

The Use of Cookies on Finnish Municipal Websites

UNIVERSITY OF TURKU
Department of Computing
Master of Science (Tech) Thesis
Software Engineering
May 2024
Samuel Leinonen

UNIVERSITY OF TURKU
Department of Computing

SAMUEL LEINONEN: The Use of Cookies on Finnish Municipal Websites

Master of Science (Tech) Thesis, 50 p.
Software Engineering
May 2024

Privacy can be seen as a fundamental and universal human right. Therefore privacy is very important and its importance has only grown as time has passed and more and more of our lives are spent online. People generate a lot of data when they use websites on the internet and this data can be used by big companies like Google in order to better advertise to them. Municipal websites are no exception to this and they also generate a lot of data when they are used. This data can be collected through the use of analytics and marketing cookies that are on the websites.

This thesis studies the use of cookies on Finnish Municipal websites by analyzing a small subset of 25 websites out of the 309 websites. The goal is to find if the websites have a proper cookie consent notice and if the cookies that they use have been categorized correctly. If they are not categorized correctly, the users will have a problem when they are giving consent to the use of cookies as the websites might use some cookies that the user unknowingly consents to.

5 out of the 25 websites are found to have problems with the cookie consent notice and 10 out of the 25 have problems with the categorization of cookies. This means that there are underlying problems with the use of cookies. The use of Google's 3rd party cookies for analytics is especially prominent. This means that many of the websites are collecting data without a clear and explicit consent from the end user although that is required by the GDPR. Also, a regular non-technical user will not be able to tell if a website is using a 3rd party service to gather data only by looking at the cookie consent notice.

Keywords: online privacy, web-based service, municipal websites, data leaks, GDPR

Contents

1	Introduction	1
2	Background	4
2.1	Privacy	4
2.1.1	Data Collection and Tracking	4
2.1.2	Web Analytics	5
2.2	Cookies	7
2.2.1	Cookie Structure	8
2.2.2	Cookie Categorization	9
2.2.3	Tracking Users with Cookies	11
2.3	GDPR	12
2.3.1	The history of privacy legislation in the EU	14
2.3.2	User Consent	15
2.4	Cookie Consent Interfaces	17
2.4.1	Design Choices	18
2.4.2	Dark Patterns	19
3	Methods	22
4	Results	25
4.1	Cookie Notices	25

4.1.1	Good Cookie Notices	27
4.1.2	Dark Patterns and Non-GDPR Compliant Cookie Cotices	28
4.2	Cookies Used on the Websites	31
5	Discussion	45
6	Conclusions	48
	References	51

1 Introduction

Privacy can be seen as a fundamental and universal human right. Therefore privacy is very important and its importance has only grown as time has passed and more and more of our lives are spent online [1].

Data collection is a part of privacy that has become more important as people spend increasing amounts of time on the internet. People generate a lot of data when they use the internet and this data can be used by big companies like Google in order to better advertise to them [2]. Data collection can also cause additional risks if the data is stolen in a data breach.

Municipal websites are something that almost every resident has to use at least occasionally. They contain important information about public services and events that are happening in the area. The users of the websites can also use them to find help for health related issues. Search terms and navigational paths can reveal sensitive information about the user and they might not want that information to be leaked to third parties, especially without consent.

This thesis focuses on the cookies that are used on the Finnish municipal websites. We analyze a small subset containing 25 of the 309 Finnish municipal websites to see what kind of cookie consent notices they have and how they have categorized the cookies that they are using.

There are two research questions that this thesis tries to answer. These questions are applied to the Finnish municipal websites.

- **RQ1:** How much control does the user have on the privacy settings?
- **RQ2:** Have the cookies been categorized properly?

With the first question this thesis wants to find out if the users can properly control their privacy on the websites. Essentially this question has two requirements for the websites. Do the websites have a good cookie consent notice and does the option on the cookie notice actually affect the cookies that are used on the website. The GDPR requires the websites to ask consent in order to use cookies and this is generally done through a cookie consent notice that the user sees when they first enter the website. These cookie consent notices are not standardized in any way but they should allow a "freely given, specific, informed and unambiguous" consent from the user [3].

The second question concerns categories of cookies such as "strictly necessary cookies", "functionality cookies", "analytical cookies" and "marketing cookies". The main focus of this thesis is on the analytics and marketing cookies since those are typically third-party cookies that are a bigger privacy risk than the necessary cookies that are often first-party [4].

However, it is possible that the cookies are not categorized correctly. For example a website can claim that the cookies that they use for analytics purposes are necessary cookies. This thesis will compare the categorization on the websites to see if it is consistent and if there are any outliers. Cookies from popular third-parties such as analytics cookies from Google analytics are easy to recognize and therefore it is easy to see what category it should be under.

The rest of the thesis is structured as follows. Chapter 2 contains background information

about topics that are relevant to this thesis. Chapter 3 outlines the research methods used in this thesis. The results of the study, which consists of analyzing a set of websites, are discussed in Chapter 4. Chapter 5 has a discussion on the effects of the results and some potential reasons and solutions to the problems. Lastly, Chapter 6 brings the thesis to a conclusion.

2 Background

2.1 Privacy

Privacy can be seen as a fundamental and universal human right. Privacy is not only an individual right, but it also has a social value. Historically privacy has been regarded as an element of liberty, the right to be free from intrusions by the state. [1]

Privacy can also be seen as the ability of individuals to seclude both themselves and information about themselves. Online privacy has become more and more important as the use of internet has become common for everyday use. However, the use of internet generates a lot of personal data which has led to various privacy concerns over the years [2]. Tracking the actions of the users throughout the internet is also a privacy concern.

2.1.1 Data Collection and Tracking

Data collection can happen either with the user's awareness and consent or it can happen without the user knowing about it. Ideally, each user should have control over who can access their personal data, what data is being collected and what the data is used for. This can only happen if the user has control over their data, which is often achieved by legislation as that forces the companies who can gather data to change their approach.

Cookies have been used to track users for a while. Earlier it was possible to use them without the user's knowledge or consent but now with current legislation websites have to both disclose the use of cookies and obtain user consent to use them. In addition, browsers have started to add countermeasures to tracking with cookies which has also made them less popular. This means that companies might want to use some other methods that are less regulated and harder to detect to collect user data.

There are also other ways to track users like for example browser fingerprinting. The browser sends a lot of information to the website about the user's settings like the screen resolution and the browser version to name a few. Most browsers can be given a unique "fingerprint" by using this data and the browser and by extension the user can be tracked across the internet. This is currently completely unregulated unlike cookies so it is harder for users to stay private if websites decide to use it. [5]

In addition to these methods used by the websites, users often willingly and freely give away their personal information online. The main reason is that some services online require the user to share their personal information so there is a trade-off between benefits and privacy concerns. This is especially true on social media websites, which in turn causes those websites to gain access to even more personal data. [6]

2.1.2 Web Analytics

The most popular web analytics tool that is currently used is Google Analytics. It is used by 84% of all the websites that use any web analytics tool. [7] The way Google Analytics operates is that it is free to use for the developers who might want to use some analytics on their website [8]. This means that the barrier to entry is very low as there is no required investment to start using it.

This means that all the data is being stored and processed by Google and in exchange

Google shares some of that data with the website maintainers for analytics purposes. Because Google controls the data, it has some implications for the user's privacy. Google gets access to data such as the user's IP address, as well as the pages they visit on the website and any searches they do. These can sometimes contain personal data that the user does not want or expect to get stored by a 3rd party. [7]

With the access to the IP addresses and other unique browser identifiers, Google can then track the user's visits on other websites that also use Google analytics by obtaining the same IP address from those websites. Using this data, Google can create and enrich their existing data profiles on the users.

Google Analytics is part of the Google Marketing Platform therefore it is a part of Google's advertising toolset [8]. Google will use the data that they acquire from website to make their advertising and marketing better. This is a privacy concern and web developers might want to use a different platform to better protect the privacy of their users.

Matomo is an open source alternative to Google Analytics. There are two ways to use Matomo. The website can either host it themselves, meaning that they will be in control of 100% of the data collected from the users. Using Matomo locally means that the website has complete ownership of the data that is collected from the users. This is great for privacy as the data stays in control of only one entity and does not get transferred to another party. Alternatively, they can use a cloud version where Matomo will handle the data storage in exchange for money. Because the service costs money unlike Google Analytics, Matomo will not use the data for their own purposes as they have already been compensated. [7]

Matomo has multiple options that can make using it more privacy-friendly. For example Matomo has tools that can be used to anonymize personal data gathered from the users such as IP addresses. Additionally, Matomo can be configured to not process any personally identifiable information and the tracking cookies can be made to expire much earlier than

the 2 years that Google Analytics would store them.

Other analytics companies that are relevant in this thesis are Siteimprove and Hotjar. Siteimprove markets itself as an alternative to Google Analytics. It is cloud-based and cannot be run locally like Matomo but it would be a better option for privacy than Google. It is also easier to use effectively compared to Google Analytics. [9]

Hotjar is trying to fill a slightly different niche than Google Analytics. Google Analytics is a tool for quantitative analytics and Hotjar is a tool for qualitative analytics [10]. Hotjar even recommends that it is used in conjunction with Google Analytics as they are doing slightly different things.

Google Analytics records what the users are doing on the website while Hotjar tells why the users are doing what they are doing. For example Hotjar can be used to create heatmaps based on what parts of the website the users are engaging the most and what parts are not getting much user interaction. [11]

2.2 Cookies

Cookies can be classified into first-party cookies and third party cookies. First-party cookies are created by the website that the user visits and third-party cookies are created by third-party entities. First-party cookies cannot compromise user privacy in the same way as third-party cookies because they are only active on the websites that created them. [12]

Cookies are often used to implement essential features on the website such as tracking login state or shopping carts. However, cookies can also be used to track the user's actions on the website. Web analytics companies can use third-party cookies to track users even if the user does not directly visit their website. If the user visits a site that has content from a

third party and then later visits another website that has content from that same third party, they can track the user across the websites [13].

2.2.1 Cookie Structure

A cookie is a formatted string that consists of key-value pairs that are separated by a semi-colon [14]. Each cookie has a required name-value pair and it can additionally have zero or more additional attribute-value pairs. The cookies are created and sent by the server to the user agent, which can be for example a web browser. The cookie is then stored by the user agent. The cookie can have the following attributes. [13]

Name: The name attribute contains the name that the cookie has received from the server. The name identifies the cookie to a particular server [14].

Value: The Value attribute contains the data that the cookie is transmitting between the server and the browser. The data can be in plain text but this would be a privacy and security concern as the data is not always sent over a secure channel. The data should be encrypted or otherwise obfuscated to ensure privacy and security. [13], [14]

The value can also be used to track users. For example Google Analytics uses the Value field to create a unique client ID for each browser. This means that the user can be tracked if they visit multiple websites that use Google Analytics. The ID is unique to the browser so if the user changed their browser or device, they will receive a new ID. [15]

Domain and Path: The domain and path attributes define the scope of the cookie. The domain attribute specifies the hosts to which the cookie will be sent to. For example, if the value of the domain attribute is "example.com" the cookie will be included in the cookie Header when the browser makes a request to "example.com". The value in the path attribute must exist in the URL of the web site that is being requested by the browser.

If the domain and path values are not set, they will default to the domain and path of the requested resource. Subdomains are not included in this case but they are if the domain is set manually. [13]

Expires and Max-Age: The expires attribute contains the date and time when the cookie will expire. The browser is not required to store the cookie this long however and it can be deleted earlier for memory or privacy concerns. The Max-age attribute indicates the maximum lifetime of the cookie in the number of seconds until it expires. If the cookie has both the Expires and Max-age attributes, the Max-age has precedence. If the cookie has neither, it will be retained until the current session is over. [13] These attributes are currently limited to 400 days in Google Chrome when earlier the expiration could be set as far into the future as possible [16].

Secure and HttpOnly: The secure attribute limits the cookie to secure channels that are defined by the browser. The HttpOnly attribute limits the scope of the cookie to just HTTP requests. This means that the cookie cannot be accessed by non-HTTP APIs such as a web browser API that would expose the cookie to scripts. [13]

2.2.2 Cookie Categorization

Cookies can be categorized based on their origin, duration and purpose. Origin refers to whether a cookie is a first-party or third party cookie. Cookies can be either session cookies or persistent cookies depending on their expiration duration. However, categorizing for purpose is not as easy.

The most common way to categorize the purpose of cookies is to divide them into four categories: *strictly necessary* cookies, *functionality* cookies, *performance* cookies and *targeting/advertising* cookies. These four categories were originally proposed by UK International Chamber of Commerce (UK ICC) and they are widely used, although sometimes

with slightly different names. [17]

These category names can be confusing for users. For example targeting cookies can be harder for users to understand than advertising cookies [18]. Similarly some might think that functionality cookies are needed for the website to function although that only applies to strictly necessary cookies [19].

Strictly mandatory or essential cookies are cookies that are required for the website to function. They do not require consent from the user but the user should still be informed that they are being used and for what purpose they are used for [4]. They are typically first-party cookies.

Functionality cookies are used for remembering user choices on the website, such as language preferences or personalization of the website [4], [17]. They offer additional functionality to the user, but are not essential.

Performance cookies are cookies that are used to collect information about the browsing habits and preferences of users for the purpose of improving the website. They are typically third party cookies, for example Google Analytics. Advertising cookies are used to observe the users' browsing activities and to help deliver targeted ads. These are typically third party cookies. [4]

In general essential and functionality cookies are more useful to the user. Performance and especially advertising cookies are more useful for the website than the user and because of this the user does not need to enable them to use the website properly [17]. This means that it would be more beneficial for the website owners to categorize their cookies more favourably for them by having cookies used for advertising listed as necessary ones instead.

It is not always clear how some cookies should be categorized. Additionally, website

owners and users and might all have different assessments on what counts as a mandatory cookie. This is a big problem and it is why there needs to be legislation such as the GDPR to make rules on what is allowed and what is not.

2.2.3 Tracking Users with Cookies

Because cookies can be used to recognize returning users, they can also be used for tracking. Third-party cookies are widely used to track users across websites and to serve targeted ads to them. The majority of third-party cookies are set by advertising and tracking services. [20]

Many browsers have set countermeasures against cross-site tracking by third-party cookies. For example, Firefox has partitioned all third-party cookie access since 2022. This partitioning means that cookies set by a third party on one website are distinct from cookies set by that same third-party on other websites. As a result, the user cannot be tracked by those third-party cookies. Google Chrome is the only major browser that does not restrict third-party cookies by default, although it plans to restrict them by late 2024. [20]

First-party cookies can also be used to track users. The obvious use case is same-site tracking, which is not as invasive to privacy as tracking the user across multiple different sites but it can still be used to obtain information from the activity of the users on the websites that they use, for example social media or news websites. [20]

A big issue arises when the third-parties collaborate with the first-parties by setting first-party tracking cookies through third-party scripts. These tracking cookies can be shared to multiple different websites to be used as first-party cookies. With this it is possible to track users across multiple domains by only using first-party cookies for tracking. [20]

Additionally, these different third-party trackers can share information with each other.

This makes tracking users more simple because every third-party tracker does not have to directly collaborate with each first-party to publish their cookies but can instead use tracking cookies set by other third-parties to monitor user activity. [20]

Countermeasures for first-party tracking are more difficult to create compared to third-party tracking. Most strictly necessary and functionality cookies are set by first parties so for example blocking all first-party cookies will negatively impact the user experience unlike blocking all third-party cookies. Some possible countermeasures are to use machine learning to detect which first-party cookies are used for tracking and selectively only block those ones. [20]

2.3 GDPR

The EU Charter of Fundamental Rights stipulates that EU citizens have the right to protection of their personal data. The GDPR (General Data Protection Regulation) is an EU legislation that regulates the handling of personal data and it has been applied since 25th of May 2018. The goal of the legislation is to "strengthen individuals' fundamental rights in the digital age and facilitate business by clarifying rules for companies and public bodies in the digital single market". [3]

The GDPR applies to personal data which is defined in the Article 4 of the GDPR as "any information relating to an identified or identifiable natural person" and "an identifiable natural person is one who can be identified, directly or indirectly" [3]. This data includes pseudonymized data that could be linked to a person using additional information. This means that GDPR applies any time that identification is possible, even if a person is not currently identifiable. [21]

Transparency is a key point for the GDPR. Articles 12 requires that anyone who processes

personal data needs to inform the data subject that their data is being collected and they need to present the information in “a concise, transparent, intelligible, and easily accessible form, using clear and plain language”. Article 13 further clarifies what needs to be shared with the data subject which includes the data controller’s contact data, the purpose and legal basis for the data collection and the data subject’s rights regarding their personal data. This essentially means that every website needs to have a privacy policy that contains this information. [3], [22]

Data protection by design and by default: Article 25 states that should implement appropriate technical and organisational measures to implement data-protection principles such as data minimisation in an effective manner. Article 32 further clarifies these technical measures that are required including pseudonymisation and encryption of personal data. Also only personal data which are necessary for each specific purpose of the processing should be processed meaning that the data controllers have to be specific on what data they collect. [3]

Based on an EU court ruling from 2014, the right to erasure or the right to be forgotten was included in the GDPR article 17. This means that the data subject can ask the data controller to delete their personal data if one of the points mentioned in article 17 applies. Most notably, if the legal basis for the data collection was user consent, the data subject can withdraw their consent and the data controller has to delete the obtained data unless they have a different legal basis to retain the data. [3], [23]

The GDPR does explicitly mention cookies as one of the ways to identify people online: ”natural persons may be associated with online identifiers provided by their devices, applications, tools and protocols, such as internet protocol addresses, cookie identifiers or other identifiers”. This is because many cookies contain unique user identifiers which can be used to match personal information to the individual. The GDPR does not differentiate between first- and third-party cookies but treats both equally. [21]

The GDPR applies to all companies to process the personal data of EU citizens even if those companies exist and process the data outside of the EU as is stated by article 3(1): "This Regulation applies to the processing of personal data in the context of the activities of an establishment of a controller or a processor in the Union, regardless of whether the processing takes place in the Union or not"[3].

2.3.1 The history of privacy legislation in the EU

Before the GDPR the EU had two different directives. The Data Protection Directive that was passed in 1995 was the first EU-wide legislation that regulated the processing of personal data. It defined basic definitions that are still relevant today such as defining personal data as "any information relating to an identified or identifiable natural person". [23]

The Privacy and Electronic Communications Directive, also known as ePrivacy directive was passed in 2002 and it built upon the earlier directive especially when concerning internet data traffic. It introduced the concept of informed consent where the user would have to be informed of what data is being stored locally on their device, most notably cookies. This directive was further amended in 2009 to address technological advances and the informed consent was replaced by explicit consent meaning that the users would have to explicitly consent before any information could be stored on their devices. [23]

The biggest drawback of these directives is that they are not laws or regulations, which the GDPR is. Directives instead only set specific goals that the nation states are supposed to implement through their national laws. This means that the data protection laws could vary a lot depending on the member state where each EU citizen lives.

The GDPR is an EU-wide law. Since directives were enacted through national laws, the privacy standards varied throughout the EU. The enactment of the GDPR aimed to estab-

lish high privacy standards in all of the EU to regulate the processing of personal data. Recital 10 of the GDPR mentions that the level of protection should be equivalent in all member states [3]. However, each member state is still allowed to set more restrictive privacy laws if they so wish. The GDPR only defines a minimum level of data protection and privacy laws. [23]

Additionally the GDPR introduces fines as sanctions unlike the previous directives. Any company that fails to comply with the GDPR can face fines up to 20 million Euros or 4% the total worldwide annual turnover of the preceding financial year, whichever is higher [3]. Each member state is also allowed to set additional fines for infringements that are not covered by the GDPR as long as they are effective, proportionate and dissuasive. [23]

2.3.2 User Consent

A website needs to have a legal basis in order for data collection and processing to be lawful. Article 6 of the GDPR lists 6 different valid options for a legal basis: consent of the data subject, contractual obligation, legal obligation, vital interests of the data subject, public interest or the data controller's legitimate interest [3]. Of these options, asking for user consent is the most common option when using cookies although legitimate interest is used too.

The ePrivacy Directive that predated GDPR also required user consent for cookies with the exception of strictly necessary ones, but consent was not as clearly defined as it is with GDPR. Websites would often include cookie banners either containing just an OK button or a notice that the website uses cookies. GDPR has clear rules that consent cannot be implicit but instead "should be given by a clear affirmative act establishing a freely given, specific, informed and unambiguous indication of the data subject's agreement to the processing of personal data relating to him or her" [3]. Silence, pre-ticked boxes or

inactivity should not be counted as consent. [21]

Consent should be given before cookies are created. This means that websites cannot create nonessential cookies by default and then delete them if the user declines the use of cookies. Additionally, the GDPR states that websites cannot refuse serving users that decline nonessential cookies as "Consent should not be regarded as freely given if the data subject has no genuine or free choice or is unable to refuse or withdraw consent without detriment." [21]

Legitimate interest is another legal basis that the data controller might use instead of user consent. The data controller must prove that their interest in processing data outweighs the data subject's interest for privacy in order to establish legitimate interest as a legal basis for data collection. This can be hard to prove so asking for consent is a safer option as that leaves no ambiguity or room for different interpretations. [23]

An example of possible ambiguity is that recital 47 of the GDPR states that "The processing of personal data for direct marketing purposes may be regarded as carried out for a legitimate interest. However, recital 70 states that if personal data is used for the purpose of direct marketing, the data subject should have the right to object to such processing and "That right should be explicitly brought to the attention of the data subject". [3], [23]

Sometimes the data controller can ask for both consent and claim legitimate interest for the same purpose. The user can then have an option to decline consent and object to the legitimate interest as seen in Figure 2.1. This can feel deceptive as the user might not notice the legitimate interest and might think that declining consent is enough.

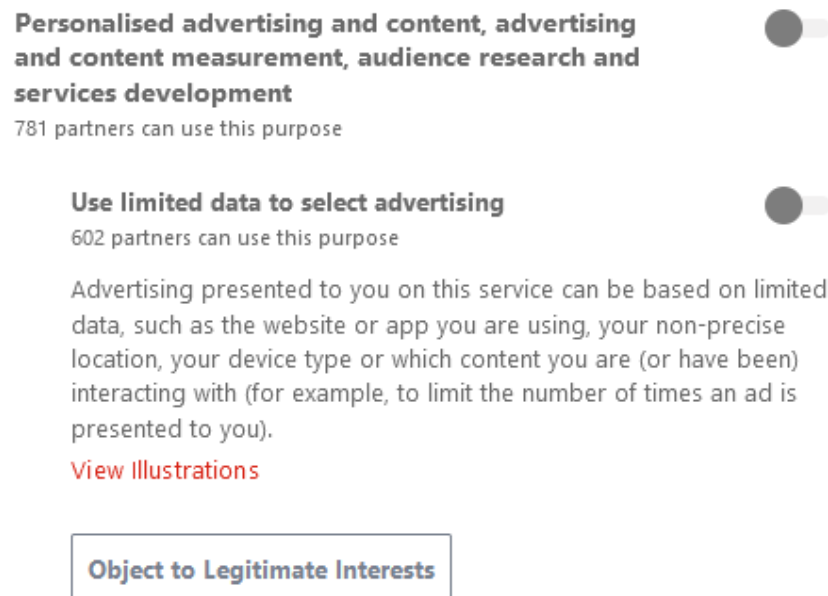


Figure 2.1: The user has an option to choose both consent and object to legitimate interest for the same use case.

2.4 Cookie Consent Interfaces

A cookie consent interface is an interface that user will see when they visit a website for the first time. While the GDPR requires that consent must be asked from users, it does not regulate how the consent interface should look like. As a result, organizations use many different designs in their implementations. The interface design has a great effect on the users' ability to understand what they are consenting to and it also influences their choices. [19]

The interface should be transparent to the users and it should be expressed in clear and understandable language. Recital 39 of the GDPR states that "It should be transparent to natural persons that personal data concerning them are collected, used, consulted or otherwise processed and to what extent the personal data are or will be processed" [3].

There are many different consent management platforms (CMP) that are widely used by

different websites to get a premade cookie consent interface for the website. This is easier than manually creating a new one and many of the available CMPs have options that are already GDPR compliant. However, some of the options that the CMPs have are intended to be used in other parts of the world that do not have as strict privacy laws so the website owner has to make sure that the CMP is configured properly.

2.4.1 Design Choices

Prominence of the interface: The interface can either be blocking or non-blocking and both have some advantages over the other. A blocking interface does not allow the user to use the website before they accept or set their cookie preferences. This is great for users who care about their privacy as they will always see the cookie consent notice but it can be annoying for users who do not put much value into the privacy aspect and only want to use the website. A non-blocking interface also allows the user to enter and use the website without ever reading or accepting the cookie consent notice. [19]

Presence of in-line options: There are two main ways that the cookie options can be accessed. One option is to have buttons for allowing all cookies and allowing only necessary cookies, with a third button for a more comprehensive cookie settings page if the user wants to enable some additional cookies. The other option is to have all the cookies categories in-line which makes setting cookie preferences faster. The downside is that there usually is not enough space in the cookie banner to include any definitions or explanations for the cookie categories so uninformed users might misunderstand the purposes of the categories as they would have to guess what they mean. [19]

Enabling decision reversal: Sometimes the user wants to change their cookie preferences [19]. While this can be done by manually deleting the cookies and refreshing the website, it can be hard for less technical users. This is not a big issue for most users, and therefore

websites generally do not have the option shown very prominently. A common way to do this is to have a cookie options button in one of the bottom corners of the page. That way the option exists for users who want it but it will not likely annoy users who do not need it.

2.4.2 Dark Patterns

There are several design choices that can be used to trick the user into accepting more cookies than they intend to. This type of misleading design is known as *dark patterns*. The following dark patterns are commonly used in cookie consent interfaces. [19]

Unequal paths: The interface has unequal interaction paths for the most and least privacy-protective options [19]. This means that for example accepting all cookies can be accomplished with a single button click, but accepting only necessary cookies requires navigating into a different menu. Choosing the privacy-protective option becomes a hassle for the user or the user might accidentally not set their cookie preferences correctly, especially if this is combined with confusing button layouts. An example can be seen in Figure 2.2.

Bad defaults: The interface has default options. For example, the user wants to decline non-necessary cookies, but advertising cookies are enabled by default and the user has to manually remove them as there is no default option for that. Having pre-ticked boxes like this is also in violation of the GDPR [24].

Confusing buttons: The interface has unintuitive placements of buttons for confirming users' cookie preferences and allowing all cookies [19]. For example after the user has selected their cookies, the interface has the "confirm my choices" and "allow all" button placements reversed so the user accidentally presses the allow all button. This can be seen in Figure 2.3. The colour and contrast used for the buttons can also be used to confuse users.

Data and Cookie Consent

In order to provide a more personal user experience, we and our 219 partners use technology such as cookies to store and/or access device information.

By clicking "Accept" you consent to these technologies which will allow us and [our partners](#) to process non-sensitive data such as IP address, unique ID, and browsing data for the purposes of serving personalized ads and content, measuring ads and content, gaining audience insights, gathering analytics on website activity, and developing and improving products.

Your choices on this site will be applied only for this site. You can change your settings at any time, including withdrawing your consent, by going to the [Privacy Policy](#) page of this site.



Figure 2.2: Accepting all cookies only requires one click but accepting less requires more clicks.

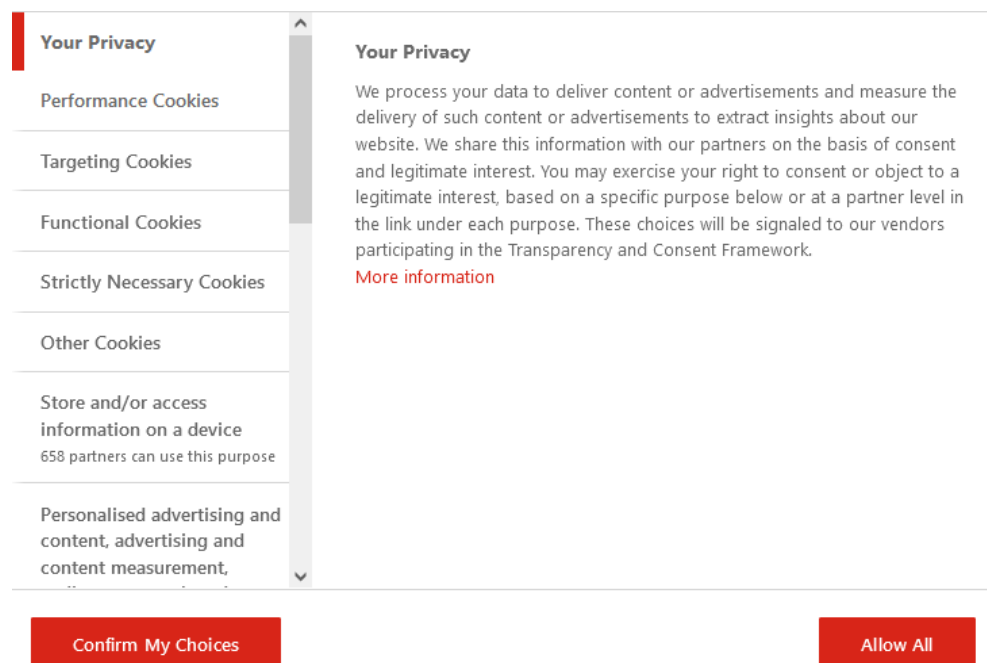


Figure 2.3: Allow All button is where the user would expect the Confirm My Choices button to be

There are also some other dark patterns like having no actual choices at all but that would be in violation of the GDPR. However, that used to be common before GDPR was implemented which shows that legislation is needed if actual changes are expected to happen.

3 Methods

Section 2 was done as a literature review. It had background information that will be important for the rest of this thesis. Privacy, cookies, GDPR and cookie consent interfaces are all important for this thesis.

The main section of this thesis is divided into two parts. First part is a study on the cookie notices that are used on the selected websites. The idea is to see if the cookie notices inform the user of the cookies that are going to be used on the website and if the user has control to choose their consent to different types of cookies. The websites are going to be presented anonymously as the intention is not to single out some specific websites.

Analyzing dark patterns is a part of this and the patterns that are considered in this thesis are mentioned in Section 2.4.2. While some of the dark patterns are not currently prohibited by the GDPR, others like pre-ticked boxes are prohibited [25].

The second part of the study is to actually look at the cookies that are used by the websites and see if it matches what the website claims to use in the cookie notice. There are some tools that are used to help with this.

Chrome developer tools shows all the cookies that are currently used by the website in the application tab. An example can be seen in Figure 3.1 with the domain removed for the first-party cookies to keep the website anonymous. The developer tools are also used to

clear the cookies from the browser so that the different options in the cookie notice can be explored.

Name	Val...	Domain	Path	Expires / ...	Size	HttpOnly	Secure	SameSite	Partition ...	Priority
AEC	AQ...	.google.com	/	2024-11-...	61	✓	✓	Lax		Mediu...
APISID	Gc...	.google.com	/	2025-06-...	40					High
HSID	AfI...	.google.com	/	2025-06-...	21	✓				High
NID	51...	.google.com	/	2024-11-...	234	✓	✓	None		Mediu...
SAPISID	-U...	.google.com	/	2025-06-...	41		✓			High
SEARCH_SAMESITE	Cg...	.google.com	/	2024-11-...	23			Strict		Mediu...
SID	g.a...	.google.com	/	2025-06-...	156					High
SIDCC	AK...	.google.com	/	2025-05-...	79					High
SOCS	CA...	.google.com	/	2025-06-...	18		✓	Lax		Mediu...
SSID	Af...	.google.com	/	2025-06-...	21	✓	✓			High
__Secure-1PAPISID	-U...	.google.com	/	2025-06-...	51		✓			High
__Secure-1PSID	g.a...	.google.com	/	2025-06-...	167	✓	✓			High
__Secure-1PSIDCC	AK...	.google.com	/	2025-05-...	91	✓	✓			High
__Secure-1PSIDTS	sid...	.google.com	/	2025-05-...	94	✓	✓			High
__Secure-3PAPISID	-U...	.google.com	/	2025-06-...	51		✓	None		High
__Secure-3PSID	g.a...	.google.com	/	2025-06-...	167	✓	✓	None		High
__Secure-3PSIDCC	AK...	.google.com	/	2025-05-...	90	✓	✓	None		High
__Secure-3PSIDTS	sid...	.google.com	/	2025-05-...	94	✓	✓	None		High
__Secure-ENID	19...	.google.com	/	2025-06-...	306	✓	✓	Lax		Mediu...
_ga	GA...	/	/	2025-06-...	29					Mediu...
_ga_DK1QNK31S	GS...	/	/	2025-06-...	53					Mediu...
_gat_gtag_UA_22137896_1	1	/	/	2024-05-...	24					Mediu...
_gid	GA...	/	/	2024-05-...	31					Mediu...

Figure 3.1: Example cookies that are used by one of the websites.

Name	Scope	Domain	Partition Key	SameSite	Category	Platform	HttpOnly	Secure	Value	Path	Expires / Max-Age
SEARCH_SAMESITE	Third Party	.google.fi		Strict	Functional	Google			CgQI-5o	/	2024-10-19T19:34:1
AEC	Third Party	.google.fi		Lax	Functional	Google Ads		✓	AQTF6H	/	2024-10-19T19:34:1
__Secure-ENID	Third Party	.google.fi		Lax	Uncategorized	Unknown		✓	19.SE=Q	/	2025-05-23T11:52:0
SID	Third Party	.google.fi			Marketing	Google			g.a000jQ	/	2025-06-11T10:07:0
__Secure-1PSID	Third Party	.google.fi			Marketing	Google Ads		✓	g.a000jQ	/	2025-06-11T10:07:0
HSID	Third Party	.google.fi			Marketing	Google			A9aAgJlf	/	2025-06-11T10:07:0
SSID	Third Party	.google.fi			Marketing	Google		✓	AFRI41XI	/	2025-06-11T10:07:0
APISID	Third Party	.google.fi			Marketing	Google			GcztNMz	/	2025-06-11T10:07:0
SAPISID	Third Party	.google.fi			Marketing	Google		✓	-Uw6Kya	/	2025-06-11T10:07:0
__Secure-1PAPISID	Third Party	.google.fi			Marketing	Google Ads		✓	-Uw6Kya	/	2025-06-11T10:07:0
NID	Third Party	.google.fi		None	Marketing	Google	✓	✓	513=Cd2	/	2024-10-22T18:28:5
__Secure-3PSID	Third Party	.google.fi		None	Marketing	Google Ads	✓	✓	g.a000jQ	/	2025-06-11T10:07:0
__Secure-3PAPISID	Third Party	.google.fi		None	Marketing	Google Ads		✓	-Uw6Kya	/	2025-06-11T10:07:0
_gid	First Party				Analytics	Google Analytics			GA1.2.12	/	2024-05-25T04:17:0
_ga	First Party				Analytics	Google Analytics			GA1.1.17	/	2025-06-28T04:17:0
_ga_DK1QNK31S	First Party				Analytics	Google Analytics			GS1.1.17	/	2025-06-28T04:17:0
COMPASS	Third Party	.docs.google.com		None	Uncategorized	Unknown		✓	apps-spr /spreadsh		2024-05-24T04:54:0
__utma	Third Party	.google.com			Analytics	Google Analytics			1732723 /get/vid		2025-06-10T18:08:0

Figure 3.2: Privacy Sandbox Analysis Tool example.

The Privacy Sandbox Analysis Tool is an extension to the Chrome developer tools that has some additional features [26]. The cookies from the same website as the previous example can be seen in Figure 3.2. This extension shows the category of the cookie whether it is a functional, analytics or a marketing cookie and it shows which platform the cookie belongs to, for example Google.

This extension also shows a lot more cookies than the regular DevTools. This is because it also shows cookies that are being blocked by the browser in every request. These are shown with a yellow background. These are not shown in the regular developer tools if the cookie is blocked in every request. It will be shown if it is only sometimes blocked.

4 Results

There are 309 municipalities in Finland. Of these 108 call themselves a city but since 1977 there has not been any difference in legal rights or obligations between cities and regular municipalities so the difference is mainly in population and size. Each municipality can decide to call itself a city if they feel like they have met the requirements [27]. There are no official requirements and the decision is made by the municipality. [28]

Since the data collection for this thesis is done manually, 309 different websites are too much to take into the data set. In total, 25 different municipalities were chosen. This was done by taking the 10 biggest cities and then filling in smaller cities and municipalities from other regions of Finland. The full list is included in Table 4.1 along with their respective websites.

4.1 Cookie Notices

The cookie notice is the main way how the website communicates the use of cookies to the user. It is important that it is easily understandable and that the user knows what they are consenting to. 11 of the websites have an English-language cookie notice in addition to a Finnish one, including some that do have an English version of the rest of the website.

Municipality	Website
Helsinki	www.hel.fi
Espoo	www.espoo.fi
Tampere	www.tampere.fi
Vantaa	www.vantaa.fi
Oulu	www.ouka.fi
Turku	www.turku.fi
Jyväskylä	www.jyvaskyla.fi
Kuopio	www.kuopio.fi
Lahti	www.lahti.fi
Pori	www.pori.fi
Joensuu	www.joensuu.fi
Lappeenranta	www.lappeenranta.fi
Vaasa	www.vaasa.fi
Rovaniemi	www.rovaniemi.fi
Lieksa	www.lieksa.fi
Forssa	www.forssa.fi
Vihti	www.vihti.fi
Ilmajoki	www.ilmajoki.fi
Juva	www.juva.fi
Sotkamo	www.sotkamo.fi
Kaustinen	www.kaustinen.fi
Pyhtää	www.pyhtaa.fi
Mynämäki	www.mynamaki.fi
Lempäälä	www.lempaala.fi
Ii	www.ii.fi

Table 4.1: Municipalities that were chosen for the data set.

4.1.1 Good Cookie Notices

A majority of the cookie notices were banners at the bottom of the screen. 16 websites used a banner, with 15 of them located at the bottom of the screen and 1 at the top of the screen. The next most common type was a pop-up window in the middle of the screen.

The most common type looks like the banner in Figure 4.1. There are some variations of similar type but they have 3 buttons: one for accepting all cookies, another to accept only necessary cookies and an additional and usually smaller button to adjust consent for more specific cookie types. 10 of the 16 banners used this three button layout.



Figure 4.1: The most common cookie banner.

The second most common banner is the type where all the specific consent selections are right on the banner instead of requiring an additional click to access. This can be seen in Figure 4.2. There are two different variants with similar design one with 3 buttons and another with 2 buttons.

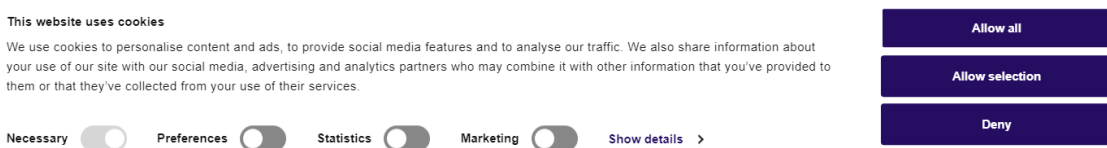


Figure 4.2: Options included in banner, 3 buttons

The 2-button variant can be seen in Figure 4.3. The 2-button variant has the accept all and only necessary buttons, and when the user selects one or more of the possible options, the only necessary cookies button will change to something like "save settings" so that the user can understand that the functionality of the button has changed. This 2-button variant can be more easily understood by the user as there are less buttons to press, especially if

colour contrast is being used to trick the user into accepting all cookies. In total, 2 of the five are 3-button variants and 3 of them are 2-button variants.

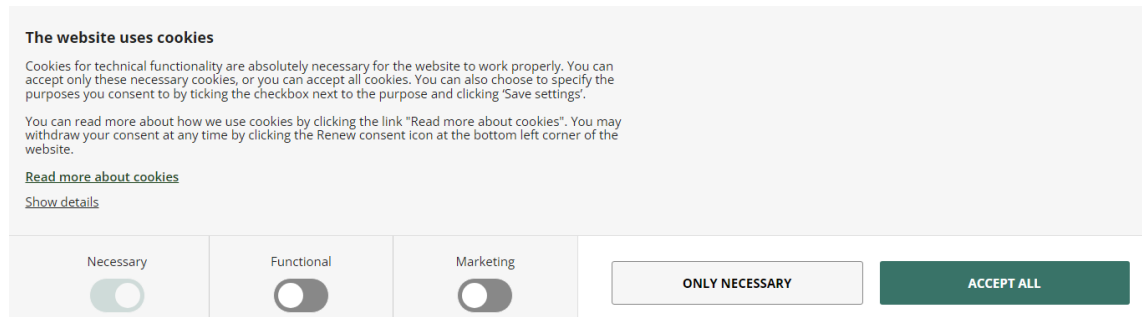


Figure 4.3: Options included in banner, 2 buttons

The second type of cookie notice aside from banners is a pop-up in the middle of the screen. There are 5 websites with pop-up notices, 2 of them are blocking and 3 are non-blocking. They are visually similar and an example can be found in Figure 4.4. Blocking in this case means that the user cannot navigate the website at all without interacting with the cookie pop-up. Although the non-blocking variants allow navigation through the website they are very obstructive compared to the banners which can be easily ignored therefore even the non-blocking version.

4.1.2 Dark Patterns and Non-GDPR Compliant Cookie Notices

Even some of the cookie notices that were discussed earlier used some mild dark patterns to make the user more likely to accept all cookies. In total, only 8 of the 25 websites do not have anything that could be seen as a dark pattern or otherwise deceptive design. However, only the most relevant dark patterns that are mentioned in Section 2.4.2 are considered in this thesis. The most common was the use of colour to highlight the "accept all" button which can be seen in Figures 4.4 and 4.3.

A few of the websites used language such as "decline all cookies" or "deny cookies" instead of "only necessary cookies" even though those options would enable the necessary

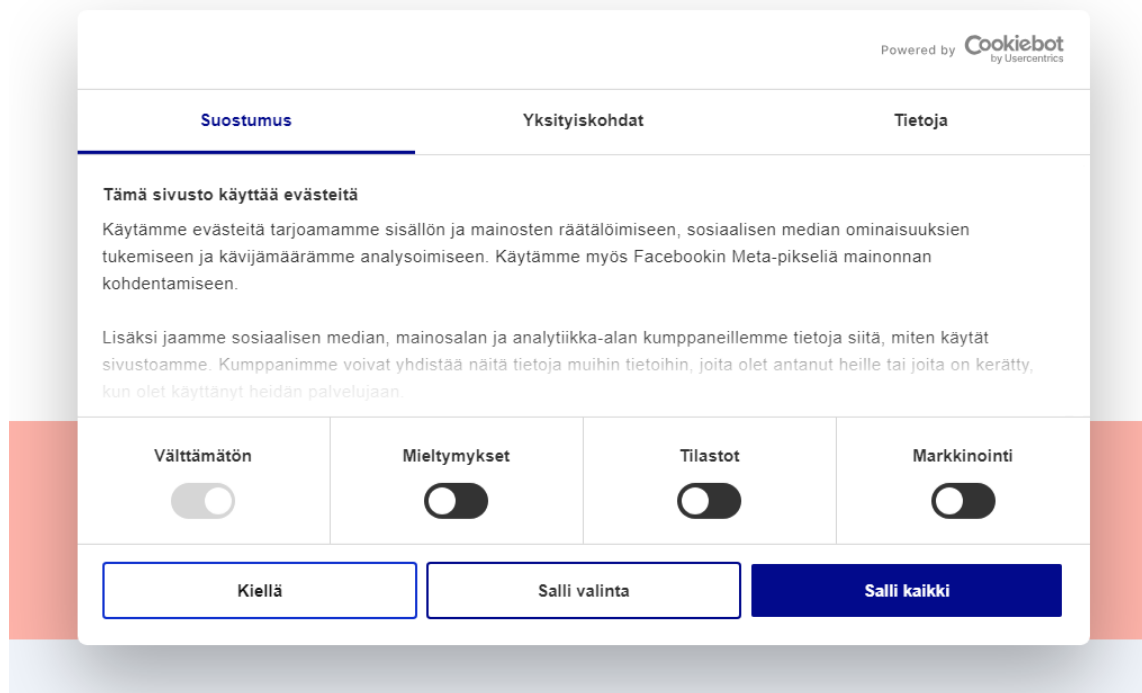


Figure 4.4: Cookie notice pop-up

cookies. This type of language is slightly misleading and it can be hard for the user to understand if there is a difference between choosing to decline all cookies or manually declining everything except the necessary cookies from the additional options.

One of the websites had pre-ticked boxes which is in violation of the GDPR and it can be seen in Figure 4.5 [25]. However, the options were not pre-ticked with every browser. For example, using Firefox in the private browsing mode would cause the options to be unticked by default where they would be ticked without private browsing or using another browser like Microsoft Edge or Google Chrome. Chrome's incognito mode did not have an effect on the options. In any case, the default option was to have them pre-ticked in most cases which makes it not okay by GDPR standards.

Next we will be looking at the websites that have a cookie notice with no options. This can be seen in Figure 4.6. This means that the only "options" available to the user are to either click the "ok" button or leave the website and leaving the website is not considered to be

Tämä sivusto käyttää evästeitä

Käytämme evästeitä tarjoamamme sisällön ja mainosten räätälöimiseen, sosiaalisen median ominaisuuksien tukemiseen ja kävijämäärämme analysoimiseen. Lisäksi jaamme sosiaalisen median, mainosalan ja analytiikka-alan kumppaneillemme tietoja siitä, miten käytät sivustoamme. Kumppanimme voivat yhdistää näitä tietoja muihin tietoihin, joita olet antanut heille tai joita on kerätty, kun olet käyttänyt heidän palvelujaan. Hyväksymällä evästeet sallit tietojen siirron EU/ETA alueen ulkopuolelle.

Vain välttämättömät evästeet	Salli valinta	Salli kaikki evästeet		
<input checked="" type="checkbox"/> Välttämätön	<input checked="" type="checkbox"/> Mieltymykset	<input checked="" type="checkbox"/> Tilastot	<input checked="" type="checkbox"/> Markkinointi	Näytä tiedot ▾

Figure 4.5: A cookie notice with pre-ticked boxes.

a valid choice for declining consent under the GDPR. Additionally, it is likely that the website has already created cookies for the user even before they click the button, which makes the choice meaningless.

One of these websites does include a link to a privacy policy page where they have a more detailed cookie policy. However, they only state that they use Google Analytics on the website and give a very vague message that the cookies can be "removed from the browser settings" which is obviously not enough under GDPR regulation. This is because the GDPR requires that consent is given as a clear affirmative action. However, a statement like this used to be common before the GDPR as that would have been enough.

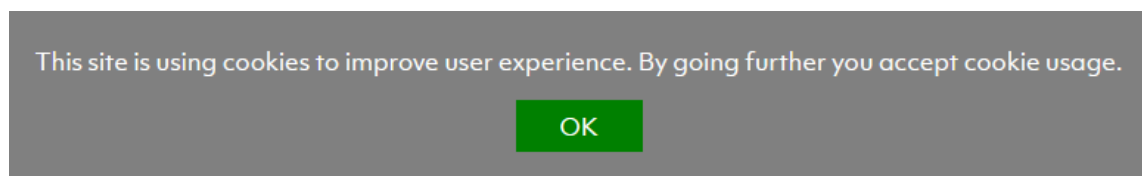


Figure 4.6: A cookie notice with no options.

The last two websites that were analyzed did not have any cookie notices and also had no mention for the use of cookies on the website at all. This is also a violation of the GDPR, assuming that the websites do use some cookies. The user visiting these websites

might assume that they do not use cookies at all even though the websites might use them without disclosing the use.

4.2 Cookies Used on the Websites

This section takes a look at the cookies that each website sets and then compares the set cookies to what would be expected by the cookie notice in the previous section. This means that the websites with more detailed cookie notices are likely to require less analysis compared to the ones with weaker cookie notices. This analysis will be done anonymously and the websites will be referred to with just a number.

We will start the analysis from the last websites mentioned in the previous section that did not have a cookie notice. Website 1 uses a list of cookies that can be found in Figure 4.7. All of the cookies on this list are first-party cookies with the domain set to the website of the municipality. The p11_language cookie is a cookie that is used to store the language of the browser and the font-size cookie determines the size of the text so they could be categorized as either functional or necessary cookies. The font-size cookie can have a value of "font-small" or "font-large" or no value for medium sized text.

Name	Value
_ga	GA1.1.401293854.171380...
_ga_BZW50FZ6Z5	GS1.1.1713809257.1.1.17...
font-size	font-small
p11_language	fi

Figure 4.7: List of cookies used by one of the websites

Both of the cookies starting with _ga are used by Google Analytics for statistics purposes. They can be used to store and count pageviews. However, the browser is assigned a unique ID when the user first visits the website. This can then be used to track the users future visits to the website. Google also handles the data storing and processing so Google also

has access to the data collected with Google Analytics.

Website 1 included a lot of third party cookies. For example if the website had an embedded Youtube video, it would load a lot of 3rd party cookies from youtube.com. An example can be found in Figure 4.8. The same would happen with cookies from Facebook. Some of these would be used for example authentication purposes but some are used as marketing cookies.

Name	Value	Domain	Path	Expires /...
VISITOR_PRIVACY_METADA...	CgJG...	.youtube.com	/	2025-05...
VISITOR_PRIVACY_METADA...	CgJG...	.youtube.com	/	2024-10...
YSC	n6IK...	.youtube.com	/	Session
__Secure-3PSID	g.a0...	.youtube.com	/	2025-05...
__Secure-3PSIDTS	sidts...	.youtube.com	/	2025-04...
__Secure-3PAPISID	cKR2...	.youtube.com	/	2025-05...
__Secure-3PSIDCC	AKEy...	.youtube.com	/	2025-04...
LOGIN_INFO	AFm...	.youtube.com	/	2025-05...
VISITOR_INFO1_LIVE	CNz...	.youtube.com	/	2024-10...
SID	g.a0...	.youtube.com	/	2025-05...
__Secure-1PSID	g.a0...	.youtube.com	/	2025-05...
HSID	AEO...	.youtube.com	/	2025-05...
SSID	AcA...	.youtube.com	/	2025-05...
APISID	_WBi...	.youtube.com	/	2025-05...
SAPISID	cKR2...	.youtube.com	/	2025-05...
__Secure-1PAPISID	cKR2...	.youtube.com	/	2025-05...
PREF	f7=4...	.youtube.com	/	2025-05...
__Secure-YEC	Cgtf...	.youtube.com	/	2025-05...
__Secure-1PSIDTS	sidts...	.youtube.com	/	2025-04...
SIDCC	AKEy...	.youtube.com	/	2025-04...
__Secure-1PSIDCC	AKEy...	.youtube.com	/	2025-04...

Figure 4.8: 3rd party cookies from Youtube, yellow blocked by Chrome

In addition to the list found in Figure 4.7, the website also uses third party cookies set by Google. These cookies are directly imported from google.com so the cookies depend on if the user has visited Google or if they are logged in to Google services. An example list of cookies can be seen in Figure 4.9. Some of the other websites would also use cookies from Google's subdomain such as accounts.google.com but these would often get automatically blocked by Chrome.

There are two cases that cause the 3rd party cookie to get blocked. The cookie gets blocked

Name	Value	Domain	Path	Expires /...	Size	HttpOnly	Secure	SameSite
AEC	AQTF...	.google.com	/	2024-10...	62	✓	✓	Lax
APISID	_WBib...	.google.com	/	2025-05...	40			
HSID	AmPc...	.google.com	/	2025-05...	21	✓		
NID	513=...	.google.com	/	2024-10...	243	✓	✓	None
SAPISID	cKR2r...	.google.com	/	2025-05...	41		✓	
SEARCH_SAMESITE	CgQl-...	.google.com	/	2024-10...	23			Strict
SID	g.a00...	.google.com	/	2025-05...	156			
SIDCC	AKEy...	.google.com	/	2025-04...	77			
SSID	Ad5lu...	.google.com	/	2025-05...	21	✓	✓	
__Secure-1PAPISID	cKR2r...	.google.com	/	2025-05...	51		✓	
__Secure-1PSID	g.a00...	.google.com	/	2025-05...	167	✓	✓	
__Secure-1PSIDCC	AKEy...	.google.com	/	2025-04...	88	✓	✓	
__Secure-3PAPISID	cKR2r...	.google.com	/	2025-05...	51		✓	None
__Secure-3PSID	g.a00...	.google.com	/	2025-05...	167	✓	✓	None
__Secure-3PSIDCC	AKEy...	.google.com	/	2025-04...	90	✓	✓	None
__Secure-ENID	19.SE...	.google.com	/	2025-05...	285	✓	✓	Lax

Figure 4.9: Google 3rd party cookies used by multiple websites.

if its SameSite attribute is set to Lax, the request is made from a different site and it was not initiated by a top-level navigation, such as clicking a link that navigates to a new URL. This causes some of the cookies to become blocked for example if they would be set only after the user consents to the use of cookies through the cookie notice. [29]

This mainly happens because the SameSite attribute gets defaulted to Lax if it is not set to anything. However, the feature did not always work this way. Before February 2020 the SameSite attribute would be set to none by default and the 3rd party cookies would work always. [29] Another case is if the request's URL domain does not match the cookie's domain exactly. Majority of the google subdomain cookies get blocked for this reason.

Website 2 is another site without a cookie notice and it is very similar to website 1 in terms of cookies used. It also uses Google Analytics and other 3rd party cookies from Google. As for other first party cookies there are cookies to change the contrast and to make the text bigger which both have boolean values.

Both of these websites are clearly in violation of the GDPR because they use optional cookies without first obtaining user consent. The websites need to create some kind of cookie notice to inform the user of the essential cookies and then either ask consent for

the analytics and marketing cookies or remove them completely. [12]

Many of the websites included in this thesis used the 3rd party cookies from Google. In total 9 of the 25 websites used Google 3rd party cookies. 3 of these were websites with either no cookie notice or no options. The remaining 6 did have proper options set in the cookie notice but only 3 had the Google 3rd party cookies tied to the marketing cookies setting. This means that 3 of them had activated the 3rd party cookies when entering the website, even before the user interacted with the cookie notice at all.

We will look at the two websites with a "no options" cookie banner next. Website 3 mentions on their web page that they are using Google Analytics. However, it does not look like Google Analytics is active even though it is the only thing that is explicitly mentioned on the website. Other 3rd party cookies from Google are active though. The only first party cookie is a PHP session id cookie which is an essential cookie.

Website 4 does not have any cookies from Google. It only uses first party cookies which can be seen in Figure 4.10. The `_pk_id` and `_pk_ses` are analytics cookies from Matomo, which is an alternative to Google Analytics. The ID cookie lasts for 13 months and it stores a unique user ID for the browser and the session cookie lasts for 30 minutes.



Name
COOKIE_SUPPORT
GUEST_LANGUAGE_ID
JSESSIONID
LFR_SESSION_STATE_89397
_pk_id.101.703b
_pk_ses.101.703b
cookieInfo
oc-6b76408e-a3d5-4dab-9...
oc-6b76408e-a3d5-4dab-9...

Figure 4.10: First party cookies.

Clicking the "ok" button on the optionless cookie banner creates the cookieInfo cookie.

However, it does not look like the cookie has any purpose other than telling the website that the button has been clicked because all the other cookies on the website are loaded by default. The same happens with the previous website except that there is no additional cookie created. This means that the user's action to accept the cookies is completely meaningless because the website uses the same cookies no matter what.

All of the remaining websites have a cookie notice so the effects of the options will be analyzed too. Websites 5, 6, and 7 all have a proper looking cookie notice, but they all have Google's third-party cookies active even if the user does not interact with the cookie notice and therefore has not accepted the use of cookies. An example from website 5 can be seen in Figure 4.11. Interestingly, on website 5 the Google third-party cookies are only active on the homepage of the website. If the path of the URL changes for example by navigating the page through a link, the cookies are no longer active. For all the other websites they were active with any path.

Name	Value
AEC	AQT...
APISID	_WB...
HSID	AmP...
NID	513...
SAPISID	cKR...
SEARCH_SAMESITE	CgQ...
SID	g.a0...
SIDCC	AKE...
SSID	Ad5l...
__Secure-1PAPISID	cKR...
__Secure-1PSID	g.a0...
__Secure-1PSIDCC	AKE...
__Secure-1PSIDTS	sidts...
__Secure-3PAPISID	cKR...
__Secure-3PSID	g.a0...
__Secure-3PSIDCC	AKE...
__Secure-3PSIDTS	sidts...
__Secure-ENID	19.S...
cookiebot-consent--marketing	0
cookiebot-consent--necessary	1
cookiebot-consent--preferences	0
cookiebot-consent--statistics	0

Figure 4.11: Google third party cookies active with only necessary cookies.

The cookie notice for website 5 has options to enable analytics and marketing cookies. However, these do not have any effects on this website as they only use cookies from Google and Youtube. The cookies from Google are active by default and changing the options does not change that. However, the Google Analytics cookies are set as third-party cookies with the domain as `developers.google.com` and that causes them to be automatically blocked. If the website has an embedded Youtube video on the current page it will set third-party marketing cookies from both Youtube and Google regardless of the settings that the user has selected.

Website 6 is very similar to website 5. The options do exist in the cookie notice but they do not do much since the third party marketing cookies from Google are enabled by default. Website 7 claims that it only uses essential cookies but it has the third-party cookies from Google active. All three are obviously not following the GDPR. The options on the cookie notices look fine but they are not configured properly.

The remaining 3 websites, which will be referred as websites 8,9 and 10, that use Google's third party marketing cookies do not have them all turned on by default. Let us look at website 8, for example. When the user first enters the website, it does not have any active cookies. Accepting necessary cookies only creates a cookie that stores the information that the cookie notice was accepted or declined.

Accepting only Analytics or only marketing cookies are interesting cases on website 8. The list with only analytics cookies turned on can be seen in Figure 4.12 and the list with only marketing cookies can be seen in Figure 4.13. There is a lot of overlap between the lists. Unique to the analytics list are one Google Analytics cookie and the `nmstat` cookie, which is an analytics cookie by Siteimprove. However, the analytics list also contains multiple cookies that are used for the purposes of marketing that are not in the marketing list.

The marketing list does contain a few additional marketing cookies named DSID, IDE and ar_debug. The DSID cookie is used by Google to identify signed users on non-Google sites and the IDE cookies are used to personalize the ads [30]. These cookies have doubleclick.net as the domain. Doubleclick is an advertisement company that was acquired by Google and it has been merged into the Google Marketing Platform so these cookies are also essentially by Google [31]. The marketing cookies list does also include one of the Google Analytics cookies which is used for analytics purposes.

Name
AEC
APISID
HSID
NID
SAPISID
SEARCH_SAMESITE
SID
SIDCC
SSID
__Secure-1PAPISID
__Secure-1PSID
__Secure-1PSIDCC
__Secure-1PSIDTS
__Secure-3PAPISID
__Secure-3PSID
__Secure-3PSIDCC
__Secure-3PSIDTS
__Secure-ENID
NID
._ga
._ga_M5YJ6NNBZE
cookiehub
nmstat

Figure 4.12: Website 8 with only analytics cookies turned on.

Website 9 has options for functional, statistics and marketing cookies. It does have more detailed explanations for the sections and the marketing section has cookies from Google, Youtube and Meta. However, accepting either the statistics or marketing cookies on their own does nothing. But if both of them are accepted at the same time they will both be

Name
DSID
IDE
ar_debug
AEC
APISID
HSID
NID
SAPISID
SEARCH_SAMESITE
SID
SSID
__Secure-1PAPISID
__Secure-1PSID
__Secure-3PAPISID
__Secure-3PSID
__Secure-ENID
_ga_M5YJ6NNBZE
cookiehub

Figure 4.13: Website 8 with only marketing cookies turned on.

active. The cookies used are the same as the ones for website 8 in Figure 4.12.

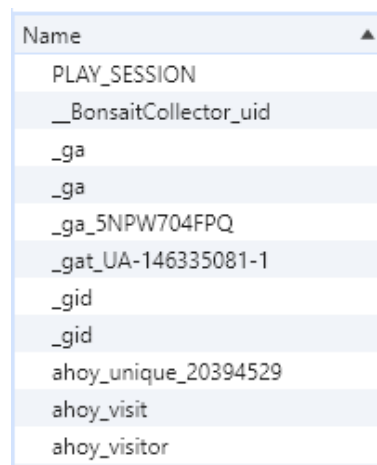
The functional cookies for website 9 include cookies from Giosg. They are used for the website's chat-box functions. It also includes an id cookie that allows the website to recognize the user on repeat visits.

Website 10 contains analytics cookies from Matomo, but they are only active when also the marketing cookies are accepted. The other way is not true like on website 9. If the marketing cookies are accepted, they will work on their own.

On some parts of the website marketing cookies from Google are active even if the marketing cookies are turned off. They are not active on the main homepage and on some subpages but they are active on some other subpages. There does not seem to be a pattern on which pages they are active and this is likely some kind of configuration issue rather than the website using the cookies maliciously.

Website 11 has analytics cookies active right away even though it does have options to decline analytics or marketing cookies in the cookie notice. However, it does not use Google's marketing cookies unlike the earlier websites. Marketing cookies from Youtube are active though, and they are active whenever the webpage contains a video from Youtube.

The website also contains 3rd party cookies from Powr, which is a website that sells plugins that make creating websites easier. But because website 11 contains third party cookies from Powr the user's visit will be visible to Powr. The cookies used on website 11 can be found in Figure 4.14. The list has the Google analytics cookies doubled as one of them is a first party cookie belonging to the website and another is a third party cookie from Powr. This list is if the user has not visited the Powr website. There would be more third party cookies from Powr if the user had previously visited that website.



Name ▲
PLAY_SESSION
__BonsaitCollector_uid
_ga
_ga
_ga_5NPW704FPQ
_gat_UA-146335081-1
_gid
_gid
ahoy_unique_20394529
ahoy_visit
ahoy_visitor

Figure 4.14: Google Analytics cookies both from the website and a third party.

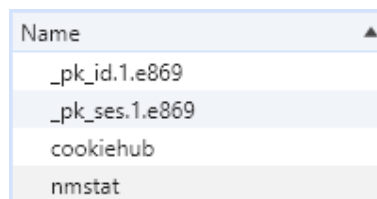
The options in the cookie notice do not do anything as the cookies are already active by default. Accepting the cookies does not add any additional cookies either. The options only give an illusion of choice to the user when in reality the only choice is to accept all cookies.

Website 12 has only one functional cookie based on the cookie notice called LastVisited-

PublicPage which stores the last page that the user has visited on the website. However, this cookie is treated as an essential cookie so it is always active regardless of the user's cookie settings.

The website uses analytic cookies from 3 different companies, Google, Hotjar and Matomo. The cookies from Google Analytics and Hotjar are activated by turning on the statistics cookies as expected. But the cookies by Matomo are only activated only if both the statistics and the marketing cookies are accepted even though they are only listed under the statistics cookie option.

Website 13 Only has necessary cookies and "other" cookies. The other cookies section only contains one uncategorized cookie called `in-session`. The necessary cookies include analytics cookies as can be seen in Figure 4.15. This is definitely unusual considering that the cookies are used for analytics but they are explicitly listed as essential in the cookie notice.



Name
<code>_pk_id.1.e869</code>
<code>_pk_ses.1.e869</code>
<code>cookiehub</code>
<code>nmstat</code>

Figure 4.15: Analytics cookies active by default.

Website 14 has a lot of information in the cookie notice, much more than the earlier websites. It shows the description, purposes and the data collected for each service. Figure 4.16 has an example for Matomo. It shows that the purpose is for analytics and optimization and that it does collect various types of data including the IP address. This is great for the user because this type of notice shows very clearly what data is being collected and there is no possible ambiguity.

This website does not have an option for analytic cookies. Instead, the analytic cookies are bundled into the functional cookies. This is better than having them set as essential as

Matomo (self hosted)

Functional ⓧ ^

Description of Service
This is an self hosted web analytics platform.

Data Purposes
This list represents the purposes of the data collection and processing.

Analytics Optimization

Technologies Used
This list represents all technologies this service uses to collect data.

Cookies Device fingerprinting

Data Collected
This list represents all (personal) data that is collected by or through the use of this service.

Browser language Browser type Device operating system Device type

Geographic location IP address Number of visits Referrer URL

Screen resolution Usage data Visited sub-pages

Retention Period
The retention period is the time span the collected data is saved for the processing purposes. The data needs to be deleted as soon as it is no longer needed for the stated processing purposes.

- The data will be deleted as soon as they are no longer needed for the processing purposes.

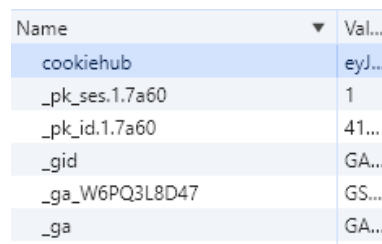
Figure 4.16: Data collected using Matomo on Website 14.

the user does have an option to opt out, but labeling analytics cookies as functional can be misleading.

In terms of this study, the remaining websites are much less interesting than the previous ones since they are handling privacy issues properly. The remaining websites do not have any issues regarding the cookies that they use and the cookies that they claim to use. Some of them still use either analytics cookies or some 3rd party marketing cookies but they are locked under the right option in the cookie consent notice.

Websites 15, 16, 17, 18, 19, 20 and 21 all have a feature that allows the user to easily change their cookie settings after the initial setup by having a persistent icon in the bottom-left of the screen that can be clicked. This makes it easier for less technical users to change their cookie preferences if they want to. Clicking the icon will bring up the same cookie notice that the user gets when they visited the website for the first time.

Websites 22, 23, 24 and 25 are similar but they do not have the option to change cookies easily. The only cookies that are set when the user chooses only necessary cookies are a cookie to remember the user's consent settings and possibly a cookie to store the user's language option on the website. Website 22 is using Matomo for analytics and Website 23 is using both Google Analytics and Matomo at the same time as can be seen in Figure 4.17.



Name	Val...
cookiehub	eyJ...
_pk_ses.1.7a60	1
_pk_id.1.7a60	41...
_gid	GA...
_ga_W6PQ3L8D47	GS...
_ga	GA...

Figure 4.17: Both Google Analytics and Matomo cookies used on the same website.

The total amount of times that each 3rd party analytics or marketing platform was used by the websites studied in this thesis can be seen in Figure 4.18. In total, only 5 of the 25

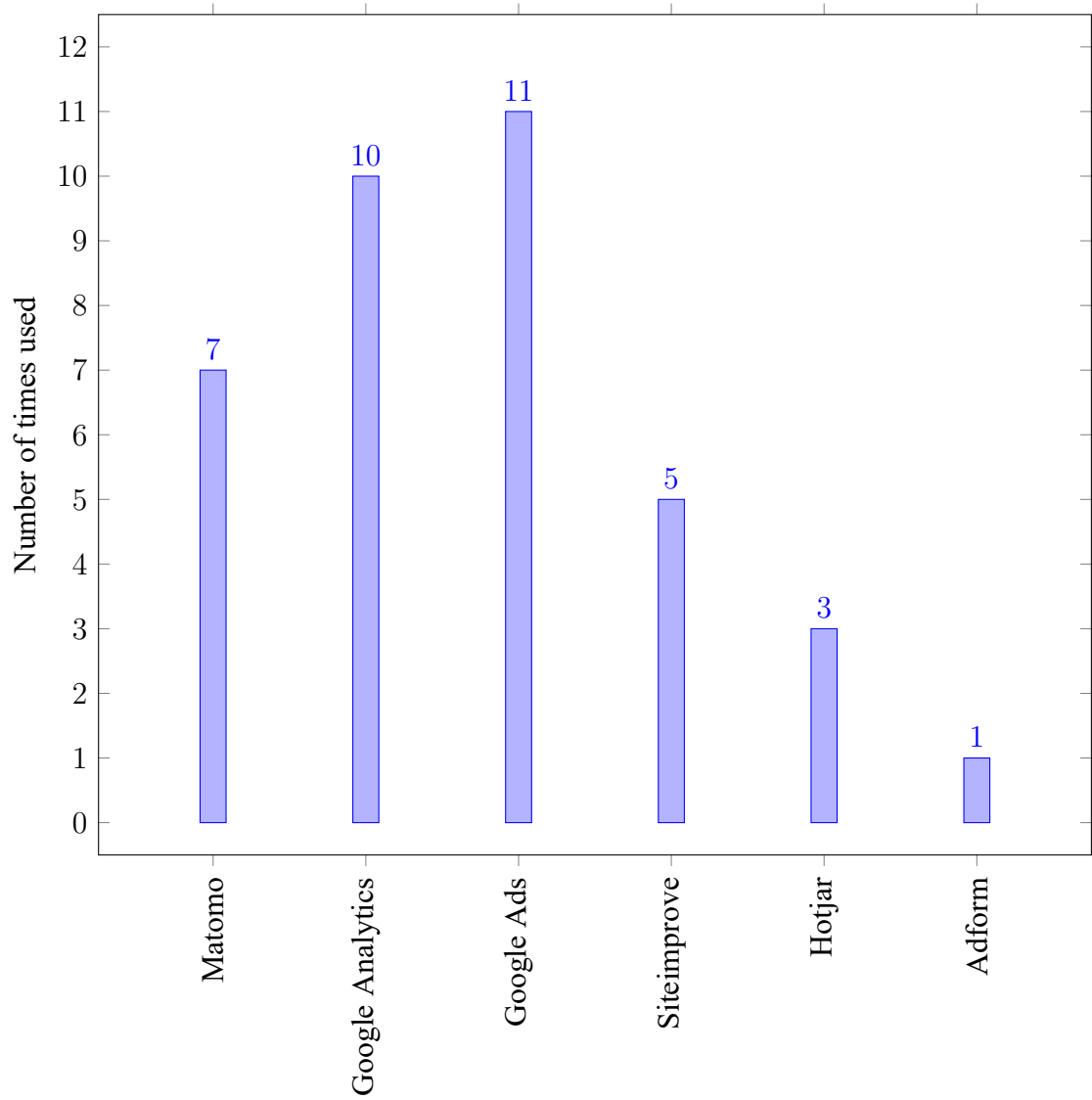


Figure 4.18: 3rd party platforms used for Analytic and Marketing cookies on the websites.

websites did not use any cookies from a 3rd party platform. 16 websites used at least one of the analytics options with some websites using multiple at the same time. Siteimprove was mostly used together with either Matomo or Google analytics with only one website deciding to exclusively use Siteimprove. All 3 websites that used Hotjar also used Google Analytics but this is typical for Hotjar as we discussed in Section 2.1.2.

Google was the most popular choice for marketing cookies with only one one website opting to use Adform instead. Many of the websites that used marketing cookies also used analytics cookies.

5 Discussion

It is evident from the results that there are problems on the websites. There are two main issues. The first issue is that not every website has a proper cookie consent notice, even though they use cookies 5 of the 25 websites included in this thesis do not have a proper cookie consent notice.

The second issue is that even if the website has a cookie notice it might use cookies without asking the user's consent and without informing the user. Especially the use of cookies by Google, both marketing and analytics, is very common. In total 10 out of the 25 websites had this issue of miscategorizing cookies.

Because of these issues, the people who use these websites might be giving their personal data to third parties without their consent. The municipal websites can be used for example to search information about health related problems and those searches can contain very personal information. In any case, the choice of consent should definitely belong to the user and ideally the user should be able to trust the websites that they are not using any cookies without the consent of the user.

A few of the websites did not have a cookie notice and some had a very old notice without any options that look like they were created before the GDPR. There clearly has not been enough pressure to force the municipalities to change the cookie notices even after many years. These older cookie notices were found only on the smaller municipalities, not on

any of the bigger cities. The most likely reason for this is that the websites for the bigger cities have more users compared to the smaller ones and therefore they are more likely to be notified of any problems. The bigger cities also have more available resources to fix potential issues that they are notified of.

The old cookie notices are most likely not malicious but they are either a lack of knowledge, a lack of technical skills to create new ones or it could just be a mistake that has been forgotten about. It is possible that the maintainers of the website are not even aware of the problem as the problem might not be obvious. It is very easy for the website maintainers to implementing something like Google Analytics even if they do not fully understand it. Using these cookies from a 3rd party is easy and it is possible that the websites that decide to use them might not understand the privacy implications for their users when these 3rd party cookies are used.

Figure 4.16 shows a very detailed description of what the user will be consenting to if they consent to the use of cookies. Creating a detailed description like this is great for the user, but it can also be very helpful for the website itself. If the developers of the website put a lot of effort into understanding the technology and the cookies used on the website, they can categorize the cookies better. There should also be less old cookies that are kept accidentally on the website even though they have been removed from the cookie consent notice and should have also been removed from the website.

The developers and the maintainers of the websites should be aware of the cookies that they are using on the website. They should also be doing internal reviews more often. It has been six years since the GDPR was first applied and that should have been long enough for each website to have a proper cookie consent notice.

Because we did the study by manually analyzing the websites, the chosen sample size is small. Therefore it is impossible to make an accurate statement about all of the 309

municipal websites. However, the results are still valid and they show that many websites could make improvements to their cookie policy and their use of cookies.

The timing of this study is also not the best. Google Chrome will very soon be restricting the use of 3rd-party cookies more than they currently are [32]. These restrictions are planned to be implemented in Q3 of 2024 and they could fix some, but not all, of the issues with websites using 3rd party cookies. If a similar study is done even a few years later the results could be different.

6 Conclusions

The main purpose of this thesis was to find out how cookies are being used on Finnish municipal websites. The municipal websites are websites that almost everyone living in Finland visits occasionally so understanding the privacy implications of cookie usage on them is interesting. At the start of this thesis we proposed two research questions:

- **RQ1:** How much control does the user have on the privacy settings?
- **RQ2:** Have the cookies been categorized properly?

It is hard to give a conclusive answer to the first research question. The results from this study show that the choices that the user can make regarding their privacy can often be meaningless because on some websites the cookie options do not matter.

An informed and technically capable user can verify the cookies used by the website by looking at them through the DevTools. However, for a regular user who has to trust what each website promises it really is a gamble. Some websites do properly follow the cookie settings that they give to the user but there are enough websites that disregard them either by being malicious or negligent. There were definitely some websites where it felt like some cookies were active accidentally when they should not have been active without user consent.

Therefore, the answer to the first research question is that on some websites the user has a

lot of control. However, on some other websites the user thinks that they have control on the privacy settings but in reality they do not. This is a big problem and the uncertainty is definitely the major issue. It would be better if the user would know if a website has any privacy issues

The answer to the second question is that the situation is similar to the first question. For the websites with no options all the cookies are considered to be necessary and any analytics or marketing cookies would be misclassified on those websites. There were several websites that had Google's 3rd party cookies active by default and the consent options would not affect them. It is possible that these cookies are active accidentally and no one has noticed them, especially when it was the website of a smaller municipality.

In general, the larger cities had better and more accurate categorization of cookies. In total 10 out of the 25 analyzed websites had significant problems but only 2 of the 10 largest cities had similar problems. We believe that the main reason for this is that the big cities have a lot of residents, who use the website and more users looking at the website means that someone is going to find and report any problems and the problems would likely have been fixed already. Larger cities also have more resources that they can spend on developing their website. In comparison, some of the websites for smaller municipalities did not even have a cookie notice while using analytics cookies, which is a very clear violation of the GDPR and definitely would not last long on a larger website.

There are three issues that need to be solved by the municipalities. First of all the websites that either do not have a cookie notice or have an old notice without options that predates the GDPR should definitely update them. Second, using Google's 3rd party cookies that include advertising cookies without the user's consent should stop.

Google Chrome is going to limit the use of 3rd party cookies from Q3 of 2024 so it is possible that this problem is going away on its own very soon [26]. And the third problem

that needs to be corrected is categorizing analytics cookies as necessary. Especially the websites that used Google Analytics and had it active by default pose a risk to the user's privacy and these cookies should definitely only be active if the user consents to the use of analytics cookies.

References

- [1] *European data protection supervisor; data protection*, https://www.edps.europa.eu/data-protection/data-protection_en, Accessed: 2024-04-08.
- [2] J. P. Choi, D.-S. Jeon, and B.-C. Kim, “Privacy and personal data collection with information externalities,” *Journal of Public Economics*, vol. 173, pp. 113–124, 2019, ISSN: 0047-2727. DOI: <https://doi.org/10.1016/j.jpubeco.2019.02.001>. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0047272719300131>.
- [3] *European union, general data protection regulation*, <https://eur-lex.europa.eu/eli/reg/2016/679>, Accessed: 2023-10-11.
- [4] *Termsfeed, does the gdpr require the listing of individual cookies by name?* <https://www.termsfeed.com/blog/does-gdpr-require-listing-individual-cookies-names/>, Accessed: 2024-02-04.
- [5] *Smartframe, browser fingerprinting: Everything you need to know*, <https://smartframe.io/blog/browser-fingerprinting-everything-you-need-to-know/>, Accessed: 2024-03-25.
- [6] L. Baruh, E. Secinti, and Z. Cemalcilar, “Online Privacy Concerns and Privacy Management: A Meta-Analytical Review,” *Journal of Communication*, vol. 67, no. 1, pp. 26–53, Jan. 2017, ISSN: 0021-9916. DOI: 10.1111/jcom.12276. eprint: <https://academic.oup.com/joc/article-pdf/67/1/26/22321172/>

- jcnlcom0026.pdf. [Online]. Available: <https://doi.org/10.1111/jcnlcom.12276>.
- [7] *Matomo, privacy*, <https://matomo.org/privacy/>, Accessed: 2024-05-14.
- [8] *Google marketing platform, google analytics*, <https://marketingplatform.google.com/about/analytics/>, Accessed: 2024-05-14.
- [9] *Siteimprove, should you consider a google analytics alternative?* <https://www.siteimprove.com/glossary/google-analytics-alternative/>, Accessed: 2024-05-23.
- [10] *Medium, hotjar vs google analytics: A comprehensive comparison*, <https://medium.com/@adamwilsonwebmaxy/hotjar-vs-google-analytics-a-comprehensive-comparison-2fc61932d254>, Accessed: 2024-05-23.
- [11] *Hotjar, hotjar vs google analytics*, <https://www.hotjar.com/blog/hotjar-vs-google-analytics/>, Accessed: 2024-05-23.
- [12] O. Pantelic, K. Jovic, and S. Krstovic, “Cookies implementation analysis and the impact on user privacy regarding gdpr and ccpa regulations,” *Sustainability*, vol. 14, no. 9, 2022, ISSN: 2071-1050. DOI: 10.3390/su14095015. [Online]. Available: <https://www.mdpi.com/2071-1050/14/9/5015>.
- [13] A. Barth, *HTTP State Management Mechanism*, RFC 6265, Apr. 2011. DOI: 10.17487/RFC6265. [Online]. Available: <https://www.rfc-editor.org/info/rfc6265>.
- [14] A. Cahn, S. Alfeld, P. Barford, and S. Muthukrishnan, “An empirical study of web cookies,” in *Proceedings of the 25th International Conference on World Wide Web*, ser. WWW ’16, Montréal, Québec, Canada: International World Wide Web Conferences Steering Committee, 2016, pp. 891–901, ISBN: 9781450341431. DOI: 10.1145/2872427.2882991. [Online]. Available: <https://doi.org/10.1145/2872427.2882991>.

- [15] *Analytics mania, a guide to google analytics client id*, <https://www.analyticsmania.com/post/google-analytics-client-id/>, Accessed: 2024-02-20.
- [16] *Chrome developer blog, cookie expires and max-age attributes now have upper limit*, <https://developer.chrome.com/blog/cookie-max-age-expires/>, Accessed: 2023-11-09.
- [17] X. Hu, N. Sastry, and M. Mondal, “Cccc: Corralling cookies into categories with cookiemonster,” in *Proceedings of the 13th ACM Web Science Conference 2021*, ser. WebSci ’21, Virtual Event, United Kingdom: Association for Computing Machinery, 2021, pp. 234–242, ISBN: 9781450383301. DOI: 10.1145/3447535.3462509. [Online]. Available: <https://doi.org/10.1145/3447535.3462509>.
- [18] A. Abhyankar, L. Barit, S. Jiwani, R. Sasheendran, M. Sotelo, and L. Cranor, “The recipe for the perfect batch: Assessing new cookie category terms,” in *Eighteenth Symposium on Usable Privacy and Security (SOUPS)*, 2022.
- [19] H. Habib, M. Li, E. Young, and L. Cranor, ““okay, whatever”: An evaluation of cookie consent interfaces,” in *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*, ser. CHI ’22, , New Orleans, LA, USA, Association for Computing Machinery, 2022, ISBN: 9781450391573. DOI: 10.1145/3491102.3501985. [Online]. Available: <https://doi.org/10.1145/3491102.3501985>.
- [20] S. Munir, S. Siby, U. Iqbal, S. Englehardt, Z. Shafiq, and C. Troncoso, “Cookie-graph: Understanding and detecting first-party tracking cookies,” in *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security*, ser. CCS ’23, , Copenhagen, Denmark, Association for Computing Machinery, 2023, pp. 3490–3504, ISBN: 9798400700507. DOI: 10.1145/3576915.3616586. [Online]. Available: <https://doi.org/10.1145/3576915.3616586>.
- [21] I. Sanchez-Rola, M. Dell’Amico, P. Kotzias, *et al.*, “Can i opt out yet? gdpr and the global illusion of cookie control,” in *Proceedings of the 2019 ACM Asia Con-*

- ference on Computer and Communications Security*, ser. Asia CCS '19, Auckland, New Zealand: Association for Computing Machinery, 2019, pp. 340–351, ISBN: 9781450367523. DOI: 10.1145/3321705.3329806. [Online]. Available: <https://doi.org/10.1145/3321705.3329806>.
- [22] M. Degeling, C. Utz, C. Lentzsch, H. Hosseini, F. Schaub, and T. Holz, “We value your privacy ... now take some cookies: Measuring the gdpr’s impact on web privacy,” in *Proceedings 2019 Network and Distributed System Security Symposium*, Internet Society, 2019. DOI: 10.14722/ndss.2019.23378. [Online]. Available: <http://dx.doi.org/10.14722/ndss.2019.23378>.
- [23] M. Kretschmer, J. Pennekamp, and K. Wehrle, “Cookie banners and privacy policies: Measuring the impact of the gdpr on the web,” *ACM Trans. Web*, vol. 15, no. 4, Jul. 2021, ISSN: 1559-1131. DOI: 10.1145/3466722. [Online]. Available: <https://doi.org/10.1145/3466722>.
- [24] I. Van Ooijen and H. U. Vrabec, “Does the gdpr enhance consumers’ control over personal data? an analysis from a behavioural perspective,” *Journal of consumer policy*, vol. 42, pp. 91–107, 2019.
- [25] T. Kollmer and A. Eckhardt, “Dark patterns: Conceptualization and future research directions,” *Business & Information Systems Engineering*, vol. 65, no. 2, pp. 201–208, 2023.
- [26] *Google, introducing the privacy sandbox analysis tool (psat)*, <https://developers.google.com/privacy-sandbox/blog/psat-announcement>, Accessed: 2024-05-23.
- [27] *Finlex, kuntalaki*, <https://www.finlex.fi/fi/laki/ajantasa/2015/20150410>, Accessed: 2024-04-15.

-
- [28] *Kuntaliitto, kaupunkien ja kuntien lukumäärät ja väestötiedot*, <https://www.kuntaliitto.fi/kuntaliitto/tietotuotteet-ja-palvelut/kaupunkien-ja-kuntien-lukumaarat-ja-vaestotiedot>, Accessed: 2024-04-08.
- [29] *Auth0, samesite cookie attribute changes*, <https://auth0.com/docs/manage-users/cookies/samesite-cookie-attribute-changes/>, Accessed: 2024-05-23.
- [30] *Google privacy & terms*, <https://policies.google.com/technologies/cookies/>, Accessed: 2024-05-06.
- [31] J. Lee, “The google-doubleclick merger: Lessons from the federal trade commission’s limitations on protecting privacy,” *Communication Law and Policy*, vol. 25, no. 1, pp. 77–103, 2020.
- [32] *Google, third-party cookies restricted by default for 1% of chrome users*, <https://developers.google.com/privacy-sandbox/blog/cookie-countdown-2024jan>, Accessed: 2024-05-24.