UNIVERSITY OF TARTU

Institute of computer science

Cyber Security Curriculum

Priit Põdra

Web tracking in the most popular Estonian websites

Master's thesis (24 EAP)

Supervisor: Arnis Paršovs PhD

Tartu 2021

Web tracking in the most popular Estonian websites

Abstract:

Every day we open our computers, laptops, or mobile phones to browse the web. We visit different websites and open various links or look for items. After some time, a separate website offers us a picture of the same thing we were looking for. That means we are being tracked and delivered tailored advertisements depending on our previous interests and location based on cookies content. What makes the situation complex is that they are not permitted to do that.

The work aimed to study how are visitors of popular Estonian websites tracked and how their privacy is affected. For that, all the cookies were identified by category and type. We determine if popular Estonian webpages comply with the ePrivacy Directive to understand if visitors of popular Estonian websites were tracked without consent. Finally, we calculate which are effective defense methods against third-party tracking.

This study has been based on 22 popular Estonian websites ranked by Amazon Alexa.com. These websites were divided into five categories: banking, education, e-commerce, news, and services, and for the crawl OpenWPM, a framework of Princeton University was used.

The results showed that 64% of the popular Estonian websites use third-party cookies, and most of these websites track visitors without their consent.

Keywords:

Web tracking, cookies, ePrivacy Directive, third to third-party tracking, web beacon, Ghostery, Do Not Track, private browsing mode.

CERCS:

T120 - Systems engineering, computer technology

Veebijälgimine Eesti populaarseimatel veebisaitidel

Lühikokkuvõte:

Iga päev avame me oma arvutid, sülearvutid või mobiiltelefonid, et külatada erinevaid veebilehti. Seda tehes avame me erinevaid linke või otsime midagi konkreetset. Mõne aja möödudes külastades mõnda teist veebilehte pakub see meile sama asja, mida me olime just vaadanud

eelmisel veebilehel. See tähendab, et meid jälgitakse ja pakutakse meile suunatud reklaame, mis põhinevad meie eelnevatel otsingutel ja asukohal. Mis teeb kogu selle olukorra keeruliseks on see, et nad ei tohiks seda teha.

Selle töö eesmärk oli teada saada, kuidas jälgitakse populaarsete Eesti veebilehtede külastajaid ja kuidas see mõjutab nende privaatsust. Selleks tuvastati kõikide kasutatavate küpsiste tüübid ja kategooriad. Tegime kindlaks, kas populaarsed Eesti veebilhed vastavad e-privaatsuse direktiiviga kehtestatud nõuetele, et teada saada, kas külastajaid jälgitakse ilma nende nõusolekuta. Viimaks selgitasime välja, millised on efektiivsed meetodid kolmandate osapoolte jälgimise peatamiseks.

See uuring põhineb kahekümnekahel populaarsel Eesti veebisaidil, mille on pingeritta seadnud Amazon Alexa.com. Need veebilehed jagati viide kategooriasse: pangandus, haridus, kaubandus, uudised ja teenused ja tulemuse saamiseks kasutati OpenWPM programmi, mis on välja töötatud Princeton ülikooli poolt.

Tulemused näitasid, et 64% populaarsetest Eesti veebisaitidest kasutavad kolmanda osapoole küpsiseid ja nende lehtede külastajaid jälgitakse ilma nende nõusolekuta.

Võtmesõnad:

Veebijälgimine, küpsised, ePrivaatsuse direktiiv, kolmandalt osapoolelt kolmandale osapoolele jälgimine, veebimajakad, Ghostery, Do Not Track, privaatne veebisirvimine.

CERCS:

T120 – Süsteemitehnoloogia, arvutitehnoloogia

Table of contents

1 Introduction 9

1.1 Research question and tasks 9

1.2 Contributions 10

1.3 Structure 10

2 Literature review and background 11

2.1 Web tracking 11

2.1.1 Web tracking privacy implications 14

2.2 Cookies 15

2.2.1 Third-party cookies 16

2.2.2 Third-party HTTP requests 17

2.3 ePrivacy Directive 18

2.4 Anti-tracking privacy measures 19

2.4.1 Do Not Track header 19

2.4.2 Mozilla private browsing mode 20

2.4.3 Ghostery browser extension 21

3. Web tracking study 22

3.1 Research goal and questions 22

3.2 Websites analyzed 23

3.3 Methodology 24

3.3.1 Data collection 24

3.3.2 Data analysis 25

4. Results 27

4.1 Cookies 27

4.1.1 Overview of cookies and comparison by type 27

4.1.2 Cookies by categories 32

4.1.3 Third-party cookies 34

4.1.4 Third party HTTP requests 36

4.1.5 Third to third-party HTTP redirects 39

4.2 Compliance to EU Privacy Directive 40

4.2.1 Cookie lifetime 42

4.3 Effectiveness of anti-tracking privacy measures 44

4.3.1 DO Not Track header 44
4.3.2 Private browsing feature and Ghostery browser extension 44
4.4 Summary of the results 47
5. Discussion 48
6. Conclusion 49
6.1 Answers to research tasks 49
6.2 Limitations 50
6.3 Recommendations for the future research 50
References 51
Appendix 1: Alexa TOP 50 popular websites 58
Appendix 2: Top-20 tracking companies and their third-party domains 60
Appendix 3: Cookie table 62
Appendix 4: HTTP request of web beacon 84
Appendix 5: OpenWPM crawl settings 87
License 89

Table of figures

Figure 1: Tracking mechanisms, implications, and defenses [8].

Figure 2: Overview and conceptualization of web tracking [11].

Figure 3: Cookie syncing between websites.

Figure 4: Example of Do Not Track.

Figure 5: Mozilla private browsing window.

Figure 6: Example of Mozilla add-on Ghostery.

Figure 7: High level overview of OpenWPM [59].

Figure 8: Example of Mozilla web developer extension storage.

List of tables

Table 1: Popular Estonian websites [2].

Table 2: Comparison of cookies by categories.

Table 3: Summary of banking category cookies.

Table 4: Summary of education category cookies.

Table 5: Summary of e-commerce category cookies.

Table 6: Summary of services category cookies.

Table 7: Summary of news category cookies.

Table 8: List of third-party cookies.

Table 9: Third-party tracking companies and their domains.

Table 10: Summary of third-party HTTP requests.

Table 11: Top5 of domains, where third-party requests are made.

Table 12: Example of show facepile attribute [65].

Table 13: Example of HTTP request from www.tallinn.ee website.

Table 14: Example of HTTP request from www.telia.ee website.

Table 15: OpenWPM example of third-to-third party trackers.

Table 16: Summary of popular websites by ePrivacy Directive.

Table 17: OpenWPM summary of cookies with duration more than 5 years.

Table 18: Summary of Ghostery and private browsing cookies.

Table 19: Summary of tracking in popular Estonian websites.

List of graphs

Graph 1: Summary of banking category cookies by type.

Graph 2: Summary of education category cookies by type.

Graph 3: Summary of e-commerce category cookies by type.

Graph 4: Summary of services category cookies by type.

Graph 5: Summary of news category cookies by type.

Graph 6: Summary of popular websites by cookie duration time.

List of acronyms and definitions

B2C	Business-to-consumer marketing
DNT	Do Not Track
ENISA	European Union Agency for Network and Information Security
EU	European Union
HTTP	Hypertext Transfer Protocol
IP	Internet Protocol address
TLD	Top level domain
TOP	Table of popularity
URL	Uniform resource locator
US	United States

1 Introduction

Estonian citizens believe that their browsing information is used responsibly. 74% think that they have nothing to hide, and 88% believe that they cannot use services on websites if they do not agree with conditions. They do not know what is done with the information collected about them and how their privacy is affected. [1] After every visit to a popular website, it is possible to calculate the number of average daily visits and pageviews over the past month and point out popularly visited Estonian websites [2].

The aim of this work is to find how are visitors of popular Estonian websites tracked and how their privacy is affected. This work provides an analysis of cookies by type and category. It is possible to point, how big is the percentage of tracking and whether popular Estonian websites meet the requirements of the ePrivacy Directive.

First of all, this thesis provides an overview of cookies used by popular Estonian websites and their comparison by type. After that, cookies are divided into categories. It is possible to show, how big is the average percentage of tracking cookies and which category out of five has the biggest number of cookies. We point out all the third-party cookies, analyze third-party and third to third-party HTTP requests to find different tracking mechanisms.

Secondly, we find if visitors of popular Estonian websites are tracked without their consent, and this can be done by examining websites compliance to the ePrivacy Directive. In the last part of this work, the effectiveness of anti-tracking measures Ghostery, Do Not Track, and private browsing mode is calculated.

This study has been based on 22 popular Estonian websites ranked by Amazon Alexa.com. These websites were divided into five categories: banking, education, e-commerce, news, and services, and for the crawl OpenWPM, a framework of Princeton University was used.

1.1 Research question and tasks

The main research question in this work is: *How are visitors of popular Estonian websites tracked, how their privacy is affected, and which are effective defense methods against third-party tracking?*

To get the answer to research question, several tasks have to be solved:

1. **Evaluate the usage of first and third-party cookies.** Overview what type, and category cookies are used.
2. **Find if websites comply with the ePrivacy Directive.** Are visitors of popular Estonian websites tracked without user consent?
3. **Find which privacy enhancement and tracking defense techniques guarantee safer browsing.** Are Ghostery, Do Not Track and private browsing mode effective against third-party tracking?

1.2 Contributions

The main contribution in the context of this thesis is that there has not been any research on web tracking on popular Estonian websites. It is possible to track website visitor internet behavior by collecting cookies data. It can happen either when consent is given or by using monitoring instruments [3]. The second contribution is the observation of popular Estonian webpages compliance to the ePrivacy Directive. Privacy is a problem in the web because most of the web services follow their users to obtain as much personal information as possible without users' consent [4]. The third contribution is measuring the effectiveness of anti-tracking defense methods. The aim is to measure which publicly available anti-tracking methods can reduce tracking.

1.3 Structure

This study consists of 6 chapters. Chapter 1 presents an overview of the addressed problem. Chapter 2 describes the background of web tracking, privacy implications, and anti-tracking methods. Chapter 3 is web tracking study. The data collection and analysis methods are covered in this section. Details of research results will be shown in Chapter 4. Chapter 5 concentrates on the discussion of the findings. Chapter 6 concludes the work, points out limitations and gives recommendations for future research.

2 Literature review and background

This literature and background chapter has five subsections. First introduces the background of web tracking and what are the privacy implications of web tracking. Second subsection gives an overview of cookies and third-party cookies. The third-party HTTP requests section opens the background of web tracking using web beacons. Furthermore, this chapter explains, what is the main content of the ePrivacy Directive and gives overview of three anti-tracking privacy measures.

2.1 Web tracking

Several articles confirm that tracking is a problem that needs attention. But what is web tracking? *"Tracking means it can be an act or the process of following something or someone". Websites mostly gather technical data: IP address, screen determination, or the browser used at the website visit. If cookies are used, there are two main assignments: either make browsing experience better or collect browsing info and share with other counterparts*. [5]

Already in the year 2001, some conclusions were made about cookies and tracking cookies, that are actual 20 years later:

1. People do not know about cookies and how websites might use them.
2. People know how cookies can track them and are unconcerned about it.
3. People do not know which cookies they will accept and accept all of them.
4. People want to assume that they are protected from bad usage of cookies, and regulations help them. [6]

How do understand being tracked, and how to defend from that? The first indication on the website is a cookie banner or cookie notice that the user is tracked. The opposite situation occurs when the banner is not used and tracking still occurred (by looking at installed cookies). There can be a wrong understanding that all websites are tracking website visitors. [7]

Web track was initially developed to facilitate better marketing, and it is the same with most websites. There are several mechanisms and implications of how tracking can happen. As cybersecurity is about defense, there are several defense methods to minimize the impact of mechanisms and implications. There are several options, how to write about web tracking. As

shown in Figure 1, there are several tracking auditing tools and tracking mechanisms on what to focus. [4]

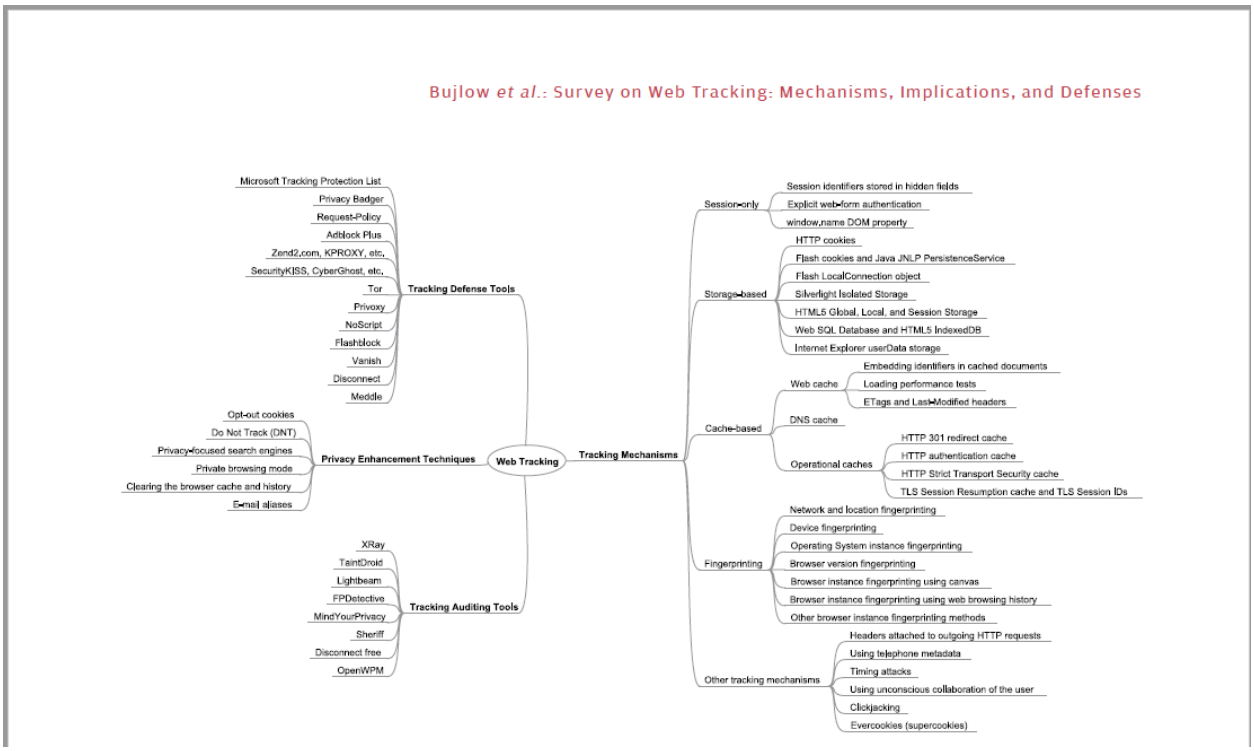


Figure 1: Tracking mechanisms, implications, and defenses [4]

For whom the collected tracking data is necessary:

1. Advertisement companies collect information to influence and produce tailored ads to users. The main aim is to find out users' interests in different categories.
2. Law enforcement and intelligence agencies collect information to perform the tasks assigned to them.
3. Website owners are interested in their website analytics. It also helps to produce a tailored website. [8]

A similar finding is that advertisement companies value more collected user data than user privacy because their business receives tracking information, and does not care, if there would be a data leakage [9]. Tracking is not always a bad thing: it is more like a relationship between consumers and online advertising companies. Advertisement companies get the most of it, but if consumers are not satisfied, there is always a possibility to delete the cookies [5].

Some companies do not offer third-party cookies; they offer “cross-device” tracking. It means search can identify the same person across many different devices. Operating system ID and IP addresses are just a few examples of what to gather. If website visitor just looked up lawnmowers on Google and is now browsing Facebook, they can get lawnmowers ad in that browsing environment. [10]

A high-level overview and conceptualization of web tracking have been done, where privacy, technology and commercial part have important role [11].

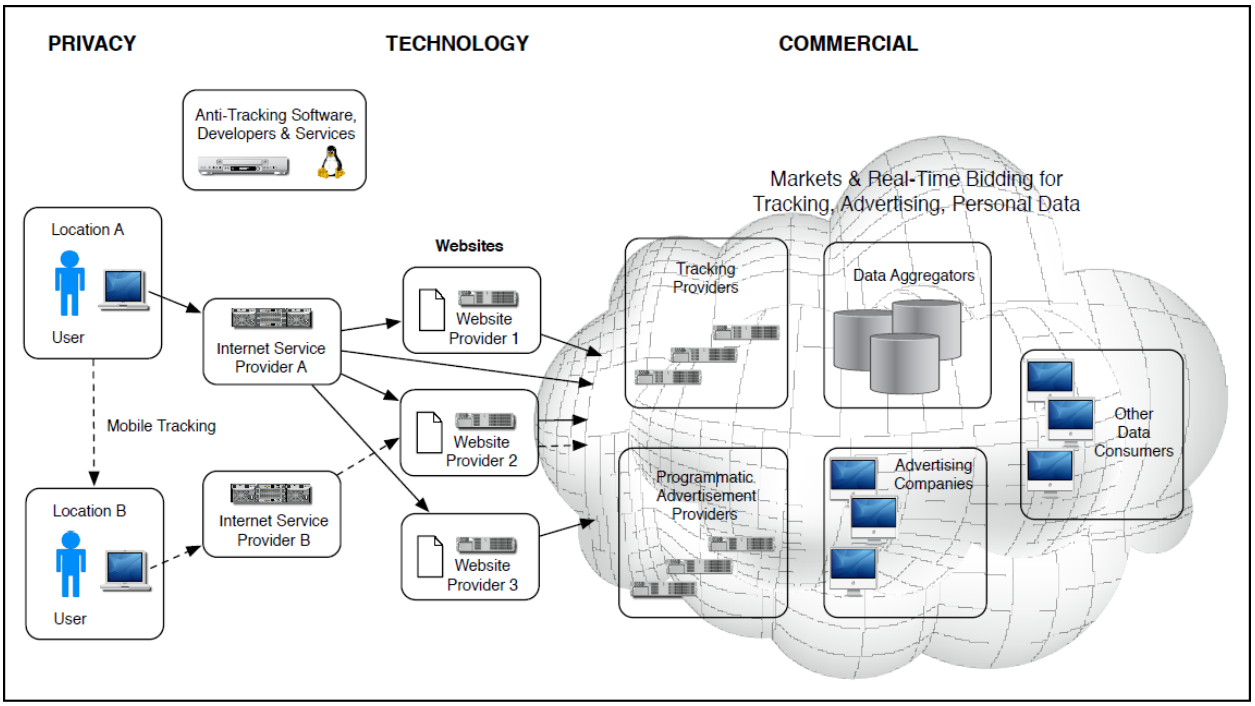


Figure 2: Overview and conceptualization of web tracking [11]

Another example is the news category. News websites are at the top of using tracking cookies. News websites have used cookies already from the beginning of 1800. Several times it is pointed that news websites are having the highest cookie coefficient [7]. For example, most popular Finnish websites have trackers, and news websites have the most significant number of trackers [12].

2.1.1 Web tracking privacy implications

"*The definition of online privacy is the level of privacy protection an individual has while connected to the Internet. It covers the amount of online security available for personal and financial data, communications, and preferences"* [13].

Timothy Libert wrote an article in New York Times about privacy issues and data collection violations. Especially problematic is the situation with news websites. The first analysis revealed that almost 50 different companies track personal data to obtain as much personal information as possible for commercial purposes [14]. Analysis of news websites' privacy policies revealed that newsreader's data was shared with third parties. Third-party cookies on these websites were used to collect information about user activities, and the result of that is targeted advertisements on these news and other websites [15].

Web tracking can be a threat to personal privacy. Companies know about visitors and can create a complete online profile. Even if browsing websites, you trust, the visitor's information can be gathered and sold to another counterpart [16]. Primarily personal data, but what is the private information different parties are interested? Personalization in the web tracking context is name, location, age, gender, and a unique computer identifier [17]. As analyzing the HTTP headers, the user's personal identifiable private data like age or gender can be revealed [18].

If the visitor first visits the website, he gets a cookie with a unique ID, but several other websites use the same third-party cookie and then it is possible to expose the unique ID and user data to third-party [3]. On average, 74% of websites accept cookies from third parties before any user consent, and it was mostly done by big players (Google), and with US background. Almost 30 % of websites do not provide a cookie banner and install some cookies, and only 7% wait for user's consent [19].

Private data leakage and identifiability are the two main privacy issues in web tracking that can lead to harmful web tracking scenarios. The most common identification method uses the IP address to track a visitor over websites visited or using cookies to track over websites called. [5]

The authors decided to analyze the third-party tracking to determine if there is a risk of leaking user privacy. Instead of blocking all access tracking, they choose to pay more attention to the

websites that use a certain words connected to users' privacy (like cookies, cards, passwords). They resulted in a tracking rate drop from 71% to 24%. They concluded, that when there is a third-party tracker, there is a question of privacy, and online ethics and the bigger is the third-party tracker' network, the greater is the potential to track users. [20].

2.2 Cookies

What are cookies? Cookies were created to make a stateful browser-server interaction possible in a stateless protocol. Cookies, also known as HTTP (Hypertext Transfer Protocol) cookies, are created and updated by the server, stored by the browser, and transmitted between the browser and the server [21]. Usually, cookies are created by web developers.

Cookies can be divided by type. First-party cookies are set by the website host (domain), and data is not passed from one website to another. Third-party cookies are the opposite; it means information is passed from one website to another, and the website has given access to collect data. As a result, the third-party cookies can be an invasion of privacy. [22]

By categories, we can classify cookies. The most common categories that were used in the context of this work are:

1. Strictly Necessary cookies: Cookies are essential for the provision of the website and its service.
2. Performance cookies: Provide statistical information on website usage.
3. Functionality cookies: Provide enhanced functionality for all website functions.
4. Targeting/Advertising cookies: Create profiles or personalized content. Set mainly by third parties and with the highest privacy risks to visitors. [23]

In the vast majority of instances, cookies are harmless but unable to guarantee privacy. Cybercriminals and fraudsters can use cookies to monitor user online activities if there is a case of unprotected information [24].

Cookies can divide by lifetime. Usually, the lifetime of the cookie is relatively short, about six months. This is a time browser should invalidate the cookies. User can use cookie expiration time

to determine the difference between a transient cookie and a tracking cookie. Cookies can divide into two categories:

1. Non-persistent cookies: mostly session cookies that are deleted after the session.
2. Persistent cookies. Stored in browser memory with expiration time [25].

Estonian websites were mentioned in a study "Towards a global perspective on web tracking". Estonian websites are among TOP6 EU countries with the most extensive number of cookies with a long validity period. A significant number of third-party cookies on websites have long validity period, that violates the EU law [26]. Cookies with a far-in-time expiration time are most commonly tracking cookies, and 80% of third-party cookies last one month or more [19]. The median lifetime of the first-party cookie is one day, and the median lifetime of a third-party cookie is six months (181 days) [27].

There are several categories where using cookies is common. Using cookies in e-commerce makes sense as their aim is a better user experience. In e-commerce, cookies are mostly used for:

1. Storing log-in information: Gives benefits.
2. Shopping carts: Shopping choices are stored and remembered.
3. Retargeting campaigns: Purchase was not successfully delivered [28].

2.2.1 Third-party cookies

"Third-party cookies are created by domains that are not the website (or domain) that you are visiting. These are usually used for online advertising purposes and placed on a website through adding scripts or tags. A third-party cookie is accessible on any website that loads the third-party server's code. Third-party cookies can be seen as an invasion of privacy". [29]

Some websites are built up, so that if you want to visit that website, you might have to accept third-party cookies, and not accepting can be a reason for site downgrade (visitor can use not all functionalities). From that, there is a big privacy question, what is more valuable: privacy or accessibility? [30].

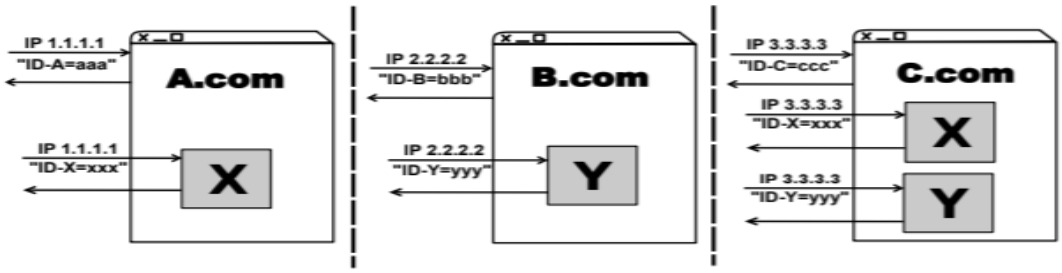


Figure 3: Cookie syncing between websites [31]

Englehardt et al. describe how syncing using third-party cookies occurs. Visitor accesses three different websites with three different IP. If the third website uses same cookies as other two, then it recognizes that this is the same user [31].

There are 20 TOP third-party tracking companies and their third-party domains (Appendix 2) [32]. US-based study about third-party cookies revealed that news websites are more reliant on third parties than non-news websites [33]. Overview of news websites in different countries resulted that 95% of news sites use third-party content, and they use them in such a considerable amount that the number exceeds trackers of 500 most popular websites in the same country. In every news site, US-based technology companies (Google, Facebook, Amazon, and Twitter) have their interests in collecting data. The first EU-based technology company had cookies only on 7% of websites. [34]

46% of Alexa.com popular websites were tracked, and they have at least one third-party tracker. At least five third-party trackers tracked 29% of the websites. What made them worry was that Google, as a third-party, collects cookie information from 25% of the Alexa.com popular websites. [35]

2.2.2 Third-party HTTP requests

One of the third-party tracking methods is the web beacon. A web beacon can “work” together with a cookie. It is a technique made for checking if visitors have accessed some content and are usually not visible to the visitor [36].

All the information gathered may also be exchanged from one third-party to another third-party making it one of the most intrusive monitoring mechanisms to compromise user's privacy. A tracker will use cookie syncing to monitor the user's visits to the websites where it is included and the websites where its partners (third-party) are included. There are two possible ways: cookie syncing or cookie forwarding. The difference is that either cookie is set or not. If yes, then it is cookie forwarding [37].

In 2007 there was a Facebook beacon scandal. The main problem was that Facebook was tracking users, even if they log out of Facebook. If a Facebook user bought something, then friends saw it [38]. Websites use B2C customer segmentation to target audience emotions. Marketers use four methods to track potential clients:

1. Behavioral data: How products on website are browsed.
2. Geographical data: By location tracking, personalized offers are sent.
3. Psychographic data: Customers are targeted based on their lifestyle.
4. Demographic data: Focusing on personal data makes it easy to do offers [39].

2.3 ePrivacy Directive

"*The EU e-Privacy Directive is a part of the European Union's strive to enhance online privacy for its citizens. Websites that are either owned by EU businesses or directed towards EU citizens must inform visitors that cookies are in use, how these cookies are used, and obtain consent before cookies can be used"*.[40]

EU Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 describes how website owner may use cookies. To comply with the ePrivacy Directive website has to have:

1. Privacy policy: It should inform how their personal information (identify an individual) is collected, used, and stored. An important thing to notice that policy has to have good visibility.
2. Have a cookie policy: It should notify users which cookies were used.
3. Have a banner: Using cookies to track.
4. Receive consent from the user: The user has the possibility to accept or decline cookies.

5. Provide an opt-out method: Possibility or knowledge to users how they can quickly delete all the cookies [41]

For example, the ePrivacy Directive says: before consent, no cookie - except technical cookies - can be installed [42]. Strengthening fines is a way how ePrivacy regulation could reduce tracking and improve privacy protection [7].

There was a study where question of what the ePrivacy regulation must do better? And suggestion was that if ePrivacy required privacy policies of websites and the scope of data, this would increase transparency and create a market for privacy [43].

2.4 Anti-tracking privacy measures

People rarely protect their online privacy and one of best defense methods for deleting cookies. They also mentioned Declining cookies, DNT, private browsing and Opt-out websites. People do not see a problem if their personal information is collected and shared with other counterparts [44].

Mozilla has worked out a solution called Total Cookie Protection. Its main work is to maintain a separate "cookie jar" for every website. If before, there was one big "jar" for every third-party, then in the future, there would be "jar" for every website. The main aim is to prevent cross-site scripting, offer better privacy and a better browsing experience [45]. Another measure is Mozilla tracking protection. Website visitors can use three different settings: standard, strict, and custom:

1. Standard: standard setting that blocks third-party tracking cookies in private windows
2. Strict: blocks third-party tracking cookies in all windows, but some websites might break
3. Custom: Allows to choose what to block.[46]

2.4.1 Do Not Track header

DNT is a browser option that is nothing more than a request. [35]. DNT allows users to opt-out tracking if they do not want to be tracked [47]. If the user does not want to be tracked, this information is sent via the HTTP header if the user has declared it in a browser. DNT supports all browsers [48].

In Mozilla user can activate DNT from settings, and it is possible to send websites a signal that the user does not want to be tracked (Figure 4).

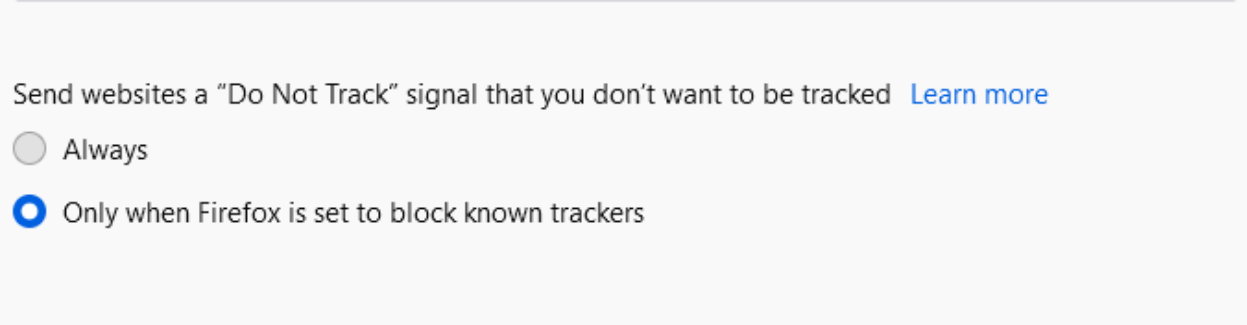


Figure 4: Example of Mozilla DNT

Third-party content can be a mass surveillance instrument, and it includes many privacy concerns (collecting data about website size, personalized visiting experience), and these were not directly used for tracking. Simple browser settings change like blocking cookies or using DNT to make it always safe and bypassed. A solution would be using blacklists [49]. Englehardt made Dissertation pointed out that DNT does not have impact on tracking, and it is ineffective [50].

2.4.2 Mozilla private browsing mode

It is not easy even for users who have more knowledge of protection methods to protect them from third-party tracking. They suggest that altering and obfuscating, blocking requests to third-party tracking services, removing user identifiers and cookies from the proposals, or using anonymity services would be the solution. However, these services are not always allowed [51. All major browser vendors include private browsing modes into their browsers under various names. In Mozilla, a feature called "Private Browsing" using a new private window. In private mode, typically, cookies and other browser persistence mechanisms are disabled. No browser history and writing of caching information to disk is prevented [8].

User can activate private browsing mode from Mozilla settings, and it helps to keep online privacy (Figure 5).

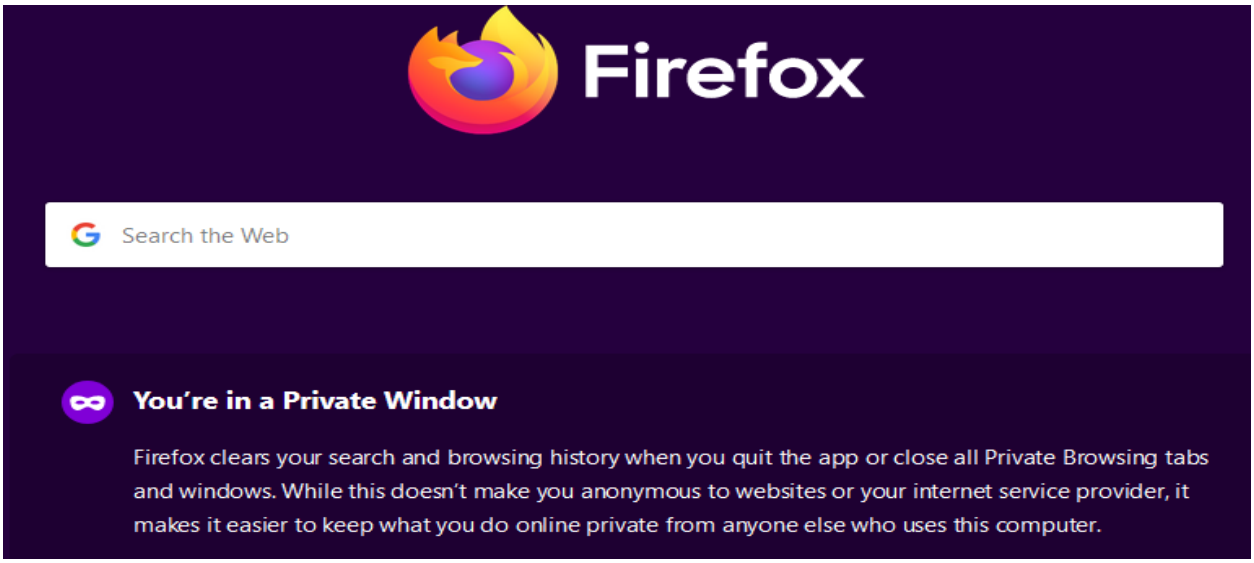


Figure 5: Mozilla private browsing window

In November 2020, project Cover your tracks was started to test online browser tracking. It allows using tools against trackers and provides privacy for everyone. In the beginning, they did a test to find if browser was leaking, and the answer was yes- web trackers were following their movements online. They have worked out simple suggestions, how the user can find a balance between privacy and convenience, and one of the suggestions is using private browsing mode [52].

2.4.3 Ghostery browser extension

Ghostery gives a possibility to block ads, stop trackers, and speed up websites. The main idea is a safer browsing experience and the possibility for a more protected online privacy [53]. Ghostery gives the user the opportunity to block tracking and decide which of those tracking mechanisms to block or permit. Users can decide based on tracking mechanisms or base on specific categories [54].

Several research was done about third parties and how browser extensions block third-party tracking. Third-party blocking extensions are very popular, and their methods overlap. Browser add-on Ghostery had the best results in the third-party blocking category with some inclusions. One of their conclusions was that the more prominent and known company, the more significant was the possibility of being blocked [55].

"By using the Ghostery browser extension for blocking trackers, you might be able to better protect your personal privacy and re-take control of your personal data. And, along the way, you might just speed up the performance of your web browser because you won't be waiting for all those annoying tracker scripts that invisibly track you in the background to load" [56].

Ghostery can be found from the Mozilla add-on section (about: addons). It allows to block ad-s and protect privacy.

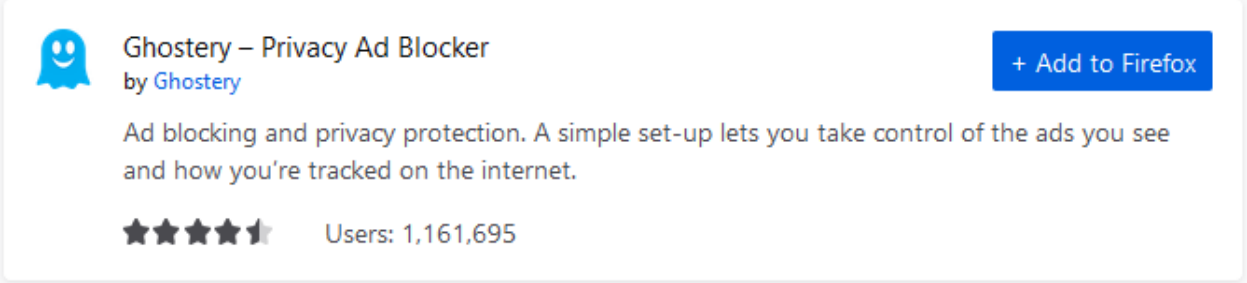


Figure 6: Example of Mozilla add-on Ghostery

3. Web tracking study

In Chapter 3 research goal and questions explain, what is the aim of this study and which data collection and analysis methods are covered to achieve this goal.

3.1 Research goal and questions

The goal of this study is to analyze what web tracking mechanisms are used in 22 most popular Estonian websites and to find answers:

- What type of cookies are set by the websites?

- What third-party tracking mechanisms are used in the websites?

- Do the websites comply to EU Privacy Directive?

- How effective are anti-tracking privacy measures? (such as DNT, Private Browsing mode and Ghostery browser add-on)

3.2 Websites analyzed

Place in the top was sorted out from Alexa.com TOP50 visited Estonian websites (Appendix 1). First crawl showed that many popular websites are back-end websites (no cookies). From there *.ee* or connected with Estonia websites were sorted out. They were categorized by their content and used IP address was contacted by OpenWPM crawl. Category was chosen by the web content and it was divided into five categories: news, banking, education, e-commerce and web services (Table 1). E-commerce and services were differentiated by selling aspect. If there is possible to buy something, then it belongs to e-commerce category. Alexa.com was chosen, as in previous research it has been pointed out, that tracking occurs more on the higher ranked websites [57].

In the Alexa.com popular websites list was www.microsoftonline.ee , but OpenWPM generated a general error code and in manual check that website does not exist [2]. In general, 22 websites were added for analysis.

Place in TOP 50	Website URL	Used IP address	Category
3	http://www.delfi.ee	185.20.100.194	News
4	https://www.postimees.ee	185.154.221.151	News
7	https://www.swedbank.ee	193.203.196.77	Banking
13	https://www.harjuelu.ee	212.47.220.55	News
15	https://www.ekool.eu	172.67.73.80	Education
18	https://www.ut.ee	193.40.5.73	Education
19	https://www.seb.ee	78.24.199.2	Banking
20	https://www.microsoftonline.ee*	-	
21	https://www.err.ee	194.36.162.172	News
22	https://www.auto24.ee	217.159.201.10	E-commerce
23	https://www.neti.ee	195.50.209.244	Web services
26	https://www.ohtuleht.ee	185.20.102.226	News
28	https://www.telia.ee	217.159.238.116	E-commerce

29	https://www.eki.ee	193.40.113.42	Education
30	https://www.zone.ee	185.31.240.3	Web services
32	https://www.ope.ee	217.146.78.179	Education
34	https://www.tootukassa.ee	195.50.195.21	Services
36	https://www.lhv.ee	91.224.189.34	Banking
38	https://www.online.ee	194.126.119.77	Web services
41	https://www.kv.ee	90.190.106.69	E-commerce
42	https://www.elu24.ee	185.154.221.150	News
43	https://www.tallinn.ee	80.235.77.44	Web services
47	https://www.andmorefashion.com	185.192.14.125	E-commerce

Table 1: Popular Estonian websites

3.3 Methodology

3.3.1 Data collection

For the collection, Ubuntu 20.04.1 is being used in Windows10 machine and additionally OpenWPM version 0.12.0, which uses Mozilla Firefox 82.0.2. For the collection of data OpenWPM – Open web privacy measurement framework is used . This framework is part of the web transparency and accountability project of Princeton University, and it allows data collection for privacy studies on a large scale [58]. Figure 6 gives the overview of OpenWPM [59]. What has changed is that, if before there was a possibility to use several browsers, then now it is only possible to use Mozilla.

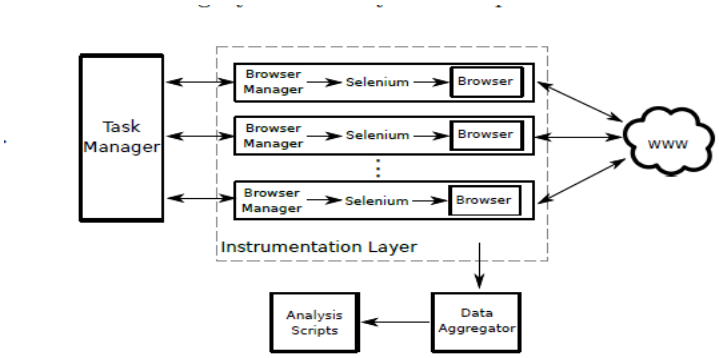


Figure 7: High level overview of OpenWPM [59]

Mozilla web developer extension storage helps to find out cookies placed on the computer and their metadata (Figure 7).

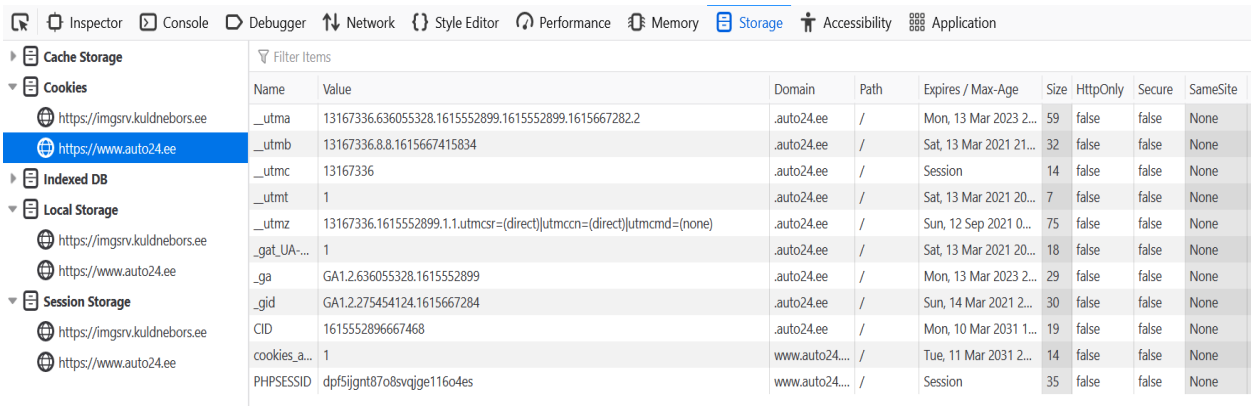


Figure 8: Example of Mozilla web developer extension storage

3.3.2 Data analysis

After the web crawl, data for the analysis was inserted to SQLlite (built for x86_64-little_endian-llp64, running on x86_64, Qt Version 5.12.8, SQLite Version 3.33.0). First of all, a total number of cookies were examined. As some of the cookies had the same name and value, unique cookies were sorted out. After that, cookies from 22 popular websites were sorted by hostname, cookie name, expiration time, first or third-party, purpose, and finally categorized. After that, every cookie category summary was pointed, and the average was calculated.

The first analysis of the results showed that the OpenWPM crawl does not show cookies accepted by the user after the consent. If the cookies were installed after the user consent, it was marked

into the table as + and a with number of cookies (7+7). If the website did not show the cookie category and its purpose, it was treated like an unknown cookie (with category number 5). The purpose of the cookie was then established with the help of Open-Cookie-Database. This database was built to describe and categorize all major cookies [60]. Later analysis revealed that some of the cookies are website specific, and CookiePro [61] was used. The reason for that was that they have more than 31 million cookies categorized.

Secondly, from the cookie results, third-party cookies were sorted out in a way that if the hostname did not match with TOP level domain from Alexa.com, it was treated as a third-party cookie. First-party requests, third-party requests and total % of third-party requests from total requests were written out from HTTP requests. The HTTP request part is a resource type, which helps determine which resource was fetched [62]. In the HTTP redirect section old request URL and new request URL were compared. If there was a difference in the TOP domain, it was treated as a third-to-third party tracker.

All the popular websites in the alexa.com Estonia list were again visited with Mozilla 86.0.1 (64 bit). The reason for that was that the websites crawl in OpenWPM was also done with Mozilla. Information about the banner, privacy info, cookie policy, user consent, and opt-out method were saved, and from first page was made a screenshot. The aim was to compare if they correspond to five points noticed in the ePrivacy Directive. Before the visit, all the Mozilla browser cookies were deleted. Analysis of popular Estonian websites will help to determine if website visitors are noticed about tracking. Timestamp of the OpenWPM was 2021-02-21T10:34:23.452Z.

For the anti-tracking defense methods, the browser Mozilla 86.0.1 (64-bit) was used. Mozilla add-on Ghostery and extension private browsing were used. All the third-party trackers were listed, and the Ghostery add-on installed to compare popular web sites storage cookies. Same method was used in the private browsing mode. DNT was tested with OpenWPM and in settings value DNT was set to true.

4. Results

In this chapter the results of the study are presented. All the cookies gathered during the OpenWPM crawl and from Mozilla are given in Appendix 3.

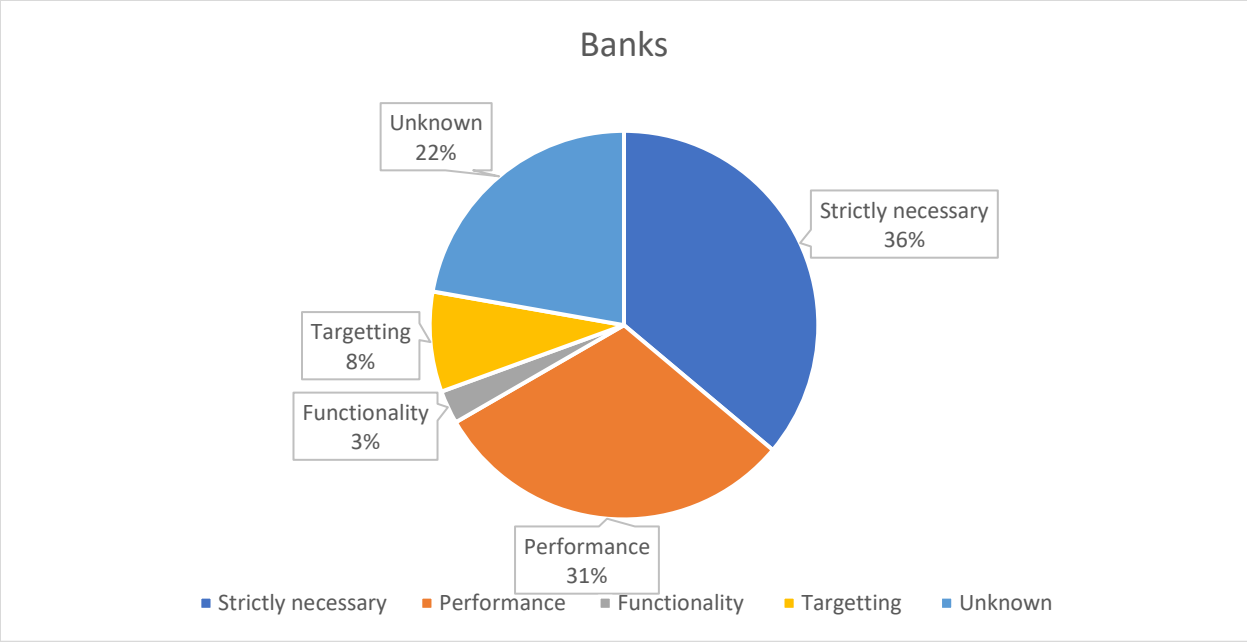
4.1 Cookies

4.1.1 Overview of cookies and comparison by type

Banking:

Site	Number of cookies by OpenWPM	Number of cookies by Mozilla	Overall remarks
www.swedbank.ee	7	7+7	OpenWPM crawl and Mozilla Browser cookie check confirmed, that only technical cookies were installed during first visit. After cookie agreement, non-technical cookies were added. As other banks had initially session cookies also, then Swedbank had only persistent cookies. Most of the cookies were unknown, because Swedbank did not notice them in their website
www.seb.ee	4	4+6	OpenWPM crawl and Mozilla cookie check confirmed, that only technical cookies were installed during first visit. After cookie agreement, non-technical cookies were added.
www.lhv.ee	5	5+7	OpenWPM crawl and Mozilla cookie check confirmed, that only technical cookies were installed during first visit. After cookie agreement, non-technical cookies (7) were added. LHV has pointed out all the cookies they use but does not write any information about *client_id* cookie.
Overall	16	16+20	

Table 3: Summary of banking category cookies



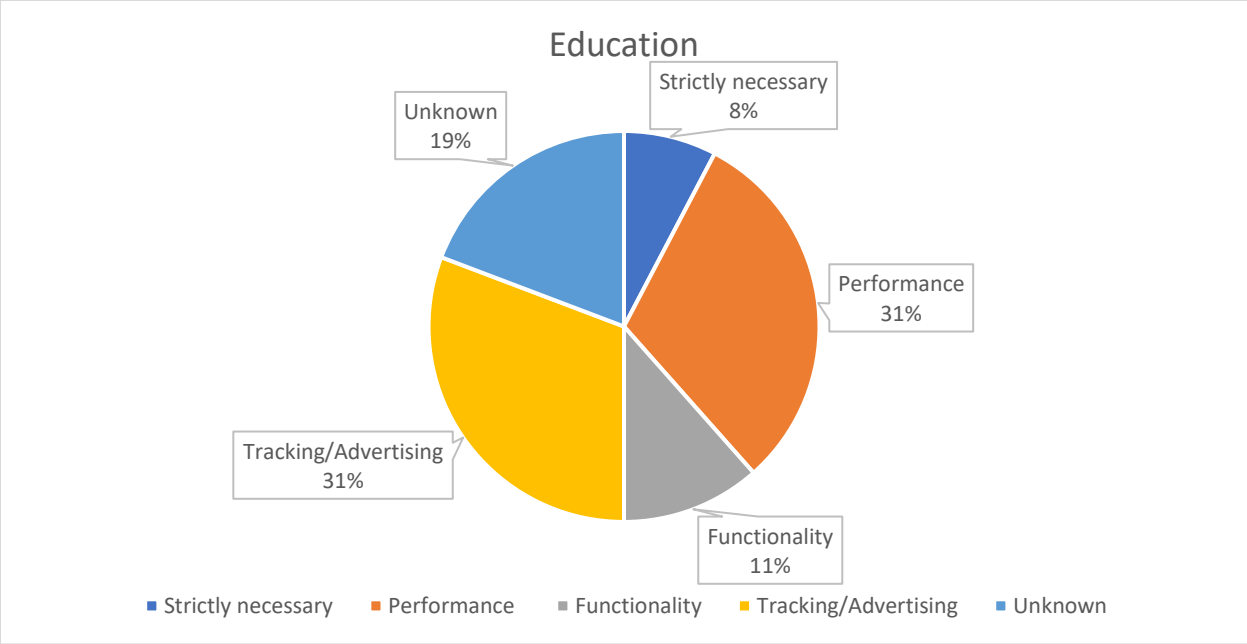
Graph 1: Summary of banking category cookies by type

LHV, Swedbank and SEB had in total 36 cookies (Graph 1). Most of the cookies were strictly necessary 36%, performance cookies 31%, then unknown cookies 22%, targeting cookies 8% and functionality cookies 3%.

Education:

Site	Number of cookies by OpenWPM	Number of cookies by Mozilla	Overall remarks
www.ut.ee	11	8+1	Almost all the cookies (including third-party) installed before the cookie consent. At homepage there was no information about the cookies. Three tracking cookies *xtc, uvc and loc* were missing. After the cookie consent *simple_cookie_compliance_dismissed* (value 1) was added.
www.ekool.eu	10	7	OpenWPM did not identify: _ga; _gat_gtag_UA; _gid were added by browser
www.eki.ee	2	2	OpenWPM and Mozilla match 100%
www.ope.ee	3	3	OpenWPM and Mozilla match 100%
Overall	26	20+1	

Table 4: Summary of education category cookies



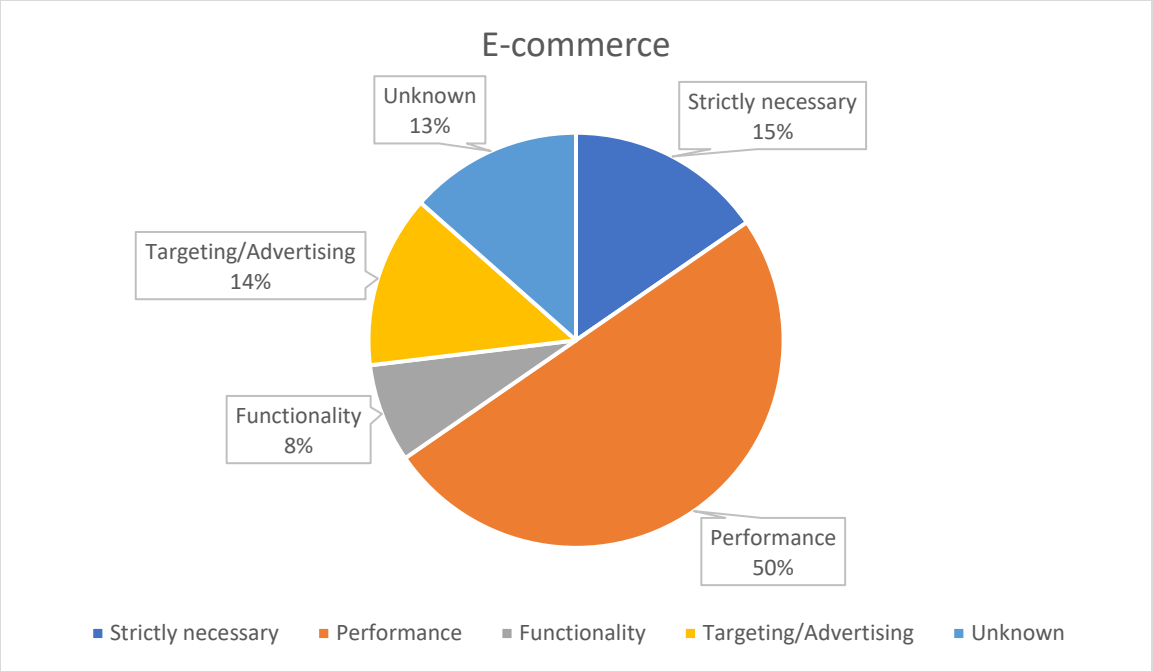
Graph 2: Summary of education category cookies by type

Summary: Tartu University, Ekool, EKI and Ope.ee had in total 26 cookies. Most of the cookies were tracking cookies by 31%. Then performance cookies 31%, unknown cookies 19%, functionality cookies 12% and strictly necessary 8% (Graph 2).

E-Commerce:

Site	Number of cookies by OpenWPM	Number of cookies by Mozilla	Overall remarks
www.telia.ee	7	6+15	15 was added after cookie consent. OpenWPM did not show *_gid*, but Mozilla did (initial cookies). At home page, there are different *alphachat* cookies, but *alphachat-test* is not written.
www.auto24.ee	10	10+1	After the cookie agreement, cookie cookies_agreed was added. OpenWPM and Mozilla match 100%. There are not written out google analytics cookies _ga,_gat, _gid explanations
www.kv.ee	13	11	OpenWPM doesn't recognize Hotjar 3rd party cookies. Same cookies as Postimees, as it belongs to Eesti Meedia
www.andmorefashion.com	6	9	OpenWPM did not recognize Hotjar 3rd party cookies.
Overall	36	36+16	

Table 5: Summary of e-commerce category cookies



Graph 3: Summary of e-commerce category cookies by type

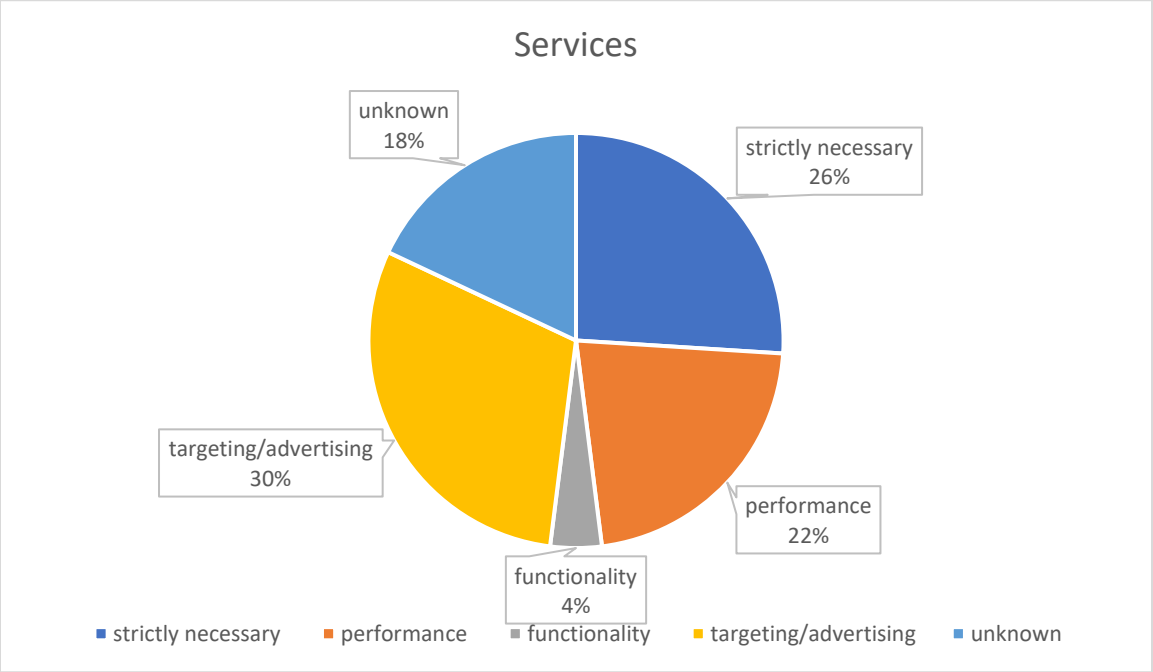
Summary: E-commerce has 52 cookies in total. 50% of them are performance cookies. After that strictly necessary cookies by 15%, unknown 13%, targeting by 13% and functionality cookies by 8% (Graph 3).

Services:

Site	Number of cookies by OpenWPM	Number of cookies by Mozilla	Overall remarks
www.neti.ee	8	6	Using the same cookies as www.telia.ee. Mozilla did not show 3rd party cookies *Gtest* and *Gdyn*
www.zone.ee	8	6	Mozilla did not have third-party *AWSALB* and *AWSALBCORS* cookie.
www.online.ee	4	4	Results match 100%.
www.tallinn.ee	17	13+1	After the cookie agreement TallinnEU cookie was added to the list. Mozilla had two language cookies keel_nimi value eng and keel_nimi value est. OpenWPM did not show Facebook cookie fbp. Mozilla also shows connections to Twitter. OpenWPM did not show any Twitter cookies.
www.tootukassa.ee	13	11+1	eu-cookie-compliance-tootukassa cookie was added after the consent.

Overall	50	40+2

Table 6: Summary of services category cookies



Graph 4: Summary of services category cookies by type

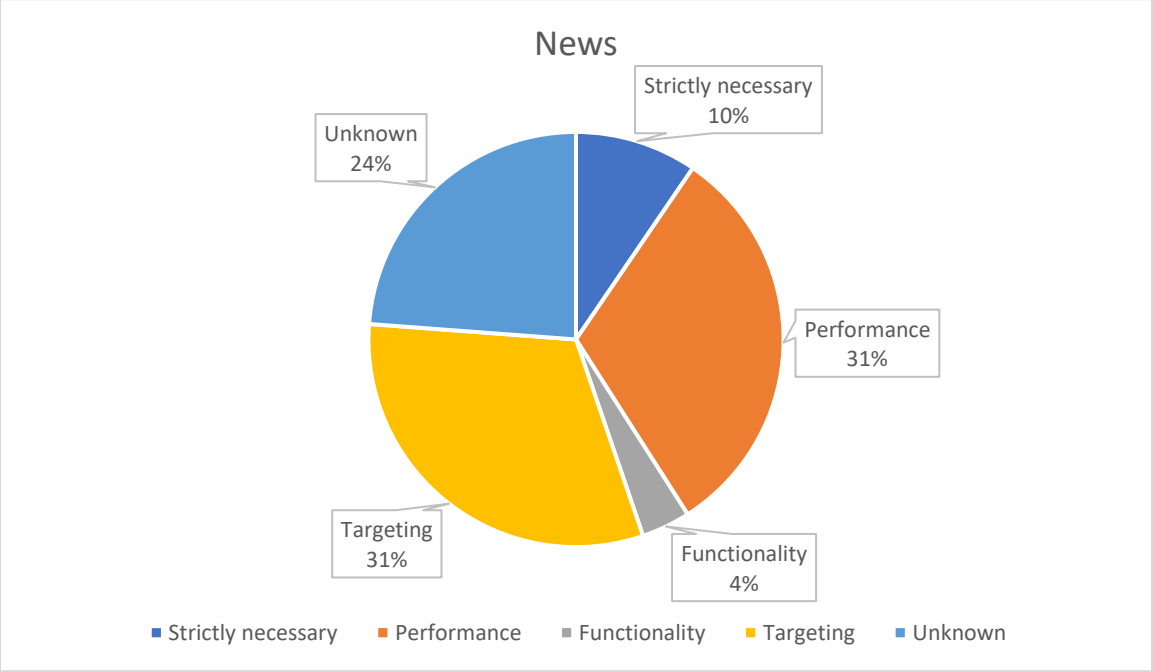
Summary: Services have 50 cookies in total. 30% of them are targeting/advertising cookies, 26% are strictly necessary cookies, performance cookies 22%, unknown 18% and 4% are functionality cookies (Graph 4).

News:

Website	Number of cookies by OpenWPM	Number of cookies by Mozilla	Overall remarks
www.delfi.ee	25	22	__gfp_64b is a tracking cookie (4)
www.postimees.ee	19	13	__gfp_64b is a strictly necessary cookie (1)
www.harjuelu.ee	15	11	Cookies ts; x-cdn, akavpau_ppsd, ts_c were missing
www.elu24.ee	19	11	Has the same cookies as Postimees, as they have the same owner (Eesti Meedia)
www.ohtuleht.ee	13	14	OpenWPM did not show Google cookies *cX_G, cx_P and cX_S.*

www.err.ee	14	14	Mozilla ERR does not show *Gdyn and Gtest* cookies (third-party).
Overall	105	85	

Table 7: Summary of news category cookies



Graph 5: Summary of news category cookies by type

Summary: News have 105 cookies in total. Targeting and performance cookies 31%, unknown 24 %, strictly necessary 10% and functionality cookies 4% (Graph 5).

4.1.2 Cookies by categories

Although the number of popular websites was different in categories, it is still possible to determine category with the lowest and highest cookies. Table 2 shows that news category websites have the biggest number of cookies, 105, followed by e-commerce with 52. The lowest number of cookies had education, 26. If visiting news websites, then it is two times bigger possibility that cookies are placed in browser compared with other categories.

Category	Total number	Strictly necessary cookies (%)	Performance cookies (%)	Functional cookies (%)	Tracking cookies (%)	Unknown cookies (%)
Banking	36	36%	31%	3%	8%	22%
Education	26	8%	31%	12%	31%	19%
News	105	10%	31%	4%	31%	24%
E-commerce	52	18%	41%	9%	16%	16%
Services	50	15%	50%	8%	13%	13%
Average		17,4%	36,8%	7,2%	19,8%	18,8%

Table 2: Comparison of cookies by categories

Comparison by categories:

Strictly necessary cookies: The banking category had 36% strictly necessary cookies, and this was the only category where all the technical cookies were added after the consent.

Performance cookies: Services have the biggest number of 50 %. The E-commerce has 36%. E-commerce category depends on the analytics, what visitors have looked, and saved in the shopping basket.

Functional cookies: Education has the biggest number of 12%, and banking has the lowest number of 3%. The average number shows, that websites do not use many functional cookies.

Tracking cookies: Two categories out of five have a result of 30% or more. Education category tracking cookies come mostly from www.ut.ee. The banking category has the lowest number of tracking cookies, and none of them were third-party tracking cookies.

Unknown cookies: News category had 24% of unknown category cookies.

Summary: Most of the cookies used in popular Estonia websites are performance cookies, with an average of 36,8%. It is almost 20% of possibility to be tracked by visiting popular Estonian websites, either by first or third-party cookies. As websites do not point out which cookies they use, they can lose the transparency and trust because of the big number of unknown cookies.

4.1.3 Third-party cookies

From the total number of cookies (396), 83 (21%) was third-party tracking cookie. In Table 8, all the unique (sorted) third-party cookies are pointed out. OpenWPM crawl resulted in 64% of the websites (14/22) use third-party cookies.

Site URL	Host	Third party cookie name
Delfi.ee	.hit.gemius.pl	*Gtest*
Delfi.ee	.hit.gemius.pl	*Gdyn*
Delfi.ee	.cxense.com	*gckp*
Delfi.ee	.google.com	*NID*
Postimees.ee	.doubleclick.net	*test_cookie*
Postimees.ee	.doubleclick.net	*IDE*
Postimees.ee	.adform.net	*C*
Postimees.ee	.adform.net	*uid*
Postimees.ee	.hit.gemius.pl	*Gtest*
Postimees.ee	.hit.gemius.pl	*Gdyn*
Harjuelu.ee	.paypal.com	*x-cdn*
Harjuelu.ee	.www.paypal.com	*akavpau_ppsd*
Harjuelu.ee	.paypal.com	*ts*
Harjuelu.ee	.paypal.com	*ts_c*
Ekool.eu	.nr-data.net	*JSESSIONID*
Ut.ee	.addthis.com	*uvc*
Ut.ee	.addthis.com	*xtc*
Ut.ee	.addthis.com	*loc*
Ut.ee	.doubleclick.net	*test_cookie*
Err.ee	.hit.gemius.pl	*Gtest*
Err.ee	.hit.gemius.pl	*Gdyn*
Neti.ee	.hit.gemius.pl	*Gtest*
Neti.ee	.hit.gemius.pl	*Gdyn*
Ohtuleht.ee	.google.com	*NID*
Ohtuleht.ee	.hit.gemius.pl	*Gtest*
Ohtuleht.ee	.hit.gemius.pl	*Gdyn*
Zone.ee	widget-mediator.zopim.com	*AWSALB*
Zone.ee	widget-mediator.zopim.com	*AWSALBCORS*
Tootukassa.ee	6168205.global.siteimproveanalytics.io	*AWSELB*
Tootukassa.ee	6168205.global.siteimproveanalytics.io	*AWSELBCORS*

Kv.ee	.hit.gemius.pl	*Gtest*
Kv.ee	.hit.gemius.pl	*Gdyn*
Elu24.ee	.doubleclick.net	*test_cookie*
Elu24.ee	.doubleclick.net	*IDE*
Elu24.ee	.adform.net	*C*
Elu24.ee	.adform.net	*uid*
Elu24.ee	.hit.gemius.pl	*Gtest*
Elu24.ee	.hit.gemius.pl	*Gdyn*
Tallinn.ee	.google.com	*NID*
Tallinn.ee	balticlivecam.com	*wppas_pvbl*
Tallinn.ee	.balticlivecam.com	*__asc*
Tallinn.ee	.balticlivecam.com	*__auc*
Tallinn.ee	.balticlivecam.com	*_ga*
Tallinn.ee	.balticlivecam.com	*_gid*
Tallinn.ee	.balticlivecam.com	*_gat*
Andmorefashion.com	.doubleclick.net	*test_cookie*

Table 8: List of third-party cookies

With whom is our data shared? Out of the 14 websites with third-party cookies, 50% (7) have hit.gemius.pl cookie, and 6 pages share data with Google. These results are similar to Trevisan et al., information is shared with US big player Google and EU service providers (mostly Gemius). Gemius is an international research and technology company providing web analytics and ad serving [63].

Results also match with 20 TOP third-party companies noticed by Falahrastegar et al., [32], as Google and AddThis are in the list. In this list, there are also Facebook and Twitter. Facebook changed its *_fbp pixel* cookie from third-party to first party [64]. Twitter had several HTTP requests, but no cookies installed. Table 9 describes the companies host address belongs to.

Company	**Domain**
Gemius	hit.gemius.pl
Google	google.com; doubleclick.net
Cxense	cxense.com
Adform	adform.net
Paypal	paypal.com
New Relic	Newrelic.com
AddThis	Addthis.com
Zopim	Zendesk.com
BalticLiveCam	balticlivecam.com
Gandi SAS	Gandi.net

Table 9: Third-party tracking companies and their domains

4.1.4 Third party HTTP requests

A third-party HTTP request is one that is sent toward an URL that does not share the same hostname as the website the user intentionally visits [35].

Website URL	HTTP requests	HTTP requests by third-party	Total %
http://www.delfi.ee	246	64	26%
https://www.postimees.ee	124	72	58%
https://www.swedbank.ee	36	0	0%
https://www.harjuelu.ee	220	22	10%
https://www.ekool.eu	50	23	46%
https://www.ut.ee	119	15	13%
https://www.seb.ee	71	8	11%
https://www.microsoftonline.ee	1	0	0%
https://www.err.ee	284	75	26%
https://www.auto24.ee	107	11	10%
https://www.neti.ee	44	19	43%
https://www.ohtuleht.ee	200	99	50%
https://www.telia.ee	42	15	36%
https://www.eki.ee	24	0	0%
https://www.zone.ee	32	13	40%
https://www.ope.ee	9	2	22%
https://www.tootukassa.ee	77	8	10%

https://www.lhv.ee	42	1	2%
https://www.online.ee	27	5	19%
https://www.kv.ee	125	38	30%
https://www.elu24.ee	135	68	50%
https://www.tallinn.ee	141	96	68%
https://www.andmorefashion.com	136	23	17%
TOTAL	**2292**	**677**	**30%**

Table 10: Summary of third-party HTTP requests

As shown from the table (Table 10) www.ohtuleht.ee has the biggest number of third-party requests, 99. www.tallinn.ee has 96 and after that news category websites www.delfi.ee , www.postimees.ee and www.err.ee are followed. Biggest percentage of third-party requests compared to total requests is www.tallinn.ee (68%). 4 websites out of 22 have more or equal than 50% of total requests. Out of the 3 are news media websites (www.postimees.ee, www.ohtuleht.ee, www.elu24.ee).

As shown in Table 11, most third-party HTTP requests are made to Google and its subsidiaries (82%), Facebook followed by 45%. 32% of the webpages do requests to Gemius, and 23% have HTTP requests to Twitter and Adform.

Number of Websites	Percentage	Third-party	Category	Domain
19/22	82%	Google	Analytics/Marketing	Google.com
10/22	45%	Facebook	Marketing	Facebook.com
7/22	32%	Gemius	Analytics	Gemius.com
5/22	23%	Twitter	Marketing	Twitter.com
5/22	23%	Adform	Marketing	Adform.net

Table 11: Top5 of domains, where third-party requests are made

By categories, most third-party domains are contacted by news websites, 400. News followed by web services, 129. E-commerce category has 95 contacts to third-party and followed by education

with 40 and banks, 9. When visiting news media websites, then there is big possibility that data is shared with third parties.

Web beacon:

Part of the HTTP request is the type of resource, and one of them is the web beacon. Web beacons are used to gather visiting information. From popular websites 8 out of 22 (36%) were using web beacons (Appendix 4).

For example, HTTP request was from www.tallinn.ee to www.facebook.com. Analysis of the header shows that visitor information is gathered. Among technical settings is an attribute data-show-facepile (Table 12 and 13), that tells to show profile photos when friends like the same thing.

Setting	HTML5 Attribute	Description	Default
show_facepile	data-show-facepile	Show profile photos when friends like this	true

Table 12: Example of show facepile attribute [65]

From the privacy perspective Facebook has gathered information about yourself and friends, and now is telling the visitor, which friends have already liked the same thing. Information about website, user agent, language are delivered to third-party [65].

referrer	headers
https://www.facebook.com/v9.0/plugins/page.php?adapt_container_width=true&app_id=1531182050350940&channel=https%3A%2F%2Fstaticxx.facebook.com%2Fx%2Fconnect%2Fxd_arbiter%2F%3Fversion%3D46%23cb%3Df315dbf472221a6%26domain%**3Dwww.tallinn.ee**%26origin%3Dhttps%253A%252F%252Fwww.tallinn.ee%252Ff2d4ba42d7db5f8%26relation%3Dparent.parent&container_width=269&height=300&hide_cover=false&href=https%3A%2F%2Fwww.facebook.com%2Ftallinnalinn%2F&locale=et_EE&sdk=joey&show_facepile=true&show_posts=true&small_header=true	[["Host","www.facebook.com"],["**UserAgent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0)Gecko/20100101Firefox/80.0"],["Accept","*/*"],["**Accept-Language","en US,en;q=0.5"],["Accept-Encoding","gzip,deflate,br"],["Referer","https://www.facebook.com/v9.0/plugins/page.php?adapt_container_width=true&app_id=1531182050350940&channel=https%3A%2F%2Fstaticxx.facebook.com%2Fx%2Fconnect%2Fxd_arbiter%2F%3Fversion%3D46%23cb%3Df315dbf472221a6%26domain%**3Dwww.tallinn.ee%**26origin%3Dhttps%253A%252F%252Fwww.tallinn.ee%252Ff2d4ba42d7db5f8%26relation%3Dparent.parent&container_width=269&height=300&hide_cover=false&href=https%3A%2F%2Fwww.facebook.com%2Ftallinnalinn

	%2F&**locale=et_EE**&sdk=joey&**show_facepile=true**&show_posts=true&small_header=true"],["ContentType","multipart/form-data; boundary=----------805425890245514646929642016 92"],["Content-Length","4910"],["Origin","https://www.facebook.com"],["Connection","keep-alive"]]

Table 13: Example of HTTP request from www.tallinn.ee website

B2C customer segmentation

As shown in Table 14, Telia is using web beacon to target clients with segment B2C. Information about user-agent, language and website are delivered.

referrer	headers
https://plumbrapi.telia.ee/api/browser/data/peekon?accountId=e94012bjk6eut9q98706q281jl&batchId=29269280-b0c0-c1ec-c807-69e6d33e0278	**[["Host","plumbrapi.telia.ee"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64;rv:80.0)Gecko/20100101 Firefox/80.0"],[**"Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip,deflate, br"],["Referer","https://sso.telia.ee/lib/webapi?webapi=true&locale=et"],["Content-Type","text/plain;charset=UTF-8"],["Content-Length","427"],["Origin","https://sso.telia.ee"],["Connection","keep-alive"],["Cookie","LANGUAGE_EXT=et_EE**;segment=b2c;** shoppingCartIcon=0"]]

Table 14: Example of HTTP request from www.telia.ee webpage

4.1.5 Third to third-party HTTP redirects

There were three websites, where third to third-party tracking technique was used (table 15):

www.delfi.ee: Partners were csynr.cxense.com, and dmp.adform.net. From the header it is seen that technique called cookie matching was used. Header /serving/cookie/match says that visitor with certain ID has visited website. Main purpose of that is targeted ad serving [66].

www.postimees.ee and www.elu24.ee, both websites are using third party cookie forwards, as the cookie has been set up. Partners were track.adform.net and www.facebook.com. In these cases, the aim is to share browsing history.

N	Old request URL	New request URL	Header
1	https://csynr.cxense.com/?cxsite=1145189970857384309&partnerId=csr&redir=https%3A%2F%2Fdmp.adform.net%2Fserving%2Fcookie%2Fmatch%2F%3Fparty%3D1077%26cid%3D%24%7BUUID%7D%26cxckp%3Dcx%3Ap8knv1epeyom3gqefs67gds5u%3A2a7084v65tra5	https://dmp.adform.net/serving/cookie/match/?party=1077&cid=3vztuz7u9u8bn239t6oyw78yzs	[["Date","Sun, 21 Feb 2021 10:32:15GMT"],["Location","https://dmp.adform.net/serving**/cookie/match/?**party=1077&cid=3vztuz7u9u8bn239t6oyw78yzs"],["StrictTransportSecurity","maxage=31600"],["ContentLength","109"],["Server","Jetty(9.4.28.v20200408)"]]
2	https://track.adform.net/adfserve/?CC=1&bn=40578808;1x1inv=1;srctype=3;ord=464213972	https://www.facebook.com/tr?id=1190842377642537&ev=LOCALMEDIA&noscript=1&rnd=52651	[["server","nginx"],["date","Sun, 21 Feb 2021 10:32:40 GMT"],["content-type","text/html;charset=utf8"],["cache-control","no-cache, no-store, must-revalidate,notransform"],["pragma","nocache"],["expires","1"],["location","//www.facebook.com/tr?id=1190842377642537&ev=LOCALMEDIA&noscript=1&rnd=52651"],["access-control-allow-origin**","*"],["setcookie","uid=139394781009857 1289**; domain=adform.net; expires=Thu, 22-Apr-2021 09:32:40 GMT; path=/"],["p3p","CP=\"NOI DSP COR NID CURa ADMa DEVa TAIa PSAa PSDa OUR LEG NAV INT\""],["strict-transport-security","max-age=31536000;includeSubDomains"],["X-Firefox-Spdy","h2"]]
3	https://track.adform.net/adfserve/?CC=1&bn=40578808;1x1inv=1;srctype=3;ord=1163921781	https://www.facebook.com/tr?id=1190842377642537&ev=LOCALMEDIA&noscript=1&rnd=36154	[["server","nginx"],["date","Sun, 21 Feb 2021 10:40:12GMT"],["contenttype","text/html; charset=utf-8"],["cache-control","no-cache, no-store, must-revalidate, notransform"],["pragma","nocache"],["expires","1"],["location","//www.facebook.com/tr?id=1190842377642537&ev=LOCALMEDIA&noscript=1&rnd=36154"],["access-control-allow-origin**","*"],["set-cookie","uid=670837272229838892;** domain=adform.net; expires=Thu, 22-Apr-2021 09:40:12 GMT;path=/"],["p3p","CP=\"NOI DSP COR NID CURa ADMa DEVa TAIa PSAa PSDa OUR LEG NAV INT\""],["strict-transport-security","maxage=31536000; includeSubDomains"],["X-Firefox-Spdy","h2"]]

Table 15: OpenWPM example of third-to-third party trackers

4.2 Compliance to EU Privacy Directive

Observation of websites showed (Table 16) that 63% are compliant with EU Privacy Directive. 23% of the websites had something missing, and 14% do not have any of the conditions filled. 77% had banner installed, and 86% had privacy information noticed. 68% had a cookie policy and 72% asked for user consent. 68% of the pages had an opt-out method noticed.

Some of the pages had misleading information: www.delfi.ee had a cookie policy page and link to aki.ee/et/kupsised (how to delete cookies), but that leads to data protection inspectorate, and that only says, which cookies www.aki.ee site uses. www.ekool.eu had https://ekool.eu/terms/privacy_et.html privacy page, where they explain privacy terms, but it is written that it was valid until 26.10.2014.

What makes the website transparent is information about all the cookies used. Several websites had pointed out the category, but not the specific cookies that were used. Some of the websites pointed out third-party companies that they cooperate. Education had 3 out of 4,websites by the categories, where almost none of the conditions was fulfilled. Banking category had all the conditions filled and biggest news websites also (except www.harjuelu.ee).

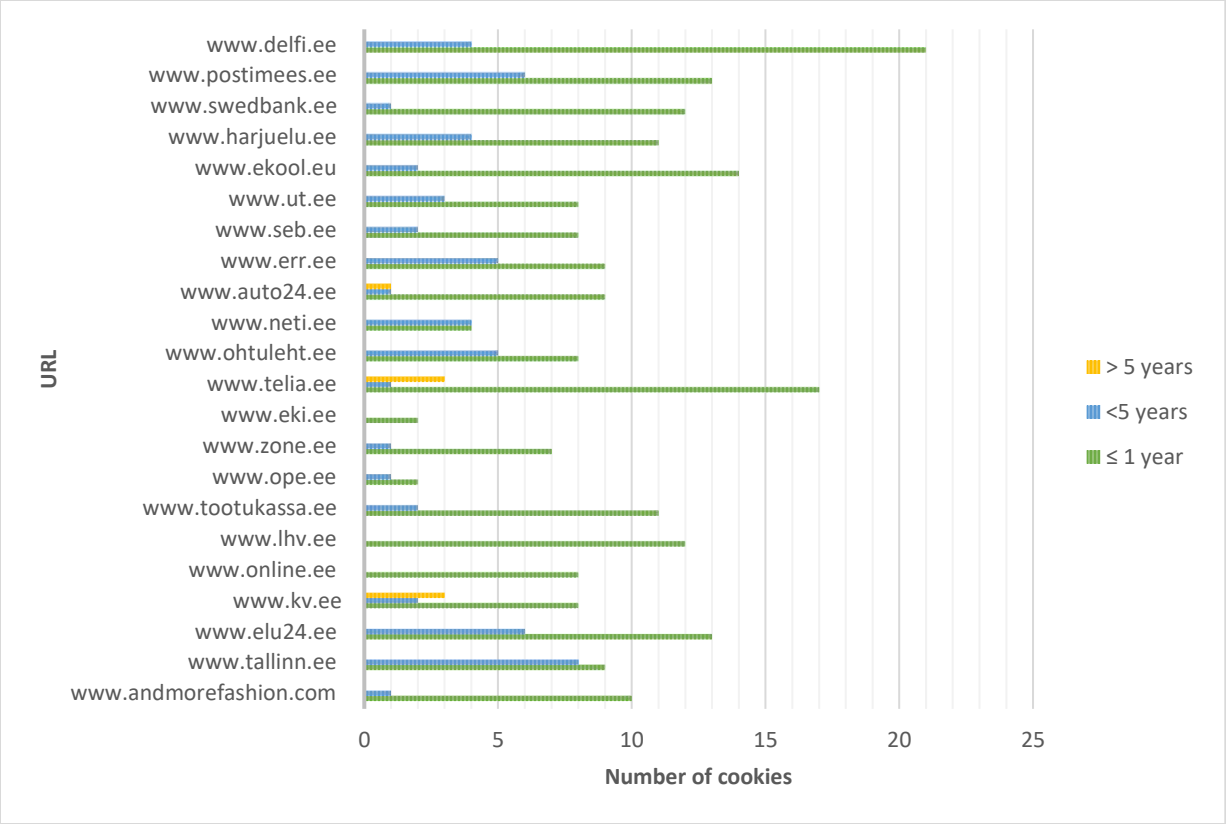
Site URL	Banner	Privacy info/policy	Cookie policy	User Consent	Opt-out method
http://www.delfi.ee	✓	✓	✓	✓	✓
https://www.postimees.ee	✓	✓	✓	✓	✓
https://www.swedbank.ee	✓	✓	✓	✓	✓
https://www.harjuelu.ee		✓			
https://www.ekool.eu					
https://www.ut.ee	✓	✓	✓	✓	✓
https://www.seb.ee	✓	✓	✓	✓	✓
https://www.err.ee	✓	✓	✓	✓	
https://www.auto24.ee	✓	✓	✓	✓	✓
https://www.neti.ee	✓	✓	✓	✓	✓
https://www.ohtuleht.ee	✓	✓	✓	✓	✓
https://www.telia.ee	✓	✓	✓	✓	✓
https://www.eki.ee					
https://www.zone.ee		✓			✓
https://www.ope.ee					

https://www.tootukassa.ee	✓	✓		✓	
https://www.lhv.ee	✓	✓	✓	✓	✓
https://www.online.ee	✓	✓	✓	✓	✓
https://www.kv.ee	✓	✓	✓	✓	✓
https://www.elu24.ee	✓	✓	✓	✓	✓
https://www.tallinn.ee	✓	✓	✓	✓	✓
https://www.andmorefashion.com	✓	✓			

Table 16: Summary of popular websites by ePrivacy Directive

4.2.1 Cookie lifetime

According to the ePrivacy Directive, cookies should not last longer than 12 months, but in practice, they could remain on in device much longer if no actions are taken [67]. Cookies by duration were divided into 3 categories: Less or equal to 1, under 5 years and over 5 years (Graph 6). Out of 396 cookies, 150 cookies had an expiry time of more than a year (37,8%), and out of 396, 57 (14%) were session cookies. 48,2% of cookies were with expiry time less or equal to 1 year. Session cookies number was similar to Cahn et al., [68] they had 17%.



Graph 6: Summary of popular websites by cookie duration time

ePrivacy Directive gives a suggestion that cookie expiry time should be no longer than 1 year. Research showed that 19 popular website out of 22 (86%) had set cookies, with an expiry time more than a year. Three popular website out of 22 (14%) had an expiry time of more than 5 years (Table 17).

URL	Cookie name	Value	Expiry
.auto24.ee	*CID*	1613903758327840	2031-02-19T10:36:01.000Z
www.telia.ee	*CookieConsent*	-3	2060-01-01T00:00:00.000Z
.telia.ee	*plumbr-supported*	telia.ee	2031-02-19T10:39:25.000Z
www.kv.ee	*cookies_notified*	1	2031-02-19T10:39:50.000Z

Table 17: OpenWPM summary of cookies with duration more than 5 years

A total number of third-party cookies was 83. Out of 83, 8 (10%) cookies were session cookies. Analysis of third-party cookies exposed that their average "expiry" date is 24-10-2022 (180 days). There were some differences between OpenWPM and Mozilla cookies. For example, cookie *segment* (www.telia.ee) shows that it is session, OpenWPM shows expiry after 1 year. Cookie *consent* (www.telia.ee) shows expiry until 2030 and OpenWPM shows until 2060.01.01. Mozilla

doesn't show Gdyn and Gtest cookies (third-party) on www.err.ee. Their expiry time is 5 years, www.err.ee policy claims that all the cookies are for maximum expiry of two and a half years.

4.3 Effectiveness of anti-tracking privacy measures

4.3.1 DO Not Track header

First was tested DNT in OpenWPM, what is one of the privacy enhancement techniques. When it was switched on, it did not show any remarkable results. The cookie number was reduced from 396 to 350 (8%).

4.3.2 Private browsing feature and Ghostery browser extension

As seen from table 18, Ghostery showed results of blocking 98% of third-party trackers. Similar results were received with Private Browsing mode 91%. Both methods did not block balticlivecam.com cookie *wppas_pvbl*. Reason for that can be similar to Bujlow et al.,[4] that if it is a cookie of an unknown company, then it is bigger chance, that it is not blocked.

Visit ID	Host	Cookie name	Mozilla Ghostery	Cookies left	Mozilla Private browsing	Cookies left
Delfi.ee	.hit.gemius.pl	Gtest	x	_edid _edt cp_user_package_t dcid	x	_edid _edt cp_user_pac kage_t dcid
Delfi.ee	.hit.gemius.pl	Gdyn	x		x	
Delfi.ee	.cxense.com	gckp	x		x	
Delfi.ee	.google.com	NID	x		x	
Postimees.ee	.doubleclick.net	test_cookie	x	Pm_c_list	x	Pm_c_list
Postimees.ee	.doubleclick.net	IDE	x		x	
Postimees.ee	.adform.net	C	x		x	
Postimees.ee	.adform.net	uid	x		x	
Postimees.ee	.hit.gemius.pl	Gtest	x		x	
Postimees.ee	.hit.gemius.pl	Gdyn	x		x	

Harjuelu.ee	.paypal.com	x-cdn	x	PHPSESSID	x	Adrotate-293 Adrotate-294 PHPSESSID tk_lr tk_or tk-r3d
Harjuelu.ee	.www.paypal.com	akavpau_ppsd	x		x	
Harjuelu.ee	.paypal.com	ts	x		x	
Harjuelu.ee	.paypal.com	ts_c	x		x	
Ekool.eu	.nr-data.net	JSESSIONID	x	_cfduid; e_kool_login_session; tmp_show_mobile_version XSRF-token	x	_cfduid; e_kool_login_session; tmp_show_mobile_version; XSRF-token
Ut.ee	.addthis.com	uvc	x	has_js; simple_cookie_compliance	x	_gcl_au; has_js
Ut.ee	.addthis.com	xtc	x		x	
Ut.ee	.addthis.com	loc	x		x	
Ut.ee	.doubleclick.net	test_cookie	X		x	
Err.ee	.hit.gemius.pl	Gtest	x	attld; statUniqueId	x	attld; statUniqueId
Err.ee	.hit.gemius.pl	Gdyn	x		x	
Neti.ee	.hit.gemius.pl	Gtest	x		x	
Neti.ee	.hit.gemius.pl	Gdyn	x		x	
Ohtuleht.ee	.google.com	NID	x		x	
Ohtuleht.ee	.hit.gemius.pl	Gtest	x		x	
Ohtuleht.ee	.hit.gemius.pl	Gdyn	x		x	
Zone.ee	widget-mediator.zopim.com	AWSALB	x	pll_language	x	_zlcmid; pll_language
Zone.ee	widget-mediator.zopim.com	AWSALBCORS	x		x	

Tootukassa.ee	6168205.global.siteimproveanalytics.io	AWSELB	x	Fontsize, has_js; layout; lineheight; responsive; SESS…; SSESS…;	x	Fontsize, has_js; layout; lineheight; responsive; SESS…; SSESS…; nmstat
Tootukassa.ee	6168205.global.siteimproveanalytics.io	AWSELBCORS	x		x	
Kv.ee	.hit.gemius.pl	Gtest	x	pm_c_list; cookies_notified; kv_web; saved_searches	x	cookies_notified; kv_web; saved_searches
Kv.ee	.hit.gemius.pl	Gdyn	x		x	
Elu24.ee	.doubleclick.net	test_cookie	x		x	pm_c_list;
Elu24.ee	.doubleclick.net	IDE	x		x	
Elu24.ee	.adform.net	C	x		x	
Elu24.ee	.adform.net	uid	x		x	
Elu24.ee	.hit.gemius.pl	Gtest	x		x	
Elu24.ee	.hit.gemius.pl	Gdyn	x		x	
Tallinn.ee	.google.com	NID	x	Sess_admin and all the Tallinn cookies.		Sess_admin and all the Tallinn cookies.
Tallinn.ee	balticlivecam.com	wppas_pvbl				
Tallinn.ee	.balticlivecam.com	__asc	x			
Tallinn.ee	.balticlivecam.com	__auc	x			
Tallinn.ee	.balticlivecam.com	_ga	x		x	
Tallinn.ee	.balticlivecam.com	_gid	x		x	
Tallinn.ee	.balticlivecam.com	_gat	x		x	
Andmorefashion.com	.doubleclick.net	test_cookie	x		x	

Table 18: Summary of Ghostery and private browsing cookies

4.4 Summary of the results

Summary is based on findings from this thesis and based on popular Estonian websites (Table 19). Aim of this summary is to summarize, which websites can affect website visitor privacy. If visitors visit news category websites www.delfi.ee , www.postimees.ee and www.elu24.ee then they are noticed about tracking, but their privacy is affected by using third-party cookies, web beacons and third to third-party tracking. Their browsing history and purchases are exchanged with third-party counterparts and it can be done over several websites. www.harjuelu.ee, www.ekool.eu, www.zone.ee, www.eki.ee and www.ope.ee do not notice about tracking and they use third-party cookies and have requests to 3rd party. www.seb.ee, www.swedbank.ee, www.lhv.ee and www.auto24.ee are websites that comply with ePrivacy Directive and have no third-party trackers.

Anti-tracking privacy measures Ghostery and private browsing mode block over 90% of third-party trackers and it is possible to say, that they are effective. Do Not Track did not show remarkable results.

URL	3rd party cookies	3rd party requests	Web beacon	ePrivacy directive	3rd to 3rd party tracking
www.delfi.ee	4	26 %	✓	✓	✓
www.postimees.ee	6	58%	✓	✓	✓
www.swedbank.ee	-	0%		✓	
www.harjuelu.ee	4	10%			
www.ekool.eu	1	46%	✓		
www.ut.ee	4	13%	✓	✓	
www.seb.ee	-	11%		✓	
www.err.ee	2	26%		✓	
www.auto24.ee	-	10%		✓	
www.neti.ee	2	43%		✓	
www.ohtuleht.ee	3	50%		✓	
www.telia.ee	-	36%	✓	✓	
www.eki.ee	-	0%			
www.zone.ee	2	40%			
www.ope.ee	-	22%			
www.tootukassa.ee	2	10%		✓	
www.lhv.ee	-	2%		✓	
www.online.ee	-	19%	✓	✓	
www.kv.ee	2	30%		✓	
www.elu24.ee	6	50%	✓	✓	✓
www.tallinn.ee	7	68%	✓	✓	
www.andmorefashion.com	1	17%		✓	

Table 19: Summary of tracking in popular Estonian websites

5. Discussion

This Chapter focuses on explaining and evaluating how results of the study are related to literature background.

In this research, we observed tracking and user privacy in popular Estonian websites. Aim of this master thesis was to see which cookies are installed, if we visit popular Estonian websites and how it affects our privacy. Analysis of cookies, compliance to EU privacy Directive, and defense methods were described on the background and methodology part.

As shown from the results, an average of 18,8% of the cookies was unknown. That could mean websites lack of transparency, and all these could be tracking cookies. That would raise the tracking average to 30%, and visitor privacy could be affected more if there is no notice. The only category where the websites comply and have no third-party trackers is banking.

As said by Timothy Libert [31]: "From a privacy perspective, news websites are among the worst on the web". This work showed that news websites have the most significant number of cookies, the biggest number of tracking cookies and third-party tracking cookies, the biggest number of third-party HTTP requests, and there is third-to-third party tracking. From the same author, it is possible to point out that financial considerations' overweight visitor privacy. And there is a lack of information on how they benefit from tracking.

The comparison of OpenWPM and Mozilla cookie numbers shows that in news, services, and education category OpenWPM delivered 20% more cookies. Banking and E-commerce category match with cookie numbers but with different cookies. There were some certain cookies that OpenWPM did not recognize, for example, analytics company Hotjar cookies. Twitter had several HTTP requests, but no cookies installed.

37% of the popular websites do not comply with the ePrivacy Directive. Several EU member states that impose fines as a punishment method decrease tracking by 36%. The Directive aims to reduce tracking and protect websites visitor privacy, not to collect fines [7]. Still, if that method works, then Estonian authorities should consider that.

From the TOP20 (appendix 2), tracking companies and their third-party domains, popular Estonian websites had connections to Facebook, Twitter, Google, AddThis, Amazon, and Yahoo. There are

companies that gather most of the tracking information and give advantage if third-party tracking support should disappear from 2022. In this list, there is also Facebook. Facebook changed its *_fbp* pixel cookie from third-party to first party because of the tracking future changes [33]. Results of tracking defense confirmed that the bigger the company the stronger possibility to be blocked.

6. Conclusion

This chapter answers to research tasks, addresses the limitations, and gives recommendations for the future research.

6.1 Answers to research tasks

1. **Evaluate the usage of first and third-party cookies.**

The tracking analysis revealed that the biggest number of first and third-party cookies have education and news (31%). All categories together had an average of 19,8% of first and third-party tracking cookies. The news category has the biggest number of cookies, 105. Visiting news sites, results two times bigger possibility that cookies are stored.

64% of the sites use third-party cookies, and 100% of the websites with third-party cookies installed third-party cookies before user consent. From the total number of cookies, 21% were third-party tracking cookies.

2. **Find, if websites comply with EU Privacy Directive.**

63% of popular websites are compliant with EU Privacy Directive. 23% of the websites had something missing, and 14% of the websites do not have any of the conditions filled. Cookie expiry results showed, that 48,2% of the cookies were with expiry time less or equal to one; out of them, 14% were session cookies and 37,8% have an expiry time of more than a year.

3. **Find, which privacy enhancement and tracking defense techniques guarantee safer browsing.**

Mozilla add-on Ghostery showed results of blocking 98% of third-party cookies. Similar results were received with Private Browsing mode, 91%. From here, we can conclude that Ghostery and Private browsing mode can guarantee safer and more private browsing.

6.2 Limitations

The main methodological and measurement limitation is that OpenWPM would not interact with sites in ways a real user might, and logging into websites does not do actions such as scrolling or clicking links, and OpenWPM supports only the Firefox browser. Comparison with other browsers would be needed [8]. OpenWPM has been used for extensive data studies, where websites crawl can be 1 million websites. Web crawl of this work showed, that out of 396, several cookies had duplicated values, which had the same host, name, value, and the time stamp. This could result incorrect results.

6.3 Recommendations for the future research

There are many other tracking mechanisms, tracking defense, auditing, and enhancement tools, to discover web tracking [4]. In this work privacy measurement tool OpenWPM was used, that gives a lot of other tracking possibilities. Another option would be detecting browser fingerprinting or flash cookies. Comparison with other tracking auditing tools would give more detailed overview, which cookies are used.

As from 2022 third-party cookie tracking is about to change, it would be possible to compare cookies results crawled in this work and after 2022 and conclude if tracking average is under 19,8%. Another research option from the security aspect is cookie attributes. Cookie attributes would tell, how developers have protected cookies and how privacy could be affected.

References

[1] Eesti Rahvusringhääling, “Uuring: Eesti elanikud usaldavad kergekäeliselt oma privaatseid andmeid erinevatele institutsioonidele,” [Online]. Available: https://novaator.err.ee/257853/uuring-eesti-elanikud-usaldavad-kergekaeliselt-oma-privaatseid-andmeid-erinevatele-institutsioonidele [accessed 16.11.2020].

[2] Alexa, An Amazon.com company, “Top Sites in Estonia,” [Online]. Available: https://www.alexa.com/topsites/countries/EE [accessed 06.11.2020].

[3] J.Jokinen, “Personal Internet Privacy and Surveillance Implementation and evasion of user tracking”, [Master’s thesis], Technology, communication and transport Master's Degree Programme in Information Technology. [Online]. Available: https://www.theseus.fi/bitstream/handle/10024/146658/Jokinen_Juha.pdf;jsessionid=76E31134C810F1418A2FF1AAF5C6E6E1?sequence=1 [accessed 06.11.2020].

[4] T. Bujlow, V. Carela-Español, J. Solé-Pareta and P. Barlet-Ros, "A Survey on Web Tracking: Mechanisms, Implications, and Defenses," in Proceedings of the IEEE, vol. 105, no. 8, pp. 1476-1510, Aug. 2017, https://doi:10.1109/JPROC.2016.2637878. [accessed 06.11.2020].

[5] K.Aswini, S.S.Vidhya, “A survey on importance of user awareness of web tracking”. International Journal of advance research in science and engineering. Vol. No 5, Issue No.07. (2016).

[6] D.M Kristol (2001). HTTP Cookies: Standards, Privacy, and Politics. ACM Transactions on Internet Technology, 1(2), 151–198. 10.1145/502152.502153.

[7] Rodríguez, E, R, T. “Tracking Cookies in the European Union, an Empirical Analysis of the Current Situation”. MSc. Management of Technology. Delft University of Technology Faculty of Technology, Policy and Management Section Organization & Governance. [Online] 2018. Available: http://resolver.tudelft.nl/uuid:50d04f83-222d-479e-8ddd-661d2243857a. [accessed 11.11.2020].

[8] N.Schmücker, (2011). Web Tracking SNET 2 Seminar Paper-Summer Term 2011.

[9] H. Metwalley, S. Traverso and M. Mellia, "Unsupervised Detection of Web Trackers," 2015 IEEE Global Communications Conference (GLOBECOM), San Diego, CA, USA, 2015, pp. 1-6.

[10] Diginomica. "If the future of privacy is cookieless, what happens to the ad tech industry when third-party cookies go away,?" [Online]. Available: https://diginomica.com/if-future-privacy-cookieless-what-happens-ad-tech-industry-when-third-party-cookies-go-away [accessed 13.03.2021].

[11] Ermakova, T & Fabian.B & Bender.B & Klimek.K. (2018). Web Tracking – A Literature Review on the State of Research. 10.24251/HICSS.2018.596.

[12] J.Kyrölä, "Tapping the web to measure its privacy". Seminar paper. Proceedings of the Seminar in Computer Science: Internet, Data and Things (CS-E4000), Spring 2018.

[13] Winston & Strawn LLP. "What is the Definition of Online Privacy,?" [Online]. Available: https://www.winston.com/en/legal-glossary/online-privacy.html [accessed 03.03.2021].

[14] The New York Times. "Opinion This article is spying on you," [Online]. Available: https://www.nytimes.com/2019/09/18/opinion/data-privacy-tracking.html [accessed 11.11.2020].

[15] P.C.Adams. "Agreeing to Surveillance: Digital News Privacy Policies". Journalism & Mass Communication Quarterly 2020, Vol. 97(4) 868–889, [Online]. Available: https://journals.sagepub.com/doi/10.1177/1077699020934197 [accessed 02.05.2021].

[16] CPO Magazine. "Invasion of privacy: Tracking your online behavior across the web," [Online]. Available: https://www.cpomagazine.com/data-privacy/invasion-of-privacy-tracking-online-behavior-across-web [accessed 03.05.2021].

[17] K.Martin,(2015). "Privacy Notices as Tabula Rasa: An empirical investigation into how complying with a privacy notice is related to meeting privacy expectations online,". 1. Journal of Public Policy & Marketing. 34. 10.1509/jppm.14.139. [Online]. Available: https://journals.sagepub.com/doi/10.1509/jppm.14.139 [accessed 02.05.2021].

[18] B. Krishnamurthy and C. E. Wills, "On the leakage of personally identifiable information via online social networks," in *Proc. 2nd ACM Workshop Online Soc. Netw. (WOSN)*, 2009. pp. 7–12

[19] M.Trevisan, S.Traverso, E.Bassi, & M.Mellia. “4 Years of EU Cookie Law: Results and Lessons Learned. Proceedings on Privacy Enhancing Technologies,” 2019(2) 126-145. [Online]. Available: https://doi.org/10.2478/popets-2019-0023 [accessed 04.03.2021].

[20] S. Yu, D. V. Vargas and K. Sakurai, "Effectively Protect Your Privacy: Enabling Flexible Privacy Control on Web Tracking," 2017 Fifth International Symposium on Computing and Networking (CANDAR), Aomori, Japan, 2017, pp. 533-536, [Online]. Available: https://doi:10.1109/CANDAR.2017.26. [accessed 17.03.2021].

[21] ENISA. “Bittersweet cookies. Some security and privacy considerations,”. [Online]. Available: https://www.enisa.europa.eu/publications/copy_of_cookies [accessed 06.11.2020].

[22] Cookiepedia. “Types of cookies,” [Online]. Available: https://cookiepedia.co.uk/types-of-cookies (accessed 08.11.2020).

[23] Cookiepedia. “How we classify cookies,” [Online]. Available: https://cookiepedia.co.uk/classify-cookies (accessed 08.11.2020).

[24] Azim Mohammed. “Cookies, privacy, and cybersecurity,” [Online]. Available: https://medium.com/@azimmohamed2014/cookies-privacy-and-cybersecurity-41c2fc1799b8. [accessed 08.11.2020].

[25] J. Ruohonen and V. Leppänen, "Whose Hands Are in the Finnish Cookie Jar,?" 2017 European Intelligence and Security Informatics Conference (EISIC), Athens, Greece, 2017, pp. 127-130.

[26] Nayanamana Samarasinghe, Mohammad Mannan. “Towards a global perspective on web tracking”, Computers&Security, Volume87, 2019, 101569, ISSN0167-4048.

[27] G. Acar, C. Eubank, S. Englehardt, M. Juarez, A. Narayanan, and C. Diaz. “The web never forgets: Persistent tracking mechanisms in the wild”. In Proceedings of CCS, 2014.

[28] Customer Think. “Are Cookies Good or Bad for Your eCommerce Health,”? [Online]. Available: https://customerthink.com/are-cookies-good-or-bad-for-your-ecommerce-health. [accessed 02.02.2021].

[29] CookiePro. “What is a Third-Party Cookie,”? [Online]. Available: https://www.cookiepro.com/knowledge/what-is-a-third-party-cookie/ [accessed 01.05.2021].

[30] I.D Mitchell. “Third-Party Tracking Cookies and Data Privacy.” SSRN Scholarly Paper. Rochester, NY: Social Science Research Network, April 25, 2012. [Online]. Available: https://doi.org/10.2139/ssrn.2058326 [accessed 11.11.2020].

[31] S.Englehardt, & D.Reisman, & C.Eubank, & P.Zimmerman, & J.Mayer, & A.Narayanan, & E.Felten. (2015). “Cookies that give you away,”. The surveillance implications of web tracking. Proceedings of the 24th International Conference on World Wide Web. 289-299.

[32] M.Falahrastegar, & H.Haddadi, & S.Uhlig,& R.Mortier. (2014). “Anatomy of the Third-Party Web Tracking Ecosystem”.

[33] T.Libert and R.Binns, 2019. “Good News for People Who Love Bad News”: Centralization, Privacy, and Transparency on US News Sites. In Proceedings of 11th ACM Conference on Web Science, Boston, MA, USA, June 30-July 3, 2019 (WebSci ’19), 10 pages.

[34] T.Libert., L.Graves., & R.K.Nielsen, (2018). “Changes in third-party content on European news websites after GDPR” (Reuters Institute for the Study of Journalism Reports: Factsheet). Reuters Institute for the Study of Journalism.

[35] T.C.Li., H. Hang., M.Faloutsos., P.Efstathopoulos. (2015) “TrackAdvisor: Taking Back Browsing Privacy from Third-Party Trackers”. In: Mirkovic J., Liu Y. (eds) Passive and Active Measurement. PAM 2015. Lecture Notes in Computer Science, vol 8995. Springer, Cham. [Online]. Available: https://doi.org/10.1007/978-3-319-15509-8_21 [accessed 03.03.2021].

[36] Cnet. “Nearly undetectable tracking device raises concern,”. [Online]. Available: https://www.cnet.com/news/nearly-undetectable-tracking-device-raises-concern/]. [accessed 11.11.2020].

[37] I.Fouad., N.Bielova., A.Legout.,& N.Sarafijanovic-Djukic. “Missed by Filter Lists: Detecting Unknown Third-Party Trackers with Invisible Pixels,”. Proceedings on Privacy Enhancing Technologies, 2020(2), 499-518. [online]. Available: https://doi.org/10.2478/popets-2020-0038 [accessed 03.03.2021].

[38] Wired. “Facebook is always watching you,” [Online]. Available: https://www.wired.com/2007/12/facebooks-is-al [accessed 01.05.2021].

[39] Leadspace. “B2B Segmentation vs. B2C Segmentation: A Look at the Differences,” 2020. [Online]. Available: https://www.leadspace.com/b2b-segmentation-vs-b2c-segmentation [accessed 01.05.2021].

[40] Termsfeed. “EU Cookies Directive,” [Online]. Available: https://www.termsfeed.com/blog/eu-cookies-directive/ [accessed 12.01.2021].

[41] Privacy Policies. “The Must-know Guide to the EU Cookie Directive,” [Online]. Available: https://www.privacypolicies.com/blog/eu-cookies-directive/ [accessed 03.02.2021].

[42] Cookiebot. “GDPR and cookie consent | Compliant cookie use,” [Online]. Available: https://www.cookiebot.com/en/gdpr-cookies/ [accessed 02.02.2021].

[43] Ghostery. “Study: Google is the Biggest Benificiary of the GDPR,” [Online]. Available: https://www.ghostery.com/study-google-is-the-biggest-beneficiary-of-the-gdpr [accessed 03.05.2021].

[44] S.C Boerman., S.Kruikemeier., & F.J. Zuiderveen Borgesius, (2018). “Exploring Motivations for Online Privacy Protection Behavior: Insights From Panel Data,”. Communication Research. [Online]. Available: https://doi.org/10.1177/0093650218800915 [accessed 11.11.2020].

[45] Mozilla Security blog. “Firefox 86 Introduces Total Cookie Protection,” [Online]. Available: https://blog.mozilla.org/security/2021/02/23/total-cookie-protection/ [accessed 01.04.2021].

[46] Optimanova. “Safari and Firefox on the rise against third-party tracking,” [Online]. Available: https://www.optimanova.com/safari-and-firefox-on-the-rise-against-third-party-tracking [accessed 01.03.2021].

[47] Wikipedia.”Do Not Track,” [Online]. Available: https://en.wikipedia.org/wiki/Do_Not_Track [accessed 02.05.2021]

[48] 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," in Logic Journal of the IGPL, vol. 25, no. 1, pp. 18-29, Feb. 2017.

[49] V. Dudykevych and V. Nechypor, "Detecting third-party user trackers with cookie files," 2016 Third International Scientific-Practical Conference Problems of Info communications Science and Technology (PIC S&T), Kharkiv, 2016, pp. 78-80. [Online]. Available: https://doi:10.1109/INFOCOMMST.2016.7905341 [accessed 04.03.2021].

[50] S.T Englehardt. “Automated discovery of privacy violations on the web”. 2018. Princeton University. Princeton University Doctoral Dissertations, 2011-2020 Computer Science. [Online]. Available: http://arks.princeton.edu/ark:/88435/dsp01hq37vr346 [accessed 10.11.2020].

[51] J. Purra and N. Carlsson, "Third-Party Tracking on the Web: A Swedish Perspective," 2016 IEEE 41st Conference on Local Computer Networks (LCN), Dubai, United Arab Emirates, 2016, pp. 28-34.

[52] A Project of the Electronic Frontier Foundation. “Cover your tracks,” [Online]. Available: https://coveryourtracks.eff.org/learn [accessed 17.03.2021].

[53] Ghostery. www.ghostery.com [Online]. [accessed 01.05.2021].

[54] D.Bouhnik, & G. Carmi, (2018). Interface Application Comprehensive Analysis of Ghostery. 10.13140/RG.2.2.34446.82249.

[55] G. Merzdovnik et al., "Block Me If You Can: A Large-Scale Study of Tracker-Blocking Tools," 2017 IEEE European Symposium on Security and Privacy (EuroS&P), Paris, France, 2017, pp. 319-333. [Online]. Available: https://doi:10.1109/EuroSP.2017.26. [accessed 02.02.2021].

[56] Cpomagazine. “Invasion of privacy tracking online behavior across web,”. [Online]. Available: https://www.cpomagazine.com/data-privacy/invasion-of-privacy-tracking-online-behavior-across-web [accessed 02.05.2021].

[57] Z.Yang & C.Yue (2020). “A Comparative Measurement Study of Web Tracking on Mobile and Desktop Environments,”. Proceedings on Privacy Enhancing Technologies. 2020. 24-44. 10.2478/popets-2020-0016.

[58] Englehardt, S & Narayanan, A. (2016). “Online Tracking: A 1-million-site Measurement and Analysis,”. 1388-1401. 10.1145/2976749.2978313.

[59] S.Englehardt., C.Eubank., P.Zimmerman, D.Reisman, & A.Narayanan. (2016). “OpenWPM: An automated platform for web privacy measurement,”.

[60] Github. “Open cookie database,” [Online]. Available: https://github.com/jkwakman/Open-Cookie-Database [accessed 11.11.2020].

[61] Cookiepro. 2021. [Online]. Available: https://www.cookiepro.com [accessed 01.02.2021].

[62] Developer.google.com. “Webextentions,”. [Online]. Available: https://developer.mozilla.org/en-US/docs/Mozilla/Addons/WebExtensions/API/webRequest/ResourceType [accessed 05.05.2021].

[63] Gemius. [online]. Available: https://www.gemius.com/about-us.html [accessed 02.05.2021].

[64] Clearcode. “What Facebook’s First-Party Cookie Means for AdTech,” [Online]. Available: https://clearcode.cc/blog/facebook-first-party-cookie-adtech/ [accessed 01.02.2021].

[65] Ytci. “Facebook plugin.html,” [Online]. Available: http://www.ytci.com/wp-content/uploads/2014/05/Facebook-Plugin.html [accessed 01.04.2021].

[66] Developers.google.com. “Cookie Matching,” [Online]. Available: https://developers.google.com/authorized-buyers/rtb/cookie-guide [accessed 01.04.2021].

[67] Gdpr.eu. “Cookies, the GDPR, and the ePrivacy Directive,” [Online]. Available: https://gdpr.eu/cookies [accessed 01.04.2021].

[68] A.Cahn, & S.Alfeld,&P.Barford, & S.Muthukrishnan. “An Empirical Study of Web Cookies,”. 891-901. 10.1145/2872427.2882991.

Appendix 1: Alexa TOP 50 popular websites

Place in TOP50	Domain
1	https://www.Google.com
2	https://www.Youtube.com
3	https://www.Delfi.ee
4	https://www.Postimees.ee
5	https://www.Vk.com
6	https://www.Bongacams.com
7	https://www.Swedbank.ee
8	https://www.Google.ee
9	https://www.Wikipedia.org
10	https://www.Ok.ru
11	https://www.Facebook.com
12	https://www.Google.ru
13	https://www.Harjuelu.ee
14	https://www.Mail.ru
15	https://www.Ekool.eu
16	https://www.Reddit.com
17	https://www.Aliexpress.ru
18	https://www.Ut.ee
19	https://www.Seb.ee
20	https://www.Microsoftonline.ee*
21	https://www.Err.ee
22	https://www.auto24.ee
23	https://www.neti.ee
24	https://www.rambler.ru
25	https://www.lenta.ru
26	https://www.ohtuleht.ee
27	https://www.aliexpress.com

28	https://www.telia.ee
29	https://www.eki.ee
30	https://www.zone.ee
31	https://www.Reverso.net
32	https://www.Ope.ee
33	https://www.Yahoo.com
34	https://www.Tootukassa.ee
35	https://www.Chaturbate.com
36	https://www.Lhv.ee
37	https://www.Rutracker.org
38	https://www.Online.ee
39	https://www.Roblox.com
40	https://www.Echo.msk.ru
41	https://www.Kv.ee
42	https://www.Elu24.ee
43	https://www.Tallinn.ee
44	https://www.Presaver.com
45	https://www.Glosbe.com
46	https://www.Zoom.us
47	https://www.Andmorefashion.com
48	https://www.Seasonvar.ru
49	https://www.Fishki.net
50	https://www.Twitch.tv

Appendix 2: Top-20 tracking companies and their third-party domains.

Company	Domain
AddThis	addthis.com, addthiscdn.com, addthisedge.com
AOL	aol.com, advertising.aol.com, atwola.com, advertising.com, adsonar.com, Tacoda.com, pictela.net, huffingtonpost.com, huffpost.com, huffpo.net, mapquestapi.com, 5min.com, aolcdn.com, goviral-content.com, srvntrk.com blogsmithmedia.com, mirabilis.com, mqcdn.com
Adobe	omniture.com, 2o7.net, demdex.net
Amazon	amazonaws.com, images-amazon.com, cloudfront.net
AudienceScience	revsci.net, wunderloop.com
Baidu	baidu.com, baidustatic.com, hao123.com, hao123img.com, bdstatic.com, bdimg.com, hao123.com.br
Burst Media	burstnet.com, blinkx.com
ComScor	voicefive.com, scorecardresearch.com, securestudies.com, sitestat.com
Conversant (ValueClick)	conversantmedia.com, awltovhc.com, kdukvh.com, qksrv.net, apmebf.com, ftjcfx.com, tqlkg.com, yceml.net, dotomi.com, mediaplex.com, lduhtrp.net
Facebook	facebook.com, facebook.net, fbcdn.net
Google	doubleclick.net, youtube.com, blogblog.com, android.com, ajax.googleapis.com, googlesyndication.com, doubleclick.com, youtube.googleapis.com, blogger.com, channelintelligence.com, content.googleapis.com, googletagmanager.com, 2mdn.net, youtube-nocookie.com, bloggercomments.googlecode.com, eedburner.com, fonts.googleapis.com, googleusercontent.com, ytimg.com, , blogspot.com, gmodules.com, goo.gl, googlevideo.com, wordtechnews.blogspot.com, invitemedia.com, googleadservices.com, gstatic.cn, ggpht.com, orkut.com, googleadsserving.cn, gstatic.com, recaptcha.net, google-analytics.com, javaplugins.googlecode.com, urchin.com, googleapis.com, maps.googleapis.com, googlecode.com, translate.googleapis.com,

	google.com, www.googleapis.com, googlecommerce.com
Nielson	mrworldwide.com, nielson.com
Quantcast	quantcast.com, quantserve.com
RadiumOne	radiumone.com, gwallet.com, po.st
247 Real Media	realmediadigital.com, realmedia.com, rmlacdn.net, 247realmedia.co.kr
Sina	sinajs.cn, sinaimg.cn, leju.com, weibo.com, sinauda.com, sinajs.js, wcdn.cn, sinahk.net, sina.com.cn, sinacdn.com, appsina.com, sinahk.net
Sizmek	serving-sys.com, peer39.net, republicproject.com
Twitter	twitter.com, twimg.com
Yahoo	yahoo.com, flickr.com, yieldmanager.com, bluelithium.com, overture.com, yahooapis.com, staticflickr.com, yldmgrimg.net, maktoob.com, xtendmedia.com, yahoo.net, sstatic.net, yimg.com, zenfs.com

Appendix 3: Cookie table

NEWS category

Host	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
www.delfi.ee	testcookie	Session	First party		5
.delfi.ee	cp_user_package_t	3 months	First party		5
.delfi.ee	delfi-adid	365 days	First party		5
www.delfi.ee	evid_set_0020	1 minute	First party		5
.delfi.ee .ohtuleht.ee .err.ee	__gfp_64b	3 years	First party	This is reported to be a cookie to manage the acceptance of other cookies by the visitor to the site	1
.hit.gemius.pl(www.delfi.ee) .hit.gemius.pl(www.postimees.ee .hit.gemius.pl(www.ohtuleht.ee) Hit.gemius.pl (www.err.ee) Hit.gemius.pl (elu24.ee)	Gtest	5 years	Third party/Gemius	This cookie file is used to prevent displaying survey questionnaires to internet users too frequently	4
.hit.gemius.pl (www.delfi.ee) .hit.gemius.pl (www.postimees.ee) hit.gemius.pl (www.ohtuleht.ee) hit.gemius.pl (www.err.ee) Hit.gemius.pl (elu24.ee)	Gdyn	5 years	Third party / Gemius	This cookie file is used to prevent displaying survey questionnaires to internet users too frequently	4

.delfi.ee	dcid	365 days	First party		5
.delfi.ee	__utma	2 years	First party	ID used to identify users and sessions	2
.delfi.ee	__utmb	30 minutes	First party	Used to distinguish new sessions and visits	2
.delfi.ee	__utmc	session	First party	Used only with old Urchin versions of Google Analytics and not with GA.js. Was used to distinguish between new sessions and visits at the end of a session	2
.delfi.ee	__utmz	6 months	First party	Contains information about the traffic source or campaign that directed user to the website	2
.delfi.ee	__utmt	10 minutes	First party	Used to monitor number of Google Analytics server requests	2
.delfi.ee	cX_T	Few seconds	First party	Set briefly, and then immediately deleted. This cookie holds a random value and is just used to find the top-level domain of the site	5
.delfi.ee	cstp	Few seconds	First party	Used to throttle cookie syncs with ad partners	5
.delfi.ee	cX_S	Session	First party	Site specific user session – single session	5
.delfi.ee	cX_P	365 days	First party	Site specific user session – across sessions	5
.cxense.com	gckp	365 days	Third Party / cxense.com	For building user profile information across all sites in the Cxense network	4
www.delfi.ee	enreachresp_0020	2 hours	First party		5
www.delfi.ee	evid_0020	3 months	First party		5
.delfi.ee	evid_0020-synced	1 month	First party		5
www.delfi.ee	adptset_0020	2 hours	First party		5
.delfi.ee	cX_G	365 days	First party	Global ID mapping different ids together into one ID	5

www.delfi.ee	enr_cxense_throttle	7 days	First party		5
.google.com	NID	6 months	Third party/Google	This cookies is used to collect website statistics and track conversion rates and Google ad personalization	4

Host	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
www.postimees.ee www.ohtuleht.ee www.err.ee www.elu24.ee	_cb_ls	365 days	First party	This cookie is from a legacy data migration process	2
.postimees.ee .elu24.ee	pm_c_list	2 years	First party	Cookies necessary for Postimees to function, signing in, gallery voting, etc	1
.postimees.ee .elu24.ee	__gfp_64b	3 years	First party	This is reported to be a cookie to manage the acceptance of other cookies by the visitor to the site	4
www.postimees.ee www.ohtuleht.ee www.err.ee www.elu24.ee	_cb	2 years	First party	This cookie stores a visitor's unique identifier for Chartbeat tracking	2
www.postimees.ee www.ohtuleht.ee www.err.ee www.elu24.ee	_chartbeat2	365 days	First party	This cookie stores timing information about when this visitor has visited site. This enables to distinguish between new, returning, and loyal visitors	2
www.postimees.ee www.ohtuleht.ee www.err.ee www.elu24.ee	_cb_svref	30 minutes	First party	This cookie stores the original referrer for this site visitor	3
www.postimees.ee www.elu24.ee	GoogleAdServingTest	Session	First party		5
.doubleclick.net (postimees.ee) .doubleclick.net (elu24.ee)	test_cookie	15 minutes	Third party/ Doubleclick.net	to determine if the website visitor's browser supports cookies	4

.postimees.ee .elu24.ee	__gads	1 year	First party	Showing of adverts on the site, for which the owner may earn some revenue	4
.postimees.ee .harjuelu.ee .ohtuleht.ee www.err.ee .elu24.ee	_ga	2 years	First party	To measure traffic on website. No sensitive information is saved	2
.postimees.ee .harjuelu.ee .ohtuleht.ee www.err.ee .elu24.ee	_gid	1 day	First party	The cookie is used to store information of how visitors use a website and helps in creating an analytics report of how the website is doing	2
.doubleclick.net(postimees.ee) .doubleclick.net (elu24.ee)	IDE	365 days	Third party/Doubleclick.net	Used for serving targeted advertisements that are relevant across the web.	4
.adform.net(postimees.ee) Adform.net (elu24.ee)	C	Some seconds	Third party/adform.net	Identifies if user's browser accepts cookies	4
.adform.net (postimees.ee) .adform.net (elu24.ee)	uid	365 days	Third party/adform.net	adform.net Unique identifier	4
.postimees.ee .elu24.ee	_gat_UA-78678198-1	1 minute	First party	Unique identity number of the account or website it relates to	2
.postimees.ee .elu24.ee	_fbp	90 days	First party	Facebook pixel tracking	4
www.postimees.ee www.ohtuleht.ee www.elu24.ee	_chartbeat4	1 minute	First party	This is a short-lived cookie that carries leftover engagement data from one page to the next	2

Host	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
www.harjuelu.ee	PHPSESSID	Session	First party	PHP session cookie associated with embedded content from this domain.	1
.paypal.com	x-cdn	Session	Third party	Is set by PayPal and minimizes delays in loading web page content by reducing the physical distance between the server and the user.	1
.www.paypal.com	akavpau_ppsd	Session	Third party	The cookie is used in context with transactions on the website - The cookie is necessary for secure transactions.	1
.paypal.com	Ts	3 years	Third party	Used in context with the PayPal payment-function on the website. The cookie is necessary for making a safe transaction through PayPal.	1
.paypal.com	ts_c	3 years	Third party	The cookie is necessary for making a safe transaction through PayPal.	1
.harjuelu.ee .www.err.ee	_gat	1 minute	First party	These cookies are used to collect information about how visitors use our website	2
.harjuelu.ee	tk_tc	Session	First party		5
.harjuelu.ee	tk_r3d	Three days	First party		5
.harjuelu.ee	tk_or	5 years	First party		5
.harjuelu.ee	tk_lr	365 days	First party		5
www.harjuelu.ee	adrotate-293	1 day	First party		5
.harjuelu.ee	__asc	30 minutes	First party	This cookie is used to collect information on consumer behavior	4
.harjuelu.ee	_auc	365 days	First party	This cookie is used to collect information on consumer behavior	4

Host	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
.ohtuleht.ee	_gat_optionalTracker	1 minute	First party	Google Analytics Cookies	4
.ohtuleht.ee	_gat_defaultTracker	1 minute	First party	Google Analytics Cookies	4
.google.com	NID	1 year	Third party/ Google	This cookie is used to collect website statistics and track conversion rates and Google ad personalization	4

Host	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
.err.ee	_gat_errKokku	1 minute	First		5
.err.ee	AttId	30 days	First		5
.err.ee	statUniqueId	30 days	First		5

Banking category

Host	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
www.lhv.ee	JSESSIONID	Session	First party	Used to display the content of the website correctly.	1
.lhv.ee	LANGUAGE	365 days	First party	Used to remember the last language used.	1
www.lhv.ee	LHV_LOGIN_TYPE_EE	365 days	First party	Used to save the user login method.	1
www.lhv.ee	ROUTE	Session	First party	Used for smooth operation of the internet bank.	1

www.lhv.ee	Client_ID	365 days	First party		5
www.lhv.ee	_fbp	90 days	First party	Used to display ads on Facebook.	4
www.lhv.ee	_gat	1 minute	First party	These cookies are used to collect information about how visitors use our website.	2
www.lhv.ee	_ga	24 hours	First party	Used to distinguish users.	2
www.lhv.ee	_gid	1 day	First party	The cookie is used to store information of how visitors use a website and helps in creating an analytics report of how the website is doing.	2
www.lhv.ee	alphachat_active session	90 days	Third party/ Alpha Blues	Used to remember user preferences and previous history in order to provide a better user experience.	3
www.lhv.ee	alphachat-test	Session	Third party/Alpha Blues	Used so that the customer service representative knows what the bot and the user were discussing when handing over the conversation. Contains user information entered in the chat window.	5
www.lhv.ee	COOKIES_CONSENT	365 days	First party	Used to track which cookies have been accepted.	1

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
.seb.ee	responsive	Session	First party	Used to adjust the website's layout depending on browser display size.	1
www.seb.ee	seblanguage	365 days	First party	Used to remember last used language.	1

www.seb.ee	sebsession	Session	First party	Used to correctly present the content of a website.	1
www.seb.ee	has_js	Session	First party	Used to ensure correct hyphenation.	1
www.seb.ee	gpv_pn	10 minutes	First party	Contains the previous page name to allow monitoring of user flow.	2
www.seb.ee	gpv_pu	1 day	First party	Stores the page name of the previous page.	2
www.seb.ee	s_cc	Session	First party	Adobe Site Catalyst cookie, determines whether statistical cookies are enabled in the browser.	2
www.seb.ee	s_fid	3 years	First party	Adobe Site Catalyst cookie, used to generate a unique id to identify a unique user.	2
www.seb.ee	s_vi	3 years	First party	Adobe Site Catalyst cookie, used to identify unique visitors, with an ID and timestamp	2
www.seb.ee	SEBConsents	365 days	First party	List of user accepted cookies.	1

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
.swedbank.ee	hanza	365 days	First party	Security tracking cookie.	1

.swedbank.ee	language	365 days	First party	Language settings of customer.	1
www.swedbank.ee	lastApp	365 days	First party	Last open application in iBank.	1
www.swedbank.ee	spa	365 days	First party		5
www.swedbank.ee	windowHeight	365 days	First party		5
www.swedbank.ee	windowWidth	365 days	First party		5
www.swedbank.ee	TSe5a42406029	Session	First party		5
www.swedbank.ee	AMCV_AB_	2 years	First party	This cookie is used to identify a unique visitor.	2
www.swedbank.ee	AMCVS_AB_	Session	First party	This cookie is used to identify a unique visitor.	2
www.swedbank.ee	COOKIE_CONSENT	90 days	First party		5
www.swedbank.ee	s_cc	Session	First party	Adobe Site Catalyst cookie, determines if cookies enabled in the browser.	2
www.swedbank.ee	scpa	90 days	First party	It is used to be aware of content preferences on this website	4

www.swedbank.ee	scp	90 days	First party	It is used to be aware of content preferences on this website	4
www.swedbank.ee	TS0…….	Session	First party		5

EDUCATION category

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
.www.ut.ee www.ope.ee www.ekool.eu	_gid	1 day	First party	ID used to identify users for 24 hours after last activity	2
.www.ut.ee www.ope.ee	_gat	1 minute	First party	Used to monitor number of Google Analytics server requests when using Google Tag Manager	2
.www.ut.ee www.ope.ee www.ekool.eu	_ga	2 years	First party	These cookies are used to throttle the request rate and to distinguishes unique users by assigning a randomly generated number as a client identifier.	2
www.ut.ee	has_js	Session	First party	Drupal uses this cookie to indicate whether or not the visitor browser has JavaScript enabled	3
www.ut.ee	__atuvc	2 years	First party	This cookie is associated with the AddThis social sharing widget, it stores an updated page share count	3
www.ut.ee	_atuvs	2 years	First party	This cookie is associated with the AddThis social sharing widget, it stores an updated page share count.	3
.ut.ee	_fbp	90 days	First party	Used by Facebook to deliver advertising. The cookie contains an encrypted Facebook user ID and browser ID.	4

.addthis.com	uvc	365 days	Third party/Addthis.com	Tracks how often a user interacts with AddThis	4
.addthis.com	xtc	365 days	Third party /Addthis.com	Anonymously tracks user behavior on the websites that allow a user to share pages on social media using the AddThis tool.	4
.addthis.com	loc	365 days	Third party /Addthis.com	Geolocate, within a State, the people sharing content via social media sites.	4
.doubleclick.net	Test.cookie	15 minutes	Third party/doubleclick.net	Tests if the user's browser supports cookies, on behalf of Google Inc. Advertising platform DoubleClick.	4

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
.ekool.eu	XSRF-TOKEN	Two hours	First	This cookie is written to help with site security in preventing Cross-Site Request Forgery attacks.	1
.ekool.eu	__cfduid	1 month	First	It is used to override any security restrictions based on the IP address the visitor is coming from. It does not contain any user identification information.	1
.ekool.eu	ekool_login_session	Few seconds	First party		5
.ekool.eu	tmp_show_mobile_version	1 day	First party		5
login.ekool.eu	Locale	1 week	First party		5
.nr-data.net	JSESSIONID	Session	Third party/New relic (Rebel ltd)		4

.ekool.eu	_gat_gtag_UA	1 minute	First party	Google analytics	4
www.google.com	NID	365 days	Third party/Google	These cookies are used to collect website statistics and track conversion rates and Google ad personalization	4

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
Portaal.eki.ee	9b4fc36747be05bd12ac96eb96821813	Session	First party		5
Portaal.eki.ee	jfcookie[lang]	365 days	First party		5

E-COMMERCE category

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
www.telia.ee	shoppingCartIcon	Session	First party	A cookie required to make a purchase from the online store, determining how many products the customer has in the shopping cart and allowing the customer to move to the shopping cart.	1
www.telia.ee	Segment	365 days	First party	A cookie that contains information about whether a visitor is a private or business customer so that the website can display the respective content.	1
www.telia.ee	LANGUAGE_EXT	365 days	First party	A cookie that stores the language of the application selected by the user	1

www.telia.ee	BIGipServer~DR~pood.telia.ee_http_pool	Session	First party	A cookie that helps balance the load on the web server by directing traffic	1
www.telia.ee	CookieConsent	39 years	First party	Stores the user's cookie consent state for the current domain.	5
www.telia.ee	alphachat-test	Session	First party	Used so that the customer service representative knows what the bot and the user were discussing when handing over the conversation. Contains user information entered in the chat window.	5
sso.telia.ee	plumbr-supported	Session	First party	A cookie that checks whether a web browser supports the use of cookies	2
.telia.ee	Plumbr-supported	10 years	First party		2
www.telia.ee	alphachat-active-session	1 day	First party	A cookie that contains information about a chat session on a website.	1
www.telia.ee .kv.ee .ivonikkolo.ee	Dc_gtm_UA....	1 minute	First party	Cookie used to restrict Google Analytics queries to increase the efficiency of network calls.	2
www.telia.ee www.andmorefashion.com .kv.ee .ivonikkolo.ee	_fbp	90 days	First party	Cookie used to forward biddings of promotional products, such as from third-party advertisers, in real time.	4
www.telia.ee www.auto24.ee	_gat	1 day	First party	Cookie used to restrict Google Analytics queries to increase the efficiency of network calls.	2
www.telia.ee auto24.ee .kv.ee .ivonikkolo.ee	_ga	2 years	First party	Google Analytics identifies users by unique identifiers, i.e. 'Client IDs'.	2
www.telia.ee	_gcl_au	3 months	First party	Google AdSense uses to test the effectiveness of ads on websites that use it.	4

www.telia.ee	_vis_opt_s	3 months and 1 week	First party	Cookie that measures the number of times a web browser is opened and closed. This cookie tracks user-created sessions.	2
www.telia.ee	_vis_opt_test_cookie	Session	First party	Cookie that monitors whether cookies are allowed in a web browser. It also helps track the number of user sessions	2
www.telia.ee	_vwo_ds	1 month	First party	Cookie that stores session-based information	2
www.telia.ee	_vwo_sn	30 minutes	First party	Cookie that stores session-based information	2
www.telia.ee	_vwo_uuid_v2	366 days	Third party	Cookie that measures traffic of unique visits to a website.	2
www.telia.ee	_vwo_uuid	10 years	Third party	Cookie that creates a unique identifier for each visitor and it is used to create segments for reporting.	2

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
www.auto24.ee	PHPSESSID	Session	First party	If the User has disabled the installation of cookies through the settings of his / her browser, the Services may not function in their full functionality.	1
.auto24.ee	CID	10 years	First party	If the User has disabled the installation of cookies through the settings of his / her browser, the Services may not function in their full functionality.	1
.auto24.ee	_utma	2 years	First party	ID used to identify users and sessions	2
.auto24.ee	_utmb	30 minutes	First party	Used to distinguish new sessions and visits.	2

.auto24.ee	_utmc	Session	First party	Used to distinguish between new sessions and visits at the end of a session.	2
.auto24.ee	_utmt	10 minutes	First party	Used to monitor number of Google Analytics server requests	2
.auto24.ee	_utmz	6 months	First party	Contains information about the traffic source or campaign that directed user to the website.	2
www.telia.ee .auto24.ee .kv.ee .ivonikkolo.ee	_gid	1 day	First party	ID used to identify users for 24 hours after last activity	2
.auto24.ee	cookies_agree	1 day	First party		5

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
.kv.ee	__gfp_64b	3 years	First party	This is reported to be a cookie to manage the acceptance of other cookies by the visitor to the site.	1
www.kv.ee	cookies_notified	10 years	First party		5
www.kv.ee	Saved_searches	6 months	First party		5
www.kv.ee	Kv_web	Session	First party		5
www.kv.ee	lang	Session	First party		5

.hit.gemius.pl	Gtest	5 years	Third party/ ls.hit.gemius.pl	This cookie file is used to prevent displaying survey questionnaires to internet users too frequently.	4
.hit.gemius.pl	Gdyn	5 years	Third party/ ls.hit.gemius.pl	This cookie file is used to prevent displaying survey questionnaires to internet users too frequently.	4
r1.kv.ee	OAID	365 days	Third party/ OpenX	This cookie is used by the ad server software to manage which ads are placed on our website, and to capture clicks on those ads.	4

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
.ivonikkolo.ee	_gcl_au	3 months	First party	Experimenting with advertisement efficiency across websites using their services.	4
.doubleclick.net	Test_cookie	15 minutes	Third party/doubleclick.net	Tests if the user's browser supports cookies, on behalf of Google Inc.	4
www.andmorefashion.com	_hjid	365 days	First party	It is used to persist the random user ID, unique to that site on the browser.	2
www.andmorefashion.com	_hjFirstSeen	30 minutes	First party	To track the beginning of the user's journey for a total session count.	3
www.andmorefashion.com	_hjAbsoluteSessionInProgress	30 minutes	First party	To track the beginning of the user's journey for a total session count.	3
www.andmorefashion.com	_hjIncludedInPageviewSample	30 minutes	First party	Whether that visitor is included in the data sampling defined by pageview limit.	3

www.andmorefashion.com	_hjIncludedInSessionSample	30 minutes	First party	Whether that visitor is included in the data sampling defined by site's daily session limit.	3

SERVICES category

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
www.neti.ee	__gfp_64b	3 years	First party	This is reported to be a cookie to manage the acceptance of other cookies by the visitor to the site.	1
.neti.ee	_utma	2 years	First party	ID used to identify users and sessions	2
.neti.ee	_utmb	30 minutes	First party	Used to distinguish new sessions and visits. This cookie is set when the GA.js javascript library is loaded and there is no existing __utmb cookie. The cookie is updated every time data is sent to the Google Analytics server.	2
.neti.ee	_utmc	Session	First party	Used only with old Urchin versions of Google Analytics and not with GA.js. Was used to distinguish between new sessions and visits at the end of a session.	2
.neti.ee	_utmt	10 minutes	First party	Used to monitor number of Google Analytics server requests	2
.neti.ee	_utmz	6 months	First party	Contains information about the traffic source or campaign that directed user to the website. The cookie is set when the GA.js javascript is loaded and updated when data is sent to the Google Anaytics server	2
.hit.gemius.pl	Gtest	5 years	Third party / Gemius	This cookie file is used to prevent displaying survey questionnaires to internet users too frequently	4

.hit.gemius.pl	Gdyn	5 years	Third party /Gemius	This cookie file is used to prevent displaying survey questionnaires to internet users too frequently	4

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
www.zone.ee	Pll_language	365 days	First party	This cookie name is associated with the Polylang plug-in for WordPress powered websites. it stores a language preference for the visitor to support multi-lingual websites.	3
.zone.ee .tallinn.ee .tootukassa.ee	_ga	2 years	First party	These cookies are used to throttle the request rate and to distinguishes unique users by assigning a randomly generated number as a client identifier.	2
.zone.ee .tallinn.ee .tootukassa.ee	_gid	1 day	First party	ID used to identify users for 24 hours after last activity	2
.zone.ee	Zte2095	1 day	First party	These cookies are used to enhance the performance and functionality of Mailchimp Sites	2
.zone.ee	_dc_gtm_UA-XXXXXXXX	1 minute	First party	Used to monitor number of Google Analytics server requests	1
widget-mediator.zopim.com	AWSALB	7 days	Third party/Zopim	ID Load Balancer	4
widget-mediator.zopim.com	AWSALBCORS	7 days	Third party/Zopim	ID Load Balancer	4
.zone.ee	__zlcmid	365 days	First party	Live chat widget on Slack contact page (ZopIM)	4

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
www.online.ee	Suhtlus	1 month 2 days	First party		5
www.online.ee	BIGipServer~DR~online_http_pool	Session	First party		5
sso.telia.ee	TestThirdPartyCookie	Session	First party		5
.telia.ee	alphachat-test	Session	First party	Used so that the customer service representative knows what the bot and the user were discussing when handing over the conversation. Contains user information entered in the chat window.	5
.telia.ee	plumbr-supported	1 minute	First party	A cookie that checks whether a web browser supports the use of cookies	2
www.telia.ee	alphachat-active-session	1 day	First party	A cookie that contains information about a chat session on a website.	1
.telia.ee	LANGUAGE_EXT	365 days	First party	A cookie that stores the language of the application selected by the user.	1
.telia.ee	segment	Session	First party	cookie that contains information about whether a visitor is a private or business customer so that the website can display the respective content	1

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)

.www.tallinn.ee	TallinnFontSize	3 years	First party	to remember the selected font size	1
	TallinnAccNoStyle	3 years	First party	style removal feature for the visually impaired	1
	TallinnAccNoColor	3 years	First party	color roll-off symbol for the visually impaired	1
	TallinnAccLineHeight	3 years	First party	line spacing height indicator for the visually impaired	1
	TallinnAccEmphasis	3 years	First party	a sign of emphasis on choice for the visually impaired	1
	TallinnAccContrast	3 years	First party	contrast feature selected for the visually impaired	1
www.tallinn.ee	Sess_admin	Session	First party	to manage website sessions	1
.tallinn.ee	_gat	1 minute	First party	These cookies are used to collect information about how visitors use our website. We use this information to compile reports and to help us improve the website.	2
.google.com	NID	6 months	Third party/Google	Cookies are used to collect website statistics and track conversion rates and Google ad personalization	4

balticlivecam.com	wppas_pvbl	Session	Third party/ balticlivecam		4
balticlivecam.com	__asc	30 minutes	Third party/ balticlivecam		4
balticlivecam.com	__auc	365 days	Third party/ balticlivecam		4
balticlivecam.com	_ga	2 years	Third party/ balticlivecam	ID used to identify users	4
balticlivecam.com	_gid	1 day	Third party/ balticlivecam		4
balticlivecam.com	_gat	1 minute	Third party/ balticlivecam		4

Domain	Cookie Name	Expiration time	First party or third party/owner	Purpose	Category (1-5)
.www.tootukassa.ee	SSESS34b76d9a97fc1d776265da602ace6715	Session	First party		5
.www.tootukassa.ee	SESS34b76d9a97fc1d776265da602ace6715	Session	First party		5
www.tootukassa.ee	has_js	Session	First party	Used to ensure correct hyphenation	3

www.tootukassa.ee	Responsive	Session	First party	Used to adjust the website's layout depending on browser display size.	1
.tootukassa.ee	nmstat	2 years 9 months	First party	It enables site owners to gather usage statistics about their websites.	2
.tootukassa.ee	_gat_gtag_UA_67689823_2	1 minute	First party	Google Analytics	4
6168205.global.siteimproveanalytics.io	AWSELB	Session	Third party/ Gandi SAS	ID Load Balancer	4
6168205.global.siteimproveanalytics.io	AWSELBCORS	Session	Third party/ Gandi SAS	ID Load Balancer	4
www.tootukassa.ee	lineheight	Session	First party		5
www.tootukassa.ee	layout	Session	First party		5
www.tootukassa.ee	fontsize	Session	First party		5

Appendix 4: HTTP request of web beacon

referrer	headers
https://www.delfi.ee/	[["Host","ee.hit.gemius.pl"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0"],["Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip, deflate, br"],["Referer","https://www.delfi.ee/"],["Origin","https://www.delfi.ee"],["Connection","keep alive"],["Cookie","Gdyn=KlxY7RGGQMQGr9bckILD8liissGMf1oaL6nxmG7pcjCZBLInmGRoSvgYwvKxGsRP5G7tGKGGqC08bGw8EoG2GxsK3Fy_9FSG"]]
https://www.postimees.ee/	[["Host","www.facebook.com"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0"],["Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip, deflate, br"],["Referer","https://www.postimees.ee/"],["Content-Type","multipart/form-data; boundary=---------------------------201352507713345676013622182345"],["Content-Length","92902"],["Origin","https://www.postimees.ee"],["Connection","keep-alive"]]
https://login.ekool.eu/iframe/	[["Host","bam.eu01.nr-data.net"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0"],["Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip, deflate, br"],["Referer","https://login.ekool.eu/iframe/"],["Content-Type","text/plain;charset=UTF-8"],["Content-Length","33"],["Origin","https://login.ekool.eu"],["Connection","keep-alive"],["Cookie","JSESSIONID=5afadd8e854348f8"]]
https://www.ut.ee/et	[["Host","m.addthis.com"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0"],["Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip, deflate, br"],["Referer","https://www.ut.ee/et"],["Content-Type","text/plain;charset=UTF-8"],["Content-Length","0"],["Origin","https://www.ut.ee"],["Connection","keep-alive"],["Cookie","uvc=1%7C8; loc=MDAwMDBFVUVFMzcyMzk0MjA0ODAwMDAwMDBDSA=="]]
https://sso.telia.ee/lib/webapi?webapi=true&locale=et	[["Host","plumbrapi.telia.ee"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0"],["Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip, deflate,

	br"],["Referer","https://sso.telia.ee/lib/webapi?webapi=true&locale=et"],["Content-Type","text/plain;charset=UTF-8"],["Content-Length","427"],["Origin","https://sso.telia.ee"],["Connection","keep-alive"],["Cookie","LANGUAGE_EXT=et_EE; segment=b2c; shoppingCartIcon=0"]]
https://www.telia.ee/era	[["Host","plumbrapi.telia.ee"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0"],["Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip, deflate, br"],["Referer","https://www.telia.ee/era"],["Content-Type","text/plain;charset=UTF-8"],["Content-Length","990"],["Origin","https://www.telia.ee"],["Connection","keep-alive"],["Cookie","LANGUAGE_EXT=et_EE; segment=b2c; shoppingCartIcon=0; alphachat-test=1"]]
https://sso.telia.ee/sso2/login_suhtlus.jsp?goto=https%3A%2F%2Fwww.online.ee%2FinitSSO.php&lang=1&loginURI=login_suhtlus.jsp&RequestID=5cdef626620a8efdc0d49213173592f2a02a90dc&IssueInstant=2021-02-21T12%3A39%3A20%2B02%3A00&ProviderID=https%3A%2F%2Fwww.online.ee%3A443%2F%3FRealm%3D%2F&RelayState=bc4e838e15aa0e9ac40bb1befce63a34c6e702aa	[["Host","plumbrapi.telia.ee"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0"],["Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip, deflate, br"],["Referer","https://sso.telia.ee/sso2/login_suhtlus.jsp?goto=https%3A%2F%2Fwww.online.ee%2FinitSSO.php&lang=1&loginURI=login_suhtlus.jsp&RequestID=5cdef626620a8efdc0d49213173592f2a02a90dc&IssueInstant=2021-02-21T12%3A39%3A20%2B02%3A00&ProviderID=https%3A%2F%2Fwww.online.ee%3A443%2F%3FRealm%3D%2F&RelayState=bc4e838e15aa0e9ac40bb1befce63a34c6e702aa"],["Content-Type","text/plain;charset=UTF-8"],["Content-Length","676"],["Origin","https://sso.telia.ee"],["Connection","keep-alive"]]
https://www.elu24.ee/	[["Host","www.facebook.com"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0"],["Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip, deflate, br"],["Referer","https://www.elu24.ee/"],["Content-Type","multipart/form-data; boundary--------------------------547811312389310062245936693 4"],["Content-Length","80596"],["Origin","https://www.elu24.ee"],["Connection","keep-alive"]]
https://www.elu24.ee/	[["Host","www.facebook.com"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0"],["Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip, deflate, br"],["Referer","https://www.elu24.ee/"],["Content-Type","multipart/form-data; boundary----------

	142751941035256637308657619 0"],["Content-Length","80596"],["Origin","https://www.elu24.ee"],["Connection","keep-alive"]]
https://www.facebook.com/v9.0/plugins/page.php?adapt_container_width=true&app_id=153118205035094 0&channel=https%3A%2F%2Fstaticxx.facebook.com%2Fx%2Fconnect%2Fxd_arbiter%2F%3Fversion%3D46%23cb%3Df315dbf472221a6%26domain%3Dwww.tallinn.ee%26origin%3Dhttps%253A%252F%252Fwww.tallinn.ee%252Ff2d4ba42d7db5f8%26relation%3Dparent.parent&container_width=269&height=300&hide_cover=false&href=https%3A%2F%2Fwww.facebook.com%2Ftallinnalinn%2F&locale=et_EE&sdk=joey&show_facepile=true&show_posts=true&small_header=true	[["Host","www.facebook.com"],["User-Agent","Mozilla/5.0 (X11; Linux x86_64; rv:80.0) Gecko/20100101 Firefox/80.0"],["Accept","*/*"],["Accept-Language","en-US,en;q=0.5"],["Accept-Encoding","gzip, deflate, br"],["Referer","https://www.facebook.com/v9.0/plugins/page.php?adapt_container_width=true&app_id=1531182050350940&channel=https%3A%2F%2Fstaticxx.facebook.com%2Fx%2Fconnect%2Fxd_arbiter%2F%3Fversion%3D46%23cb%3Df315dbf472221a6%26domain%3Dwww.tallinn.ee%26origin%3Dhttps%253A%252F%252Fwww.tallinn.ee%252Ff2d4ba42d7db5f8%26relation%3Dparent.parent&container_width=269&height=300&hide_cover=false&href=https%3A%2F%2Fwww.facebook.com%2Ftallinnalinn%2F&locale=et_EE&sdk=joey&show_facepile=true&show_posts=true&small_header=true"],["Content-Type","multipart/form-data; boundary=---------------------------805425890245514646929642016 92"],["Content-Length","4910"],["Origin","https://www.facebook.com"],["Connection","keep-alive"]]

Appendix 5: OpenWPM crawl settings

```
from automation import CommandSequence, TaskManager

# The list of sites that we wish to crawl

NUM_BROWSERS = 1

sites =
["http://www.delfi.ee","https://www.postimees.ee","https://www.swedbank.ee","https://www.harjuelu.ee","https://www.ekool.eu","https://www.ut.ee","https://www.seb.ee","https://www.microsoftonline.ee","https://www.err.ee","https://www.auto24.ee","https://www.neti.ee","https://www.ohtuleht.ee","https://www.telia.ee","https://www.eki.ee","https://www.zone.ee","https://www.ope.ee","https://www.tootukassa.ee","https://www.lhv.ee","https://www.online.ee","https://www.kv.ee","https://www.elu24.ee","https://www.tallinn.ee","https://www.andmorefashion.com",]

# Loads the default manager params

# and NUM_BROWSERS copies of the default browser params

manager_params, browser_params = TaskManager.load_default_params(NUM_BROWSERS)

# Update browser configuration (use this for per-browser settings)

for i in range(NUM_BROWSERS):

    # Record HTTP Requests and Responses

    browser_params[i]["http_instrument"] = True

    # Record cookie changes

    browser_params[i]["cookie_instrument"] = True

    # Record Navigations

    browser_params[i]["navigation_instrument"] = True

    # Record JS Web API calls

    browser_params[i]["js_instrument"] = True

    # Record the callstack of all WebRequests made

    browser_params[i]["callstack_instrument"] = True

    # Record DNS resolution

    browser_params[i]["dns_instrument"] = True

    browser_params[i]['display mode'] = "headless"

# Launch only browser 0 headless

browser_params[i]["display_mode"] = "headless"

# Update TaskManager configuration (use this for crawl-wide settings)

manager_params["data_directory"] = "~/Desktop/"
```

```
manager_params["log_directory"] = "~/Desktop/"
# Instantiates the measurement platform
# Commands time out by default after 60 seconds
manager = TaskManager.TaskManager(manager_params, browser_params)
# Visits the sites
for site in sites:
 # Parallelize sites over all number of browsers set above.
  command_sequence = CommandSequence.CommandSequence(
    site,
    reset=True,
    callback=lambda success, val=site: print("CommandSequence {} done".format(val)),  )
  # Start by visiting the page
  command_sequence.get(sleep=3, timeout=60)
  # Run commands across the three browsers (simple parallelization)
  manager.execute_command_sequence(command_sequence)
# Shuts down the browsers and waits for the data to finish logging
manager.close()
```

License

Non-exclusive license to reproduce thesis and make thesis public

I, **Priit Põdra,**

1. herewith grant the University of Tartu a free permit (non-exclusive license) to reproduce, for the purpose of preservation, including for adding to the DSpace digital archives until the expiry of the term of copyright,

Web tracking in the most popular Estonian websites,

supervised by Arnis Paršovs PhD.

2. I grant the University of Tartu a permit to make the work specified in p. 1 available to the public via the web environment of the University of Tartu, including via the DSpace digital archives, under the Creative Commons license CC BY NC ND 3.0, which allows, by giving appropriate credit to the author, to reproduce, distribute the work and communicate it to the public, and prohibits the creation of derivative works and any commercial use of the work until the expiry of the term of copyright.

3. I am aware of the fact that the author retains the rights specified in p. 1 and 2.

4. I certify that granting the non-exclusive license does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Priit Põdra

14/05/2021