful client authentication. A session token is normally composed of a string of variable width and it could be used in different ways, like in the URL, in the header of the http requisition as a cookie, in other parts of the header of the http request, or yet in the body of the http requisition [134].

This method of attack compromises these session token by stealing or predicting a valid session token to gain unauthorized access to the Web Server. The most common methods of attack are:

- "Predictable session token;
- Session Sniffing;
- Cross-site scripting
- Man-in-the-middle attack
- Man-in-the-browser attack" [134]

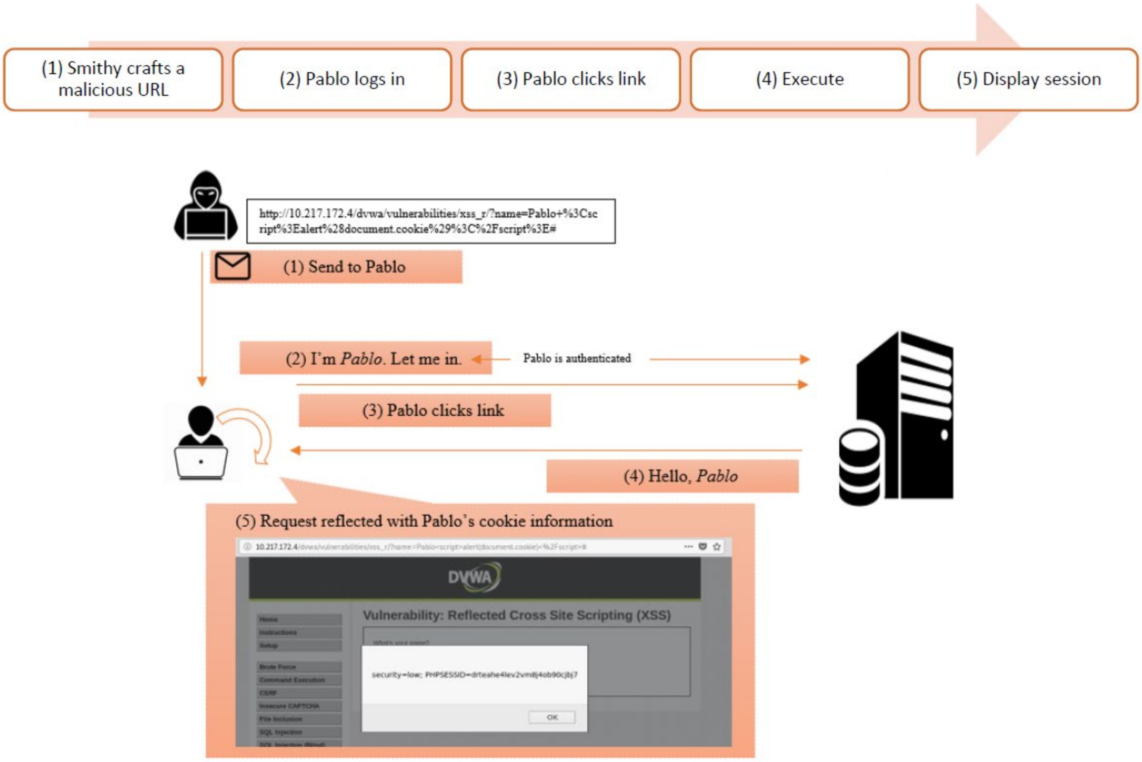
This attack vector is selected for demonstration based on our research that strictly necessary cookies reveal privacy details.

6.2.3.1 Session hijacking using Cross-site Scripting

To prove that session ID or session cookie contains personal data, an experiment was conducted as described below.

Test Case: OTG-INPVAL-001 Testing for Reflected Cross Site Scripting, OTG-SESS-001 Testing for Bypassing Session Management Schema [134].

The complete setup and workflow of the XSS is described in the Figure 27. Each setup is detailed and has been followed:

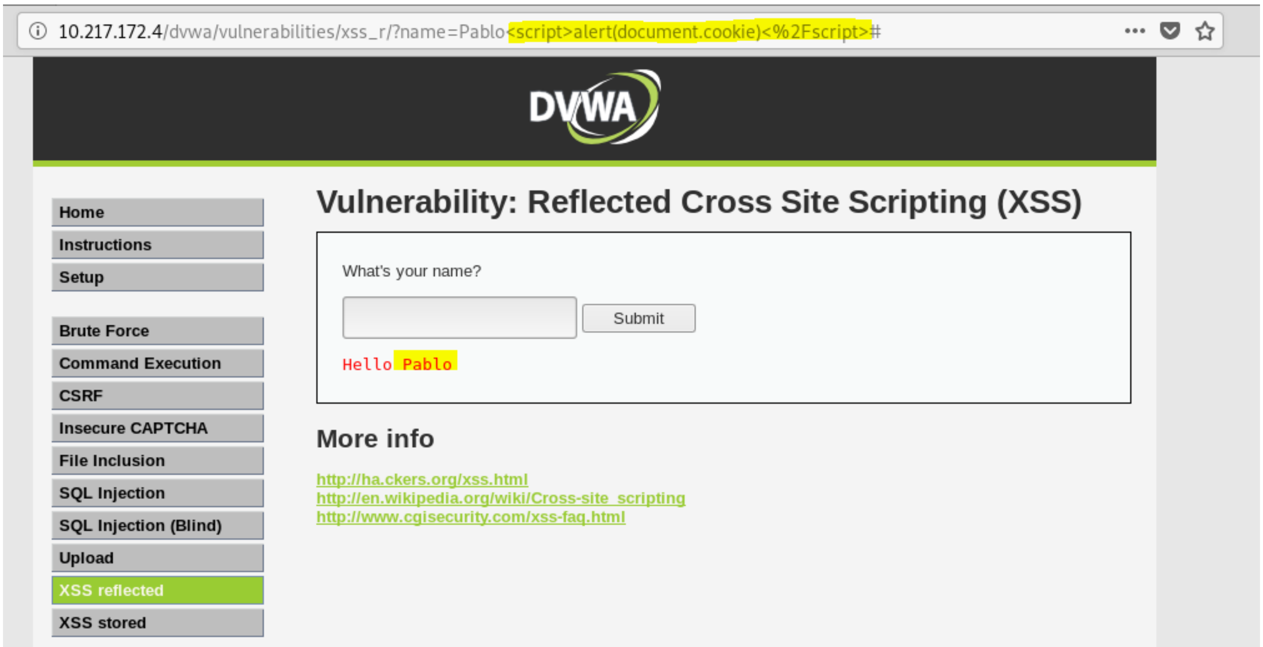

(1) Smithy crafts a malicious URL
(2) Pablo logs in
(3) Pablo clicks link
(4) Execute
(5) Display session
(1) Send to Pablo
(2) I'm *Pablo*. Let me in.
Pablo is authenticated
(3) Pablo clicks link
(4) Hello, *Pablo*
(5) Request reflected with Pablo's cookie information
DVWA
Vulnerability: Reflected Cross Site Scripting (XSS)
OK

```
http://10.217.172.4/dvwa/vulnerabilities/xss_r/?name=Pablo+%3Cscript%3Ealert%28document.cookie%29%3C%2Fscript%3E#
```

```
security=low; PHPSESSID=drteahe4lev2vm8j4ob90cjbj7
```

Figure 27 XSS Workflow

The experiment begins with the XSS attempt and nothing seems suspicious from the client's browser as shown in Figure 28.



10.217.172.4/dvwa/vulnerabilities/xss_r/?name=Pablo<script>alert(document.cookie)<%2Fscript>#

Vulnerability: Reflected Cross Site Scripting (XSS)

What's your name?

Hello Pablo

More info

http://ha.ckers.org/xss.html
http://en.wikipedia.org/wiki/Cross-site_scripting
http://www.cgisecurity.com/xss-faq.html

Figure 28 Reflected XSS

The experimentation network setup can be checked on Figure 29. The next step involves stealing the cookies. The Browser Exploitation Framework2 or BeEF is used, which is an application made for pen testing the web browser. For this, an Ubuntu container (10.217.172.5) was used as the user Pablo, Kali is the attacking machine with BeEF and user Smithy.

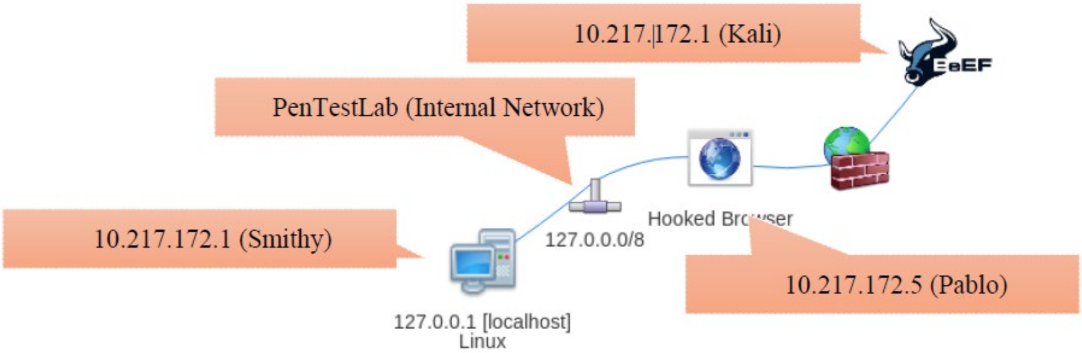
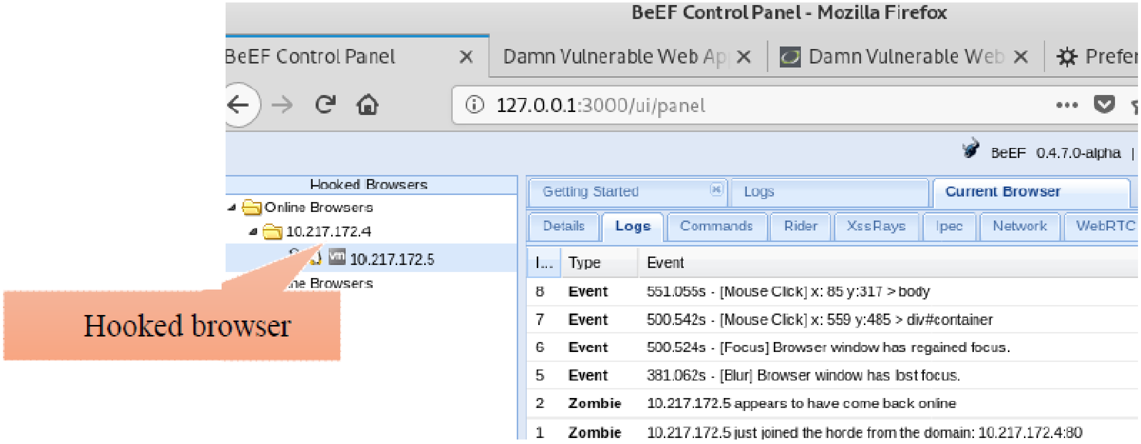


Figure 29 BeEF Network Map

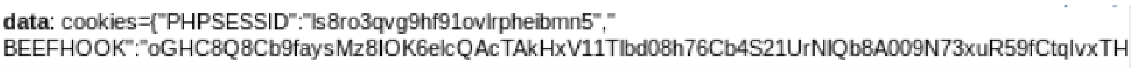
Beef was activated from the Kali Linux Machine. The hook URL was followed which is generally the Kali local loop address to open the BeEF console from the Web Browser. The overview of the BeEF control panel is shown on Figure 30 where the browser hooked.



I...	Type	Event
8	Event	551.055s - [Mouse Click] x: 85 y:317 > body
7	Event	500.542s - [Mouse Click] x: 559 y:485 > div#container
6	Event	500.524s - [Focus] Browser window has regained focus.
5	Event	381.062s - [Blur] Browser window has lost focus.
2	Zombie	10.217.172.5 appears to have come back online
1	Zombie	10.217.172.5 just joined the horde from the domain: 10.217.172.4:80

Figure 30 BeEF Control Panel

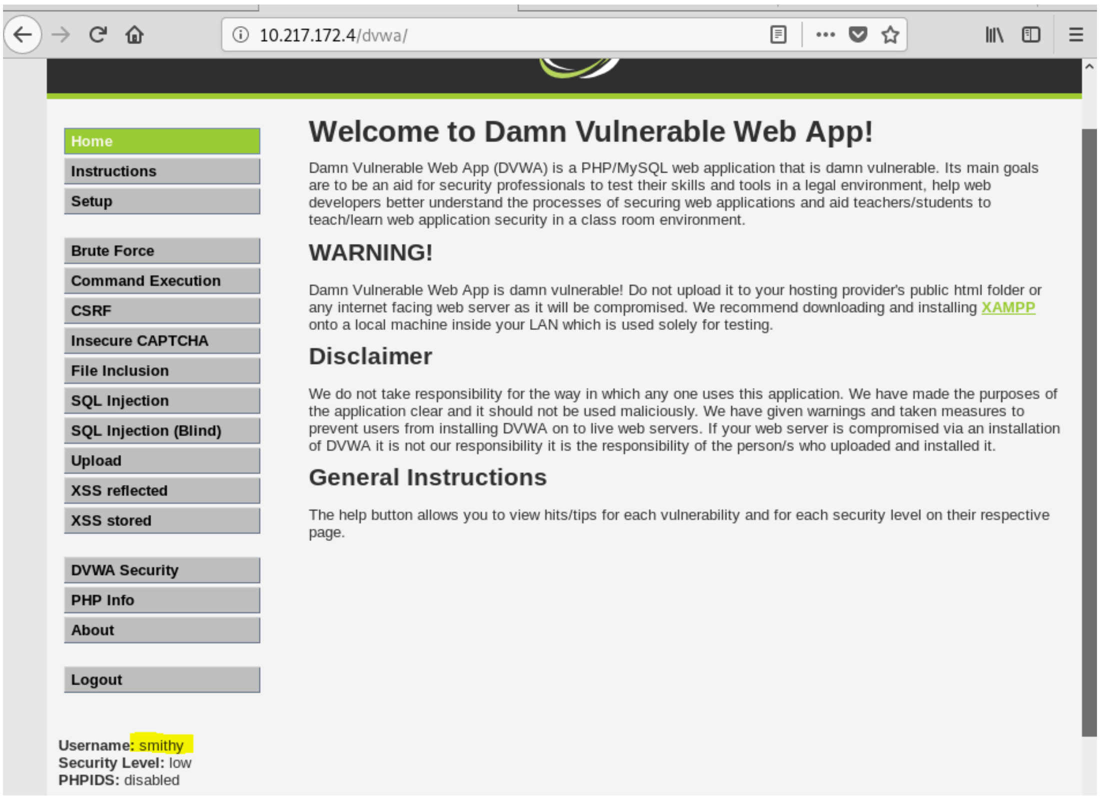
On the BeEF exploit the Apache Cookie Disclosure vulnerability was chosen and executed. After running the exploit, the client's cookie containing the Session ID is disclosed as shown in Figure 31 below:



```
data: cookies={"PHPSESSID":"ls8ro3qvg9hf91ovlrpheibmn5","
BEEFHOOK":"oGHC8Q8Cb9faysMz8lOK6elcQAcTAkHxV11Tlbd08h76Cb4S21UrNlQb8A009N73xuR59fCtqlvxTH
```

Figure 31 XSS Apache Cookie Disclosure Results (Session ID)

Next, smithy was logged in as in Kali Linux into the Damm Vulnerable Web App (DVWA) as seen in Figure 32 below:



Welcome to Damn Vulnerable Web App!

Damn Vulnerable Web App (DVWA) is a PHP/MySQL web application that is damn vulnerable. Its main goals are to be an aid for security professionals to test their skills and tools in a legal environment, help web developers better understand the processes of securing web applications and aid teachers/students to teach/learn web application security in a class room environment.

WARNING!

Damn Vulnerable Web App is damn vulnerable! Do not upload it to your hosting provider's public html folder or any internet facing web server as it will be compromised. We recommend downloading and installing XAMPP onto a local machine inside your LAN which is used solely for testing.

Disclaimer

We do not take responsibility for the way in which any one uses this application. We have made the purposes of the application clear and it should not be used maliciously. We have given warnings and taken measures to prevent users from installing DVWA on to live web servers. If your web server is compromised via an installation of DVWA it is not our responsibility it is the responsibility of the person/s who uploaded and installed it.

General Instructions

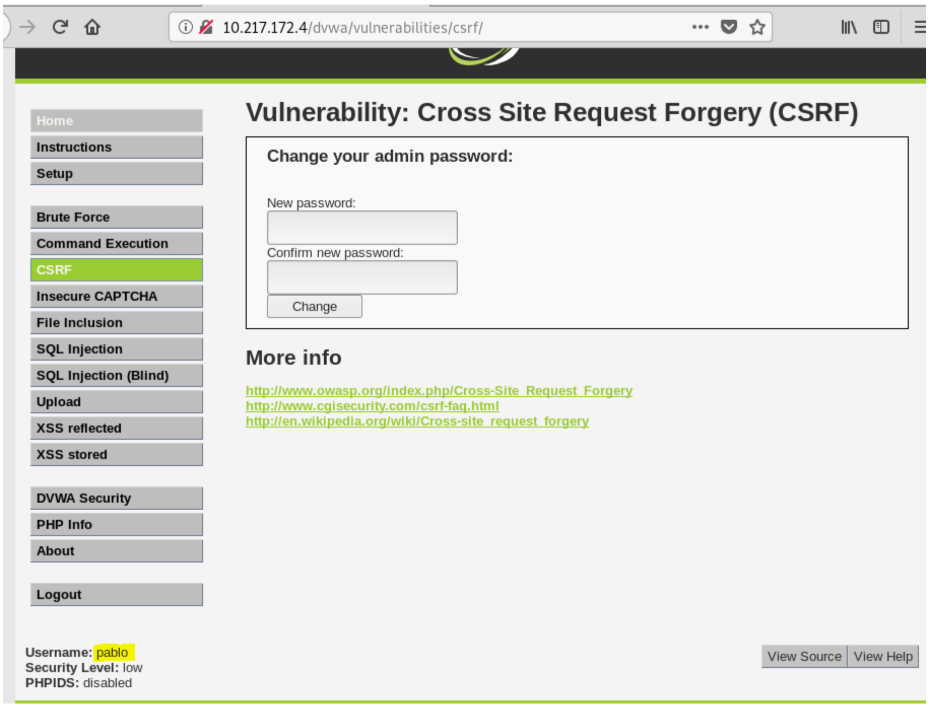
The help button allows you to view hits/tips for each vulnerability and for each security level on their respective page.

Username: smithy
Security Level: low
PHPIDS: disabled

Figure 32 DVWA User Smithy

BURP was used to intercept parameters. Then the session id is replaced with the one hijacked from Pablo. PHPSESSID=ls8ro3qvg9hf91ovlrpheibmn5

It can be observed that Pablo is the user logged in as per Figure 33 below



Vulnerability: Cross Site Request Forgery (CSRF)

Change your admin password:

New password:

Confirm new password:

More info

http://www.owasp.org/index.php/Cross-Site_Request_Forgery
http://www.cgisecurity.com/csrf-faq.html
http://en.wikipedia.org/wiki/Cross-site_request_forgery

Username: pablo
Security Level: low
PHPIDS: disabled

Figure 33 Smithy's Stolen Pablo Session

By using the session ID of Pablo, a hacker can login the access level as Pablo. The session cookies are piece of data which can give access to entire information of the user whose session ID is stolen. Highlighting the Yahoo! Case scenario, the session ID gave access to vital personal information such as names, email addresses, phone numbers, and birthdays. The personal information of the users which are not public domain. In the example of Pablo are for the company to process the data and deliver services.

The session online identifiers are unique and can “pretend to be someone else” [135]

Keeping in mind the experimentation and analysis performed it can be said that session ID or session cookie are personal data. As a recommendation cookie ID should have provisions in the GDPR as personal data.

7 Conclusions and Recommendations

The growth of the internet economy has bought with it several challenges to users' privacy and data protection, particularly with the emergence of new tracking technologies such as online identifiers. This has allowed companies to collect and process personal data of individual users.

The usage of online identifiers has been a practice since the development of the first browser. The function of online identifiers, also known as cookies has evolved from being responsible for maintaining the current user session to tracking a user's online behaviour. There are several malicious uses of cookies which have emerged and are listed below:

- Tracking online user behaviour
- Stealing cookies to gain unauthorized access
- Unauthorized stalking and interception of communication
- Automated user profiling

The data stored on the web server, particularly, a user's personal data is of interest to the malicious agents of internet. Hence, the above malicious usage of cookie are primary motivators among the business websites, hackers and other cyber criminals to exploit the flaws in these online identifiers.

The key current regulation in force for data protection and privacy in the European Union is the GDPR. This legal framework has streamlined personal data protection and privacy by implementing a higher data protection standard for user privacy, mandating user consent for tracking their personal data through cookies and automated profiling as well as giving users, for the first time the right to access their own tracked data. However, there are certain gaps between how the GDPR was meant to be implemented and how it has actually been implemented. Some of these key gaps are summarized below:

- Ambiguous cookie consent notices
- Exclusion of first-party cookies from regulation even though private authentication data is involved

- Inclusion of third-party cookies into the code as first-party cookies
- No provision for the users to revoke consent
- No provision for users to make acceptable choices with respect to online identifiers

The GDPR has brought in more compliance to safeguard consumer rights and data with more checks on data controllers. Other prominent changes heralded by the GDPR were clarity on what is considered personal data. Cookie identifiers which previously required no consent for usage on websites, became mandatory with GDPR enforcement. The strictly necessary cookie was excluded from consent, where most websites changed to integrate their tracking codes into the first party cookies. The strictly necessary cookies maintain the session and state of the web page hence, were required for functioning. Therefore, the consent prompt that was adopted was a very simple text popup informing the users about the usage of cookies on websites without much detail and choice could be seen from the experimentation.

In comparison to the existing ePrivacy Directive, the upcoming ePR better aligns with the GDPR. The ePR also known as the cookie law, adds greater importance to user consent, profiling, privacy by design and default settings of browsers to accept or deny certain online identifiers. In doing so, the ePR addresses certain gaps in the GDPR which are summarized below:

- Transparency in cookie consent
- Transparency in cookie choices
- Browser settings for cookie choice
- Protection of communication data

Session cookies are deemed strictly necessary cookies for website functioning and they are exceptions to both the GDPR and the upcoming ePR laws.

However, this thesis has established that session cookies remain most susceptible to cookie based cyber-attacks. Session cookies contain a session ID which are of interest to the malicious agents of internet. The session ID can be hacked / forged or stolen in multiple ways, an example of which was demonstrated in Chapter 4. They are responsible for establishing the session including authentication as the session ID would be matched with the one from the web server.

The vulnerabilities of session based online identifiers to cyber-attacks is summarized below in Table 6:

Online Identifier	**Susceptibility**
Session Cookies (First- Party Cookies)	• Predictable Algorithm • Transmission over insecure protocol • Limited fingerprinting and active scanning of brute force attempts • Short length of session ID • Longer Expiration of cookies

Table 6 Online Identifier with their Susceptibility

It can be assumed that a session ID, or a cookie ID is personal data just like the IP addresses of a user as it can uniquely identify the browser-user explicitly if reverse engineered.

In the light of the above vulnerability analysis of session cookies and the identified gaps in both the GDPR and the existing draft of the ePR to address these vulnerabilities, the author would like to make some recommendations that would enhance cookie security regulations. These are listed below:

1. Adding the Privacy by design (Article 10 of ePR) tenet, a key component to user privacy management which was removed in the last draft [95] which offered users the convenience to select the cookie choices acceptable to them (First Party, Third Party or Advertisement cookies).

2. The standardization of custom privacy settings like those offered by browser consortiums such as the Do Not Track initiative and Platform for Privacy Preferences (P3P). Standardization of such initiatives would promote more widespread adoption of these settings thereby promoting user privacy and security.

3. First-party cookies that store the authentication information in session ID, should be included in the context of personal data into legislature, which will offer better technical implementations for safeguarding the session management.

7.1 Future Work

The future work to this research would be to analyse advertisers (online behaviour tracking platforms) and their communication using online identifiers and the compliance of such platforms with the GDPR and the upcoming ePR. Development of online platforms for the automated analysis of privacy policies, consent and online identifiers to check compliance with the GDPR and ePR.

References

[1] R. Tirtea, C. Castelluccia and D. Ikonomou, "Bittersweet cookies. Some security and privacy considerations," European Network and Information Security Agency, 2011.

[2] S. Hill, "Are cookies crumbling our privacy? We asked an expert to find out," 2015. [Online]. Available: https://www.digitaltrends.com/computing/history-of-cookies-and-effect-on-privacy/.

[3] J. Giannandrea and L. Montulli, "Persistent Client State:," 1994.

[4] D. K. a. L. Montulli, "RFC 2109," 1997. [Online]. Available: https://www.ietf.org/rfc/rfc2109.txt.

[5] A. Cahn, S. Alfeld, P. Barford and S. Muthukrishnan, "An Empirical Study of Web Cookies," in *International World Wide Web Conference Committee*, Montreal, 2016.

[6] D. Kristol, "HTTP Cookies: Standards, Privacy, and Politics," Lucent Technologies, New Jersey, 2001.

[7] F5 Networks, "Cookies, Sessions, and Persistence," 19 January 2018. [Online]. Available: https://www.f5.com/services/resources/white-papers/cookies-sessions-and-persistence.

[8] B. Krishnamurthy and C. E. Wills, "Generating a Privacy Footprint on the Internet," in *ACM SIGCOMM Conference on Internet Measurement*, Rio de Janeriro, 2006.

[9] E. Friberg and NordVPN, "Super cookies: definition and removal," 17 May 2018. [Online]. Available: https://nordvpn.com/blog/super-cookies-going-global/.

[10] J. Kastrenakes, "FCC fines Verizon $1.35 million over 'supercookie' tracking," 7 March 2016. [Online]. Available: https://www.theverge.com/2016/3/7/11173010/verizon-supercookie-fine-1-3-million-fcc.

[11] P. Eckersley, "How Unique Is Your Web Browser?," Electronic Frontier Foundation,.

[12] European Data Protection Supervisor, "Regulation 2016/679," [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679&from=En.

[13] European Data Protection Supervisor, "Directive 95/46/EC," [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A31995L0046.

[14] European Data Protection Supervisor, "Directive 2002/58/EC," [Online]. Available: https://eur-lex.europa.eu/eli/dir/2002/58/oj.

[15] European Data Protection Supervisor, "Directive 2009/136/EC," 2009. [Online]. Available: https://edps.europa.eu/node/3082.

[16] European Data Protection Supervisor, "Proposal ePrivacy Regulation," August 2018. [Online]. Available: https://ec.europa.eu/digital-single-market/en/proposal-eprivacy-regulation.

[17] Who tracks me, "GDPR What happened?," [Online]. Available: https://whotracks.me/blog/gdpr-what-happened.html.

[18] ICANN, "Registering Domain Names," [Online]. Available: https://www.icann.org/resources/pages/register-domain-name-2017-06-20-en.

[19] ICANN, "History of WHOIS," [Online]. Available: https://whois.icann.org/en/history-whois.

[20] Insurance Architects, "The impact of GDPR on fighting cybercrime," [Online]. Available: https://www.insurance-architects.be/en/news/impact-gdpr-fighting-cybercrime.

[21] Emerald Publishing, "Challenges for online privacy: the use of cookies in social media," p. Challenges for online privacy: the use of cookies in social media.

[22] D. Kristol, "RFC 2965," 2000. [Online]. Available: https://tools.ietf.org/html/rfc2965.

[23] J. J, "HTTP Cookie Weakness, Attack Methods and Defense Mechanisms: A systematic literature review," 2018.

[24] A. F. M. J. Tappenden, "Cookies: A Deployment Study and the Testing Implications," *ACM Transactions on the Web (TWEB),* vol. 3, no. 3, p. 49, 2009.

[25] K. P. M. W. ,. K. P. Brahim Zarouali, "“Do you like cookies?” Adolescents' skeptical processing of retargeted," *Computers in Human Behavior,* p. 9, 2016.

[26] I. Sanchez-Rola, X. Ugarte-Pedrero, I. Santos and P. G. Bringas, "The web is watching you: A comprehensive review of web-tracking techniques and countermeasures," *Logic Journal of the IGPL,* vol. 25, no. 1, p. 18–29, 2017.

[27] B. Krishnamurthy and C. Wills, "Privacy diffusion on the web: a longitudinal perspective," in *18th international conference on World Wide Web (WWW)*, 2009.

[28] F. Roesner, T. Kohno and D. Wetherall, "Detecting and Defending Against Third-Party Tracking on the Web," in *9th USENIX Conference on Networked Systems Design and Implementation*, Berkeley, 2012.

[29] J. R. Mayer and J. C. Mitchell, "Third-Party Web Tracking: Policy and Technology," in *IEEE Symposium on Security and Privacy*, 2012.

[30] I. Ayandi, A. Serhrouchni, G. Pujolle and N. Simoni, "HTTP Session Management: Architecture and Cookies Security," *IEEE,* 2011.

[31] J. Ruohonen and V. Leppänen, "Whose Hands Are in the Finnish Cookie Jar?," in *European Intelligence and Security Informatics Conference*, 2018.

[32] C. Bader, E.-L. Castefelt and L. Gunnarsson, "The receipe for cookies - a study about cookies & the GDPR-Law," University of Boras, Boras, 2017.

[33] 8027, "Profiling and Online Behavioural Advertisement Under the GDPR," University of Oslo, Oslo.

[34] T. Urban, D. Tatang, M. Degeling, T. Holz and N. Pohlmann, "The Unwanted Sharing Economy: An Analysis of Cookie Syncing and User Transparency under GDPR," 2018.

[35] M. Degeling, C. Utz, C. Lentzsch, H. Hosseini, F. Schaub and T. Holz, "We Value Your Privacy ... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy," in *Network and Distributed Systems Security (NDSS) Symposium 2019*, San Diego, 2019.

[36] European Council, "Recital 20 Interference with end-users' terminal equipment," May 2018. [Online]. Available: https://indivigital.com/resources/eprivacy/recital-20/.

[37] A. Barth, "RFC 6265 - HTTP State Management Mechanism," Internet Engineering Task Force (IETF), [Online]. Available: http://www.faqs.org/rfcs/rfc6265.html.

[38] Nanyang Technological University | NTU, "HTTP State & Session Management," [Online]. Available: https://www.ntu.edu.sg/home/ehchua/programming/webprogramming/HTTP_StateManagement.html.

[39] J. Dizdarevic, F. Carpio, A. Jukan and X. Masip-Bruin, "A Survey of Communication Protocols for Internet of Things," *ACM Computer,* vol. 1, no. 1, p. 30, 2019.

[40] D. Nickull, D. Hinchcliffe and J. Governor, Web 2.0 Architectures, O'Reilly, 2009.

[41] Network Working Group, "RFC2616," [Online]. Available: https://tools.ietf.org/html/rfc2616.

[42] R. Fielding, J. Gettys, J. Mogul, H. Frystyk and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1 - RFC 2068," Network Working Group, January 1997. [Online]. Available: https://www.ietf.org/rfc/rfc2068.txt.

[43] I. Gricorik, High Performance Browser Networking, O'Reilly Media, 2013 .

[44] M. Belshe, R. Peon and E. M. Thomson, "Hypertext Transfer Protocol Version 2 (HTTP/2)," Internet Engineering Task Force (IETF), May 2015. [Online]. Available: https://tools.ietf.org/html/rfc7540.

[45] Technical Info, "Web Based Session Management," Technical Info.

[46] E. R. Fielding, "RFC7231," 2014. [Online]. Available: https://tools.ietf.org/html/rfc7231.

[47] A. Barth, "RFC6265," [Online]. Available: https://tools.ietf.org/html/rfc6265.

[48] PHP, "setcookie — Send a cookie," [Online]. Available: https://www.php.net/manual/en/function.setcookie.php.

[49] Hacking Articles, "Beginner Guide to Understand Cookies and Session Management," [Online]. Available: https://www.hackingarticles.in/beginner-guide-understand-cookies-session-management/.

[50] Coding Horror, "Protecting Your Cookies: HttpOnly," 28 August 2008. [Online]. Available: https://blog.codinghorror.com/protecting-your-cookies-httponly/.

[51] Mozilla, "HTTP cookies," [Online]. Available: https://developer.mozilla.org/en-US/docs/Web/HTTP/Cookies.

[52] HTTP Working Group, "Same-Site Cookies," 20 June 2016. [Online]. Available: https://tools.ietf.org/id/draft-ietf-httpbis-cookie-same-site-00.txt.

[53] Netsparker, "Using the Same-Site Cookie Attribute to Prevent CSRF Attacks," [Online]. Available: https://www.netsparker.com/blog/web-security/same-site-cookie-attribute-prevent-cross-site-request-forgery/.

[54] L. Dubrawsky, Eleventh Hour Security+, Burlington: Elsevier Inc., 2010.

[55] Martin, Daniek; NGS Secure, "Development and Implementation of Secure Web Applications," Centre for the Protection of National Infrastructure, 2011.

[56] Me and My Shadow, "Flash Cookies, Local Storage and Web Beacons," [Online]. Available: https://myshadow.org/flash-cookies-local-storage-and-web-beacons.

[57] VERINT, "Cookie Policy - Verint," [Online]. Available: https://www.verint.com/our-company/legal-documents/cookies-overview/.

[58] Practical Ecommerce, "An Introduction to Flash Cookies; How to Manage Them," [Online]. Available: https://www.practicalecommerce.com/An-Introduction-to-Flash-Cookies-How-to-Manage-Them.

[59] J. Angwin and M. Tigas, "Zombie Cookie: The Tracking Cookie That You Can't Kill," 14 January 2015. [Online]. Available: https://www.propublica.org/article/zombie-cookie-the-tracking-cookie-that-you-cant-kill.

[60] O. Sörensen, "Zombie-cookies: Case studies and mitigation," *8th International Conference for Internet Technology and Secured Transactions,* 2013.

[61] ENISA, "Privacy considerations of online behavioural tracking," 19 October 2012.

[62] R. Sheldon, "Supercookies take a bite out of enterprise desktop security," [Online]. Available: https://searchenterprisedesktop.techtarget.com/tip/Supercookies-take-a-bite-out-of-enterprise-desktop-security.

[63] SAP, "Cookie Manipulation," [Online]. Available: https://help.sap.com/saphelp_nw73/helpdata/de/72/7a1cf208fd47169ee4bd58fded9408/content.htm?no_cache=true.

[64] S. Sivakorn, J. Polakis and A. D. Keromytis, "HTTP Cookie Hijacking in the Wild: Security and Privacy Implications".

[65] PortSwigger, "Cross Site Scripting (Reflected)," [Online]. Available: https://portswigger.net/web-security/cross-site-scripting/reflected.

[66] International Chamber of Commerce, "ICC UK Cookie Guide," November 2012. [Online]. Available: https://www.cookielaw.org/media/1096/icc_uk_cookiesguide_revnov.pdf.

[67] European Union, "Article 4 Definations (EU GDPR)," [Online]. Available: http://www.privacy-regulation.eu/en/article-4-definitions-GDPR.htm.

[68] European Union, "Article 6 Lawfulness of processing (EU GDPR)," [Online]. Available: http://www.privacy-regulation.eu/en/article-6-lawfulness-of-processing-GDPR.htm.

[69] European Union, "Article 9 Processing of special categories of personal data," [Online]. Available: http://www.privacy-regulation.eu/en/article-9-processing-of-special-categories-of-personal-data-GDPR.htm.

[70] European Union, "Article 12 Transparent information, communication and modalities for the exercise of the rights of the data subject," [Online]. Available: Transparent information, communication and modalities for the exercise of the rights of the data subject.

[71] European Union, "Article 20 - GDPR Right to data portability," [Online]. Available: http://www.privacy-regulation.eu/en/article-20-right-to-data-portability-GDPR.htm.

[72] European Union, "Article 22 - GDPR Automated individual decision-making, including profiling," [Online]. Available: http://www.privacy-regulation.eu/en/article-22-automated-individual-decision-making-including-profiling-GDPR.htm.

[73] European Union, "Article 25 Data protection by design and by default," [Online]. Available: http://www.privacy-regulation.eu/en/article-25-data-protection-by-design-and-by-default-GDPR.htm.

[74] European Union, "Article 13 - GDPR Information to be provided where personal data are collected from the data subject," [Online]. Available: http://www.privacy-regulation.eu/en/article-13-information-to-be-provided-where-personal-data-are-collected-from-the-data-subject-GDPR.htm.

[75] European Union, "What is personal data?," [Online]. Available: https://ec.europa.eu/info/law/law-topic/data-protection/reform/what-personal-data_en.

[76] European Union, "Recital 68 - GDPR," [Online]. Available: http://www.privacy-regulation.eu/en/recital-68-GDPR.htm.

[77] Article 29 Data Protection Working Party, "Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679," 3 October 2017. [Online]. Available: https://ec.europa.eu/newsroom/document.cfm?doc_id=47742.

[78] Data Protection Working Party, "Opinion 2/2010 on online behavioural advertising," *Article 29 ,* 22 June 2010.

[79] Judgement of the Court, "Patrick Breyer vs Bundesrepublik Deutschland,," Case-law of the Court of Justice, 12 May 2016. [Online]. Available: http://curia.europa.eu/juris/document/document.jsf?text=&docid=184668&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=1116945.

[80] IAB Europe, "Working paper 04/2018," *Data Subject Requests,* 6 April 2018.

[81] European Commission, "Proposed e-Privacy regulation changes," 10 1 2017. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52017PC0010&from=EN.

[82] European Council, "Article 5 Confidentiality of electronic communications data," 4 May 2018. [Online]. Available: https://indivigital.com/resources/eprivacy/article-5/.

[83] European Council, "Article 6 Permitted processing of electronic communications data," 4 May 2018. [Online]. Available: https://indivigital.com/resources/eprivacy/article-6/.

[84] European Council, "Article 8 Protection of end-users' terminal equipment information," 4 May 2018. [Online]. Available: https://indivigital.com/resources/eprivacy/article-8/.

[85] European Council , "Article 10 Information and options for privacy settings to be provided," 4 May 2018. [Online]. Available: https://indivigital.com/resources/eprivacy/article-10/.

[86] European Council, "Recital 21 Obtaining end-users' consent to set cookies," 4 May 2018. [Online]. Available: https://indivigital.com/resources/eprivacy/recital-21/.

[87] European Council, "Recital 23 Cookie settings in browsers and other applications," 4 May 2018. [Online]. Available: https://indivigital.com/resources/eprivacy/recital-23/.

[88] European Data Protection Board, "Opinion of the Board (Art.64)," 2019.

[89] Data Protection Working Party, "Working Document 02/2013 providing guidance on obtaining consent for cookies," *Article 29,* October 2013.

[90] European Commission, "Information Provider's Guide," [Online]. Available: http://ec.europa.eu/ipg/basics/legal/cookies/index_en.htm.

[91] M. Wlosik and M. Sweeney, "What's the Difference Between First-Party and Third-Party Cookies?".

[92] IETF RFC 2965, "HTTP State Management Mechanism," Internet Society, 2000. [Online]. Available: https://www.ietf.org/rfc/rfc2965.txt.

[93] J. Penland, "Cookie Guide," 15 February 2019. [Online]. Available: https://www.whoishostingthis.com/resources/cookies-guide/.

[94] IT-Pol, "Five reasons to be concerned about the Council ePrivacy draft," 26 September 2018. [Online]. Available: https://edri.org/five-reasons-to-be-concerned-about-the-council-eprivacy-draft/.

[95] Council of European Union, "Interinstitutional File: 2017/0003(COD)," 13 March 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CONSIL:ST_7099_2019_REV_1&from=EN.

[96] Council of the European Union, "Interinstitutional File: 2017/0003(COD)," November, Brussels, 2018.

[97] Information Commissioner Officee, "When is consent appropriate?," [Online]. Available: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/consent/when-is-consent-appropriate/.

[98] New York Times, "Cookie Policy," [Online]. Available: https://www.nytimes.com/subscription/dg-cookie-policy/output.html.

[99] Data Protection Working Party - Article 29, "Opinion 04/2012 on Cookie Consent Exemption," 7 June 2012. [Online].

[100] National Geographic, "Cookie Policy of National Geographic," [Online]. Available: https://www.nationalgeographic.com/legal/cookie-policy/.

[101] D. Kontotasiou, "Consent & Cookies: How Will GDPR and the ePrivacy Regulation Impact Websites?," 12 December 2018. [Online]. Available: https://blog.convert.com/consent-cookies-how-will-gdpr-and-the-eprivacy-regulation-impact-websites.html.

[102] Clym.IO, "Clym Cookie Policy," [Online]. Available: https://www.clym.io/.

[103] Google, "Privacy Policy - Google," [Online]. Available: https://policies.google.com/privacy?hl=en.

[104] Cobham, "Privacy Policy - Cobham," [Online]. Available: https://www.cobham.com/about-cobham/cookies/.

[105] The Bitcoin Revolution, "About Cookie - Bitcoin Revolution," [Online]. Available: https://the-bitcoinrevolution.com/cookie.php.

[106] e-Estonia, "Home Page," [Online]. Available: https://e-estonia.com/.

[107] CSpace Hostings, "Home Page," [Online]. Available: https://cspacehostings.com.

[108] Alexa an Amazon.com Company, "The top Regional Estonia sites on the web in News and Media," [Online]. Available: https://www.alexa.com/topsites/category/Regional/Europe/Estonia/News_and_Media.

[109] Google, "Google Analytics Cookie Usage on Websites," [Online]. Available: https://developers.google.com/analytics/devguides/collection/analyticsjs/cookie-usage.

[110] aripaev.ee, "Majandus- ja äriuudised Eestist ja välismaalt.," [Online]. Available: https://www.aripaev.ee.

[111] Facebook, "How the Facebook pixel works," [Online]. Available: https://www.facebook.com/business/learn/facebook-ads-pixel.

[112] Similarweb, "postimees.ee Traffic Statistics," [Online]. Available: https://www.similarweb.com/website/postimees.ee#overview.

[113] Google, "Get your ad on Google today.," [Online]. Available: https://ads.google.com/intl/en/home/.

[114] Postimees, "Postimees: Värsked uudised Eestist ja välismaalt," [Online]. Available: https://www.postimees.ee.

[115] Delfi, "DELFI - Värsked uudised Eestist ja välismaalt - DELFI," [Online]. Available: https://www.delfi.ee/.

[116] Ekspress Meedia, "Ekspress Meedia," [Online]. Available: https://www.ekspressmeedia.ee.

[117] Lehepunkt OÜ, "Lehepunkt," [Online]. Available: http://www.lehepunkt.ee.

[118] ERR, "ERR News readership grows by 72.8% in first half of 2018," 13 July 2018. [Online]. Available: https://news.err.ee/846289/err-news-readership-grows-by-72-8-in-first-half-of-2018.

[119] ERR, "ERR," ERR, [Online]. Available: https://www.err.ee.

[120] J. Mayer, A. Narayanan and S. Stamm, "Do Not Track: A Universal Third-Party Web Tracking Opt Out," [Online]. Available: https://www.ietf.org/archive/id/draft-mayer-do-not-track-00.txt.

[121] Future of Privacy Forum, "What is Do Not Track?," [Online]. Available: https://allaboutdnt.com.

[122] Spread Privacy, "The "Do Not Track" Setting Doesn't Stop You from Being Tracked," DuckDuckGo, 05 February 2019. [Online]. Available: https://spreadprivacy.com/do-not-track/.

[123] L. B. A. J. J. XU, "TestingWeb applications focusing on their specialties," in *ACM SIGSOFT Software Engineering Note*, 2005.

[124] M. Snider and E. Weise, "Yahoo notifies users of 'forged cookie' breach," USA Today, 15 February 2017. [Online]. Available: https://eu.usatoday.com/story/tech/news/2017/02/15/yahoo-notifies-users-forged-cookie-breach/97955438/.

[125] BBC, "Yahoo knew of 'state-backed' hack in 2014," 10 November 2016. [Online]. Available: https://www.bbc.com/news/technology-37936219.

[126] NakedSecurity, "Yahoo staff knew they were breached two years ago," Sophos, 11 November 2016. [Online]. Available: https://nakedsecurity.sophos.com/2016/11/11/yahoo-staff-knew-they-were-breached-two-years-ago/.

[127] H. L. Newman, "HACK BRIEF: HACKERS BREACH A BILLION YAHOO ACCOUNTS. A BILLION," Wired.com, 12 December 2016. [Online]. Available: https://www.wired.com/2016/12/yahoo-hack-billion-users/.

[128] S. Musil, "Yahoo says forged cookie attack accessed about 32M accounts," Cnet, 1 March 2017. [Online]. Available: https://www.cnet.com/news/yahoo-says-forged-cookie-attack-accessed-about-32m-accounts/.

[129] R. Lawler, "Yahoo hackers accessed 32 million accounts with forged cookies," Engadget, [Online]. Available: https://www.engadget.com/2017/03/01/yahoo-hackers-accessed-32-million-accounts-with-forged-cookies/.

[130] OWASP, "Session Prediction," [Online]. Available: https://www.owasp.org/index.php/Session_Prediction.

[131] M. Kolšek, "Session Fixation Vulnerability in Web-based Application," ACROS Security, December 2002. [Online]. Available: http://www.acros.si/papers/session_fixation.pdf.

[132] The Web Application Security Consortium, "Credential and Session Prediction," 2010. [Online]. Available: http://projects.webappsec.org/w/page/13246918/Credential%20and%20Session%20Prediction.

[133] D. Endler, "Brute-Force Exploitation of Web Application Session IDs," iDefence, Chantilly, 2001.

[134] OWASP, "Session hijacking attack," 14 August 2014. [Online]. Available: https://www.owasp.org/index.php/Session_hijacking_attack.

[135] Clear Site Web Solutions Corporation, "Common Misconceptions: The cookie threat...," [Online]. Available: http://www.clear-site.com/templates/virtual_members_page.php?pageid=27.

[136] L. S. Sterling, The Art of Agent-Oriented Modeling, London: The MIT Press, 2009.

[137] T. Libert, "An Automated Approach to Auditing Disclosure of Third-Party," in *International World Wide Web Conference Committee*, Lyon, 2018.

[138] G. V. N. H. A. V. Edith G. Smit, "Understanding online behavioural advertising: User knowledge, privacy," *Computers in Human Behavior,* p. 8, 2013.

[139] D. De Lima and A. Legge, "The European Union's approach to online behavioural advertising: Protecting individuals or restricting business?," *Computer Law & Security Review,* vol. 30, p. 8, 2014.

[140] S. Englehardt, D. Reisman, C. Eubank, P. Zimmerman, J. Mayer, A. Narayanan and E. W. Felten, "Cookies That Give You Away: The Surveillance Implications of Web Tracking," in *International World Wide Web Conference Committee (IW3C2)*, Florence, 2015.

[141] D. Lyon, "Everyday Surveillance - Personal data and social classifications," *Information, Communication & Society,* vol. 5, no. 2, p. 16, 2002.

[142] European Commission, "Data Protection Factsheet," 2018. [Online]. Available: https://ec.europa.eu/commission/sites/beta-political/files/data-protection-factsheet-sme-obligations_en.pdf.

[143] OWASP, "Testing Guide v4," [Online]. Available: https://www.owasp.org/index.php/OWASP_Testing_Guide_v4_Table_of_Contents.

[144] eesti.pl, "Estonia - kultura, podróże, aktualności w Eesti.pl," [Online]. Available: https://www.eesti.pl.