ABSTRACT: Since last March 2012, when the Spanish legislation incorporated the EU Directive on Electronic Privacy, all Spanish websites collecting users’ data are required to inform clearly on the type and purpose of the data collected and to obtain the users’ consent to use their data with analytic or marketing purposes. In order to provide support to webmasters, the Spanish Data Protection Agency (AEPD) released last April the «Use of cookies guidance» similar to the UK ICO «Guidance on the rules on use of cookies and similar technologies», with some advises to help websites accomplish with the current regulations. In October 2013, more than a year and a half after the law incorporation, the results of this research show that only 51% of the 500 most visited Spanish websites (according to Alexa Rank), apply the regulations correctly, asking their users for the informed consent to collect their data. As a direct consequence, a new question arises: Who are the trackers behind these cookies? We identify and analyze in this research the most important tracking corporations (also named Data Collectors) and their direct effect with Cyberspace Data Privacy. We propose, to complete the dissertation, the need for greater responsibility by stakeholders concerned, such as Institutions, Cybermedia and other relevant websites and Internet users. Internet users have the right to be correctly informed about the potential risks they assume while surfing the web and the right to know how these data will be used. As the boundaries between physical space and Cyberspace tend to disappear, E-privacy becomes a major issue, which has a direct effect not only in the online data safety, but also in the offline life.

KEYWORDS: e-Privacy, Cookies, Big Data, Behavioral Targeting, Advertising, Transparency.

1. INTRODUCTION

Last March 2012, Spanish legislation incorporated the EU e-Privacy Directive, also known as Cookie Law. The law requires all websites engaged in economic activities and collecting users’ data through cookies or similar technologies to inform clearly on the type and purpose of the data collected and to obtain the users’ consent. We consider interesting to see if, almost two years after the implementation of the regulations, the most visited websites in Spain correctly inform their users about their data collection, with statistical or marketing purposes. We also suggest the need for identify the most important data collection companies behind the cookies operating in Spain, in order to provide greater transparency on the Web. We believe data privacy of Internet users’
is one of the major challenges we face in the coming years. As the boundaries between physical space and cyberspace tend to disappear, privacy of Internet users becomes a matter of vital importance with also direct consequences in offline word data safety\(^1\).

### 1.1. Internet usage in Spain

According to eEspaña 2013 report, developed by the Orange Foundation\(^2\), 68% of Spanish households have an Internet connection and 25 million Spanish people used the Internet in 2012. Regarding the expansion of the Mobile Web, the statistics underline that 39% of Spanish people have used a laptop or mobile phone to access the Internet.

Most Web users in Spain use the Internet to send or receive emails, search for information and access to news and online media. Half of the Internet users make more complex activities such as participating in social networks or downloading content such as games, movies or music.

The figures state that, in general terms, Spanish citizens are above the European average in many uses of the Internet. Noteworthy in this regard a significant difference in the use of services related to entertainment and creativity. For example, while 45.6% of Spanish Internet users upload content on websites to be shared, in the rest of Europe this activity is only done by a third of the Internet users.

The data we observe, as well as the prospective studies, predict an increasingly ubiquitous Web, in an increasingly connected and Internet dependent society.

We could assume, in this sense, there is a risk of diving into the rivers of information without having the necessary tools to ensure our privacy. Technological progress, in our opinion, must be hand on hand with users’ education and with the necessary mechanisms to ensure user’s privacy, web transparency and, ultimately, a more democratic Web.

We would like to emphasize the need, which is the subject of this study, to evaluate the potential risks on data privacy Internet users assume while surfing the Web. It is important, in our point of view, to be aware of the lack of information in this area\(^3\). That is the main reason we have decided to observe the degree of compliance with data

---

3. According to a research *The Impact of Cookie Deletion on Site-Server and Ad-Server* (2011) developed by the Web Analytics company Comscore, third party cookies are deleted monthly at a rate of between 30% and 40% of computers, depending on the country. This means that between 60% and 70% of Internet users do not remove cookies frequently.
protection regulations by most influential Spanish websites, as well as to identify the leading data collectors, which gather the information we generate while surfing the Web. We would like to believe this kind of studies can somehow help clarifying this complex scenario, while putting on the table some of the potential risks that we will undoubtedly have to face in the coming years.

1.2. Online Privacy and Legislation

In 2003 was implemented the European Directive 2002/58/EC\(^4\) which aimed to protect users privacy in electronic communications. In 2009, it was amended by Directive 2009/136/EC\(^5\), with a modification in the Article 5(3) of the e-Privacy Directive, which requires the consent to store or access to information stored in users' or subscribers' terminals. In other words: it is an essential requirement to obtain the consent of the Internet users to install cookies and similar technologies. European legislation started, since then, requiring the explicit or implied consent of the users, to websites using cookies that may threaten users' privacy.

The EU directive was incorporated into Spanish legislation on March 31\(^\text{st}\), 2012\(^6\) and, in order to make compliance easier for websites, Spanish Data Protection Agency (AEPD) published last April 2013 the «Use of Cookies Guidance», developed in collaboration with media and advertising agencies.

The report, which is similar to the UK ICO «Guidance on the rules on use of cookies and similar technologies», offers advice on how to comply with e-Privacy regu-

---


lations. Basically, the document explains that all websites collecting users’ data which are not strictly necessary for the correct function of the website, are required to clearly inform about the type and purpose of the data collected, and to obtain the consent of the users to exploit their data with analytic or advertising objectives.

According to the ICO Cookies Guidance\(^7\), a person shall not store or gain access to information stored, in the terminal equipment of a subscriber or user unless the next requirements are met:

a) The user is provided with clear and comprehensive information about the purposes of the storage of, or access to, that information; and

b) The user has given his or her consent.

Or, summarizing the above information, those setting cookies must:

a) Tell people that the cookies are there,

b) Explain what the cookies are doing, and

c) Obtain their consent to store a cookie on their device.

About a year and a half after the incorporation into Spanish legislation of the EU Directive, particularly in early July 2013, about half of Spain’s most important cybermedia, did not apply correctly the law yet\(^8\). It is true, however, the adaptation of websites to the new EU guidelines, as has been shown for example by reports published by Discapnet on Web Accessibility, may take some time to implement.

It is necessary in this regard to underline that article 5, paragraph 3 of Directive 2002/58/EC, allows cookies to be exempted from the requirement of informed consent, if they satisfy one of the following criteria:

- Criterion a) the cookie is used «for the sole purpose of carrying out the transmission of a communication over an electronic communications network».

- Criterion b) the cookie is «strictly necessary in order to the provider of an information society service explicitly requested by the subscriber or user to provide the service».

Accordingly, the Working Group of Article 29, in its Opinion 04/2012 on Cookie Consent Exemption\(^9\), concludes websites using these cookies are exempted to require their user’s consent:

---


• «User-input» cookies
• Authentication cookies
• User centric security cookies
• Multimedia player session cookies
• Load balancing session cookies
• UI customization cookies (language preference cookies, result display preference cookies).
• Social plug-in content sharing cookies

On the other hand, among non-exempted cookies (the websites using these cookies are required to provide information and obtain users’ consent) we can find:
• Social plug-in tracking cookies
• Third party advertising
• First party analytics

1.3. About cookies

The Regulations are applied to cookies and similar technologies used by websites to store information. This definition includes, for example, Local Shared Objects (commonly named «Flash Cookies») and web beacons (transparent 1x1 pixels images).

According to the website All about cookies, «Cookies are usually small text files, given ID tags that are stored on your computer’s browser directory or program data subfolders. Cookies are created when you use your browser to visit a website that uses cookies to keep track of your movements within the site, help you resume where you left off, remember your registered login, theme selection, preferences, and other customization functions. The website stores a corresponding file(with same ID tag)to the one they set in your browser and in this file they can track and keep information on your movements within the site and any information you may have voluntarily given while visiting the website, such as email address. […] Cookies are often indispensable for websites that have huge databases, need logins, have customizable themes, other advanced features. However, marketing is becoming increasingly sophisticated and cookies in some cases can be aggressively used to create a profile of your surfing habits».

Cookies can be classified into different categories in relation to their functions. It is important to underline that the same cookie may be included in more than one category.

The next classification is based on the analysis made by the Spanish Data Protection Agency (AEPD) and published in the «Use of cookies guidance»\(^\text{10}\).

1.3.1. **Depending on the entity who manages the cookie:**

a) **First-party cookies**
These cookies are sent to the user’s computer from a computer or domain managed by the website editor.

b) **Third-party cookies**
Third party cookies are cookies served from a domain other than the page in which it is embedded. Third-party cookies are usually not strictly necessary for the user to visit a website, as they are usually linked to a different service from the one specifically requested by the user.

1.3.2. **Depending on the time cookies remain active:**

a) **Session cookies:**
These files are deleted when the user closes the browser, once the navigation is completed. The next time the user visits the website, he or she is not recognized and is treated as a new user, since there is no file in the browser enabling the site to know if the user has visited it previously.

b) **Persistent cookies:**
These cookies remain on the hard disk of the terminal until the user removes them or they expire. The duration of the cookie depends on how the file has been programmed by the developer. These cookies have the ability to authenticate the user account, so the user do not have to provide his or her account details each new visit, as well as various options for interface customization, such as language selection, menu preferences, etc. The time these cookies remain on user’s terminal can range from a few minutes to several years.

1.3.3. **According to its purpose:**

a) **Technical cookies**
These cookies allow the users to navigate and interact with the website.

b) **Customization cookies:**

Customization cookies allow users to access the service with some particular features based on a predefined set of criteria in the user terminal such as the language, the type of browser, font size, etc.

c) Analytics cookies

These cookies allow the website owner to monitor and analyze the audiences and the users’ behavior in the site.

d) Advertising cookies:

These cookies allow web editors and online advertising companies to manage ad spaces in an efficient way.

e) Behavioral targeting cookies:

These cookies store information about users’ behavior. These data are obtained through a continued observation of users’ browsing habits. The information collected makes possible to develop a specific profile of each user and to display ads based on his or her preferences.

2. METHODS

2.1. Universe of Study

We have selected in this study the top 500 websites in Spain, according to Alexa Rank, in September, 2013.

Alexa Internet, Inc. is a subsidiary company of Amazon, which has been established as one of the most important web audiences’ measurement systems, auditing more than 30 million websites worldwide. Alexa Rank measures the popularity of websites throughout the cyberspace and offers the possibility to segment audiences by country. The rank is calculated using a combination of the estimated average daily unique visitors to the site and the estimated number of pageviews on the site over the past 3 months. The site with the highest combination of unique visitors and pageviews is ranked #1.

We have removed from the sample a total of 39 web addresses which, in our opinion, could decrease the reliability of the study, based on two reasons:

a) Duplicated addresses (27):

Like, for example, multiple extensions (http://www.google.de, http://www.google.fr, etc.), subpages (http://www.files.wordpress.com) or websites from the same owner (for instance, http://www.bookryanair.com).

b) URL addresses without a website (12):

We have removed here from the sample the URLs with redirections to other domains, or web addresses with marketing and tracking purposes, which do not lead to a website (for example, http://www.itrack.it).
The units analyzed after this filter are, therefore, a total of 461.

2.2. Analysis variables

2.2.1. Degree of compliance

The first objective of this research is to see if most important Spanish websites correctly apply the guidelines on privacy to comply with regulations concerning the use of cookies.

In order to develop this part of the study, we have analyzed the most popular websites in Spain according to Alexa Rank. In this regard we have considered appropriate at first discern the number of cookies used by each website, if they are first or third-party cookies and if the cookies are or not strictly necessary to use the service.

<table>
<thead>
<tr>
<th>Cookies</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>First-party cookies</td>
<td>0</td>
</tr>
<tr>
<td>Third-party cookies</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>0</td>
</tr>
<tr>
<td>Strictly necessary</td>
<td>Yes/No</td>
</tr>
</tbody>
</table>

We have visited the top 500 Spanish websites with each one of the most used web browsers: Internet Explorer v.10, Google Chrome v.28 and Mozilla Firefox v.22.0.

If the website uses cookies which are not strictly necessary to provide the service required by the user, we have analyzed the two variables mentioned in the Spanish regulations:

a) Information: the website has to inform their users clearly about the use and purpose of the data collected.

b) Consent: after providing the information, the user must agree the data collection.

It is important to highlight among 500 most popular websites in Spain, there are a lot of companies and institutions established outside the Spanish and EU boundaries and, therefore, the Spanish legislation does not apply. As a consequence, we have selected from the universe of study those companies and institutions which, as indicated in the «Law of Information Society Services and Electronic Commerce» (LSSI) (Ley de Servicios de la Sociedad de la Información y de Comercio Electrónico)\(^\text{11}\), have a perma-

---

nent establishment in Spain. We have divided this part of the research into six different categories: Internet services, news, businesses, government agencies, online shops and websites with content for adults.

2.2.2. Most used cookies by top Spanish websites

On this second point we have used Mozilla Collusion tool to draw a map of cookies, web beacons and similar technologies used by most visited Spanish websites. We have also used Google Chrome’s Ghostery tool to double check the data reliability.

![Picture 1. Mozilla Collusion Interface](image)

In this section we have considered convenient to analyze the 500 top Spanish websites provided by Alexa Rank (461 after the filter already mentioned), regardless of the companies or institutions nationalities. In our opinion, as they are the most visited platforms in Spain, the chances these cookies are installed on Spanish Internet users’ web browsers are very high. The main objective of this part is to detect and identify the most important Data Collectors operating in Spain.

2.2.3. Most important Data Collectors in Spain

The third part of the study is based on the analysis conducted by The Guardian: *Tracking the trackers: who are the companies monitoring us online?*\(^\text{12}\), developed with the help of its readers, in April 2012. The research, which collects information on more than 7000 websites, aims to identify the main companies that benefit from the information generated by Internet users’ behavior on the Web.

In our case, we have used the tool Mozilla Collusion to analyze the 461 most visited websites by Spanish people, in several waves during the month of September 2013. After identifying the third-companies or Data Collectors, we have managed the

---

data provided to discern how many websites among analyzed use a specific cookie, thus providing statistical information on cookies most used in Spanish cyberspace.

In the next step, we have created a detailed list of the main companies collecting data from the user’s behavior, and tried to answer the following questions:

a) What is it? Which is the company responsible for the cookie?

b) What information does it collect?

c) Are data associated to personal information could identify a concrete user?

d) How long are data stored?

e) Are data sold to third parties?

3. RESULTS

3.1. Degree of compliance

Once established the top 500 most visited Spanish websites and, after performing the filter that we talked about in the methodology (eliminating duplications and ad tracking URLs), the sample is reduced to 461 websites.

Then we have observed the origin of each website to find out if whether or not they are established in Spain. Understandably, we have ignored in this analysis websites that, although very visited by Spanish people, are outside the scope of EU and national legislation.

According to the information obtained, we can highlight a first data: among the most visited websites by Spanish users, only 59% (272) are within the scope of Spanish legislation. The remaining 41% (189 websites, most of them from the U.S.) do not have, according to the current legislations\(^\text{13}\), any legal obligation to require the informed consent to their users.

We have deleted from the universe of study those websites outside the scope of Spanish legislation. As a consequence, we have observed the degree of compliance focusing the analysis on the remaining 272 websites, which need to require informed consent to their users.

The result of the analysis show that 139 of the 272 pages we have tested (51%) correctly apply current legislation on privacy, while 133 (49%) of the analyzed websites do not apply the regulations and do not ask, when it is needed, for the informed consent to their users.

\(^{13}\) As stipulated in the LSSI, Spanish legislation will apply to service providers established in Spain. The company address in Spain must be indicated on their website and has to be able to be checked through Spanish Commerce Register. Spanish law also applies to purchases made to service providers established in another Member States of the European Union or the European Economic Area (EU countries plus Norway, Iceland and Liechtenstein).
At this point, we would like to highlight the fact that most of the 133 Spanish websites which do not apply the regulations, do it because they do not ask for the user’s consent to use their data, although they provide information about cookies. In fact, 70% of analyzed websites that do not comply with the legislation, provide some kind of information about privacy and data protection to their users.

3.1.1. Degree of compliance classified in categories

We have divided the sample into 6 categories:

a) Internet Services (127 websites)
b) News (66 websites)
c) Businesses (34 websites)
d) Government agencies (24 websites)
e) Online shops (15 websites)
f) Adults (6 websites)

a) Internet services

We have selected in the Internet Services category those websites whose products or services are intrinsic to the nature of the Web, like weblogs, hypertext directories, search engines, forums, social network services, marketplace platforms, services based on the Web (like cloud computing) and other. The results show that 72 of the 127 websites analyzed (57%) do not comply with Spanish legislation. On the other hand, 55 websites (43%) correctly apply the legislations.

b) News

Online media use more cookies in average than the rest of websites. This should not be surprising, since the most important cybermedia’s source of revenue is advertising and most of online web metrics and advertising techniques are based on the use of first and third-party cookies. Most visited Spanish websites install an average of 28 cookies in their visitor’s devices (11 first-party and 17 third-party cookies). Most visited Spanish online media, on the other hand, install an average of 54 cookies (14 first-party and 40 third-party cookies). It stands to reason that digital media are more likely to threaten the Internet users’ privacy, so we consider that this category is of particular importance.

According to the results of this research, 62% of the most important Spanish digital media correctly apply the current regulations. There is a 38%, however, do not require the informed consent to their users and, as a consequence, do not comply with the law yet.
c) Businesses

Within the most important corporate websites we have analyzed, we can find i.e. airlines, banks and IT companies. 50% of top business websites in Spain do not apply the e-privacy regulations while 50% correctly inform their users and ask for their consent to collect data.

d) Government agencies

Most of Spanish Government sites do not perform any economic activity and accordingly, do not need to ask for the informed consent to their visitors to collect their data\(^{14}\). However, there are some government platforms that use non-exempt cookies and provide products or services through their websites (via, i.e. an online shop). These sites consequently, are required to comply with the obligations stipulated in the LSSI. There are 24 government websites within the top 500 most visited websites in Spain according to Alexa Rank. 83% of these websites (20) are not involved in any economic activity, so do not need to provide the informed consent to their users. On the other hand, 17% of these government platforms sell products or services online and use non-exempted cookies. However these websites do not correctly inform their users, and do not require the users’ consent to collect their data, so we have to conclude they do not comply with the law.

e) Online shops

We have considered online shop in this research those platforms that allow users to buy products or services online and meet the characteristics inherent to this type of format (i.e. the shopping cart option). We have analyzed 15 sites within this category, including online shops from different industry sectors such as textile, furniture and decoration, food and entertainment. The results show that 67% of the analyzed samples do not apply correctly the regulations, while 33% precisely inform their visitors and customers and ask for their consent to collect data.

f) Adults

We have identified in this study 6 websites with content for adults within the scope of Spanish legislation. Only 1 of those websites (17%) complies with regulations, while the rest (83%) do not.

\(^{14}\) Spanish government agencies and administrations hardly ever have the need to require the informed consent to their users. This is because these websites are normally not involved in any economic activity (such as e-commerce or advertising). As stipulated in the LSSI if no economic activity is performed in the website, there is no need to ask for the informed consent to the users.
3.1.2. Most used cookies by top Spanish websites

With the help of Collusion tool we have generated a map of most frequently used cookies and similar technologies in Spanish websites. There is a high probability, thereby, any Internet user in Spain have some of the following cookies installed in their devices once the navigation is completed.

As mentioned, after filtering the 500 most visited Spanish websites according to set criteria (duplications and ad tracking URLs), the sample is reduced to 461 websites.

Collusion tool detects a total amount of 882 different cookies and similar technologies. In other words, after visiting the 461 most relevant sites in Spain within the same navigation session, the information generated while browsing the Internet is collected by 882 different cookies. These files send the data collected to the companies responsible of their installation and they will be frequently used with commercial or statistical purposes.

The most commonly used cookie is, by far, Google Analytics cookie –google-analytics.com–. It is followed by DobleClick cookie –doubleclick.net– (which is an online advertising company owned by Google), Facebook –facebook.com cookie–, comScore –scorecardresearch.com web beacon– and AppNexus –adnxs.xom cookie–.
Figure 2. Most used cookies by top websites in Spain

We identify below the list of 50 most used cookies in most visited websites in Spain (Table 1).

<table>
<thead>
<tr>
<th>Rank</th>
<th>Cookie name</th>
<th>Number of websites</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>google-analytics.com</td>
<td>340</td>
<td>73.75%</td>
</tr>
<tr>
<td>2</td>
<td>Doubleclick.net</td>
<td>198</td>
<td>42.95%</td>
</tr>
<tr>
<td>3</td>
<td>Facebook.com</td>
<td>166</td>
<td>36.01%</td>
</tr>
<tr>
<td>4</td>
<td>scorecardresearch.com</td>
<td>143</td>
<td>31.02%</td>
</tr>
<tr>
<td>5</td>
<td>google.com</td>
<td>109</td>
<td>23.64%</td>
</tr>
<tr>
<td>6</td>
<td>googlesyndication.com</td>
<td>68</td>
<td>14.75%</td>
</tr>
<tr>
<td>7</td>
<td>ajax.googleapis.com</td>
<td>59</td>
<td>12.80%</td>
</tr>
<tr>
<td>8</td>
<td>fonts.googleapis.com</td>
<td>52</td>
<td>11.28%</td>
</tr>
<tr>
<td>9</td>
<td>google.es</td>
<td>49</td>
<td>10.63%</td>
</tr>
<tr>
<td>10</td>
<td>adnxs.com</td>
<td>46</td>
<td>9.98%</td>
</tr>
<tr>
<td>11</td>
<td>twitter.com</td>
<td>45</td>
<td>9.76%</td>
</tr>
<tr>
<td>12</td>
<td>serving-sys.com</td>
<td>42</td>
<td>9.11%</td>
</tr>
<tr>
<td>13</td>
<td>imrworldwide.com</td>
<td>40</td>
<td>8.68%</td>
</tr>
<tr>
<td>14</td>
<td>Factbook.net</td>
<td>34</td>
<td>7.38%</td>
</tr>
<tr>
<td>15</td>
<td>googleadservices.com</td>
<td>34</td>
<td>7.38%</td>
</tr>
<tr>
<td>16</td>
<td>Quantserve.com</td>
<td>29</td>
<td>6.29%</td>
</tr>
<tr>
<td>17</td>
<td>googletagmanager.com</td>
<td>27</td>
<td>5.86%</td>
</tr>
<tr>
<td>18</td>
<td>cxense.com</td>
<td>25</td>
<td>5.42%</td>
</tr>
<tr>
<td>19</td>
<td>youtube.com</td>
<td>25</td>
<td>5.42%</td>
</tr>
<tr>
<td>20</td>
<td>Chartbeat.net</td>
<td>23</td>
<td>4.99%</td>
</tr>
<tr>
<td>21</td>
<td>Xiti.com</td>
<td>23</td>
<td>4.99%</td>
</tr>
<tr>
<td>22</td>
<td>2mdn.net</td>
<td>22</td>
<td>4.77%</td>
</tr>
<tr>
<td>Rank</td>
<td>Cookie name</td>
<td>Number of websites</td>
<td>Percentage</td>
</tr>
<tr>
<td>------</td>
<td>------------------------------------------</td>
<td>--------------------</td>
<td>------------</td>
</tr>
<tr>
<td>23</td>
<td>smartadserver.com</td>
<td>22</td>
<td>4.77%</td>
</tr>
<tr>
<td>24</td>
<td>criteo.com</td>
<td>21</td>
<td>4.56%</td>
</tr>
<tr>
<td>25</td>
<td>gstatic.com</td>
<td>21</td>
<td>4.56%</td>
</tr>
<tr>
<td>26</td>
<td>wtp101.com</td>
<td>21</td>
<td>4.56%</td>
</tr>
<tr>
<td>27</td>
<td>2o7.net</td>
<td>20</td>
<td>4.34%</td>
</tr>
<tr>
<td>28</td>
<td>twimg.com</td>
<td>20</td>
<td>4.34%</td>
</tr>
<tr>
<td>29</td>
<td>omtrdc.net</td>
<td>19</td>
<td>4.12%</td>
</tr>
<tr>
<td>30</td>
<td>yieldmanager.com</td>
<td>19</td>
<td>4.12%</td>
</tr>
<tr>
<td>31</td>
<td>atemda.com</td>
<td>16</td>
<td>3.47%</td>
</tr>
<tr>
<td>32</td>
<td>dmn506yrbagrg.cloudfront.net</td>
<td>16</td>
<td>3.47%</td>
</tr>
<tr>
<td>33</td>
<td>newrelic.com</td>
<td>16</td>
<td>3.47%</td>
</tr>
<tr>
<td>34</td>
<td>Rubiconproject.com</td>
<td>16</td>
<td>3.47%</td>
</tr>
<tr>
<td>35</td>
<td>Weborama.fr</td>
<td>15</td>
<td>3.25%</td>
</tr>
<tr>
<td>36</td>
<td>bluekai.com</td>
<td>13</td>
<td>2.82%</td>
</tr>
<tr>
<td>37</td>
<td>sascdn.com</td>
<td>13</td>
<td>2.82%</td>
</tr>
<tr>
<td>38</td>
<td>360yield.com</td>
<td>12</td>
<td>2.60%</td>
</tr>
<tr>
<td>39</td>
<td>adroll.com</td>
<td>12</td>
<td>2.60%</td>
</tr>
<tr>
<td>40</td>
<td>betrad.com</td>
<td>12</td>
<td>2.60%</td>
</tr>
<tr>
<td>41</td>
<td>d5nxx8fruw4z.cloudfront.net</td>
<td>12</td>
<td>2.60%</td>
</tr>
<tr>
<td>42</td>
<td>disqus.com</td>
<td>12</td>
<td>2.60%</td>
</tr>
<tr>
<td>43</td>
<td>mathtag.com</td>
<td>12</td>
<td>2.60%</td>
</tr>
<tr>
<td>44</td>
<td>optimizely.com</td>
<td>12</td>
<td>2.60%</td>
</tr>
<tr>
<td>45</td>
<td>Uservoice.com</td>
<td>11</td>
<td>2.39%</td>
</tr>
<tr>
<td>46</td>
<td>gigya.com</td>
<td>10</td>
<td>2.17%</td>
</tr>
<tr>
<td>47</td>
<td>ivwbox.de</td>
<td>10</td>
<td>2.17%</td>
</tr>
<tr>
<td>48</td>
<td>revsci.net</td>
<td>10</td>
<td>2.17%</td>
</tr>
<tr>
<td>49</td>
<td>247realmedia.com</td>
<td>9</td>
<td>1.95%</td>
</tr>
<tr>
<td>50</td>
<td>atdmt.com</td>
<td>9</td>
<td>1.95%</td>
</tr>
</tbody>
</table>

Table 1. Most used cookies by top websites in Spain

As we can see in the chart above, 73.75% of most visited Spanish websites use Google Analytics cookie, 42.95% use DoubleClick Advertising cookie, 36.01% use Facebook cookie and 31.02% use comScore online audience measurement web beacon.

### 3.1.3. Most important Data Collectors in Spain

This classification is based on the *Tracking the Trackers* project, developed by The Guardian, in cooperation with its readers. We have also individually analyzed the information provided on each Data Collector’s websites. We have selected at this point those
companies whose cookies are most frequently used in top Spanish websites: Google, DoubleClick, Facebook and comScore.

a) Google Analytics

**What is it?**

Google Analytics is a free tool that can be used by website publishers to better understand how people are using their website. It is also used by Google to better understand the performance of its own websites.

**What information does it collect?**

According to The Guardian study, Google Analytics cookie stores a unique identifier—so the website can recognize the user if he or she visits the website again—as well as information about the pages the browser visits; when the browser is seen on the website; how long the browser was seen on the website; the IP address (which can allow the Google Analytics to infer the browser's location), and what site the browser was looking at before arriving at the site (the referring url).

By default, this information is shown to website publishers through the Google Analytics tools and is not shared with anyone else. It is a first-party cookie.

**Are data associated to personal information could identify a concrete user?**

Data collected are not associated to any personal information according to Google.

**How long are data stored?**

If the user does not delete the cookies, Google Analytics cookies expire after two years.

**Are data sold to third parties?**

Websites may share anonymous analytics data with Google and other companies.

b) Doubleclick

**What is it?**

Doubleclick is a company owned by Google which business model is based on two sides: online advertisers and publishers. The business operates in three different branches:

1. **Ad-serving:** Online publishers use Doubleclick to display adverts on their websites.
2. **Ad delivery:** Doubleclick let advertisers control how often an ad is shown to a browser, how long it is shown for and how often it will appear.
3. **Behavioural targeting:** It is divided in two categories:
   a) **Targeting for one website owner:** an online publisher can set a Doubleclick cookie to identify the users' favorite sections on the website. Doubleclick will then select the type of adverts the users might like to see according to the browsing information collected. For example, if the user visits the sports pa-
ges of a news website, then adverts for match tickets may be more relevant. This information belongs to the website owner only.

b) Targeting in advertising networks: Google runs a service called Adsense, in which lots of different publishers pool the information they get on browsers. This helps them build up a better idea of the type of adverts someone might want to see. This is a third-party advertising cookie.

What information does it collect?
In their privacy policy, Google explains how data is recorded from a generic Doubleclick cookie. It looks like this:

- time: 06/Aug/2013 12:01:32
- ad_placement_id: 105
- ad_id: 1003
- userid: 0000000000000001
- client_ip: 123.45.67.89
- referral_url: «http://youtube.com/categories»

This data send Doubleclick information about the time and date the user see an advert. It also shows:

- userid: the unique ID number the cookie has given user’s browser
- ad_id: the unique ID of the advert
- ad_placement_id: the ID of where the advert was seen on the site
- referral_url: what page the user was on when he or she saw the advert

Doubleclick can also collect user’s geolocation information via IP address.

Are data associated to personal information could identify a concrete user?
According to Google, data are never associated to any personal information. This would be a breach of Doubleclick’s terms and conditions. The information obtained from the cookies is never combined with information that Google obtains from its other products and services. Users’ browsing behaviour will never be linked to users’ Gmail accounts, for example.

How long are data stored?
Doubleclick cookies on the browser are set to expire after a number of years. However, the override for this is clearing cookies. Newer cookies «60 days in market» and «30 days in market» are actually more valuable to advertisers as they give a better indication of what the person using that browser is interested in right now. IP addresses are anonymized after nine months and the data in cookies is anonymized after 18 months.

Are data sold to third parties?
According to Google and, as highlighted by The Guardian, data are not sold to third parties. When the service is being used by a publisher for its own purposes (not in an ad network), the publisher owns that data, not Doubleclick.
c) Facebook

What is it?

Facebook is the most used social network on the Internet. Facebook has been estimated an average of 750 million unique visitors per month in 2013. It is followed in the same category by Twitter, with 250 million unique visitors per month, and Linkedin with 110 million.

What information does it collect?

The Facebook tracker that appears in our data is not for targeted advertising. In fact, Facebook explicitly told The Guardian researchers it has no need for such a thing – the information its users willingly volunteer on the platform is a far richer resource for advertising.

Facebook’s cookies installed in user’s devices are frequently related to the Facebook’s «social plugins». These are tools that link back to Facebook in some way, such as the «like», subscribe» or «recommend» buttons.

They appear on other websites through the use of «iframes», a very common way of embedding content on to a web page. In order for this to load, Facebook’s servers will know the page, the time and date it was loaded and the browser IP address.

Facebook’s tracking cookies are used in three different ways:

1. If the user does not have a Facebook account and his or her browser has loaded a facebook.com page, then no cookie is set when the user browses a page with social plugins.
2. If the user does not have a Facebook account but has visited a facebook.com page in the past, three cookies are installed. One is for security reasons and the other two are used to track registration effectiveness. If a user later decides to create an account, the aim is to find out what convinced them to do it. This is done by recording the first and last Facebook pages the browser visited.
3. If the user has a Facebook account, a cookie is set on the browser containing a unique ID that relates back to the user’s profile. When the user visits a page with a social plugin, it will check that cookie. If the user is signed in, it will use a unique ID to show you how many of the user’s contacts in the social network have clicked on the like button and whether or not you the user has liked the page.

Are data associated to personal information could identify a concrete user?

If the user clicks on a Facebook «like» or «share» button, this is displayed on the user’s Facebook wall. In order to do this, Facebook has to match this action with the user’s account details. This is done through the log-in cookie.

How long are data stored?

Data associated with a user’s Facebook account is stored for as long as that account is active. When the user deletes an account, it is permanently removed from Facebook.
It typically takes about one month to delete an account, but some information may remain in backup copies and logs for up to 90 days.

**Are data sold to third parties?**

Information about the social plugins the user has clicked will be shown to his or her friends, both on Facebook or on the. Facebook says these information is not shared with anyone else.

d) **comScore**

**What is it?**

ScorecardResearch is part of a company called Full Circle Studies, which is owned by comScore. comScore tracks more than three million unique websites worldwide and its methodology page says it has «approximately two million worldwide consumers under continuous measurement».\(^{15}\) It provides market research data to website owners through a mixture of online surveys and the use of web beacons\(^{16}\).

**What information does it collect?**

ScorecardResearch's privacy policy says its tracking will collect information such as: when a browser visited a website, what page of the website it was, the title of the web page or the IP address.

**Are data associated to personal information could identify a concrete user?**

As with all cookies and web beacons, ScorecardResearch cannot identify an individual user who is using the computer to visit a website that contains a ScorecardResearch tag.

However, the cookie may be used to observe certain types of browsing behaviours, which are then combined with other browser data to give a picture of what people are likely to do when they surf the web.

**How long are data stored?**

The data obtained through ScorecardResearch cookies is kept for up to 90 days.

---


\(^{16}\) «Web beacons, also called web bugs and clear GIFs are used in combination with cookies to help people running websites to understand the behaviour of their customers. A web beacon is typically a transparent graphic image (usually 1 pixel x 1 pixel) that is placed on a site or in an email. The use of a web beacon allows the site to record the simple actions of the user opening the page that contains the beacon. [...]Web beacons are typically used by a third-party to monitor the activity of a site». allaboutcookies.org. *Web Beacons and other Tools*. Retrieved March, 4th, 2014 from: http://www.allaboutcookies.org/web-beacons/.
Are data sold to third parties?

According to comScore, the information collected through the web beacon is analyzed and the resulting reports are shared with their clients.

4. DISCUSSION

As stated in the controversial article published *Giving the Web a Memory Cost Its Users Privacy* published in The New York Times by John Schwartz\(^{17}\), before the creation of cookies, «every visit to a site was like the first, with no automatic way to record that a visitor had dropped by before. Any commercial transaction would have to be handled from start to finish in one visit, and visitors would have to work their way through the same clicks again and again; it was like visiting a store where the shopkeeper had amnesia».

Cookies play an essential role in nowadays web development and are a useful tool to personalize content, helping users to customize their web navigation. These types of technologies can help web editors providing a better navigation experience to their visitors remembering i.e. users’ preferred navigation language or font size, users’ log-in data or making possible e-commerce services such as the shopping cart.

However cookies can be used as well to track users’ behavior on the Internet with commercial purposes, such as web traffic measurement or personalized advertising.

Current Spanish legislation, after EU e-Privacy directive adoption, aims to make a more transparent Web. At the end, every Internet user has the right to known that his or her navigation generates valuable data to advertising companies. It is understandable websites collecting user’s data (like, for instance, cybermedia) may have the obligation to clearly inform about the user’s data they collect and the purposes those data will be used. It is, on the other hand, reasonable that users must give their consent to allow websites to collect their data. Most of cookies we have analyzed in this study are not strictly necessary to provide a communication service and, as a consequence, users should know they must be able to visit the same websites without being tracked.

There are some interesting initiatives in this field like the Do Not Track header which is trying since 2009 to standardize a technology could Internet users’ allow web navigation without being tracked by third companies, although these kinds of projects have been unsuccessful to date.

Maybe privacy is the price we have to pay for Internet free services like search engines, online media or social networks. As stated by Goodson\textsuperscript{18} «If You’re Not Paying For It, You Become The Product». Ramonet highlights in this sense the potential risks of a culture where information is becoming merchandise and social responsibility is subject to market requirements. In Ramonet’s point of view\textsuperscript{19}, one of the most important risks we face in relation to digital information, is the three main functions of mass media (which sociologists have traditionally understood as inform, educate and entertain) can be become monitor, advertise and sell.

We can guess a future in increasingly interconnected and globalized societies. Maybe Internet user’s data safety is one of the greatest challenges we will have to face in the coming years. In our opinion, a greater awareness in this field is required as well as a greater responsibility of the participants, in three different levels: institutions should have a duty to educate citizens about these issues, web editors and data collecting companies should better inform about their tracking methods and take initiatives to advocate for a greater transparency and finally, web users should understand the risks concerning their privacy they assume while surfing the Web and accept the consequences an irresponsible navigation may have for their data privacy.

5. **BIBLIOGRAPHY**


