Measuring Web Cookies in Governmental Websites

Georgios Smaragdakis

Joint work with Matthias Götze, Srdjan Matic, Costas Iordanou, and Nikolaos Laoutaris

Appeared in the proceedings of ACM WebSci‘22
How many ads have you received today?
Web Cookies

Web Browser

Web Server of www.Berlin.com

sessionID = 26062022

sessionID = 26062022
Web Cookies & Trackers

Web Browser

Web Server of www.Berlin.com

sessionId = 26062022

sessionId = 26062022
Third Party Cookies

Web Browser

Web Server of www.Berlin.com

sessionID = BigBortherXXjt

sessionID = BigBrotherXXjt
Tracking Cookies

Web Browser

sessionID = BigBortherXXjt

Web Server of www.Munich.com
Your Digital Soul
Regulatory Intervention: GDPR (et al.)

EU General Data Protection Regulation
25 May 2018

CALIFORNIA CONSUMER PRIVACY ACT

PIPEDA
Personal Information Protection and Electronic Documents Act
GDPR (et al.) and Web Cookies

Under the EU's GDPR, cookies that are not strictly necessary for the basic function of your website must only be activated after end-users have given their explicit consent to the specific purpose of their operation and collection of personal data.
Did GDPR (et al.) make any Difference?

Yes, but the problem still exists!

We Value Your Privacy ... Now Take Some Cookies: Measuring the GDPR's Impact on Web Privacy

Martin Degeling1, Christine Urz1, Christopher Lentzsch1, Henry Hosseini1, Florian Schaub1, and Thorsten Holz2
1Ruhr-Universität Bochum, Germany
2University of Michigan, Ann Arbor, MI, USA
Email: firstname.lastname@rub.de
Email: f.schaub@umich.edu

Abstract—The European Union's General Data Protection Regulation (GDPR) went into effect on May 25, 2018. Its privacy regulations apply to any service and company collecting or processing personal data in Europe. Many companies had to adjust their data handling processes, consent forms, and privacy policies to comply with the GDPR's transparency requirements. We monitored this rare event by analyzing changes on popular websites in all 28 member states of the European Union. For each country, we periodically examined its 500 most popular websites – 6,579 in total – for the presence of and updates to their privacy policy between December 2017 and October 2018. While many websites already had privacy policies, we find that in some
data access, data deletion, and data portability. Even outside Europe, online services had to prepare for the GDPR because it not only applies to companies in Europe but any company that offers its service in Europe. As a result, the GDPR is expected to have a major impact on companies across the world.

Previous work has found that about 70 to 80% of websites in the U.S. have privacy policies [26], [28]. But analysis of privacy policies has been focused on English-language policies, performing in-depth studies on their content [42], [18], [25], [39]. Cookie consent notices have just recently seen increased attention with respect to their readability [44], but their

1 Introduction

When browsing the web, users encounter the so-called "online trackers". They build their business on the massive collection and brokering of personal data. Used by websites to monetize the content they offer via personalized advertisement, online tracking is perceived as a threat to users' privacy [43, 59]. To regulate the usage of tracking technologies in the web, the EU Commission issued already in 2002 the ePrivacy Directive. It includes one of the first and strictest regulations on the usage of online tracking mechanisms [21]. As amended in 2009 [22], it requires websites to ask "prior informed consent for storage or for access to information stored on a user’s terminal equipment". In other words, a website must ask the visitor to authorize the storage and retrieval of data sent through cookies and similar tracking mechanisms before delivering and installing them. Since 2013, the so-called "Cookie Law" provided by the ePrivacy Directive has been adopted by EU Member States [18], [25, 39].
This paper: Have We Audited the Government Websites?

Why Bother?

1. Allows for large-scale surveillance
2. May be the only point of interaction
3. If legislators do not respect laws, it breaks the trust between citizens & authorities
Web Sources: more than 150k Websites
We do NOT give consent
Cookie Expiration Date
### Third party Tracker

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Domain</th>
<th>Path</th>
<th>Expires / Max-Age</th>
<th>S. H.</th>
<th>Sec.</th>
<th>S. Sa.</th>
<th>Part.</th>
<th>Pr.</th>
</tr>
</thead>
<tbody>
<tr>
<td>_ym_sid</td>
<td>60267133165...</td>
<td>mc.yandex.ru</td>
<td>/</td>
<td>2023-01-08T...</td>
<td>58</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>_ym_d</td>
<td>1556180332</td>
<td>government.ru</td>
<td>/</td>
<td>2023-01-08T...</td>
<td>28</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>_ym_isid</td>
<td>46508577615...</td>
<td>yandex.ru</td>
<td>/</td>
<td>2023-01-08T...</td>
<td>9</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>_ym_uid</td>
<td>16159143249...</td>
<td>government.ru</td>
<td>/</td>
<td>2023-01-08T...</td>
<td>15</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yuidss</td>
<td>46508577615...</td>
<td>yandex.ru</td>
<td>/</td>
<td>2023-01-08T...</td>
<td>25</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>yuidss</td>
<td>16159143249...</td>
<td>government.ru</td>
<td>/</td>
<td>2023-01-08T...</td>
<td>26</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Tracker
With cookies
Expiring 1 to 9 years in the future
Gov Websites: Cookies are very common

- South Korea (841): 98.7%
- Saudi Arabia (657): 96.7%
- Indonesia (991): 96.3%
- Australia (14659): 95.6%
- Turkey (2353): 94.6%
- Argentina (849): 94.5%
- Russia (3670): 94.2%
- UK (4986): 92.8%
- USA (24287): 88.6%
- Canada (2366): 88.6%
- France (3863): 87.2%
- Brazil (9461): 86.8%
- Mexico (2228): 85.1%
- South Africa (712): 83%
- Italy (3334): 82.9%
- Germany (4748): 82.7%
- China (1066): 76.1%
- India (24755): 68.9%
- Japan (1590): 65.9%
Gov Websites: Third party and Tracker Cookies are also very common!
Gov Websites: Percentage of Cookies that expire after 1 year or more

- China (279)
- France (705)
- South Korea (98)
- Russia (1561)
- Turkey (205)
- Japan (96)
- Mexico (424)
- USA (5377)
- Australia (1516)
- UK (734)
- Italy (511)
- Brazil (1386)
- Argentina (76)
- India (2004)
- Indonesia (137)
- Canada (510)
- Germany (161)
- South Africa (82)
- Saudi Arabia (70)

Legend:
- Domain known for Tracking
- TP Domain not known for Tracking
Why the Trackers are there?

Unfortunately, no special care is taken for official websites; the stats resemble those of general websites!

Our manual investigation attributes that to:

1. Links to Social Networks
2. Links to Videos
3. Use of public Analytics Tools
4. Use of public Software Libraries
Recomendations

1. Responsible Gov Website Development
2. Gov Cloud
3. Gov Website Audits
4. Education about Web Tracking
Datasets

Will soon be available at:
https://govcookies.github.io/
Acknowledgements

This work was supported in part by the Atracción de Talento grant (Ref. 2020-T2/TIC-20184) funded by Madrid regional government the TV-HGGs project (OPPORTUNITY/0916/ERC-CoG/0003) co-funded by the European Regional Development Fund and the Re-public of Cyprus through the Research and Innovation Foundation, the European Research Council (ERC) Starting Grant ResolutioNet; “Resolving the Tussle in the Internet: Mapping, Architecture, and Policy Making” (ERC-StG-679158), and the EU H2020 Research and Innovation programme under grant agreements No. 871370 (Pimcity).