Journal of Web Engineering, Vol. 8, No.3 (2009) 211-244

A SURVEY OF COOKIE TECHNOLOGY ADOPTION AMONGST NATIONS

ANDREW F. TAPPENDEN JAMES MILLER

Department of Electrical & Computer Engineering, University of Alberta, Canada
{andrewt, jm}@ece.ualberta.ca

Received October 30, 2008
Revised February 11, 2009

This paper presents the results of a novel survey probing the use of cookies with respect to country of origin and related web technologies. A number of significant relationships are established between the origin of the web application and cookie deployment. Cookie usage amongst five popular dynamic web application frameworks is analyzed providing a per-country breakdown of platform adoption and the establishment of a link between dynamic web technologies and first-party and sessional cookies. The prevalence of vendor-specific third-party technologies both globally and within specific countries is studied. Although global leaders emerged, a number of country-specific market leaders were discovered, suggesting that country-specific niche technologies are competing with the globally dominant technologies within specific markets. A large association is identified between third-party persistent cookie usage and a country's e-business environment—the strongest evidence that cookies are an integral part of the global e-commerce environment.

Key words: Internet, Internet Technologies, Web Applications, Cookies
Communicated by: D. Lowe & O. Pastor

1 Introduction

The Internet has quickly become an indispensable medium upon which many facets of modern economies have come to rely. The widespread adoption of the Internet has enabled the emergence of a truly global economy, ubiquitously connecting people and business from all countries. Cookies, one of the fundamental Internet technologies, were indispensable in the establishment of e-commerce, and revolutionized both the way users could interact with web-applications, and the way in which these systems were developed. Cookie technology was initially introduced as a feature within the Netscape Navigator Browser in 1994 due to requests for an internal HTTP state management mechanism. Cookies enabled e-commerce sites to provide a state-based shopping experience to users, e.g. the omnipresent shopping cart metaphor. Although development of these types of applications was possible with previous technologies, Cookies offer a simple robust built-in alternative providing the e-commerce industry with a single technology on which to build applications [1].

Cookies have quickly become ubiquitous amongst web-applications and are used by over 65% of sites according to a recent survey [2]. Cookie technology has become intertwined with most modern online technologies, and is used extensively as the basic building block for state-based web-applications. Despite this prevalence, very little research exists which specifically investigates cookies and their relationship with the technological aspects of web-applications. This paper will present the

first study that directly compares cookie usage and country of origin, providing a unique picture of the technological and economic environments associated with each country. The results of this survey establish a significant relationship between the deployment of cookies by web applications and the maturity of a country's e-commerce environment.

The remainder of the paper will be organized as follows: Section 2 will provide a brief overview of the current state-of-the-art regarding cookies and Section 3 will discuss the related work and motivations for the investigation. Section 4 will present the survey methodology and specific research questions. Sections 5 – 7 present the results of the survey and Section 8 summarizes the findings, providing conclusions and highlighting the most important results obtained within the survey.

2 Cookies: A Brief Introduction

The HTTP cookie (cookie) was first introduced as a feature within the initial public release of Netscape Navigator in September of 1994 [1]. Since its release, cookies have become an integral part of many web-based applications, and are formally known as the *HTTP State Management Mechanism* [1, 3, 4]. Cookies have existed in a state of vendor-specific implementations, and while attempts have been made to standardize the usage of this technology across technological platforms—RFC 2109 [3] and RFC 2965 [4]—standard compliance has yet to be achieved. Recently, Tappenden and Miller proposed a context-free grammar defining the current state of cookies [5]. It is believed that this definition is the most encompassing definition of a cookie in the literature; the definition is provided in Figure 1.

```
Set-cookie   = "Set-Cookie: ", cookie;
cookie       = name, "=", value, {"; ",cookie-av};
name         = attribute;
cookie-av    = ("comment=",value)
             | ("domain=", value)
             | ("max-age=", value | "expires=", date)
             | ("path=", value)
             | ("secure")
             | ("version=",{DIGIT}+)
             | ("httponly");
             (* only one of each attribute=value (av) can be included *)
```

Figure 1. Definition of the Set-Cookie header (Tappenden et al.[5])

Cookies can be divided into several categories based upon the various *attribute=value* pairs provided in their definition. The two most common categories of cookies are sessional versus persistent cookies and first- versus third-party cookies.

Sessional and persistent cookies are derived from the expiration dates associated with a cookie, the *max-age=value* and/or *expires=date* attributes. Any cookie defined to exist until a predefined future date or period of time is considered persistent. Conversely, any cookie not defined to exist into the future, or defined to exist into the past, i.e. a negative *max-age* value or past expiration date, is considered sessional.

First- and third-party cookies are derived from the nature of the site from which the cookies were deposited—the *domain=value* attribute. If the host from which the cookie originates *matches*[a] the URL of the site the user has initially requested, a cookie is first-party; otherwise, if the host does not *match* the requested URL a cookie is third-party. Third-party cookies have long been identified as a source of concern from a privacy viewpoint [1], but despite these concerns third-party cookies remain a crucial part of many online endeavours.

While many other categories of cookies can be extracted from the specification, third-party persistent cookies have received heightened interest due to the privacy concerns surrounding them. This category of cookies is formed by the intersection of third-party and persistent cookies and is widely used to track users movements across various online sites and provide user-centric advertisements.

Cookies are set by the Set-Cookie header within any HTTP response, or through the *document* object within JavaScript. Objects and scripts embedded within web applications such as images and JavaScript components often set cookies, making the task of collecting cookies non-trivial. In order to obtain all of the relevant cookies for a single web page, the entire page and all referenced objects must be loaded, and in the case of scripts, must be executed. Because of this complexity, the task of surveying the Internet for Cookie usage requires that the survey completely load and execute each member of the survey population.

3 Related Work & Motivations for Research

As the primary state token for many applications, cookies serve as the glue holding internal components together. In addition to this fundamental operation, cookies are increasingly utilized in more sophisticated situations such as for compiling user statistics. This type of operation often involves multiple hosts and spans a time period of several years. This usage of cookies is often combined with equally sophisticated third-party web applications, whose execution is hidden from end-users. Although third-party cookie deployment raises a number of privacy issues, it has become commonplace on the Internet. An increasing number of privacy-conscious users routinely delete cookies, which threatens the validity of the statistics collected by web applications [6]. Furthermore, this habitual removal of cookies was found to affect first- and third-party cookies equally [6], posing a threat to all web applications utilizing cookies, not only advertising and traffic management applications. In addition to users actively removing cookies, third-party software agents are being developed that actively reject all third-party cookies, and selectively reject first-party persistent cookies on the basis of perceived functionality (HTTP response) [7]. These trends toward selective cookie rejection pose a significant challenge to the development of modern web applications. Understanding the deployment trends of cookies within the Internet and specifically within the geographical region in which a web application is to be deployed is of increased importance due to the emergence of habitual and selective cookie deletion.

Several recently studies of cookie usage have been conducted. Security Space produces a monthly cookie usage report [8], but explicitly states that the survey only considers the request for the HTML document—it does not load or execute any images, scripts or other objects embedded within the page.

---

[a] For more information on *domain matching* refer to RFC 2109 [3] and RFC 2965 [4].

Because of this omission, the results obtained from the Security Space survey do not accurately reflect the current technological landscape of modern cookie deployment within complex multifaceted web applications. To further understand the ways in which cookies are currently being deployed across the Internet, a survey that fully loads and executes all objects within a webpage is required.

In a previous study, Tappenden and Miller [2] corrected the omissions of the Security Space cookie report providing an accurate portrayal of cookie deployment by web applications. By completely loading and executing each page surveyed, the survey was able to provide a complete picture of cookie usage, including both first- and third-party cookies as opposed to the Security Space report which can by design only detect first-party cookies set in response to the initial GET request. The Tappenden et al. study provided novel insights into cookie usage, finding an increased rate of cookie usage amongst web applications.

It is through the union of the Tappenden et al. [2] data-set with the geographical locations of each surveyed site that the conclusions in this survey are elicited. While the previous study focuses primarily upon the occurrences of various classifications of cookies, first-party, third-party, etc., this novel study focuses exclusively on the technological implications of cookie usage from both a global and national perspective. From these perspectives we will examine the usage of cookies by the most common web application frameworks such as ASP, ASP.Net, JSP, PHP, and Cold Fusion, third-party online tracking and advertising technologies, third-party web analytics (internal web application tracking), and the metric of cookie deployment statistic as a proxy for the maturity of a nations e-commerce environment. Each of these perspectives provides unique insights into the deployment of this often ignored technology.

Other studies have explicitly examined the usage of dynamic web technologies such as ASP, Cold Fusion, JSP, and PHP. A recent study conducted by Port80 Software examined the HTTP Header signatures from the Fortune 1000 Companies websites [9]. While this study provides insight into the distribution of dynamic web frameworks within the web, the sample size is limited and the technology detection mechanisms heavily favour dynamic-web technologies that come bundled with backend server software, such as ASP and JSP. Another relevant survey conducted by Doyle and Lopes examine the breadth of dynamic web technologies currently employed by the web community [10]. While the survey provides an extensive list of potential technological platforms, it provides no insight into the deployment of these technologies.

Web application and development practices continue to mature and are under constant evolution. Since its inception in 1994, the cookie has played a crucial role in the development of the current e-commercial technological landscape. While much is still changing within the web development paradigm, such as the recent developments of frameworks such as AJAX, cookies remain fundamental to web applications. For example, consider Google's Gmail web application [11]. Despite its heavy reliance on the AJAX framework, Gmail was still observed to set 21 cookies. Clearly, cookies can be ubiquitous even within an AJAX application. This type of cookie usage, observed across a wide array of web applications [2], is becoming common-place within modern applications. As these applications continue to mature, cookie usage is expected to remain vital, as it is crucial these applications work across a wide-array of system configurations and browser versions. Despite the heterogeneous nature of end-users web platforms, cookies remain a common denominator amongst all browsers [5]. Given this reality, future web endeavours must thoroughly understand the technological landscape in which

web applications exist. Cookies and web application frameworks are one crucial component of this landscape, and understanding the technological implications and conventions of cookie usage can be crucial for web developers. Especially given the various privacy laws and sensitivities of different cultures and nations, understanding the acceptable usage of cookie technology within an application's target audience is crucial. This survey provides an accurate portrayal of the current status of cookie usage globally and within specific countries. This analysis is of benefit to any who seeks to produce web applications, and anyone who seeks to understand the technological landscape in which development products or techniques are to be applied.

4 Survey Methodology

4.1 Research Questions

An explosion in the number of global Internet users has been observed, especially within regions in which the Internet has had little penetration [12]. Since 2000, Internet penetration has risen over 300% globally, and over 1000% in the least penetrated regions, Africa and the Middle East [12]. Given the increasingly global reach of Internet-based opportunities, many e-commerce companies are interested in targeting a truly global audience. These opportunities, unique to the Internet, present the ability to create a multi-national web presence with little overhead. Concepts such as the *Internationalisability* of a software application are gaining momentum and becoming a principle concern for any global web application [13]. This heightened interest in global web applications requires an increased knowledge of the global technological landscape, especially with regard to the technological platforms and normal usage patterns of web-specific technologies such as cookies. To this end, a number of specific research questions were selected to provide a better understanding of the usage of cookies both globally and within specific nations.

The following research questions motivate this study:

- **Q1.** *Are cookies equally utilized across the globe?*
- **Q2.** *Given the explicit privacy concerns related to third-party cookies, is the usage of this technology consistent between nations?*
- **Q3.** *Are cookies set to expire as soon as possible (sessional vs. persistent), and is this constant across nations?*
- **Q4.** *Which web technologies dominate, and do regional variations exist?*
- **Q5.** *How prevalent are online advertisers (using third-party persistent cookies as indicators), both globally and amongst nations?*
- **Q6.** *How are cookies used for tracking users' movements within a site, both globally and amongst nations?*
- **Q7.** *Does cookie deployment directly correlate with the maturity of a country's e-commerce environment?*

These questions directly relate to the development of current and future web applications. A thorough understanding of the best industrial practices pertaining to cookie usage globally, and within the

country for which a specific product is targeted, can be of substantial benefit in the development of web applications. The knowledge of empirical cookie usage is also of benefit to the larger web engineering community who seeks to advance the current state-of-the-art for these systems, specifically those who seek to produce high-quality support systems for web-based software endeavours. To answer these questions a survey of the most popular 100,000 sites was overlaid with the geographical origins each site, producing a pool of over 97,000 data points containing both the cookies set through the HTTP headers, images and scripts, and the geographical location from which the site originated. This pool provided ample data from which the seven research questions could be addressed.

4.2 Site Selection

The Alexa top 100,000 list [14] was used as the starting point for the survey of cookie usage in web applications. This list was selected for two primary reasons. First, the Alexa top 100,000 list represents a very large sample population, which allows for an accurate and representative sample of current Internet sites. The second reason is that the list represents the most popular websites, based upon traffic information [15]. This "popularity" measure is based upon both the page *views*, number of pages viewed on a host, and the *reach* or number of different users who access a host [15]. The Alexa ranking is a based on the geometric mean of the *reach* and *views* quantities. Essentially, this ranking provides a list of the most popular 100,000 sites accessed on the Internet, comprising the vast majority of websites online users are likely to encounter. According to Alexa [15], websites excluded from this list have less than a 0.00125% chance of being accessed by an average Internet user. Hence this list may be considered a highly representative sample of the *usable* Internet. While the sampling bias of the Alexa rating is well documented [16], it remains in use because it is the best of the currently available web ranking data[16]. Furthermore, the Alexa top 100,000 list has been used in other academic studies, such as the investigation undertaken by Reay, Beatty, Dick and Miller [17] to study the adoption of the P3P privacy policy.

4.3 Firefox Survey Extension

Due to the large sample population (100,000 sites), the survey required an automated data gathering mechanism. Hence, an extension for the Mozilla Firefox Internet browser [18] was produced. The extension accepts a list of URLs to survey, visits each URL and fully loads each page including associated images and JavaScript components. After the page is finished loading, the entire set of cookie information is recorded and stored within a database for future analysis. The data collected regarding individual cookies includes name, value, host, path, secure connection required and expiry date. These fields all directly correspond to the specification provided in Section 2. Additional information, such as creation time and associated URL was also recorded for each cookie.

To further investigate cookie usage by web applications within specific countries, each site surveyed was mapped to the country in which the domain name was registered. The geographic location of these websites was derived by determining their IP address via the *host* program, then comparing this address with a database of locations (i.e. countries) purchased from IP2Location [19]. This approach was required since the country code top-level domain is not a reliable indicator of the actual geographical location. Of the 98,004 sites surveyed, 130 nations were identified as the origin of

97,050 sites. The remaining 956 origins could not be determined. This represents less than 1.0% of the total population, and is within in the error rate associated with the IP2Location mapping [19].

4.4 Survey Implementation

Although the process was automated, supervision of the system was required due to the inability of the extension to deal with interactive client-side GUI events, such as dialog boxes. If a site were to prompt the user with a dialog box, the automated process would stall, and require external input to continue; as a policy, 'OK' was selected for all dialog boxes, and 'CANCEL' was selected for all login prompts. Other supervision was required in the case of sites crashing the web browser. If this occurred, the survey was re-started at the URL of the site that crashed the server, and the survey continued. Sites that could not be surveyed due to re-occurring crashing are not included in the results.

4.5 Analysis Tools & Statistical Tests

Due to the large pool of data gathered by the survey, SPSS was used to analyze the experimental data and generate box plots. Because of the non-parametric and ordinal nature of the data, the Mann-Whitney U test was utilized to confirm the existence of significant differences between the medians of various data populations. The associated *null* hypothesis is that the two samples are drawn from the same population, i.e. populations that have equal probability distribution functions. All null hypotheses presented within the paper were rejected if the significance of the result was below the standard Type 1 error rate, 0.05.

To further characterize the differences verified between sample populations, an effect size was calculated as an estimate of the size of the difference between the two populations. Because the data is non-parametric, Cliff's δ [20, 21], a non-parametric effect size estimate, was used to characterize the difference between two populations. Cliff's δ examines the probability that individual observations within one group are likely to be greater than the observations in the other group, given by the following equation

$$\Delta = Pr(x_{i1} > x_{j2}) - Pr(x_{i1} < x_{j2})$$

where x_{i1} is a member of population one and x_{j2} is a member of population two. This effect size has been empirically demonstrated to be superior to Cohen's *d* and Hedges' *g* when the data is non-parametric [22, 23]. Essentially, this approach considers the ordinal, rather than the interval, properties of the data. The sample estimate of this statistic, Cliff's $\hat{\delta}$, is obtained by comparing each of the values within in one group to each in the other. The calculation of this sample statistic is given by

$$\hat{\delta} = \frac{\#(x_1 > x_2) - \#(x_1 - x_2)}{n_1 n_2}$$

where x_1 and x_2 are the individual members of each sample population; and n_1 and n_2 are the number of individuals within each sample. Cliff's δ represents the degree of overlap between the two distributions and unlike Cohen's *d*, Cliff's effect size is bounded in the range [-1, 1] and takes the value of zero if the two distributions are identical. To provide a linguistic interpretation of these differences, we will adopt the suggested values of Romano et al. [24] and interpret a $\hat{\delta} > 0.147$ as *small*, $\hat{\delta} > 0.330$ as *medium*, and $\hat{\delta} > 0.474$ as *large* differences (equivalent to the levels suggested by Cohen [25] of 0.20, 0.50 and 0.80 respectively).

Two alternatives exist for calculating the Cliff's delta associated variance. The consistent estimate of variance[b] [21] is presented, allowing for the construction of two asymmetric confidence limits at the 95% confidence level around the $\hat{\delta}$ value. This estimate of variance, as stated by Cliff [21], produces highly conservative confidence intervals and hypothesis testing should not be based upon these estimates. Despite Cliff's recommendation, the effect size was utilized to compare the differences existing between groups and provide a metric for determining the relative size of the differences between populations. The risk that the tests did in fact measure beyond the 95% level is considered an acceptable risk given the nature of the analysis. For the remainder of this paper, the following interpretation will be adopted: if 0 is included within the confidence interval surrounding Cliff's $\hat{\delta}$, then the populations are considered equal; if the confidence interval is only negative then *Group 2 > Group 1*; if it is only positive then *Group 1 > Group 2*.

Finally Spearman's ρ, a non-parametric test, was selected to calculate the correlation between the ordinal ranking of each nation and the cookie usage within the nation. This correlation was chosen because of the ordinal nature of the ranking against which the dataset was compared. To further facilitate the analysis, the ρ values were converted into an equivalent Pearson's *r* correlation coefficient as prescribed by Gilpin [26]. These coefficients were then compared and attributed a measure of effect size—*small, r* = 0.1*; medium, r = 0.3;* and *large*, *r* = 0.5; according to Cohen [25, 27].

5 Global Cookie Usage

The survey outlined in Section 4 was quite prolific in the generation of usable data from which the research questions could be effectively answered. The survey generated the cookie-usage and geographical origins of over 97,000 sites with the identification of 130 nations as the origins of these sites. Given the sheer volume of data extracted by this investigation, a smaller subset of data was selected for in-depth analysis within this paper; however, a complete summary of the survey data is provided at the end of the paper in Appendix I and Appendix II.

The appendices present the data from two perspectives: Appendix I provides a breakdown of the number of sites surveyed for each country, and Appendix II focuses on the number of cookies set by sites originating from each country. Each appendix looks at six distinct cookies classifications: all, first-party, third-party, sessional, persistent and third-party persistent cookies. It is clear from the appendices that the dataset is quite large and a large discrepancy exists in the number of sites surveyed between nations. Due to this discrepancy and the sheer size of the dataset, smaller subsets were selected and analyzed. The remainder of this section will focus upon the nations from which at least 500 sites were surveyed. This subset contains 18 nations and represents the source of over 90% of the total number of sites surveyed.

5.1 Number of Sites Surveyed per Country

The number of sites surveyed per nation with at least 500 sties is presented in Figure 2; clearly the majority of the sites surveyed (44,673) had origins within the United States (US) with the next closest

[b] Kromrey et al. [22] empirically demonstrated that the choice of variance procedure is relatively unimportant across a wide range of circumstances.

country China, with less than half the number of sites (17,196). Due to the large discrepancy in populations per country, the analysis will focus upon values normalized against the number of sites surveyed for each country.

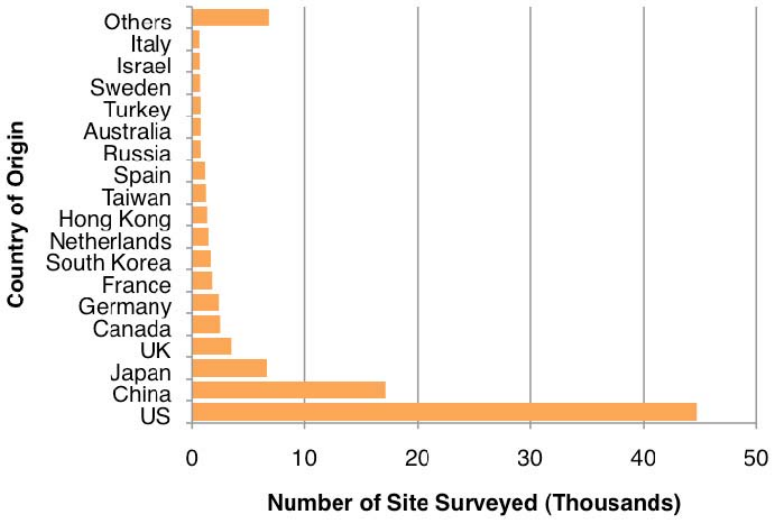


Figure 2. Country Frequency Histogram

5.2 Cookie Usage Within Each Country

The countries identified within the survey were analyzed from six perspectives, overall cookie usage, first-party, third-party, sessional, persistent, and third-party persistent cookie usage. This data is presented with respect to the mean number of sites utilizing the specific cookie categories per country in Figure 3. When comparing the countries that were the origin of at least 500 sites, it is observed that Russia, Japan and Hong Kong are outliers. Russia is observed to have a higher than average number of third-party and persistent cookies; Japan and Hong Kong are observed to have a lower than average usage of cookies. Furthermore, Japan is also significantly below average for the usage of sessional cookies, whereas the mean of Hong Kong is within the 95th percentile of sessional cookies. Both Japan and Hong Kong are within the 95th percentile first-party, third-party and persistent cookies.

To further investigate cookie usage amongst these nations box-plots for each of five types of cookies are presented in Figure 4a-f. It is clear from the shape of each of the box-plots, and confirmed by the Kolmogrov-Smirnov tests presented in Table 1, that the distribution of number of cookies per site is not normal within each country. Due to the lack of an identifiable distribution of cookie usage per country, the remainder of the paper will present a number of non-parametric statistical tests, as these tests do not rely upon the assumption that the sample is drawn from a known distribution.

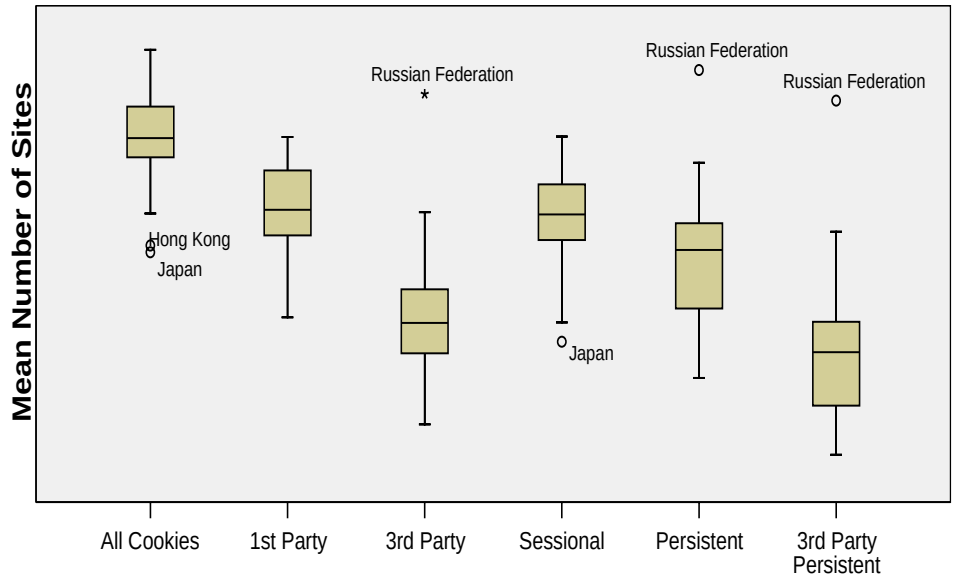
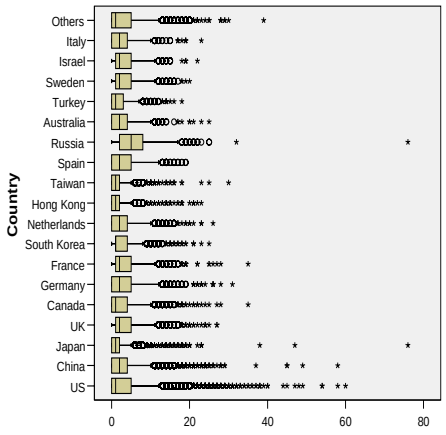


Figure 3. Mean number of sites using specific types of cookies per country

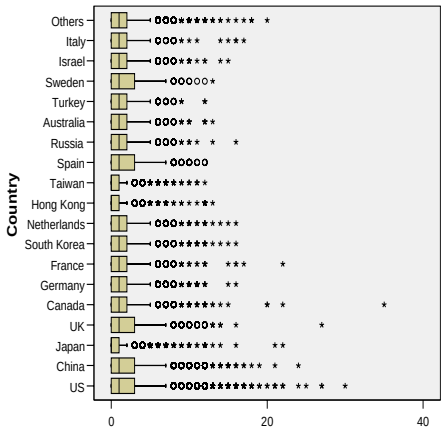
Table 1. Kolmogorov-Smirnov[c] tests for normality

	All Cookies		First-Party Cookies		Third-Party Cookies		Sessional Cookies		Persistent Cookies		Third-Party Persistent Cookies	
	K-S	Sig.	K-S	Sig.	K-S	Sig.	K-S	Sig.	K-S	Sig.	K-S	Sig.
US	.230	.000	.259	.000	.316	.000	.267	.000	.263	.000	.352	.000
China	.211	.000	.247	.000	.336	.000	.263	.000	.289	.000	.454	.000
Japan	.281	.000	.325	.000	.376	.000	.341	.000	.311	.000	.436	.000
UK	.187	.000	.239	.000	.294	.000	.262	.000	.228	.000	.340	.000
Canada	.217	.000	.253	.000	.335	.000	.257	.000	.257	.000	.375	.000
Germany	.219	.000	.285	.000	.290	.000	.251	.000	.258	.000	.336	.000
France	.191	.000	.272	.000	.261	.000	.281	.000	.208	.000	.253	.000
South Korea	.240	.000	.271	.000	.361	.000	.269	.000	.336	.000	.423	.000
Netherlands	.211	.000	.282	.000	.286	.000	.261	.000	.240	.000	.319	.000
Hong Kong	.297	.000	.303	.000	.428	.000	.333	.000	.378	.000	.477	.000
Taiwan	.295	.000	.312	.000	.422	.000	.294	.000	.392	.000	.477	.000
Spain	.194	.000	.278	.000	.307	.000	.252	.000	.246	.000	.348	.000
Russia	.151	.000	.239	.000	.195	.000	.302	.000	.145	.000	.156	.000
Australia	.212	.000	.237	.000	.315	.000	.247	.000	.255	.000	.360	.000
Turkey	.265	.000	.315	.000	.347	.000	.267	.000	.329	.000	.412	.000
Sweden	.185	.000	.244	.000	.284	.000	.251	.000	.217	.000	.314	.000
Israel	.211	.000	.252	.000	.320	.000	.229	.000	.261	.000	.377	.000
Italy	.208	.000	.275	.000	.298	.000	.259	.000	.244	.000	.325	.000
Others	.210	.000	.278	.000	.310	.000	.256	.000	.254	.000	.356	.000

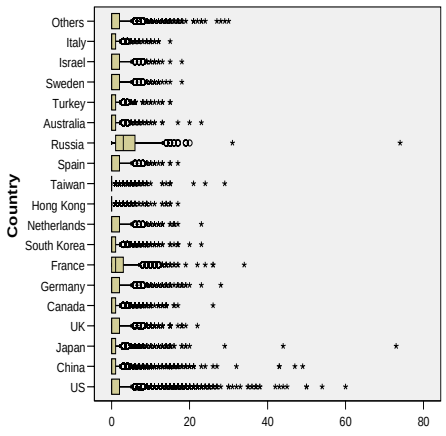
[c] with Lilliefors Significance Correction



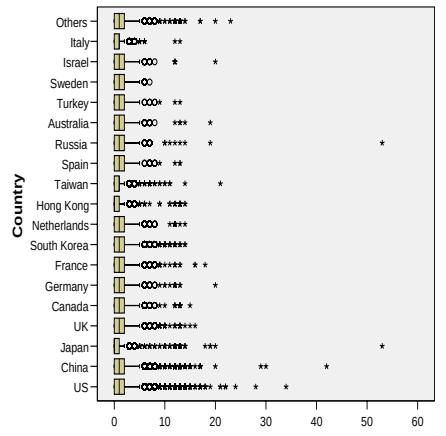
(a) **Overall Cookie Usage per Country**



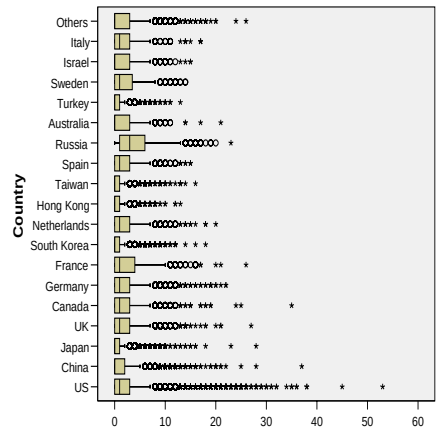
(b) **First-Party per Country**



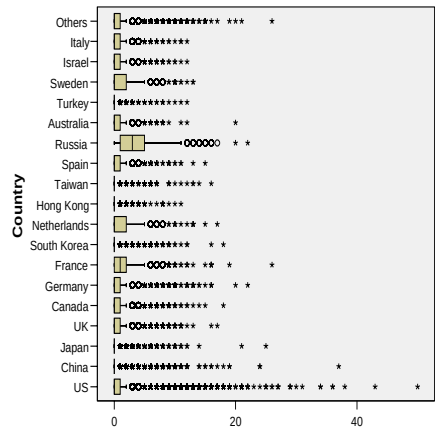
(c) **Third-Party per Country**



(d) **Sessional per Country**



(e) **Persistent per Country**



(f) **Third-Party Persistent per Country**

Figure 4. Cookie Usage per Country

As highlighted in Figure 3, sites originating from Russia use significantly more cookies than sites originating from the other countries. This trend is evident in Figures 4a-f, as the median number of sites from Russia using persistent cookies exceeds that of any other country, and is equal to or greater than the 75th percentile of all other countries with the exception of France. Similarly to persistent cookies, third-party cookies usage within Russia is also much more prominent, and the median of sites originating from Russia with respect to third-party cookie usage is only equalled by the 75th percentile of that of France, in fact all other countries have a 75th percentile less than the median of Russia. This increased usage of both third-party and persistent cookies can be attributed to the significant disparity in the frequency of third-party persistent cookies within sites originating from Russia, attested to by Figure 4f. The median frequency of Russian sites using third-party persistent cookies surpasses all other countries, and is above the 95th percentile of all but three countries—France, the Netherlands, and Sweden. Although Russia is observed to have a higher than average overall cookie adoption frequency, these effects are not consistent with respect to first-party and sessional cookie usage, shown in Figure 4b and Figure 4d respectively, where very little difference between Russian sites and sites from other countries exists.

Table 2. Mann-Whitney U test for Russia vs. all other countries

	All Cookies	First-Party	Third-Party	Sessional	Persistent	Third-Party Persistent
Mann-Whitney U	2.29E+07	3.53E+07	1.80E+07	3.47E+07	2.00E+07	1.64E+07
Z	-18.952	-2.806	-28.751	-3.534	-23.876	-33.183
Asymp. Sig. (2-tailed)	0.000	0.005	0.000	0.000	0.000	0.000
Cliff's $\hat{\delta}$	0.388	0.056	0.518	0.069	0.464	0.561
Upper Limit	0.350	0.036	0.482	0.031	0.428	0.525
Lower Limit	0.424	0.035	0.553	0.108	0.499	0.595

The Cliff's $\hat{\delta}$ effect-sizes and associated confidence intervals are presented alongside the Mann-Whitney values presented in Table 2 and in Figure 5. It is clear from the effect-sizes that the largest disparity exists between Russia and the world with respect to persistent ($\hat{\delta}$ = 0.464), third-party ($\hat{\delta}$ = 0.518), and most notably third-party persistent cookies ($\hat{\delta}$ = 0.561). Much smaller differences exist amongst all ($\hat{\delta}$ = 0.388), first-party ($\hat{\delta}$ = 0.056), and sessional ($\hat{\delta}$ = 0.069) cookies. The difference with respect to sessional and first-party cookies does not warrant the classification of *small*, leading to the conclusion that the *medium* difference observed amongst overall cookie usage is due to the *large* differences existing between Russia and other countries with respect to third-party persistent cookies. A ripple affect of this difference is observed in the *medium* and *large* differences existing in persistent and third-party cookies respectively, as third-party persistent cookies is defined as *third-party* $\cap$ *persistent*. Clearly sites originating in Russia are utilizing third-party persistent cookies at a much higher rate then other countries, these effects will be studied further in Section 6.2.

In direct contrast to the results observed for Russia, sites originating from Japan, Hong Kong, and Taiwan are observed to be consistently below average with respect to cookie usage. This discrepancy is most pronounced with respect to first-party and sessional cookies, where the 75th percentile frequency for Japan, Hong Kong and Taiwan are all equal to that of the median frequency for all other

countries with the exception of the Netherlands. While it may be tempting to view this as a geographical issue, China does not follow this trend and South Korea posts a mixed set of results.

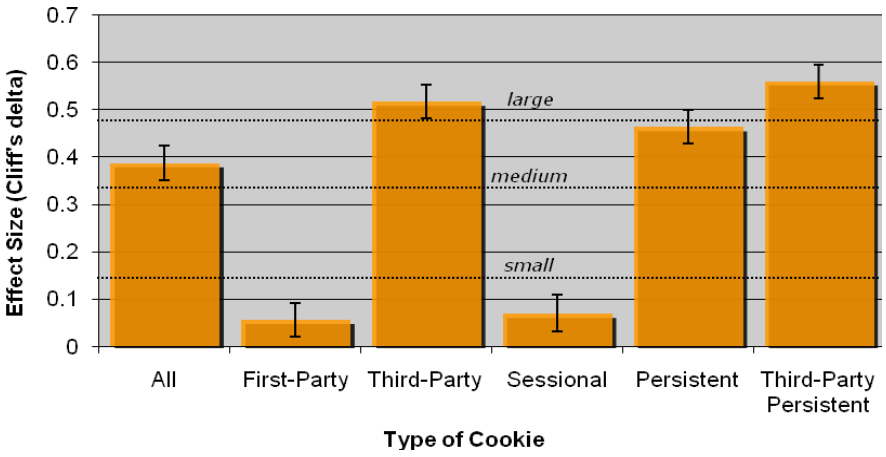


Figure 5. Relative effect sizes for Russia vs. the world

To further characterize these differences, Cliff's $\hat{\delta}$ effect-sizes and associated confidence intervals were calculated for each of the eastern Asian countries[d] with the results presented in Figure 6. Hong Kong has the overall lowest cookie usage amongst all of the countries surveyed, as shown in Figure 6, this decreased overall level is due to less than average cookies usage with respect to all five categories of cookies. Despite this general decrease in cookie usage, Hong Kong is not the minimum within each category of cookies. Japan, for example, has lower sessional cookie adoption and the second lowest overall cookie adoption rate among the countries identified as the origins of at least 500 sites surveyed. Unlike Hong Kong, this *small* difference is not as widespread across all five types of cookies and is seen as the results of decreased usage of first-party and sessional cookies. Taiwan, on the other hand, is observed to have lower overall cookie adoption; however, this is primarily attributed to the *small* decrease in third-party persistent cookie adoption.

China and South Korea present very different results, neither presenting a *small* difference in overall cookie usage, in fact the Cliff's $\hat{\delta}$ intervals for South Korea with respect to overall cookie usage encompass zero, indicating that with respect to overall cookie usage South Korea is similar to that of the rest of the world. Because of the relationship between categories, *cookies = (sessional* $\cup$ *persistent) = (first-party* $\cup$ *third-party)*, the equality observed between the global distribution and that of South Korea, with respect to all cookies, manifests itself as symmetric positive and negative disparities with respect to sessional/persistent and first-/third-party cookies. China also presents a mixed set of results; however, these results are not as symmetric as those of South Korea. With respect to third-party persistent cookies, China has a larger relative disparity than South Korea; this is very interesting especially considering the relatively large disparity amongst persistent cookies in

[d] Country classifications provided by the United Nations Statistical Division [28].

South Korea, suggesting that persistent cookie usage in South Korea is distributed more evenly between first- and third-party cookies than in China.

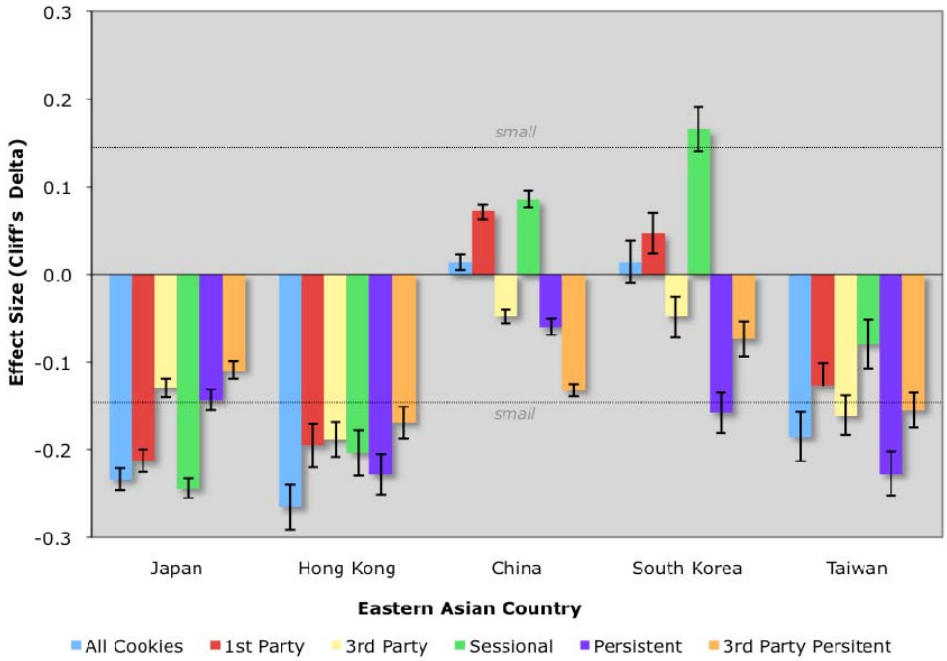


Figure 6. Eastern Asian countries vs. the world

It is clear from these findings that the usage of cookies differs greatly based upon the geographical region in which the site originates, providing a clear answer to research question *Q1*—cookies are utilized across the globe, as cookies were found to be used in every country for which more than four sites were surveyed. Although cookies are utilized globally cookie deployment was observed to vary greatly depending on the nation from which a site originates. An answer to research questions *Q2* and *Q3* are clearly illustrated by the dichotomy of cookie usage from sites originating in Russia versus those originating in Southeast Asia, and specifically Hong Kong—the distribution of first- versus third-party cookies and sessional versus persistent cookies are not consistent across nations. While numerous reasons exist for this variance, including public perceptions of cookies, cultural sensitivities to privacy evasive technologies, the legal obligations and responsibilities of web applications, and the maturity of the e-commerce environment, it is important that anyone seeking to develop web applications within these nations understand these factors in order to correctly situate the software endeavour for success.

6 COTS Cookie Deployment

From the cookies collected by the survey two distinct groups emerged, those used as part of a dynamic web application framework, and those used for online tracking and advertising. Although these cookies emerged from all categories of cookies, certain specific types of cookies were strongly related to specific technologies. The discussion of these results will be partitioned into two sections: a

discussion of dynamic web technologies and a discussion of online tracking and advertising technologies.

6.1 Dynamic Web Application Frameworks

The survey was able to identify five cookies associated directly with the state management mechanisms for five popular web application frameworks: ASP, PHP, ASP.NET, Java Servlet (JSP & J2EE), and Cold Fusion. Each of these development environments provide an internal state management mechanism relying heavily upon cookie technology. Usage of these technologies is easily identifiable by the presence of the associated technology-specific cookie: *ASPSESSIONID%*[e], *PHPSESSID, ASP.NET_SessionId*, *JSESSIONID*, and *CFID*, respectively. A closer inspection of the associated cookies, summarized in Table 3, reveals that the vast majority of the cookies were first-party, and all but one of the technologies (Cold Fusion) exclusively used sessional cookies. Cold Fusion uses a persistent cookie as the unique identifier for a user across a web-session. This cookie, unlike the cookies from the other frameworks, is not set to be removed when the browser is closed, rather the majority of *CFID* cookies (81.9%) were set to be stored in the browser for at least 30 years (10,950 days). The use of persistent cookies in association with dynamic web application frameworks is clearly in the minority, as the four highest occurring associated cookies were all nearly exclusively sessional. Although the majority of dynamic technology cookies were first-party, a percentage of the cookies were associated with third-party hosts, suggesting that third-party providers are using theses dynamic web technologies in the provision of third-party services. The extent to which these services are of value to the end-user and/or the service provider cannot be assessed from these findings, due to the plethora of specific third-party advertising and user-tracking services available on the Internet, it is assumed that these cookies represent more direct end-user service provision.

Table 3. Dynamic web application frameworks associated Cookies

Dynamic Web Technology	Cookie Name	Occurrence (Number of Sites)		First-Party (Number of Sites)		Sessional (Number of Sites)	
ASP	ASPSESSION%[e]	20875	21.3%	16917	81.0%	20874	99.9%
PHP	PHPSESSID	7379	7.5%	6028	87.1%	7052	95.6%
ASP.NET	ASP.NET_SessionID	5332	5.4%	3623	68.0%	5322	100%
JSP	JSESSIONID	5112	5.2%	3678	71.2%	5085	99.5%
Cold Fusion	CFID	1528	1.6%	1331	87.1%	1528	15.8%

All four dynamic technology cookies were found to be within the 99.9th percentile of reoccurring cookies amongst sites surveyed when ranked by cookie occurrence. In fact, ASP cookies where the most frequently retrieved cookies in this survey. PHP, ASP.NET and JSP cookies where ranked sixth, seventh, and eighth most occurring cookies. The only cookies more popular were all associated with Google Analytics[f] or ASP, making PHP, ASP.NET, and JSP the third, fourth, and fifth most widely adopted technology encountered in the survey. Cold Fusion rates where also quite high, ranking eleventh amongst the sites surveyed. Figure 7 presents the percentage of sites found using each of the

[e] % Represents a wildcard, these cookies share the same prefix *ASPSESSION* followed by a unique identifier.

[f] Google Analytics was found to use a combination of four cookies, *__utma*, *__utmb*, *__utmc*, and *__utmz*.

identifiable web application frameworks. ASP was found to be the most common technology globally, with adoption rates of 21.3%, more than 13 points above the next closest competitor, PHP with 7.5%. ASP.NET and JSP where clustered together with adoption rates of 5.4% and 5.2%, respectively, and Cold Fusion was a distant fifth with a 1.6% adoption rate. These global rates will serve as a basis for comparison against which individual country adoption can be measured.

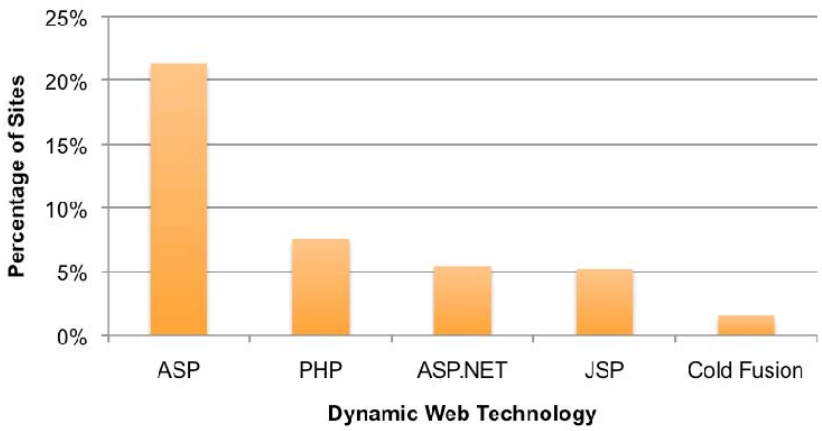


Figure 7. Global dynamic web application framework adoption

A comparison of the adoption of dynamic web content frameworks amongst the 18 countries studied is provided in Figure 8. The figure presents the cumulative percentage of each of the four dynamic technologies. The cumulative percentages range from a low of 18.4% in Japan, to a high of 69.9% in Israel. Japan having the lowest cumulative occurrence per site is inline with the results obtained presented in Figure 3, where Japan was the lower outlier with respect to sessional cookies. Although a low occurrence rate was uniform across all five technologies, Japan was found to possess the lowest occurrence rate for only ASP cookies, 7.0%, over 14 percentage points below the global average. This trend was also observed for the related ASP.NET technology for which Japan again had the lowest occurrence rate, 1.8%. Hong Kong, like Japan, had a lower than average occurrence rate across all five technologies. Despite this technology-wide depression, Russia (23.8%) was found to have a lower cumulative occurrence rate than Hong Kong (29.1%), attributed primarily to the lower-than average adoption of ASP technology. Juxtaposed to Japan and Hong Kong, the UK presented higher than average occurrence rates for all five technologies; however, this did not translate into the UK having the highest cumulative rate amongst the countries studied, in fact six other countries posted higher rates—China, South Korea, Spain, Turkey, and Sweden. Apart for Japan, Hong Kong, and the UK, all other countries displayed mixed results, suggesting that certain technologies are more and less prevalent within specific countries.

Israel had the highest cumulative occurrence rate, attributed to higher than average rates with respect to ASP, PHP, and ASP.NET technologies. On the other hand, the occurrence rates for JSP, and Cold Fusion cookies were below average. Although within Israel PHP and JSP were observed to be above average, ASP technology was observed to have the highest occurrence of any country—

48.7%. This was over 27 percentage points above the global average representing the single largest disparity encountered. The dominance of ASP within Israel is clear, as ASP accounts for more cookies than PHP, ASP.NET, JSP and Cold Fusion combined. Despite being the global leader in cookie occurrence, ASP had lower than average rates amongst 11 of the 18 countries analyzed; however, this only translated into a net result of 3 countries where ASP was not the dominant technology. The occurrence of this technology was most prevalent within Israel (48.7%), China (46.0%), and Turkey (41.4%), each of which posted rates more than 20 points above the global average (21.3%).

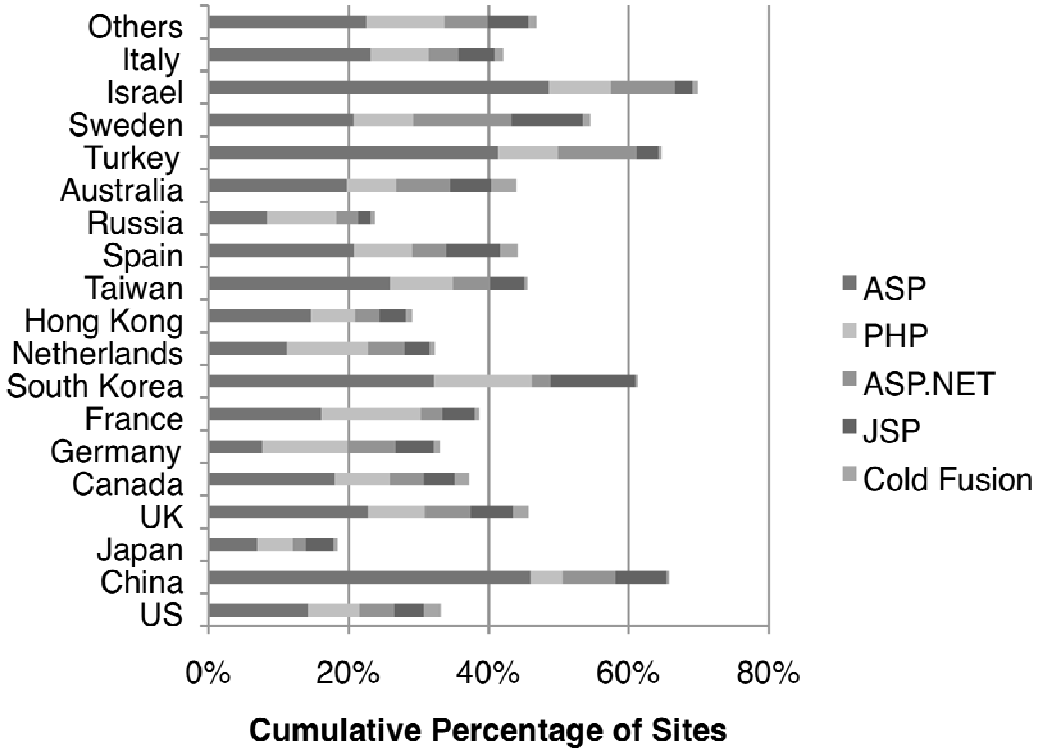


Figure 8. Cumulate comparison of dynamic web framework adoption per country

PHP, the second most encountered technology, had lower than average rates in the top three countries (ranked by number of sites surveyed), US, China, and Japan. Other than these three, PHP rates were only less than average in Hong Kong and Australia. The difference was most pronounced in China, followed closely by Japan and Hong Kong, the US appeared to be just slightly below average, less than half of one percent lower than the average. In fact, the US was observed to be below average with respect to ASP, PHP, ASP.NET, and JSP cookies, the only above average rate was with respect to Cold Fusion technology; however, it was a modest increase of less than one percentage point. The largest disparity within the US was observed amongst ASP cookies, where the US was more than 7 percentage points below the global average. All other technologies had disparities of less than 1 percent; it appears the usage of web application frameworks in the US, with the exception of ASP, is in harmony with global adoption rates.

Cold Fusion, the least occurring technology encountered by the survey, was found to be above average in only five of the countries studied: Australia, Spain, US, UK, and Canada (ordered by descending occurrence rate). Australia was observed to have the highest Cold Fusion cookie occurrence rate, more than two percentage points higher than the global average. Cold fusion usage was lowest in South Korea with an occurrence rate 0.1%. This represents the lowest occurrence rate for all dynamic web technologies across all 18 countries. Despite having such a low occurrence rate of Cold Fusion cookies, South Korea had the highest rate of JSP cookies, 6.8 percentage points about the global average, and was the second highest with respect to PHP technology, falling closely behind France, the PHP frontrunner.

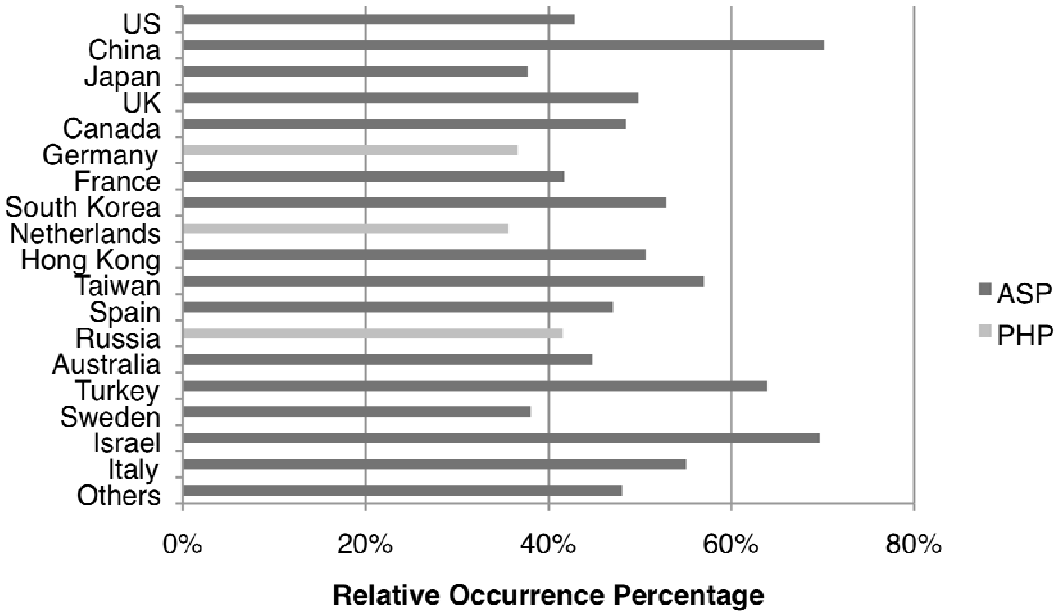


Figure 9. Relative occurrence of the dominant technological platforms per country

To answer research question *Q4*, a breakdown of the dominant platform and relative cookie occurrence rate per country is provided in Figure 9. While ASP, PHP, JSP and Cold Fusion are all utilized globally, it is clear that the most dominant technology encountered was ASP. PHP, the only other technology most frequently occurring amongst the countries analyzed, was observed within 3 countries—Germany (36.6%), Netherlands (35.7%), and Russia (41.5%); two of which, Germany and Netherlands, have the lowest relative occurrence rates amongst the 18 countries. Only three countries have a relative occurrence rate outside of the range 35-60%, China (70.1%), Turkey (64.0%), and Israel (69.7%), representing the least competitive markets with respect to dynamic web technologies.

6.2 Online Tracking & Advertising Technologies

Since the introduction of Cookies, online advertising providers have had a large stake in third-party cookie technology, specifically lobbying user-agents and the Internet Engineering Steering Group to include default support for third-party cookies [1]. Despite the lobbing, RFC 2109, and 2965 [3, 4], the

two latest cookie specifications, mandate that third-party cookies be rejected by default in any user-agent (web browser). This standard has however been generally ignored by all major user-agents—Internet Explorer, Netscape, Firefox, Opera, and Safari. Enabling third-party cookies by default allows Internet advertising networks to provided *targeted* advertisements to end-users. This *targeting* is only available due to the ability of advertisers to track Internet users across multiple sites and browsing sessions using third-party cookies. These third party cookies, often persistent in nature[g], are used extensively to track users across multiple sites and sessions, providing targeted advertisements based upon the browsing tendencies and history of end-users. Recently technologies, such as Google Analytics, have begun to use third-party embedded JavaScripts to provide tracking within a single site by setting first-party cookies. These technologies are beginning to blur the lines between first-party and third-party cookies, presenting a significant challenge for users to assess the status of their online privacy.

Third-party persistent cookies are the primary vehicle through which online advertisers have provided targeted advertisements to users across multiple sites and sessions. These advertisements are often overt; however, the user-tracking methods are not and are often concealed from the Internet user using *web bugging* or *web bugs* [29]. This process involves placing links to third-party objects, often very small images with dimensions measured in single digit number pixels. These objects act to place cookies within the user-agent, without the consent of the user and are enabled by the default web-browser configurations of the majority of browsers. To study these online advertising and tracking practices, third-party persistent cookies were analyzed with respect to the repeat occurrence of similar hosts. A per-country list of the top 5 most occurring third-party persistent hosts has been compiled for each of the 18 countries selected in Section 5 and is presented in Table 4 and graphically in Figure 10. In the compilation of the dataset, two hosts were considered equal if the second-level domain (SLD) name matched. For example, *hitbox.com* the host responsible for the largest number of third-party persistent cookies (3969) was found to range across 503 distinct domain names all sharing the common SLD, *hitbox.com*.

Russia, as indicated in Section 5.2 was found to have the highest mean of sites that utilized third-party persistent cookies. This trend is clearly aspirant in Figure 10 as Russia records the first, second, and third largest percentages for third-party persistent hosts—*rambler.ru*, *list.ru*, and *yadro.ru,* respectively. In fact, the percentage of sites within Russia that have a cookie set by the largest occurring host *rambler.ru (48.2%)*, was larger than the cumulative percentages of the top five hosts from any other country analyzed. The two least occurring top five hosts within Russia, *spylog.com*, and *adriver.ru* were the fifth and sixth largest, falling short only of the French-specific host *xiti.com*. Third-party persistent cookie usage within Russia appears to be driven by Russian-specific advertisers, as all five were located within Russia. Although the hosts did set cookies for non-Russia sites, the majority of cookies, ranging from 73-90%, were set while visiting Russian sites.

In strake contrast to the Russian results, Hong Kong was found to have the lowest percentage of sites setting third-party persistent cookies. This is reflected in the small percentage of sites for which third-party cookies are set. Hong Kong is the only country analyzed to have all five host recorded for less than 1% of the sites, suggesting that the use of third-party persistent cookies is spread across a

[g] 71.8% of all third-party cookies encountered were persistent.

wide variety of hosts, and no specific host has any substantial market share. Each of the top five hosts present in Hong Kong were found to be used in a wide variety of other countries, and no host was found to be Hong Kong specific. Taiwan, on the other hand, the county with the second lowest third-party persistent cookie rate, was found to use third-party persistent cookies at the same rate as Hong Kong. This similarity was confirmed by a Cliff's delta analysis preformed upon the two populations with respect to third-party persistent cookies ($\hat{\delta}$ = -0.007 [-0.033, 0.020]). Despite the similarity in third-party persistent cookie occurrence rates, Taiwan has a much different adoption rate than Hong Kong. Taiwan, like many of the other countries has a country-specific host that dominates the third-party persistent cookie landscape, *hotrank.com.tw*. This host is responsible for setting third-party persistent cookies on 4% of sites within Taiwan and is clearly the dominant host as the next closest host is only responsible for 0.4% of the cookies. Like the Russian-specific cookies, *hotrank.com.tw* is primarily used within Taiwan. The two populations, Hong Kong and Taiwan present two very different pictures of third-party persistent cookie usage, Hong Kong, a wide variety of globally established hosts, and Taiwan, a market with a clear country-specific market leader.

Table 4. Top 5 third-party persistent cookie hosts per country

Country	The Top 5 3rd Party Persistent Cookie Hosts (Percentage of Sites)
US	doubleclick.net (3.20%), atdmt.com (2.79%), 2o7.net (2.68%), hitbox.com (2.00%), fastclick.net (1.95%)
China	allyes.com (4.75%), information.com (0.65%), mozilla.org (0.65%), revenue.net (0.60%), admin88.com (0.56%)
Japan	valuecommerce.com (3.18%), tracer.jp (1.65%), valueclick.ne.jp (1.11%), 2o7.net (1.01%), doubleclick.net (0.92%)
UK	doubleclick.net (7.70%), atdmt.com (4.61%), hitbox.com (2.98%), estat.com (2.55%), adtech.de (2.49%)
Canada	doubleclick.net (3.76%), hitbox.com (2.06%), 2o7.net (1.86%), atdmt.com (1.82%), statcounter.com (1.39%)
Germany	ivwbox.de (9.08%), doubleclick.net (4.11%), falkag.de (3.69%), adtech.de (2.32%), www.etracker.de (1.87%)
France	xiti.com (24.17%), doubleclick.net (7.02%), cybermonitor.com (4.90%), estat.com (4.68%), weborama.fr (4.08%)
South Korea	acecounter.com (3.85%), logger.co.kr (2.39%), nasmedia.co.kr (1.86%), overture.com (1.40%), realmedia.co.kr (1.17%)
Netherlands	estat.com (7.22%), doubleclick.net (5.50%), sitestat.com (5.10%), sexcounter.com (2.19%), onestat.com (2.05%)
Hong Kong	imrworldwide.com (0.93%), casalemedia.com (0.71%), zedo.com (0.71%), doubleclick.net (0.64%), statcounter.com (0.64%)
Taiwan	hotrank.com.tw (4.16%), cnyes.com (0.39%), doubleclick.net (0.39%), paypal.com (0.31%), hitbox.com (0.23%)
Spain	doubleclick.net (10.23%), ojdinteractiva.com (8.67%), imrworldwide.com (8.10%), tradedoubler.com (2.45%), atdmt.com (1.72%)
Russia	rambler.ru (48.18%), list.ru (38.15%), yadro.ru (33.85%), spylog.com (23.44%), adriver.ru (17.58%)
Australia	imrworldwide.com (8.22%), doubleclick.net (7.18%), adsfac.net (3.66%), 2o7.net (3.26%), sextracker.com (1.96%)
Turkey	gittigidiyor.com (3.91%), statcounter.com (2.16%), mediainer.net (2.16%), iyi.net (1.62%), hbmediapro.com (1.21%)
Sweden	imrworldwide.com (12.30%), doubleclick.net (7.15%), tradedoubler.com (7.15%), research-int.se (6.29%), admeta.com (3.43%)
Israel	imrworldwide.com (6.27%), statcounter.com (3.43%), walla.co.il (2.84%), 2o7.net (1.49%), atdmt.com (1.49%)
Italy	imrworldwide.com (11.36%), doubleclick.net (7.41%), tradedoubler.com (3.79%), shinystat.com (0.02%), neodatagroup.com (0.02%)

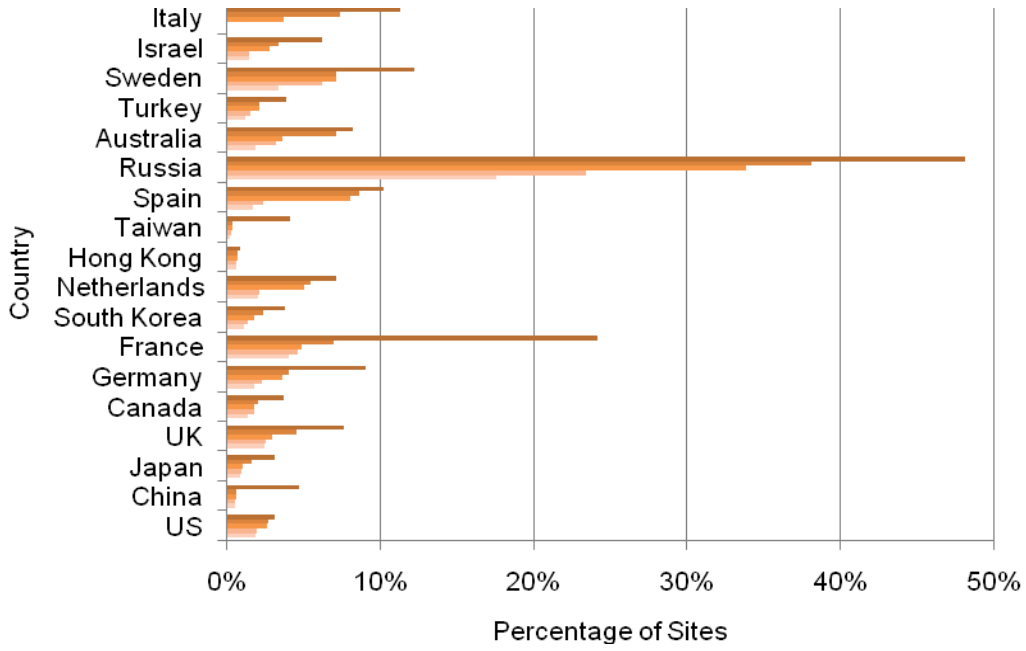


Figure 10. Top 5 third-party persistent hosts per country

Like Taiwan, China, France, and Germany have a country-specific dominant host, *allyes.com*, *xiti.com*, and *ivwbox.de*, respectively. All four of the hosts more than double the next closest competitor with respect to site occurrence percentage. Other countries such as Japan, the UK, Canada, South Korea, Turkey, Sweden, Israel, and Italy have a clear front-runner, although their dominance is not a pronounced as that found in the other countries. The Netherlands, Spain, Australia appear to have a two- or three-way competition between the most occurring hosts. Finally, the US presents a much more competitive market with less than 1.5 percentage points separating the first and fifth most prevalent hosts.

Despite the number of country-specific hosts, cookies from *doubleclick.net*, were observed to be in the top five of 13 countries. These cookies prevalence was unmatched by any other host, and was set by 2,713 (2.8%) of sites encountered by the survey, the highest globally occurring third-party persistent host. Cookies set by the next closest competitor, *hitbox.com,* were encountered on less than half of the number of sites as *doubleclick.net*, 1,221 (1.2%). Furthermore, *doubleclick.net* was found within the top five hosts for each of the G7 countries, in fact, it was the top technology for the US, UK, and Canada, and second behind the country-specific hosts in France and Germany. Japan was the only G7 country in which *doubleclick.net* was not one of the top two hosts.

Although third-party persistent cookies are deployed regardless of the country of origin, the usage of these cookies is not consistent between nations as a number of nation-specific cookies dominate locally. While third-party cookie usage was found to be universal, differences in the rates at which these cookies are deployed were observed. These findings answer research question Q5 providing insight into the tolerance of specific nations for this type of technological use. Clearly in Russia, third-

party cookies are ubiquitous, whereas in Hong Kong, developers should generally stay away from these types of technologies, as they do not appear to have widespread acceptance.

6.3 Web Analytics: Third-Party Internal Site Tracking Technologies

Third-party persistent cookies and *web-bugs* are not the only techniques employed to track users and provide content tailored to specific Internet users. A growing number of sites are beginning to use third-party embedded JavaScript to set first-party cookies, which are used to collect data about users browsing habits providing the ability to produce user-centric content, assessing which pages are most profitable. Although these goals are similar to that of advertisers, the resulting cookies are difficult to identify, as they do not share a common third-party host.

An analysis, similar to that undertaken with respect to dynamic web technology was applied to the collection of over 200,000 cookies. The distinct cookie names where ranked by occurrence and then analyzed from the perspective of the number of distinct hosts that set the cookies. This analysis revealed several reoccurring cookie names, two of which were definitively traced to third parties, namely Google Analytics [30] and Omniture SiteCatalyst [31]. Google Analytics, traceable by the presence of four reoccurring cookies *__utmc*, *__utma*, *__utmb*, and *__utmz*, was set on 8.6% (8,411) of the sites surveyed, making it the second most widely found cookie-related technology encountered by the survey. Omniture SiteCatalyst, like Google Analytics, was traceable by the presence of two unique cookies *s_cc*, and *s_sq* and was set by 1.7% (1,690) sites surveyed, ranking Omniture seventh for most encountered cookie-related web technology.

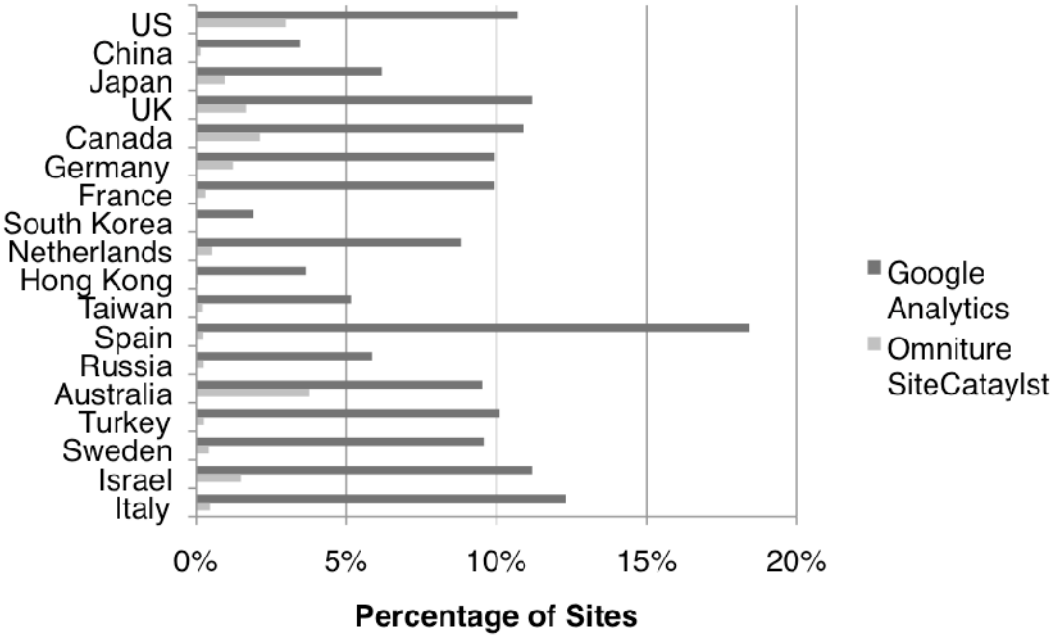


Figure 11. Web analytics technology per country

A per-country breakdown of Google Analytics and Omniture SiteCatalyst is provided in Figure 11 Google Analytics is clearly the global market leader, evident by a market share, at the lowest point, of 2.5 times that of Omniture. The market appears to be much more competitive in the Australia, Canada, Germany, Israel, Japan, UK, and US with a market share ranging from 2.5 – 8.0 times higher than that of Omniture, compared to 16.5 – 75 in the other countries. Google Analytics adoption rate ranges from a minimum of 1.9% in South Korea to a high of 18.4% in Spain. Similarly to Google, Omniture's lowest rating was in South Korea, where the technology was not encountered. Australia posted Omniture's highest rating (3.8%) and was the country with the smallest ratio of sites with Google versus Omniture cookies.

In comparison to third-party persistent cookie based technologies presented in Table 4 and Figure 10, Google Analytics is the most used user-centric technology in all but five of the countries analyzed—China, France, South Korea, Russia, and Sweden. Within China, France, and Sweden, Google ranked second falling behind the country-specific hosts of France (*xiti.com*) and China (*allyes.com*) and *imrworldwide.com* in Sweden. In South Korea, where Google's rate was the lowest of the 18 countries studied, it ranked third behind two South Korean-specific hosts (*acecounter.com*, and *logger.co.kr*). Russia, as previously noted, was very different from the other countries, and Google Analytics ranked eighth amongst user-centric technologies encountered. All eight of the cookies that ranked higher than Google were Russian-specific, further evidence that the Russian online marketplace is driven primarily by internal interactions.

Hong Kong, despite the low adoption rate of third-party advertising cookies was found to have a Google Analytics occurrence rating of 3.6%, slightly higher than that of China. Google Analytics is the first third-party technology found used on more than 1% percent of sites from Hong Kong. Further investigation revealed a second Hong Kong-specific[h] re-occurring cookie, identifiable by the name *cdb_sid*, found occurring on 3.5% of sites. Information pertaining to the technology responsible for setting the *cdb_sid* was not readily available and a direct link to any one technology could not be established. Due to the re-occurring nature of this exclusively persistent cookie and the fact that the majority of cookies (88.9%) were first-party, it is assumed that these cookies fulfil a similar role to that of Google Analytics and Omniture, however this cannot be explicitly confirmed. Further evidence suggesting that this cookie is Google's direct competitor within Hong Kong is the degree of overlap, or rather the lack there of, between sites setting these cookies. Within Hong Kong the presence of cookies from Google Analytics, Omniture, and *cdb_sid* where found to be mutually exclusive, suggesting that these cookies are dichotomous to each other. This exclusivity was also observed globally, where 98.7% of sites with these technologies were found to be mutually exclusive.

Third-party web analytics is quickly becoming an important tool within the business model of any e-commerce endeavour. Because cookies are a crucial component of these third-party software packages, the cookies are set as first-party, both sessional and persistent, to decrease the likelihood of rejection. This application raises a number of privacy questions, as these cookies are set by embedded third-party scripts; a detail that is not explicitly revealed to the user. While this study answers the research question *Q6—how are cookies used to track a users movement within a site*—they also elicit a number of further questions that are outside of the scope of this work. Similar to the privacy

[h] Found occurring in less than 0.6% in any other country.

concerns of *unverified transactions* associated with the usage of third-party cookies [1], this type of cookie usage can be seen as invasive and is currently implemented without any opt-out mechanism for the average user. Similar to the usage of third-party cookies, these issues have not crossed the tolerance threshold for the average Internet user, and as such remain in use globally.

7 Cookies: A Proxy for a Country's *E-Readiness?*

The annual *e-readiness* white paper published by the Economist Intelligence Unit "evaluates the technological, economic, political and social assets of 68 countries … and their cumulative impact on their respective information economies." [32]. The ranking is comprised of the weighted accumulation of scores from nearly 100 criterions spanning six distinct categories—connectivity, business environment, consumer and business adoption, legal and policy environment, social and cultural environment, and supporting e-services. Essentially, e-readiness provides a metric by which we can assess a country's e-business environment, and how fertile the environment is with respect to Internet-based endeavours. Of the 68 countries identified in the e-readiness report, 66 of these countries were the geographical origin of at least one surveyed site. 66 of 68 countries highlighted in the report represent 98% (96,365) of the sites surveyed and all 18 of the countries highlighted in Section 5 are ranked in the report. The e-readiness report was used to partition the survey results with respect to geographical origin of each site. These partitions are analyzed with respect to six criteria: cookie usage, first-party cookie usage, third-party cookie usage, sessional cookie usage, and persistent cookie usage, and third-party persistent usage.

To investigate the association between a country's e-readiness ranking and their cookie usage, a series of Spearman's ρ correlations were calculated and summarized in Table 5. Significant correlations were observed amongst all forms of cookie usage with the exception of first-party cookies. All of the correlation coefficients represented negative trends suggesting that the higher the e-readiness rank of the host country, the more likely the site to use cookies, especially third-party persistent cookies.

Table 5. Spearman's ρ correlations for cookie usage vs. e-readiness ranking

	Cookies	**First-Party Cookies**	**Third-Party Cookies**	**Sessional Cookies**	**Persistent Cookies**	**Third-Party Persistent Cookies**
Spearman's ρ	-.309	-.208	-.462	-.254	-.468	-.499
Pearson's *r*	-.322	-.217	-.479	-.265	-.485	-.516
Sig. (2-tailed)	.012	.093	.000	.040	.000	.000
N[i]	66	66	66	66	66	66

From the ρ values in Table 5, we see that this correlation is strongest amongst third-party persistent cookies, and weakest for general cookie usage. A *small* relationship exists between overall cookie usage and ranking ($-0.29 \leq r \leq -0.10$) [25, 27]. The classification of overall cookie usage as

[i] N=66 because of the removal of the two countries that did not have any sites surveyed, and therefore could not be included in the Spearman's ρ correlations.

small is attributed to the correlation of rank and third-party persistent cookies suppressed amongst the general population of cookies. Third-party and persistent cookies have *medium* correlations to ranking ($-0.49 \leq r \leq -0.30$) [25, 27]; the strongest relationship exists between third-party persistent cookies and e-readiness ranking. This *large* relationship ($-1.00 \leq r \leq -0.50$) [25, 27] between ranking and third-party persistent cookies suggests that the maturity of a country's e-business environment promotes the usage of third-party and persistent cookies, specifically third-party persistent cookies. These findings further support the analysis in Section 6.2; the majority of third-party persistent cookies are used within commercially viable environments to provide user-centric content. Based upon the correlations revealed in Table 5 it is clear that cookie usage, specifically third-party and persistent cookies, are related to the maturity of a nation's e-commerce environment, providing an affirmative answer to research question *Q7*—third-party persistent cookies directly correlate with the maturity of a country's e-commerce environment.

8 Conclusion

Cookies are one of the most prevalent technologies used within the Internet today. Cookies have been shown to be the backbone of many web-application frameworks, advertising networks, and web analytical packages. Furthermore, cookies are a globally accepted technology—they were observed within all countries with at least three sites surveyed. This survey presents a number of significant relationships between country of origin and cookie usage, as well as a breakdown of the most widely adopted platforms with respect to web application development, user-targeted advertising, and web-analytics. The results presented in this survey have implications for a variety of research and commercial endeavours including, but not limited to, investigations of the privacy issues surrounding third-party cookies, the country-specific geo-political and legal issues affecting cookie adoption, and country-specific technological/business environments.

The primary goal of this paper is to study global cookie usage, and to this end, cookies were successfully demonstrated to be a globally adopted technology. This global adoption was by no means uniform and large discrepancies where observed between countries, ranging from significantly lower than average usage, such as in Japan and Hong Kong, to Russia, a country with a significantly disproportionate number of third-party and persistent cookies. This wide distribution of usage amongst first- and third-party, and sessional and persistent cookies, leads to the conclusion that the motivations and factors affecting cookie usage and adoption are complex, and cannot be attributed to any one single underlying factor. A myriad of contributing factors are evidently at work.

Significant links were established by the survey with respect to the technological sources of first-party and sessional cookies. These types of cookies were observed in abundance with respect to web-application development platforms, specifically ASP, PHP, JSP, and ASP.NET. The clearest link was found with respect to sessional cookies comprising 99.0% of the technological-specific cookies, compared to only 28.7% of the non-specific cookies collected. This link was also observed amongst first-party cookies but not to the same degree—73.3% versus 49.0%. It is clear that there is a link between first-party sessional cookies and dynamic web technologies. Despite this relationship, technology-specific cookies only accounted for 37.2% of sessional cookies and 19.1% of first-party cookies, and although a clear link was found, it cannot be viewed as a definitive factor for this type of cookie usage.

The survey highlights the prevalence of vendor-specific third-party technologies both globally and within specific countries. Although the survey did find global leaders, such as *doubleclick.net* and Google Analytics, country-specific market providers such as *rambler.ru*, *allyes.com*, *xiti.com*, and *ivwbox.de* were also discovered. These country-specific providers suggest that despite a globally dominated market, there is room for country-specific competitors, as is the case within Russia, France, Germany, China, Hong Kong, and South Korea. These results are of particular relevance to existing e-commerce ventures and those looking to break into the already saturated market.

Despite the complex relationships identified between cookie usage and country, one statistically significant *large* correlation was identified between third-party persistent cookie usage and the status of a country's e-business environment, encapsulated by the e-readiness rankings [32]. This relationship, observed specifically amongst third-party persistent cookies, encapsulates the complexity of the relationship between cookie usage and country of origin, as the e-readiness ranking is a composition of nearly 100 criterions across six categories. Future areas of research include the distillation of these results across the e-readiness criteria, providing further insight into the specific microcausms of third-party persistent cookie usage. Although a number of subsequent questions remain for future study, this paper presents a conclusive correlation between cookie usage and the origin's e-commerce environment; the strongest evidence that cookies are an integral part of the global e-commerce environment.

Acknowledgements

This work is supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada and iCORE.

References

[1] D. M. Kristol, "HTTP Cookies: Standards, privacy, and politics," *ACM Transactions on Internet Technology*, vol. 1, pp. 151-198, 2001.
[2] A. Tappenden and J. Miller, "Cookies: A Deployment Study and the Testing Implications," *ACM Transactions on the Web*, vol. 3, no. 3, pp. 1–49, 2009.
[3] D. Kristol and L. Montulli, "RFC 2109: HTTP State Management Mechanism", http://www.ietf.org/rfc/rfc2109.txt accessed March 1 2006.
[4] D. Kristol and L. Montulli, "RFC 2965: HTTP State Management Mechanism", http://www.ietf.org/rfc/rfc2965.txt accessed March 1 2006.
[5] A. Tappenden and J. Miller, "A Three-Tiered Testing Strategy for Cookies," in *Software Testing, Verification, and Validation, 2008 IEEE International Conference on*. Lillehammer, Norway, 2008, pp. 131-140.
[6] comScore Inc., "Cookie-Based Counting Overstates SIze of Web Site Audiencces", http://www.comscore.com/press/release.asp?press=1389 accessed April 18 2007.
[7] C. Yue, M. Xie, and H. Wang, "Automatic Cookie Usage Setting with CookiePicker," presented at Dependable Systems and Networks, 2007. DSN '07. 37th Annual IEEE/IFIP Intl Conf on.
[8] Security Space, "Internet Cookie Report", http://www.securityspace.com/s_survey/data/man.200609/cookieReport.html accessed Oct. 2006.
[9] Port80 Software, "Industry Surveys: Top Application Servers", http://www.port80software.com/surveys/top1000appservers/ accessed Jan 2009.

[10] B. V. L. Doyle, Cristina, "Survey of Technologies for Web Application Development," *arXiv:0801.2618*, pp. 43, 2008.
[11] Google, "GMail", https://mail.google.com accessed Jan 5, 2009.
[12] Miniwatts Marketing Group, "World Internet Users and Population Stats", http://www.internetworldstats.com/stats.htm accessed Jan 2009.
[13] G. Vijayaraghavan and C. Kaner, "Bug Taxonomies: Use Them to Generate Better Tests," in *STAR EAST 2003*. Orlando, FL, 2003, pp. 1-40.
[14] Alexa Internet Inc., "Alexa Top Site Service", http://www.alexa.com/site/devcorner/top_sites accessed March 4 2006.
[15] Alexa Internet Inc., "About the Alexa Traffic Rankings", http://www.alexa.com/site/devcorner/top_sites accessed March 4 2006.
[16] L. Baker, "Alexa Bias Exposed by Top Google Engineers", http://www.searchenginejournal.com/alexa-bias-exposed-by-top-google-engineers/4487/ accessed Jan 2009.
[17] I. Reay, P. Beatty, S. Dick, and J. Miller, "A Survey and Analysis of the P3P Protocol's Agents, Adoption, Maintenance, and Future," *Dependable and Secure Computing, IEEE Transactions on*, vol. 5, pp. 151-164, 2007.
[18] Mozilla Corporation, "Firefox", http://www.mozilla.com/firefox/ accessed March 4 2006.
[19] IP2Location.com, "IP2Location", http://www.ip2location.com/ accessed November 20 2006.
[20] N. Cliff, "Dominance statistics: ordinal analysis to answer ordinal questions," *Psychological Bulletin*, vol. 114, pp. 494-509, 1993.
[21] N. Cliff, *Ordinal methods for behavioral data analysis*. New Jersey: Lawrence Erlbaum Associates, 1996.
[22] J. D. Kromrey, K. Y. Hogarty, J. M. Ferron, C. V. Hines, and M. R. Hess, "Robustness in Meta-Analysis: An Empirical Comparison of Point and Interval Estimates of Standardized Mean Differences and Cliff's Delta," in *American Statistical Association Joint Statistical Meetings*. Minneapolis: , 2005.
[23] M. R. Hess, J. D. Kromrey, J. M. Ferron, K. Y. Hogarty, and C. V. Hines, "Robust Inference in Meta-Analysis: An Empirical Comparison of Point and Interval Estimates Using the Standardized Mean Difference and Cliff's Delta," in *Annual Meeting of the American Educational Research Association*. Montreal, 2005, pp. 36.
[24] J. Romano, J. Kromrey, J. Coraggio, and J. Skowronek, "Appropriate statistics for ordinal level data," presented at Florida Association for Institutional Research, 2006.
[25] J. Cohen, "A power primer," *Psychological Bulletin*, vol. 112, pp. 155-159, 1992.
[26] A. R. Gilpin, "Table for Converstion of Kendall's Tau to Spearman's Rho Within the context of measures of magnitude of effect for meta-analysis," *Educational and Psychological Measurement*, vol. 53, pp. 87-52, 1993.
[27] J. Cohen, *Statistical power analysis for the behavioural sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
[28] United Nations Statistics Division, "Standard Country and Area Codes Classifications", http://unstats.un.org/unsd/methods/m49/m49regin.htm accessed June 2007.
[29] R. M. Smith, "The Web Bug FAQ", http://www.eff.org/Privacy/Marketing/web_bug.html accessed October 26 2006.
[30] Google, "Google Analytics", http://www.google.com/analytics/ accessed May 7 2007.
[31] Omniture Inc., "Omniture SiteCatalyst", http://www.omniture.com/products/web_analytics/sitecatalyst accessed
[32] Economist Intelligence Unit and IBM Institute for BusinessValue, "The 2006 e-readiness rankings," Economist Intelligence Unit,, 2006, pp. 23.

Appendix I: Cookie Usage Per Site According to Country

Table 6. Cookie usage per site according to country

Country	Number Of Sites		Sites Setting Cookies		Site Setting 1st Party Cookies		Sites Setting 3rd Party Cookies		Sites Setting Sessional Cookies		Sites Setting Persistent Cookies		Sites Setting 3rd Party Persistent Cookies	
United States	44,673	46%	29,589	66%	24,478	55%	16,639	37%	23,459	53%	22,822	51%	14,491	32%
China	17,196	18%	12,193	71%	10,806	63%	5,802	34%	10,741	62%	7,672	45%	3,278	19%
Japan	6,658	7%	3,431	52%	2,574	39%	1,730	26%	2,283	34%	2,678	40%	1,386	21%
United Kingdom	3,493	4%	2,628	75%	2,297	66%	1,471	42%	2,143	61%	1,980	57%	1,299	37%
Canada	2,527	3%	1,726	68%	1,402	55%	907	36%	1,357	54%	1,294	51%	790	31%
Germany	2,411	2%	1,631	68%	1,271	53%	1,053	44%	1,381	57%	1,215	50%	858	36%
France	1,837	2%	1,405	76%	1,006	55%	1,056	57%	1,000	54%	1,219	66%	992	54%
South Korea	1,716	2%	1,311	76%	1,159	68%	544	32%	1,201	70%	618	36%	392	23%
Netherlands	1,509	2%	1,016	67%	805	53%	666	44%	774	51%	809	54%	592	39%
Hong Kong	1,401	1%	705	50%	583	42%	273	19%	529	38%	427	30%	197	14%
Taiwan	1,298	1%	743	57%	641	49%	271	21%	651	50%	361	28%	185	14%
Spain	1,222	1%	860	70%	708	58%	505	41%	733	60%	641	52%	439	36%
Russian	768	1%	665	87%	499	65%	604	79%	466	61%	637	83%	595	77%
Australia	766	1%	541	71%	469	61%	284	37%	436	57%	377	49%	248	32%
Turkey	742	1%	551	74%	480	65%	239	32%	506	68%	280	38%	175	24%
Sweden	699	1%	563	81%	496	71%	307	44%	488	70%	410	59%	284	41%
Israel	670	1%	523	78%	469	70%	255	38%	476	71%	333	50%	206	31%
Italy	634	1%	457	72%	367	58%	260	41%	348	55%	352	56%	240	38%
Denmark	430	0%	353	82%	301	70%	209	49%	311	72%	276	64%	197	46%
Brazil	399	0%	266	67%	245	61%	79	20%	233	58%	158	40%	65	16%
Greece	321	0%	215	67%	188	59%	102	32%	184	57%	130	40%	75	23%
Thailand	314	0%	225	72%	201	64%	126	40%	213	68%	183	58%	84	27%
Switzerland	283	0%	199	70%	169	60%	104	37%	166	59%	139	49%	93	33%
Czech Republic	269	0%	210	78%	146	54%	170	63%	143	53%	181	67%	159	59%
Norway	261	0%	212	81%	173	66%	153	59%	195	75%	176	67%	140	54%
Belgium	234	0%	192	82%	172	74%	121	52%	149	64%	157	67%	115	49%
Poland	229	0%	167	73%	126	55%	121	53%	124	54%	137	60%	114	50%
Austria	228	0%	167	73%	127	56%	104	46%	137	60%	120	53%	95	42%
India	223	0%	135	61%	112	50%	57	26%	121	54%	57	26%	36	16%
Singapore	212	0%	131	62%	103	49%	75	35%	98	46%	93	44%	60	28%
Hungary	208	0%	169	81%	130	63%	132	63%	116	56%	145	70%	126	61%
Mexico	179	0%	103	58%	91	51%	44	25%	81	45%	66	37%	39	22%
Argentina	171	0%	126	74%	110	64%	73	43%	108	63%	93	54%	62	36%
Finland	169	0%	132	78%	104	62%	96	57%	111	66%	104	62%	91	54%
Malaysia	164	0%	99	60%	86	52%	40	24%	85	52%	59	36%	29	18%
Lithuania	155	0%	144	93%	132	85%	47	30%	124	80%	136	88%	42	27%
Saudi Arabia	132	0%	79	60%	63	48%	33	25%	72	55%	36	27%	25	19%
Viet Nam	123	0%	98	80%	90	73%	30	24%	90	73%	47	38%	17	14%
New Zealand	119	0%	91	76%	73	61%	66	55%	74	62%	75	63%	61	51%
Egypt	114	0%	83	73%	77	68%	28	25%	79	69%	31	27%	17	15%
Ireland	107	0%	85	79%	75	70%	47	44%	76	71%	60	56%	44	41%
Romania	107	0%	91	85%	75	70%	78	73%	70	65%	84	79%	76	71%
Portugal	99	0%	76	77%	67	68%	37	37%	62	63%	50	51%	31	31%

Country	Number Of Sites		Sites Setting Cookies		Site Setting 1st Party Cookies		Sites Setting 3rd Party Cookies		Sites Setting Sessional Cookies		Sites Setting Persistent Cookies		Sites Setting 3rd Party Persistent Cookies	
Bulgaria	87	0%	71	82%	61	70%	45	52%	59	68%	55	63%	42	48%
Ukraine	83	0%	44	53%	36	43%	34	41%	36	43%	38	46%	30	36%
Chile	79	0%	55	70%	49	62%	35	44%	49	62%	42	53%	34	43%
Uruguay	75	0%	36	48%	29	39%	16	21%	33	44%	23	31%	14	19%
Virgin Islands, British	70	0%	23	33%	15	21%	13	19%	16	23%	15	21%	10	14%
Venezuela	68	0%	47	69%	40	59%	17	25%	41	60%	26	38%	14	21%
South Africa	67	0%	53	79%	48	72%	37	55%	46	69%	46	69%	34	51%
United Arab Emirates	65	0%	52	80%	44	68%	22	34%	46	71%	24	37%	15	23%
Indonesia	65	0%	46	71%	34	52%	18	28%	35	54%	22	34%	11	17%
Slovenia	62	0%	52	84%	45	73%	32	52%	48	77%	40	65%	31	50%
Slovakia	61	0%	48	79%	31	51%	36	59%	27	44%	43	70%	36	59%
Serbia and Montenegro	55	0%	28	51%	23	42%	11	20%	23	42%	17	31%	9	16%
Estonia	53	0%	41	77%	27	51%	28	53%	32	60%	33	62%	27	51%
Costa Rica	43	0%	30	70%	27	63%	16	37%	22	51%	22	51%	15	35%
Croatia	43	0%	23	53%	23	53%	14	33%	21	49%	16	37%	12	28%
Philippines	32	0%	18	56%	18	56%	2	6%	17	53%	13	41%	2	6%
Colombia	30	0%	15	50%	12	40%	10	33%	11	37%	12	40%	9	30%
Panama	30	0%	20	67%	16	53%	8	27%	14	47%	16	53%	8	27%
Peru	29	0%	14	48%	11	38%	7	24%	13	45%	5	17%	5	17%
Latvia	26	0%	23	88%	20	77%	17	65%	21	81%	19	73%	16	62%
Iceland	25	0%	21	84%	20	80%	15	60%	19	76%	18	72%	13	52%
Jordan	24	0%	14	58%	10	42%	7	29%	10	42%	6	25%	4	17%
Gibraltar	23	0%	20	87%	19	83%	15	65%	18	78%	18	78%	15	65%
Kuwait	22	0%	18	82%	15	68%	4	18%	18	82%	2	9%	2	9%
Malta	22	0%	13	59%	12	55%	3	14%	12	55%	7	32%	2	9%
Dominica	21	0%	9	43%	9	43%	8	38%	8	38%	9	43%	8	38%
Belize	19	0%	16	84%	16	84%	3	16%	14	74%	5	26%	3	16%
Macao	19	0%	10	53%	7	37%	4	21%	9	47%	5	26%	2	11%
Cyprus	19	0%	14	74%	13	68%	7	37%	13	68%	10	53%	6	32%
Syrian Arab Republic	17	0%	10	59%	8	47%	3	18%	9	53%	3	18%	1	6%
Luxembourg	17	0%	14	82%	13	76%	4	24%	12	71%	8	47%	4	24%
Bermuda	16	0%	14	88%	14	88%	4	25%	14	88%	13	81%	4	25%
Iran	15	0%	6	40%	6	40%	3	20%	6	40%	1	7%	1	7%
Grenada	14	0%	9	64%	2	14%	8	57%	1	7%	9	64%	8	57%
Pakistan	14	0%	8	57%	8	57%	3	21%	7	50%	4	29%	1	7%
Dominican Republic	10	0%	8	80%	8	80%	4	40%	8	80%	6	60%	3	30%
Ecuador	10	0%	6	60%	6	60%	1	10%	5	50%	2	20%	1	10%
Puerto Rico	10	0%	7	70%	7	70%	4	40%	7	70%	4	40%	4	40%
Lebanon	10	0%	8	80%	6	60%	3	30%	7	70%	5	50%	3	30%
Morocco	9	0%	2	22%	2	22%	1	11%	2	22%	1	11%	1	11%
Qatar	9	0%	5	56%	4	44%	1	11%	5	56%	1	11%	1	11%
Netherlands Antilles	8	0%	6	75%	5	63%	3	38%	5	63%	4	50%	2	25%

Country	Number Of Sites	Sites Setting Cookies	Site Setting 1st Party Cookies	Sites Setting 3rd Party Cookies	Sites Setting Sessional Cookies	Sites Setting Persistent Cookies	Sites Setting 3rd Party Persistent Cookies
Antigua and Barbuda	7 0%	5 71%	5 71%	1 14%	5 71%	5 71%	1 14%
Bosnia and Herzegovina	6 0%	6 100%	5 83%	4 67%	6 100%	4 67%	4 67%
Tunisia	6 0%	4 67%	4 67%	1 17%	4 67%	1 17%	1 17%
Bolivia	5 0%	1 20%	1 20%	1 20%	1 20%	1 20%	1 20%
Bahamas	5 0%	2 40%	2 40%	2 40%	1 20%	2 40%	2 40%
Guatemala	4 0%	3 75%	3 75%	0 0%	3 75%	2 50%	0 0%
Sri Lanka	4 0%	3 75%	2 50%	1 25%	3 75%	1 25%	1 25%
Mauritius	4 0%	4 100%	2 50%	3 75%	3 75%	2 50%	2 50%
Yemen	4 0%	1 25%	1 25%	0 0%	0 0%	1 25%	0 0%
Libyan Arab Jamahiriya	3 0%	1 33%	1 33%	0 0%	0 0%	1 33%	0 0%
Algeria	3 0%	2 67%	2 67%	0 0%	1 33%	1 33%	0 0%
Paraguay	3 0%	2 67%	2 67%	1 33%	2 67%	1 33%	1 33%
Macedonia	0%	3 100%	3 100%	2 67%	3 100%	2 67%	2 67%
Barbados	3 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Belarus	3 0%	2 67%	2 67%	1 33%	1 33%	2 67%	1 33%
Moldova, Republic Of	3 0%	3 100%	3 100%	3 100%	3 100%	3 100%	3 100%
Bahrain	3 0%	1 33%	1 33%	0 0%	1 33%	0 0%	0 0%
Samoa	2 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Sudan	2 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Honduras	2 0%	1 50%	1 50%	0 0%	1 50%	0 0%	0 0%
Saint Kitts and Nevis	2 0%	2 100%	2 100%	0 0%	2 100%	0 0%	0 0%
Cote Divoire	2 0%	2 100%	1 50%	1 50%	1 50%	1 50%	1 50%
Nicaragua	2 0%	1 50%	1 50%	1 50%	1 50%	1 50%	1 50%
Myanmar	2 0%	1 50%	1 50%	1 50%	1 50%	0 0%	0 0%
Guinea-Bissau	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Botswana	1 0%	1 100%	0 0%	1 100%	1 100%	0 0%	0 0%
Trinidad and Tobago	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Cuba	1 0%	1 100%	1 100%	0 0%	1 100%	0 0%	0 0%
Albania	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Fiji	1 0%	1 100%	1 100%	0 0%	0 0%	1 100%	0 0%
Liechtenstein	1 0%	1 100%	1 100%	0 0%	1 100%	1 100%	0 0%
American Samoa	1 0%	1 100%	0 0%	1 100%	0 0%	1 100%	1 100%
Georgia	1 0%	1 100%	1 100%	1 100%	1 100%	1 100%	1 100%
Nigeria	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Holy See (Vatican City State)	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Iraq	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Kazakhstan	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Oman	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Afghanistan	1 0%	1 100%	1 100%	1 100%	1 100%	1 100%	1 100%
Cambodia	1 0%	1 100%	1 100%	0 0%	1 100%	0 0%	0 0%

Country	Number Of Sites	Sites Setting Cookies	Site Setting 1st Party Cookies	Sites Setting 3rd Party Cookies	Sites Setting Sessional Cookies	Sites Setting Persistent Cookies	Sites Setting 3rd Party Persistent Cookies
El Salvador	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Brunei Darussalam	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Turks and Caicos Islands	1 0%	1 100%	0 0%	1 100%	1 100%	1 100%	1 100%
Bangladesh	1 0%	0 0%	0 0%	0 0%	0 0%	0 0%	0 0%
Faroe Islands	1 0%	1 100%	0 0%	1 100%	1 100%	1 100%	1 100%
Totals	97,050 100%	65,423 67%	54,616 56%	35,593 37%	53,064 55%	47,624 49%	29,029 30%

Appendix II: Cookie Usage According to Country

Table 7. Cookie usage according to country

Country	Number of Cookies Set	Number of 1st Party Cookies	Number of 3rd Party Cookies	Number of Sessional Cookies	Number of Persistent Cookies	Number of 3rd Party Persistent Cookies
United States	140,940 50%	78,615 56%	62,325 44%	52,074 37%	88,866 63%	14,491 10%
China	47,205 17%	30,231 64%	16,974 36%	21,760 46%	25,445 54%	3,278 7%
Japan	11,012 4%	6,226 57%	4,786 43%	4,231 38%	6,781 62%	1,386 13%
United Kingdom	11,376 4%	6,735 59%	4,641 41%	4,423 39%	6,953 61%	1,299 11%
Canada	7,081 2%	4,236 60%	2,845 40%	2,680 38%	4,401 62%	790 11%
Germany	7,653 3%	3,755 49%	3,898 51%	2,704 35%	4,949 65%	858 11%
France	6,495 2%	2,831 44%	3,664 56%	2,192 34%	4,303 66%	992 15%
South Korea	4,584 2%	2,643 58%	1,941 42%	2,731 60%	1,853 40%	392 9%
Netherlands	4,482 2%	2,378 53%	2,104 47%	1,627 36%	2,855 64%	592 13%
Hong Kong	2,070 1%	1,339 65%	731 35%	1,032 50%	1,038 50%	197 10%
Taiwan	2,374 1%	1,436 60%	938 40%	1,196 50%	1,178 50%	185 8%
Spain	3,724 1%	2,131 57%	1,593 43%	1,365 37%	2,359 63%	439 12%
Russian	4,373 2%	1,250 29%	3,123 71%	1,015 23%	3,358 77%	595 14%
Australia	2,162 1%	1,346 62%	816 38%	931 43%	1,231 57%	248 11%
Turkey	1,665 1%	1,068 64%	597 36%	888 53%	777 47%	175 11%
Sweden	2,499 1%	1,344 54%	1,155 46%	940 38%	1,559 62%	284 11%
Israel	2,092 1%	1,276 61%	816 39%	980 47%	1,112 53%	206 10%
Italy	1,759 1%	1,006 57%	753 43%	600 34%	1,159 66%	240 14%
Denmark	1,600 1%	790 49%	810 51%	660 41%	940 59%	197 12%
Brazil	902 0%	708 78%	194 22%	450 50%	452 50%	65 7%
Greece	651 0%	371 57%	280 43%	292 45%	359 55%	75 12%
Thailand	1,140 0%	815 71%	325 29%	562 49%	578 51%	84 7%
Switzerland	808 0%	475 59%	333 41%	326 40%	482 60%	93 12%
Czech Republic	1,040 0%	327 31%	713 69%	279 27%	761 73%	159 15%
Norway	1,240 0%	472 38%	768 62%	501 40%	739 60%	140 11%
Belgium	879 0%	541 62%	338 38%	311 35%	568 65%	115 13%
Poland	798 0%	294 37%	504 63%	213 27%	585 73%	114 14%
Austria	757 0%	348 46%	409 54%	295 39%	462 61%	95 13%
India	405 0%	226 56%	179 44%	187 46%	218 54%	36 9%
Singapore	456 0%	253 55%	203 45%	172 38%	284 62%	60 13%
Hungary	764 0%	273 36%	491 64%	173 23%	591 77%	126 16%

Country	Number of Cookies Set		Number of 1st Party Cookies		Number of 3rd Party Cookies		Number of Sessional Cookies		Number of Persistent Cookies		Number of 3rd Party Persistent Cookies	
Mexico	336	0%	193	57%	143	43%	147	44%	189	56%	39	12%
Argentina	582	0%	338	58%	244	42%	231	40%	351	60%	62	11%
Finland	626	0%	247	39%	379	61%	204	33%	422	67%	91	15%
Malaysia	321	0%	218	68%	103	32%	155	48%	166	52%	29	9%
Lithuania	754	0%	556	74%	198	26%	267	35%	487	65%	42	6%
Saudi Arabia	239	0%	136	57%	103	43%	105	44%	134	56%	25	10%
Viet Nam	263	0%	205	78%	58	22%	150	57%	113	43%	17	6%
New Zealand	432	0%	165	38%	267	62%	171	40%	261	60%	61	14%
Egypt	224	0%	173	77%	51	23%	129	58%	95	42%	17	8%
Ireland	381	0%	222	58%	159	42%	145	38%	236	62%	44	12%
Romania	522	0%	205	39%	317	61%	104	20%	418	80%	76	15%
Portugal	346	0%	202	58%	144	42%	131	38%	215	62%	31	9%
Bulgaria	245	0%	159	65%	86	35%	117	48%	128	52%	42	17%
Ukraine	246	0%	94	38%	152	62%	57	23%	189	77%	30	12%
Chile	266	0%	115	43%	151	57%	140	53%	126	47%	34	13%
Uruguay	126	0%	78	62%	48	38%	66	52%	60	48%	14	11%
Virgin Islands, British	67	0%	37	55%	30	45%	36	54%	31	46%	10	15%
Venezuela	129	0%	79	61%	50	39%	70	54%	59	46%	14	11%
South Africa	404	0%	199	49%	205	51%	120	30%	284	70%	34	8%
United Arab Emirates	146	0%	79	54%	67	46%	79	54%	67	46%	15	10%
Indonesia	126	0%	60	48%	66	52%	80	63%	46	37%	11	9%
Slovenia	289	0%	135	47%	154	53%	108	37%	181	63%	31	11%
Slovakia	190	0%	76	40%	114	60%	55	29%	135	71%	36	19%
Serbia and Montenegro	70	0%	42	60%	28	40%	30	43%	40	57%	9	13%
Estonia	177	0%	74	42%	103	58%	74	42%	103	58%	27	15%
Costa Rica	105	0%	66	63%	39	37%	41	39%	64	61%	15	14%
Croatia	85	0%	61	72%	24	28%	43	51%	42	49%	12	14%
Philippines	62	0%	51	82%	11	18%	27	44%	35	56%	2	3%
Colombia	76	0%	31	41%	45	59%	24	32%	52	68%	9	12%
Panama	80	0%	50	63%	30	38%	30	38%	50	63%	8	10%
Peru	35	0%	15	43%	20	57%	28	80%	7	20%	5	14%
Latvia	99	0%	44	44%	55	56%	34	34%	65	66%	16	16%
Iceland	73	0%	51	70%	22	30%	39	53%	34	47%	13	18%
Jordan	30	0%	12	40%	18	60%	23	77%	7	23%	4	13%
Gibraltar	149	0%	105	70%	44	30%	52	35%	97	65%	15	10%
Kuwait	36	0%	20	56%	16	44%	26	72%	10	28%	2	6%
Malta	34	0%	24	71%	10	29%	22	65%	12	35%	2	6%
Dominica	39	0%	13	33%	26	67%	24	62%	15	38%	8	21%
Belize	31	0%	22	71%	9	29%	17	55%	14	45%	3	10%
Macao	22	0%	18	82%	4	18%	14	64%	8	36%	2	9%
Cyprus	47	0%	33	70%	14	30%	25	53%	22	47%	6	13%
Syrian Arab Republic	26	0%	12	46%	14	54%	21	81%	5	19%	1	4%
Luxembourg	43	0%	27	63%	16	37%	17	40%	26	60%	4	9%
Bermuda	82	0%	69	84%	13	16%	24	29%	58	71%	4	5%

Country	Number of Cookies Set		Number of 1st Party Cookies		Number of 3rd Party Cookies		Number of Sessional Cookies		Number of Persistent Cookies		Number of 3rd Party Persistent Cookies	
Iran	16	0%	13	81%	3	19%	9	56%	7	44%	1	6%
Grenada	17	0%	2	12%	15	88%	1	6%	16	94%	8	47%
Pakistan	22	0%	11	50%	11	50%	15	68%	7	32%	1	5%
Dominican Republic	48	0%	25	52%	23	48%	18	38%	30	63%	3	6%
Ecuador	8	0%	7	88%	1	13%	6	75%	2	25%	1	13%
Puerto Rico	37	0%	13	35%	24	65%	22	59%	15	41%	4	11%
Lebanon	20	0%	15	75%	5	25%	11	55%	9	45%	3	15%
Morocco	4	0%	2	50%	2	50%	2	50%	2	50%	1	25%
Qatar	16	0%	4	25%	12	75%	13	81%	3	19%	1	6%
Netherlands Antilles	16	0%	9	56%	7	44%	9	56%	7	44%	2	13%
Antigua and Barbuda	25	0%	23	92%	2	8%	10	40%	15	60%	1	4%
Bosnia and Herzegovina	41	0%	16	39%	25	61%	19	46%	22	54%	4	10%
Tunisia	14	0%	7	50%	7	50%	10	71%	4	29%	1	7%
Bolivia	3	0%	1	33%	2	67%	2	67%	1	33%	1	33%
Bahamas	7	0%	2	29%	5	71%	1	14%	6	86%	2	29%
Guatemala	9	0%	9	100%	0	0%	5	56%	4	44%	0	0%
Sri Lanka	6	0%	2	33%	4	67%	3	50%	3	50%	1	17%
Mauritius	13	0%	3	23%	10	77%	4	31%	9	69%	2	15%
Yemen	1	0%	1	100%	0	0%	0	0%	1	100%	0	0%
Libyan Arab Jamahiriya	1	0%	1	100%	0	0%	0	0%	1	100%	0	0%
Algeria	3	0%	3	100%	0	0%	1	33%	2	67%	0	0%
Paraguay	4	0%	2	50%	2	50%	3	75%	1	25%	1	25%
Macedonia	12	0%	9	75%	3	25%	8	67%	4	33%	2	17%
Barbados	0	0%	0		0		0		0		0	
Belarus	13	0%	7	54%	6	46%	1	8%	12	92%	1	8%
Moldova, Republic Of	20	0%	9	45%	11	55%	5	25%	15	75%	3	15%
Bahrain	1	0%	1	100%	0	0%	1	100 %	0	0%	0	0%
Samoa	0	0%	0		0		0		0		0	
Sudan	0	0%	0		0		0		0		0	
Honduras	1	0%	1	100%	0	0%	1	100 %	0	0%	0	0%
Saint Kitts and Nevis	2	0%	2	100%	0	0%	2	100 %	0	0%	0	0%
Cote Divoire	4	0%	3	75%	1	25%	3	75%	1	25%	1	25%
Nicaragua	8	0%	5	63%	3	38%	3	38%	5	63%	1	13%
Myanmar	2	0%	1	50%	1	50%	2	100 %	0	0%	0	0%
Guinea-Bissau	0	0%	0		0		0		0		0	
Botswana	3	0%	0	0%	3	100%	3	100 %	0	0%	0	0%
Trinidad and Tobago	0	0%	0		0		0		0		0	
Cuba	1	0%	1	100%	0	0%	1	100 %	0	0%	0	0%

Country	Number of Cookies Set		Number of 1st Party Cookies		Number of 3rd Party Cookies		Number of Sessional Cookies		Number of Persistent Cookies		Number of 3rd Party Persistent Cookies	
Albania	0	0%	0		0		0		0		0	
Fiji	2	0%	2	100%	0	0%	0	0%	2	100%	0	0%
Liechtenstein	3	0%	3	100%	0	0%	1	33%	2	67%	0	0%
American Samoa	3	0%	0	0%	3	100%	0	0%	3	100%	1	33%
Georgia	6	0%	3	50%	3	50%	2	33%	4	67%	1	17%
Nigeria	0	0%	0		0		0		0		0	
Holy See (Vatican City State)	0	0%	0		0		0		0		0	
Iraq	0	0%	0		0		0		0		0	
Kazakhstan	0	0%	0		0		0		0		0	
Oman	0	0%	0		0		0		0		0	
Afghanistan	7	0%	1	14%	6	86%	3	43%	4	57%	1	14%
Cambodia	1	0%	1	100%	0	0%	1	100 %	0	0%	0	0%
El Salvador	0	0%	0		0		0		0		0	
Brunei Darussalam	0	0%	0		0		0		0		0	
Turks and Caicos Islands	4	0%	0	0%	4	100%	1	25%	3	75%	1	25%
Bangladesh	0	0%	0		0		0		0		0	
Faroe Islands	2	0%	0	0%	2	100%	1	50%	1	50%	1	50%
Totals	284,073	100%	160,770	57%	123,303	43%	111,495	39%	172,578	61%	29,029	10%