Current Requirements and Implementations in the field of Web Tracking in the pre-age of E-Privacy Regulation

Christian Janßen  
*University of Oldenburg*, christian.janssen@uni-oldenburg.de

Matthias Diener  
*Carl von Ossietzky University of Oldenburg*, matthias.diener@uni-oldenburg.de

Jorge Marx Gómez  
*University Oldenburg*, jorge.marx.gomez@uol.de

Follow this and additional works at: [https://aisel.aisnet.org/ecis2021_rp](https://aisel.aisnet.org/ecis2021_rp)

**Recommended Citation**
Janßen, Christian; Diener, Matthias; and Marx Gómez, Jorge, "Current Requirements and Implementations in the field of Web Tracking in the pre-age of E-Privacy Regulation" (2021). *ECIS 2021 Research Papers*. 79.
[https://aisel.aisnet.org/ecis2021_rp/79](https://aisel.aisnet.org/ecis2021_rp/79)

This material is brought to you by the ECIS 2021 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in ECIS 2021 Research Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
CURRENT REQUIREMENTS AND IMPLEMENTATIONS IN THE FIELD OF WEB TRACKING IN THE PRE-AGE OF E-PRIVACY REGULATION

Research Paper

Janßen, Christian, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany, christian.janssen@uni-oldenburg.de
Diener, Matthias, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany, matthias.diener@uni-oldenburg.de
Marx Gómez, Jorge, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany, jorge.marx.gomez@uni-oldenburg.de

Abstract

Website operators have an interest in storing and evaluating data about their users. On this basis, websites will be optimized and marketing campaigns carried out. Cookies are a standardized method of data storage in the field of web-tracking and contain a variety of different information about a user. Often personal data is stored in cookies to identify a user. At this point, the legal requirements of the ePrivacy (Cookie) Directive and the General Data Protection Regulation must be observed. The new ePrivacy Regulation has not yet come into force and is expected in 2021. This research aims to develop legal and functional-technical requirements to meet the current challenges. As a general source for the development, a literature review and qualitative expert interviews have been carried out. For the validation, we conducted a study and implemented a web scraper in order to analyze the top 100 e-commerce websites in Germany.

Keywords: Requirements Engineering, Cookies, Web-Tracking, Study execution
1 Introduction

Website operators have an interest in storing and evaluating data about their users. Accordingly, information can be generated and used to optimize the website (Hassler, 2012). A widely used form of data storage in this field are Cookies. These small text files will be stored on a user device while surfing on a website and contains a variety of data about the user (Christl, 2014). As an example, advertising networks may use this information for personalized advertising. Different types (First Party Cookies, Third Party Cookies, Session Cookies or Persistence Cookies) and categories (strictly necessary, analytics, functionality and marketing) can be distinguished. The collection of data by website operators is in a constant conflict with increasing data protection requirements. In addition, the privacy of the user needs to be respected (Trevisan et al., 2019). For this reason, the new ePrivacy Regulation (ePR) aims to protect the privacy of citizens in electronic communications more effectively. The ongoing technical development in this area (e.g. methods of web tracking) should be taken into account. Since the new regulation has not yet come into force, the ePrivacy Directive of 2002 (ePD) and the updated version of the Cookie Directive of 2009 provide the legal basis. In the current version, users must give their consent to data processing for the use of non-essential cookies. As mentioned above, cookies contain a variety of data – especially personal data. The General Data Protection Regulation (GDPR) stipulates obligations of protection with respect to the processing and sharing of personal data between different entities. All companies from inside and outside the EU which process personal data in relation to the offering of goods or services to individuals in the EU need to respect the regulation. The protection of personal data according to Art. 5 GDPR and the related transparent communication of all processing steps as described in (Janßen and Kathmann, 2020) is essential for website operators. To refer to Recital 30 GDPR, information that are stored in cookies can be used to clearly identify a user. At present, the relationship between the ePR and the GDPR can be considered as unclear (Rauer and Ettig, 2018). Some aspects and statements cannot be clearly delimited and distinguished (Piltz, 2018). In accordance with Recital 5 and Art. 1 sect. 3 ePR of the draft, the regulation is intended to determine and supplement the GDPR. The draft is also called lex specialis, which takes up and specifies questions in the area of personal electronic communication data. Non-specific questions regulated in the ePR are taken up by the GDPR as lex generalis.

However, a large number of different requirements must already be taken into account by website operators to face the current complex legal situation. Also, technical areas, such as web tracking gained importance. Therefore, the following research questions (RQ) will be implemented: (RQ1) – What are the current legal and functional-technical requirements for saving user data on the web in the pre-age of the new ePrivacy Regulation? (RQ2) – How are the identified requirements currently realized on the top 100 e-commerce websites in Germany?

The rest of the paper is structured as follows: Section 2 describes the research methodology of this paper. Section 3 and 4 gives a short overview about cookies, Web Tracking and the legal background that is necessary for the understanding of this research. Section 5 shows the elaborated functional (FR) and non-functional (NFR) requirements and section 6 refers to the validation of the requirements in the form of a study. Finally, section 7 concludes the paper.

2 Research Methodology

The requirement elicitation, analysis, specification and validation are embedded in the Requirements Engineering discipline. In particular, Requirements Engineering is a systematic approach to specify and manage requirements with the aim of knowing the relevant requirements, satisfying the wishes and needs of stakeholders and concretizing the requirements (Pohl and Rupp, 2015). A requirement is a documented necessary characteristic or ability of a system, that solves a problem or achieves a specific goal (IEEE Standards Board, 1998). In other words, requirements are the creation basis of a system model or a system architecture. Therefore, the requirements, as qualitative criteria for good requirements, should be complete, atomic, technically, solution-neutral, consistent, testable, necessary, traceable, realizable and clear (Rupp, 2014). For a better structure and categorizing, requirements can
be divided into functional and non-functional aspects. A FR is a requirement, relating to a result or behavior that will be provided by a single function of a system (Pohl, 2008). NFRs often described as properties or qualities that a system must provide (Ebert, 2005). The following figure shows the specific research process for this work.

Figure 1. Requirement Engineering process of this research (In accordance with (Rupp, 2014))

The elevation process includes 4 steps. For this research, we selected the categories legal and functional-technical. Within the source search and determination step, we decided to collect data through two different research methods. First, a systematic literature review, that follows the methodology of (Webster and Watson, 2002) and (Cato, 2016), were carried out with the 3 process steps: 1) Definition of search strategy and items, 2) Identification and selection of relevant papers (including forward search, backward search and preliminary results discussion) and 3) Analysis of relevant papers (includes specifications of findings). Second, qualitative structured expert interviews oriented to (Flick, 2016) and (Früh, 2011) were conducted. The method is placed in the area of qualitative research as a more structured variant than conventional interviews. The focal points of the experts are perspectives and methods of action. We decided to conduct 7 expert interviews to adequately cover the topic, to be in conjunction with the systematic literature review. For the documentation and structure of the specified NFR and FR requirements as the last step, we used the template process master according to (Rupp, 2014). In their full deontic sense, the terms shall, will and should reflect degrees of accountability and responsibility. Shall requirements are mandatory in the realization. Should requirements represent the needs and wishes of different stakeholders and will requirements document future intentions.

In order to check the requirements and the current state of implementations, we decided to analyze the top 100 e-commerce websites in Germany in terms of sales. Therefore, we used concepts of Web Scraping as described in (Mitchell, 2018), (Vargiu and Urru, 2012) and (Lawson, 2015) and implemented our own web-scraper. Within the script, we used Selenium as a test framework in order to extract the necessary information about the cookie types. To structure the data, a Pandas data frame was used. Also the python library BeautifulSoup was used to determine the purposes of the considered cookies. The results of the analysis were compiled and designed in the form of a study.
3 Web-Tracking and Cookies

Within this research contribution, the understanding and interaction of different technical basics from the areas of Web-Tracking and Cookies are essential.

Web-Tracking is basically the collection of data on a website in order to extract information about the user and his behavior (Hassler, 2012). The International Working Group on Data Protection in Telecommunications gives the following definition: "[...] the collection and storage, use or exchange of data on individual online behavior over a variety of websites through the use of Cookies, JavaScript [...]" (IWGDPT, 2013). (Hassler, 2012) summarizes the goals of Web-Tracking as follows: To know, see and better understand the recipient of a message sent via internet and to adjust future messages more precisely. Therefore, different questions such as Which pages/areas of the website are the most visited?, Where did the user come from (marketing channel)?, How does he/she navigate through the website? and Where does the user leave the website? could be asked. Through this information and the determination of click paths across different websites of the user, concrete measures and goals can be derived to optimize the website and to create user profiles (Schneider et al., 2014). Various methods for data collection such as Server Logs, Tracking Pixel, Fingerprinting and Cookies can be used for these purposes (Ermakova et al., 2018).

Nowadays, cookies are a widely used and standardized concept for the collection of data. Cookies are small text files, that are stored on the user device by the web browser while surfing the internet (Christl, 2014). The files contain information about the user, such as an individual ID, which is read by the webserver each time a page is called up. This enables the creation of a continuous user session and identification. Furthermore, additional information and data can be stored in cookies which can be used for e.g. marketing or sales purposes. From this point of view, different types of cookies are used in practice (Cahn et al., 2016). 1) Session Cookies are necessary to ensure the functionality of a website and are only valid for the duration of a user session. 2) Persistent Cookies have an expiration date and are used to identify the user on a website. After expiration, the cookie will be deleted. 3) First Party Cookies are cookies that are set by the visiting domain. 4) In contrast, Third Party Cookies are set by domains that have not been visited by the user. With regard to the following chapters, it is important to classify cookies according to their purpose. We follow the approaches of (ICC UK, 2012), (Utz et al., 2019), and (Machuletz and Böhme, 2020) and use the categories strictly necessary, analytics, functionality and marketing. Cookie banners are used to obtain, consent and fulfill legal requirements. The banner contains an information text about the use of cookies and requests consent for data processing from the user (Nouwens et al., 2020). The variety in the design of cookie banners ranges from simple information text with one button to complex concepts, where the user can give his consent for individual processing operations or services. (Gray et al., 2021) also analysis different types of cookie banners in the context of HCI, design, privacy and data protection and legal research communities through “dark patterns”.

4 Legal Background

In section 3 we presented the necessary technical basics for this research. Based on this, we used legal basics from the areas of the consisting ePrivacy and Cookie Directive (2002/58/EG), the GDPR (Regulation (EU) 2016/679) and the new ePrivacy Regulation (2017/0003 COD) to set up a legal framework in this section. We also refer to existing court decisions in this sector that must be taken into account.

On the European level, the current ePrivacy Directive must be observed. The Directive sets minimum standards for data protection in electronic communication processes. Already in 2002, Art. 5 Sect. 3 of the ePrivacy Directive stipulated that the user should be provided with clear and comprehensive information about the processing of his personal data. In particular, the user must be explicitly informed that he has also the right to refuse such data processing (Walter, 2018). In this context, this kind of procedure is called Opt-out. In 2009, the previous ePrivacy Directive was amended by the new Cookie Directive. As a result, and besides the information obligation in the version of 2002, the
Consent to data processing by the user has become mandatory. These procedures are called Opt-In. In contrast to regulations, directives do not have the property of being imminently effective (Eckhardt et al., 2013). Opening clauses allow European member states to create their own implementation possibilities. In Germany, the ePrivacy directive was implemented in the German Telecommunication Act (TKG). The updated version of the Cookie Directive was not implemented because the German Telemedia Act (TMG) with its Opt-Out oriented procedures was considered as sufficient by the legislator. However, new court rulings by the European Court of Justice (C-673/17) and the Federal Constitutional Court in Germany (I ZR 7/16) decided, that a user consent is required in principle for cookies that are not strictly necessary. Therefore, Opt-In procedures must be used and Opt-Out procedures should be avoided in the future.

Cookies contain a variety of different personal and non-personal data. Because of this, the GDPR must be considered in this research work. After the ratification in May 2018 by the European member states, multiple changes and reforms became effective. A new definition of personal data according to Art. 4 Sect.1 GDPR is given as follows: Personal data means any information relating to an identified or identifiable natural person (Data Subject); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person. More precisely, cookies are addressed in Recital 30 GDPR: Natural persons may be associated with online identifies provided by their devices, applications tools and protocols [...]. This may leave traces which, in particular when combined with unique identifiers and other information received by the servers, may be used to create profiles of the natural persons and identify them. The individual Cookie identifier or the Cookie ID itself can already make it possible to identify a user. As a result, large user profiles can be created when a website is called up again. A personal reference may be assumed, even if the user is not identified by a personal date (such as the name) (Brandt, 2018). Basically, the following conclusion can be made: the more data about an individual user is stored in cookies, the more likely it is that personal data is involved (Keppeler, 2018). With regard to the lawfulness of data processing with cookies, there are three types of permissions that are recommended to be considered: Art. 6 Sect. 1 lit. a GDPR (consent), Art. 6 Sect. 1 lit b GDPR (contract) and Art. 6 Sect. 1 lit. f GDPR (legitimate interest) (DSK, 2019).

The ePrivacy Regulation is a proposal of the EU Commission to protect the privacy of citizens in electronic communication processes (Piltz, 2018). Therefore, the existing ePrivacy Directive will be replaced. The regulation will affect every company that provides a service such as an app or a website that uses cookies to collect and analyze data on user behavior. Also, services and methods of Web-Tracking will be implemented, which could not be covered by the previous ePrivacy Directive. The choice of a regulation as an instrument for the new ePrivacy regulation allows the harmonization of the law in the EU member stats and simplifies the consent requirements for data processing. Furthermore, Recital 5 GDPR and Art. 1 Sect. 3 GDPR will be supplemented and specified by the ePrivacy Regulation. Since the ePrivacy Regulation has not yet been ratified, we use the available proposal within this contribution. With regard to the Requirements Engineering process, the presented legal and functional-technical aspects are used in the next section.
5 Requirements

In section 2, we showed the data sources and research approach in order to extract our FR and NFR for the categories legal and functional-technical. As already mentioned, for the systematic literature review we used the research methodology according to (Webster and Watson, 2002) and (Cato, 2016). As databases for our literature review, we used Google Scholar, Beck Online, Juris, Orbis plus and arXiv. For the screening of the abstracts and titles, COLANDR as a text mining tool was utilized. We decided to focus on literature in the period from 2018 to 2020 because of the ongoing development in the area of legal frameworks. The subsequent table gives an overview of the search strings and paper results.

<table>
<thead>
<tr>
<th>Search String</th>
<th>Databases paper results (total findings/relevant findings)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Google Scholar</td>
</tr>
<tr>
<td>Cookies AND requirements</td>
<td>16300/2</td>
</tr>
<tr>
<td>Cookies AND GDPR</td>
<td>6690/5</td>
</tr>
<tr>
<td>Cookies AND eprivacy regulation</td>
<td>835/0</td>
</tr>
<tr>
<td>Cookies AND Anforderungen</td>
<td>16700/1</td>
</tr>
<tr>
<td>Cookies AND DSGVO</td>
<td>967/2</td>
</tr>
<tr>
<td>Cookies AND eprivacy Verordnung</td>
<td>140/0</td>
</tr>
</tbody>
</table>

Table 1. Paper results of the literature review

To supplement and validate the requirements we did 7 qualitative expert interviews that followed the stated research methodology of (Flick, 2016) and (Früh, 2011). The experts were selected on the basis of their professional background. The period and length of the interviews were no longer than 30 minutes. For the evaluation and transcription of the content, we used the tool MAXQDA. Subsequently, the generated text basis was qualitative analyzed through the objective analysis approach according to (Ginz, 2010). A content analysis within the approach allows a systematic description of the given content and a formal characterization of messages. Therefore, different phases have been passed. First, the conception phase characterizes the problem and the formation of different criteria. Second, the development phase forms different theoretical and empirically-driven categories. Third, during the testing phase, reliability and validation tests have been performed. Fourth, the application phase prepared, coded and analyzed the data. The results of the literature review and of the expert interviews have been divided into the following categories: explicit consent, voluntariness, information obligation, granularity, time limits, revocability and documentation. Additionally, we added the category ePrivacy Regulation to capture requirements in this area as well. Based on these categories a concept matrix has been developed.
### Table 2. Combined concept matrix

<table>
<thead>
<tr>
<th>Authors/Interviewees</th>
<th>Categories</th>
<th>Affected Requirements/Outlook</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Explicit consent</td>
<td>Voluntariness</td>
</tr>
<tr>
<td>(Heckmann, 2019)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Steinrötter, 2020)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Engeler and Marosi, 2019)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Kremer and Schneider, 2019)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Voigt and Kaus, 2019)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Santos et al., 2020)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Utz et al., 2019)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Machuletz and Böhme, 2020)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Degeling et al., 2019)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Taeger and Schweda, 2020)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Forgó et al., 2018)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Wilfling, 2019)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Brandt, 2018)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Keppeler, 2018)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Schwäbe, 2019)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Heidrich, 2019)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Jakobi et al., 2020)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Sanchez-Rola et al., 2019)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Nouwens et al., 2020)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Schwartmann et al., 2018)</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>(Hilgert and Hilgert, 2020)</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Interviewee 1 (Lawyer) | ✓ | ✓ | ✓ | ✓ | ✓ | NFR2, FR8 |
Interviewee 2 (Lawyer) | ✓ | ✓ | ✓ | ✓ | ✓ | NFR6, FR8 |
Interviewee 3 (Data Scientist) | ✓ | ✓ | ✓ | ✓ | ✓ | NFR4, NFR5 |
Interviewee 4 (Product Owner) | ✓ | ✓ | ✓ | ✓ | ✓ | NFR1, FR2 |
Interviewee 5 (Project Manager) | ✓ | ✓ | ✓ | ✓ | ✓ | NFR3 |
Interviewee 6 (Data Protection Officer) | ✓ | ✓ | ✓ | ✓ | ✓ | NFR1, NFR3, NFR4 |
Interviewee 7 (Business Manager) | ✓ | ✓ | ✓ | ✓ | ✓ | NFR7, NFR1, FR7, FR2 |
Based on these results and the juridical writings of section 4, the following legal and functional-technical requirements have been extracted and connected with each other. Legal requirements are represented as NFR and functional-technical requirements are represented as FR. All requirements are representing necessary “shall” conditions to be compliant with the current legal situation:

NFR 1 The consent to data processing shall be explicit.
NFR 2 The consent shall be given voluntarily.
NFR 3 The consent shall be informed.
NFR 4 The consent shall be possible in a granular manner.
NFR 5 The consent shall be obtained in advance.
NFR 6 The consent shall be revocable.
NFR 7 The consent shall be documented.

FR 1 It shall be technically ensured that the first thing a visitor of a website sees is a cookie banner. (Connected to NFR5)
FR 2 If a user has not yet interacted with the cookie banner, all tracking tools and not strictly necessary cookies shall be retained. (Connected to NFR1)
FR 3 In case of a granular decision of a user, it shall be ensured that only the selected tracking tools are executed. (Connected to NFR4)
FR 4 The cookie banner shall contain a short information text about the use of cookies. (Connected to NFR3)
FR 5 A user shall have at least two choices (accept and settings) on the cookie banner. (Connected to NFR2)
FR 6 A user shall be able to view information about certain tracking tools. (Connected to NFR3)
FR 7 A user shall have the possibility to view his or her past given consents. (Connected to NFR7)
FR 8 A user shall be provided on each page of a website to adjust his or her cookie settings. (Connected to NFR6)

Within the following section 6, we will discuss the implementation of the requirements based on the top 100 e-commerce sites in Germany.

6 Current Implementations

6.1 Study design

After the determination of the requirements in the previous section, the current status of the implementations on German e-commerce websites is examined below. A conscious decision for e-commerce websites was made since they tend to track more information than simple websites (e.g. news, community). It can be assumed, that online stores use more cookies (e.g. for personalized advertising). In addition, sections 3 and 4 made clear, that website operators need to act otherwise, they risk financial fines. The analysis regarding the state of realization of the requirements from section 5, is based on the top 100 e-commerce sites in Germany and is shown by product segments in the following figure.
Figure 2. Percentage distribution of product segments

Furthermore, it must be established if and when the requirements can be met. It was possible to build on the results of previous studies according to (Utz et al., 2019), (Schiefermair and Stabauer, 2020) and (Santos et al., 2020). Therefore, we focus in our study on a legal and functional-technical perspective. Because of this, the following steps for the study were developed and applied.

We analyzed the displayed cookie banners in order to identify the types of cookies of a website. To specify the purposes, we used the already known characteristics: Session Cookies, Persistent Cookies, First Party Cookies, Third Party Cookies, Strictly necessary Cookies, Analytical Cookies, Functional Cookies and Marketing Cookies.

The requirements of explicit consent (NFR1, FR2) are considered to be fulfilled when the cookie banner offers the possibility to give a consent. Furthermore, only strictly necessary cookies as a prerequisite for the fulfillment of the requirements are allowed to be stored before the consent of a user is given. Possible solutions at this point are a button or a non-pre-filled selection field. The following specifications need to be considered:

- No explicit action is possible. Marketing/ Analytical Cookies will be stored in advance.
- Explicit action is possible. Marketing/ Analytical Cookies will be stored in advance.
- Explicit action is possible. Marketing/ Analytical Cookies will not be stored in advance.

The consent needs to be given voluntarily (NFR2, FR5). For this reason, the cookie banner must have at least two choices. The requirement is not considered to be fulfilled, if there is only a button to confirm the processing of data. A user should have the possibility to deny the data processing or to change his opinion. Ideally and to fulfill GDPR criteria, there are three types of buttons (accept, reject and further information) to make granular settings. The following specifications need to be considered:

- No choice.
- Accept and further information.
- Accept, reject and further information.

For a legitimate consent, a user shall be fully informed in advance (NFR3, FR4, FR6). Accordingly, the information of the data processing must be shown. A link to the privacy policies should be provided on the cookie banner. The following specifications need to be considered:
- No cookie banner with information. No link to the privacy policies.
- Cookie banner with information. No link to the privacy policies.
- Cookie banner with Information. Link to the privacy policies.

In order to meet the requirement of granularity, the cookie banner must list the individual data processing operations and provide a selection field for each of them (NFR4, FR3). Only the selected processing operations may be executed. The following specifications need to be considered:

- No granular decision possible.
- Granular decisions based on categories possible.
- Granular decisions based on categories and/or tools, services possible.

If a consent has not been given yet, no non-essential cookies may be stored on the user device (NFR5, FR1). In order to check this, the stored cookies can be analyzed by the time of acceptance. Therefore, if non-essential cookies are stored without the user’s consent, the requirement is deemed not to be fulfilled. The following specifications need to be considered:

- Analytical and/or marketing cookies are stored in advance.
- Analytical and/or marketing cookies are not stored in advance.

After a user has given his consent to data processing, he shall have the possibility to revoke the consent (NFR6, FR8). If there is no specific link for revocation, the information must be provided in the privacy policies. Also, the given consents shall be documented (NFR7, FR7). The following specifications need to be considered:

- No possibility to contradict.
- Possibility to contradict (privacy policies)
- Possibility to contradict (directly on the cookie banner, link on each page)

### 6.2 Study Results

Following out the developed steps of the study design, this section shows the results with regard to the realization level of the requirements and the cookies used on the websites.

The following Table 3 use key indicators to show how many cookies were used in total on the websites. We have divided the table in two areas: in the upper area by different types and the lower area by different purposes of cookies. A total of 1797 cookies were found on the pages (Amount). In most cases, these types were First Party Cookies (78.3%) and Persistence Cookies (72.4%). On average, 18.5 cookies were set on the websites before any user interaction took place (Average value). The minimum for almost all characteristics is 0 (Minimum). The values for the maximum are significantly higher. For example, on one website 59 Third Party Cookies were placed on a user device. In addition, the highest number of cookies found on one website was 72 (Maximum). It is remarkable that the standard deviations are higher than the average value for Third Party Cookies and Marketing Cookies (Standard deviations). Consequently, the variation coefficient is greater than for the other characteristics (Variance). This means, that the data scatter more around the average value. In addition, the span measures the dispersion of the characteristics which we used in the study (Span). The distinction of the cookie types was implemented in our script. Through APIs, cookie databases such as Cookiepedia were connected and used for this reason. Around 904 identified cookies were programmed individually and did not follow a clear information structure. As a result, the exact purpose of the cookies could not be determined.
Table 3. Key indicators of the study

When looking at the frequency distribution of the individual characteristics in Figure 3, it is noticeable that the peaks (i.e. the most frequent values) are in the range of a maximum of 10 cookies for all histograms. In most cases, a small peak appears in the right part of the histograms. This means that the number of websites with more than 60 cookies is higher than the number of websites using 40 or 50 cookies. In addition, all histograms do not show a normal distribution. They have a right-skewed distribution, which means that most websites use a smaller number of cookies. Especially in the area of Third Party Cookies, it is clear that in most cases 0-5 cookies are used.

![Histograms of the individual characters](image)

Figure 3. Histograms of the individual characters

When looking at the purpose of the individual cookies, the following can be determined: 20% of the cookies can be assigned to the category marketing. 13% to the category analytical, 9% to the category technical and 2% to the category functional. 56% of the cookies could not be determined because they...
were programmed individually for the website and are assigned to an unknown purpose. With regard to the presented product segments, it can be intended that websites in the Consumer Electronics sector use the highest number of Marketing Cookies. Some distance behind, websites in the area of Pharmacy & Health followed. In contrast, websites from the sector Furniture & Housewares, Hobby & Stationery and DIY & Garden used the lowest number of cookies.

Furthermore, the following statements can be made for the presented requirements. It can be stated that on 95.9% of the examined websites a cookie banner was displayed. On 4.1% of the websites, no cookie banner was shown.

Explicit Consent (NFR1, FR2): most of the examined websites do not fulfill this requirement. In 33% of the cases, no explicit action was possible. At the same time Marketing, Cookies were set. In addition, cookie banners were already prefilled. 40% of the websites offered the possibility for an explicit action. However, cookies were also set in the background without consent. Only 27% of the websites have fully met the requirement.

Voluntariness (NFR2, FR5): Similar results are shown for this requirement. In 30% of the examinee cookie banner, the user had no possibility to make a selection. There was only a button to accept the data processing or to close the cookie banner. In contrast, 33% of the websites offered the user the option to accept the data processing or to receive further information about it. The remaining websites (37%) offered also the opportunity to refuse data processing.

Information obligation (NFR3, FR4, FR6): 95% of the viewed websites met this requirement. An information text was displayed on the cookie banner and a link to further information (e.g. privacy policies) was available. On 1% of the websites, only an information text was presented. 4% of the websites did not provide any of the requested information.

Granularity (NFR4 FR3): In terms of granularity, 45% of the websites do not allow a granular decision. Users only had the opportunity to agree to all processing operations. 28% of the websites offered a selection based on categories. For the remaining 27% of the websites, an additional tool/cookie selection could be made.

Time limits (NFR5, FR1): Before the user has completed an interaction or has given his consent, 73% of the viewed websites have already stored cookies on the user device. 27% of the websites fulfilled the requirement and did not store Analytical or Marketing Cookies in advance.

Revocability (NFR6, FR8) and Documentation (NFR7, FR7): Only 19% of the websites provided no possibility to disagree with the processing of data. On the other hand, 58% offers the option to refuse data processing via the privacy policies and the remaining websites (24%) provided a link to the cookie settings to contradict the data processing. In addition, it could be determined that all interaction processes were sufficiently documented.

7 Conclusion

7.1 Main Contribution

As motivated and presented in the previous sections, the legal challenges for website operators are enormous. Cookies, as a method of web tracking, played a significant role in the storage of personal data. Therefore, special attention has been paid to the protection of this data and the related processing.

The primary objectives of this paper can be divided into two artifacts: First, we collected a total of 7 NFR and 8 FR. Therefore, we followed a research process in accordance with (Rupp, 2014) and used the process master for a structured documentation of the requirements. As a procedure for obtaining data and information, we conducted a systematic literature review and qualitative structured expert interviews. With regard to the first research question (RQ1), a framework of requirements was established, which takes up the currently necessary implementations in this area and considered
already aspects of the new ePrivacy Regulation. Second, we developed and carried out a study and reviewed the top 100 e-commerce sites in Germany based on the state of implementation of our requirements. With regard to the second research question (RQ2), it could be determined that many of the examined enterprises showed pent-up demands. As a result, additional research is required in this field.

So far, companies can use the derived requirements as an entry point for the implementation of a compliant cookie strategy within their organization.

### 7.2 Limitation and Future Work

Like in any research contribution, this work has also some limitations. Concerning the accomplished study, it must be considered, that the focus was on large enterprises. However, SMEs can be affected more in order to implement the requirements into their organizational structures. Often, resources are partly available to complete the implementation in a timely manner. Therefore, the study should be extended and include SMEs. Also, our developed script needs some improvements to clarify the purposes of cookies, that have not yet been identified. By answering the research questions, it also became clear that further research is needed for practical implementations in this area. As a next step, we are planning to develop a practical guideline based on the requirements for companies to face future challenges.

### References


