source DEFENSE

Third-Party Digital Supply Chain Risk: Exposing the Shadow Code on Your Web Properties

A quantitative analysis of 4,300 websites

Third-Party Scripts and Shadow Code

Third-party scripts are necessary and pervasive...

Client-side JavaScript allows developers to make the browser-based portion of web applications more feature-rich and more appealing to website visitors. Scripts available from third parties allow development teams to add advanced functionality to applications without the burden of creating and maintaining them. Today, extensive libraries of such scripts are available free or at low cost from open-source software organizations, individuals and allied groups of coders, and third parties such as cloud hosting providers, social media companies, digital advertising networks, web analytics firms, and content delivery networks. The benefits of third-party scripts are so overwhelming that it is rare to find a significant commercial enterprise or government agency that doesn't use a large number in its web applications. Indeed, as we shall see in the data presented later in this report, most websites have third-party scripts on every page.

Common Uses of Client-Side JavaScript

- Widgets
- Dynamic forms
- Processing orders and payments
- Social media sharing buttons
- Displaying ads
- Displaying videos
- Site analytics
- Chatbots

...but they introduce "shadow code" risks

Third-party scripts can inject malicious "shadow code" (code that has never been inspected and validated by the site owner's IT department) into websites. When a third-party script is called by a web page, it is loaded into a browser directly from a remote server belonging to the third party (say, a social media or marketing company or an analytics tool vendor), bypassing traditional security controls such as perimeter firewalls, web application firewalls (WAFs), and network monitoring tools. If a script has been hacked or compromised by a threat actor, the shadow code comes with it.

Code reviews are not an effective measure against shadow code. Scripts are changed constantly by the third parties. In addition, many of the scripts are highly dynamic, delivering different code to browsers depending on user activities. It simply isn't possible for website owners to obtain and review every permutation of every version of every script that works with their applications.

Threat actors are very aware of the opportunities presented by shadow code. Attacks through digital supply chain partners have been on the rise since 2014. Cybercriminals and others are attracted by the ability of shadow code in third-party scripts to exfiltrate data to remote servers, redirect users to malicious websites, and to lay the groundwork for formjacking, digital skimming, and credential harvesting attacks. They are also aware that compromised scripts can provide a path into hundreds or thousands of websites, in virtually every industry that collects data or conducts transactions online, potentially yielding data from millions upon millions of website visitors.

What Makes Third-Party Scripts with Shadow Code so Dangerous?

A third-party script infected by shadow code can:

- Change content on web pages
- Add new content to web pages, including images, text, video, and form fields
- Record keystrokes
- Capture and exfiltrate website visitor credentials, payment card information, and social security numbers
- Monitor clicks on buttons and links to track website visitor behavior
- Redirect visitors to websites under the control of the attacker

Results may include:

- Data breaches
- Fraud against customers and business partners
- Compliance violations leading to regulatory fines and data breach notification costs

Can the third parties supplying the scripts be relied on to keep them safe? Most do their best, but inevitably some will be compromised. Their scripts often include open-source software that may not be reviewed or tested rigorously (or at all). They also incorporate code from other organizations, which may include code from yet other organizations, which may include code from yet another layer of organizations, any of which may include malicious shadow code. And ultimately, customers, business partners, and regulators will hold website owners responsible for incidents on their websites, regardless of the source of the code.

Fourth-, Fifth-, and Nth-Party Scripts

From the perspective of the ultimate website owner, when a third-party includes a script from another source, that is a “fourth-party script.” If the other source includes code from yet another developer, that is a “fifth-party script.” If the chain goes even farther, we can talk about “Nth-party scripts.”

In this report, we will sometimes differentiate between third- and fourth-party scripts, but for the most part “third-party scripts” will be used to refer collectively to third-, fourth-, fifth, and Nth-party scripts.

How Big is the Risk? Now You Have Data

Some cyberthreats have grave consequences, others are merely annoyances. What about third-party scripts – are they a big problem, or a small one?

Surprisingly, little data is available to help organizations make that determination. To fill that gap, during the first quarter of 2022 Source Defense scanned the top 4,300 websites by traffic worldwide and analyzed the data to provide answers to questions such as:

- How many third-party and fourth-party scripts can be found on the average website?
- How many on each web page?
- How many third-party scripts are on sensitive pages (i.e., pages that handle credentials, account numbers, payments, and other sensitive data and transactions)?
- What are the scripts being used for?
- How do the numbers and types of scripts vary by industry?

We also examine analysis findings from two real organizations, a leading brokerage and banking firm and a global hotel and hospitality company, to see how organizations can use information about third-party scripts on their websites to strengthen security.

Scripts Per Site and Scripts Per Page

Scripts per site

How many scripts appear on each website? Our analysis shows a total of 15 on average across all sites in the sample. Of these, 12 are third-party scripts and three are fourth-party scripts (see Figure 1).

It is important to note that these are averages. We've found several dozen scripts in use on many of the world's most trafficked websites.

Some of these scripts are present on every page on a site. Common examples are scripts used for website analytics and links to social media sites. Other scripts provide special functionality and appear on only one or a few pages. This might include scripts that add interactivity to forms and those that display videos.

Fourth-party scripts are less numerous than third-party scripts, but they represent an even higher level of risk to the website owner. By compromising scripts created by entities farther up the supply chain, malicious actors can circumvent the security controls of the third parties. Yet the website owner has no visibility into the defenses of those fourth parties and no ability to communicate with them. In fact, *the website owner is often unaware that the third party is using any fourth-party code at all.* Security monitoring and functionality failures are also issues with fourth-party scripts, because it is often impossible to determine when problems need to be addressed by the third party or a fourth party.

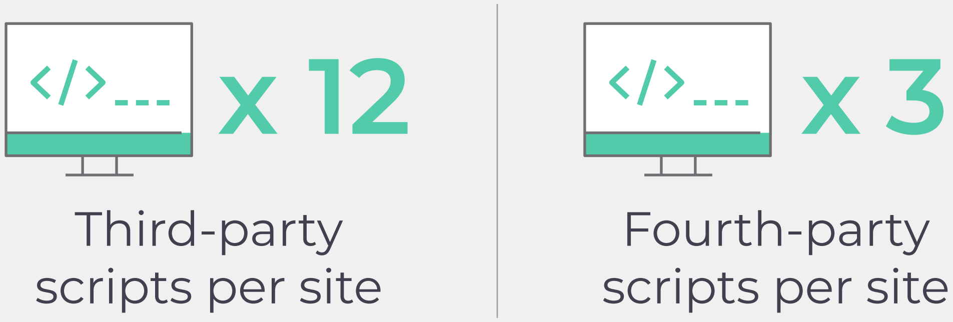

Figure 1: Third-party and fourth-party scripts per site (average for all websites in the sample)

Scripts per page

Another way we cut the data was to look at how many scripts appear on each web page. The average, across all the websites in the sample, is five: four third-party scripts and one fourth-party script (see Figure 2).

Typically, the five might include two or three of the scripts that appear throughout the site and two or three that address specific requirements on the page.

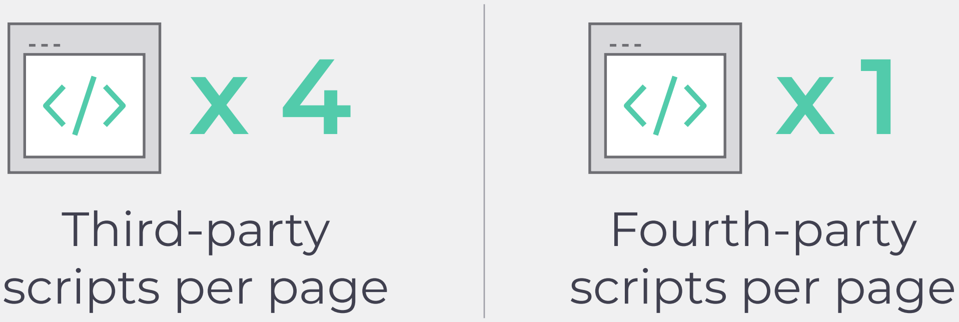

Figure 2: Third-party and fourth-party scripts per page (average for all websites in the sample)

What conclusions can we draw from this data? It shows that the challenge for the security team of an average website is not checking for shadow code on a handful of third- and fourth-party scripts, but on fifteen (and considerably more in some industries, as we will see). Even when security teams have the tools to monitor the behavior of scripts, they may need to investigate hundreds of incidents a day. Most of these checks will show no issues, but some are likely to involve malicious shadow code with the potential to cause fraud or data breaches.

Dynamic Conditions Increase Risk

Our data represents a snapshot of third- and fourth-party scripts at a moment in time. But the situation is more dynamic and more challenging.

Many organizations are continually transforming their web presence with new marketing, ecommerce, social media, customer support, and supply chain projects. They also rapidly replace existing digital suppliers and business partners with new ones. This churn means that over the course of a year the security team for an average website might need to monitor perhaps 50% or 100% more third- and fourth-party scripts than are on the site at any one time.

Scripts on Sensitive Pages and Scripts Accessing Form Data

Scripts on Sensitive Pages

Location, location, location. That's what matters in real estate, and also in shadow code. Malicious actors gain little from compromising a script that appears on a web page that only displays static information. But scripts on sensitive pages, such as login and credential capture pages, account registration pages, and payment collection pages, are another story.

Unfortunately, our data shows that 12 third- and fourth-party scripts appear on sensitive pages on the average website.

Why so many? First, those scripts that appear on every page on a website obviously will appear on sensitive pages. Second, web teams are most likely to utilize client-side scripts on pages where customers and other users log into accounts, enter PII, and make transactions. Those are the pages where interactivity, dynamic content, and excellent performance are most important. They are also the pages where the website owner has the greatest need for analytics, tag management, and other tracking and management capabilities.

Scripts on sensitive pages

Figure 3: Third- and fourth-party scripts on sensitive pages (average for all websites in the sample)

Scripts Potentially Providing Access to Form Data

There is another aspect of potential risk in client-side third- and fourth-party scripts. Many scripts include procedures that serve a legitimate purpose but can be modified to serve the interests of an attacker. For example, a script might allow form fields to be changed or added on the fly to provide website users with a more personalized experience. However, a threat actor could exploit this capability to add additional fields asking for credentials and personal information, which would then be sent to attacker's website.

Another example would be a feature that tells the website team when users click on a specific button and what happens next. Bad actors might redirect this information to their own systems to learn more about what where sensitive data is going, enabling them to target that data.

We scanned scripts on websites looking for six types of features that could be exploited by bad actors:

- **Code to retrieve form input:** code that can perform data skimming by retrieving data from forms and input fields during interactions with website visitors
- **Code to modify forms:** code that can add, remove, and change fields displayed to website visitors.
- **Button click listeners:** code that detects when a website visitor interacts with a button or field on a form.
- **Link click listeners:** code that monitors user behavior on web pages outside of forms, especially clicks on hyperlinks.
- **Form submit listeners:** code that exfiltrates data entered on a web page to a destination not intended by the website owner
- **Input change listeners:** code that observes when website visitors enter data into web pages

The results of our analysis are shown in Figure 4.

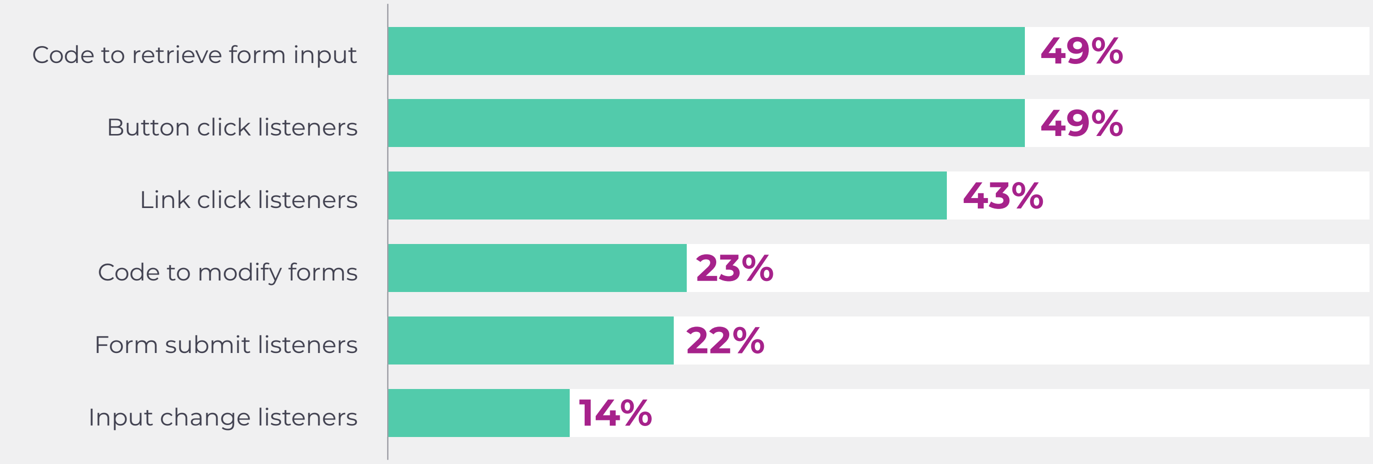

Figure 4: Percentage of sites with scripts providing potential access to form data

We found:

- Code to retrieve form input and button click listeners were present in third-party scripts on roughly half of all sites (49%).
- Link click listeners were found on more than two out of five sites (43%).
- Code to modify forms, form submit listeners, and input change listeners were found on significant numbers of sites (23%, 22%, and 14%, respectively).

Virtually every modern, dynamic website we reviewed had one or more of these features.

Industry Profiles

While overall averages are helpful, many organizations would like to benchmark themselves against others in the same industry. To see how the figures we have been discussing can vary across business types, we analyzed data from 40-50 organizations in each of four major industries: ecommerce, travel and hospitality, healthcare, and finance and financial services. The results are shown in Figures 5, 6, and 7.

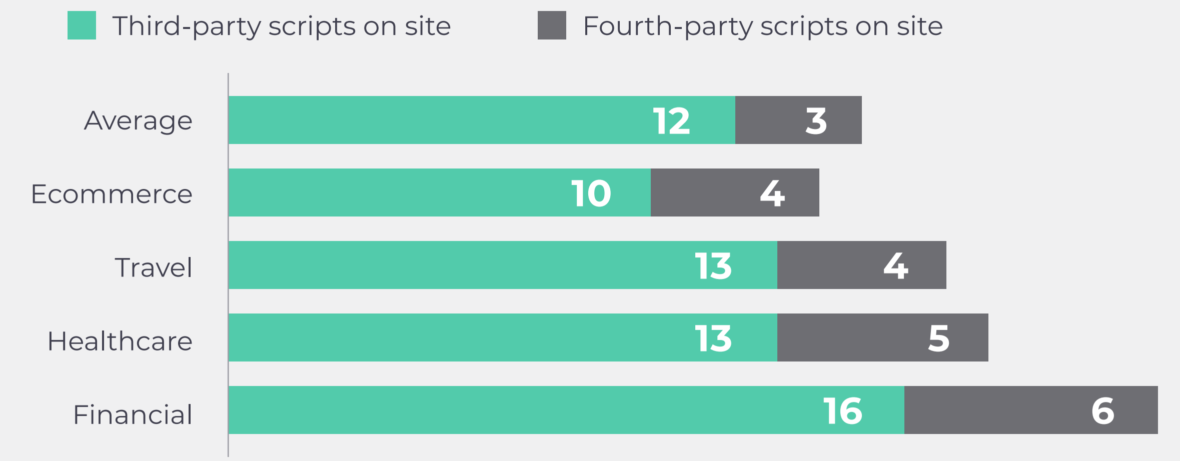

Figure 5: Third-party and fourth-party scripts on the website, by industry

Compared with the average of websites in our sample, ecommerce companies have fewer third-party scripts on their sites (10 versus 12), but more fourth-party scripts (4 versus 3). When it comes to scripts per page, they have the same number of third-party scripts (4) and one more fourth-party script (2 versus 1). The number of scripts on sensitive pages is slightly below average (11 versus 12). Overall, the profile for ecommerce companies is very close to the average.

When we look at travel and healthcare companies, risk levels are higher than average. On their sites they have more third-party scripts (13 versus 12) and more fourth-party scripts (4 and 5, respectively, versus 3). On a per-page basis, they have more third-party scripts (5 and 6 versus 4) and fourth-party scripts (2 versus 1). For scripts on sensitive pages, they have 13 and 14, respectively, versus 12.

Finance organizations have even greater challenges. For scripts on site, they have 16 third-party scripts to monitor (compared to an average of 12) and twice as many fourth-party scripts (6 versus 3). Third-party scripts per page are 7 versus the average of 4, and fourth-party scripts per page are 3 versus 1. And they have 19 scripts on sensitive pages, compared to the average of 12. In short, on these measurements the figures for finance firms range from 133% to 300% of the average across the sample.

It is interesting to note that even in finance, one of the world's most threat-aware industries, with unrivaled investment in security technology and staffing, major third-party risks still lurk on mission critical web properties.

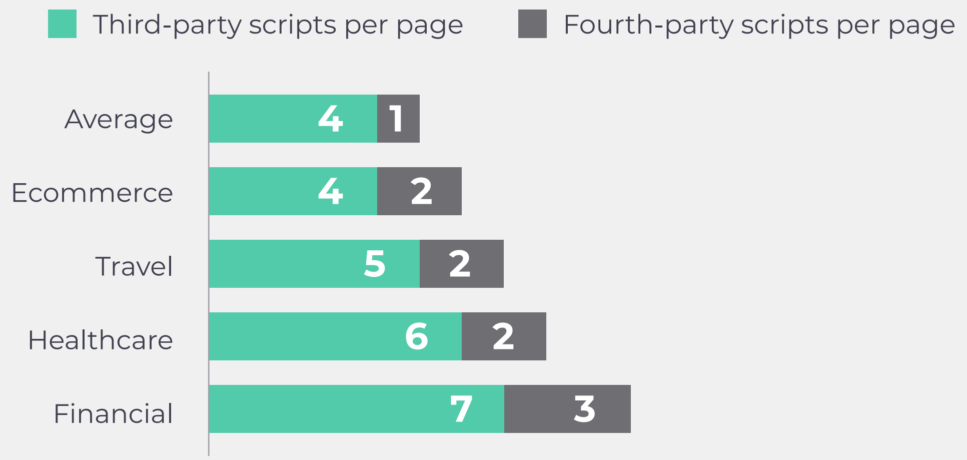

Figure 6: Third-party and fourth-party scripts per page, by industry

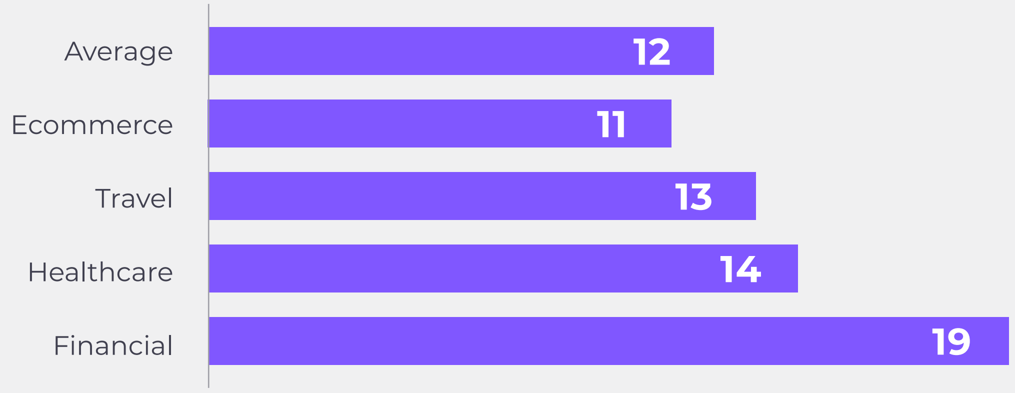

Figure 7: Third-party and fourth-party scripts on sensitive pages, by industry

The findings so far

To sum up our analysis so far:

Third- and fourth-party scripts are very widely used. The average website includes 12 third-party scripts and three fourth-party scripts. The average web page includes five scripts: four third-party scripts and one fourth-party script.

The average website has 12 third- or fourth-party scripts on pages that handle sensitive data.

A large majority of websites have third- and fourth-party scripts with features that potentially provide access to form data.

Some industries have significantly more scripts and higher risk factors than the overall averages.

This data makes it clear that managing risk inherent in third- and fourth-party scripts is both a very necessary and a very challenging task.

We will now look at two examples of specific enterprises and see how they can use data about scripts to strengthen their security.

How Can This Data Help Organizations Strengthen Their Security?

The data from our analysis has provided some clarity about the level of risks created by third-party and fourth-party scripts. But can it also be used to help organizations mitigate those risks?

To explore that question, let's look at data from two specific organizations. One is a leading brokerage and banking firm. The other is a global hotel and hospitality company.

Example #1: The Brokerage and Banking Firm

What might our example brokerage and banking firm learn by comparing its data with averages for the finance and financial services industry (see Figures 8, 9, and 10)?

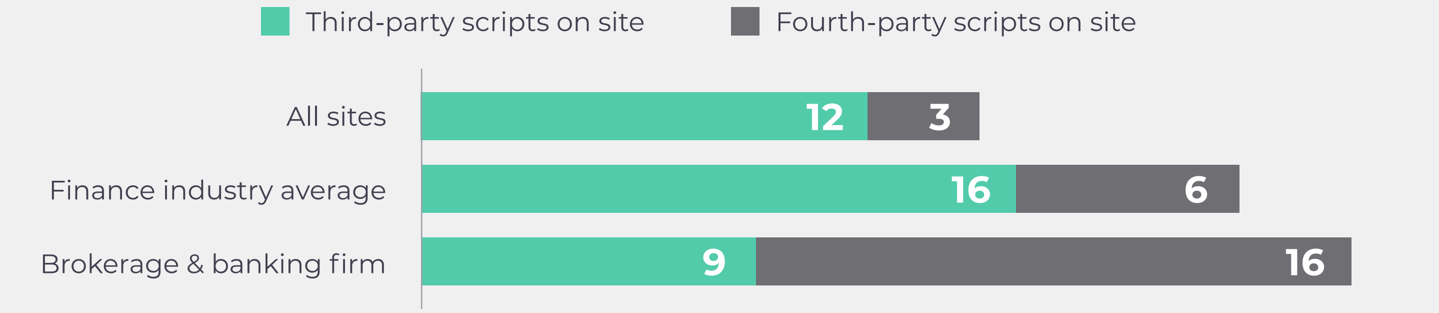

Figure 8: Scripts on the website: all sites, finance industry, and example brokerage firm

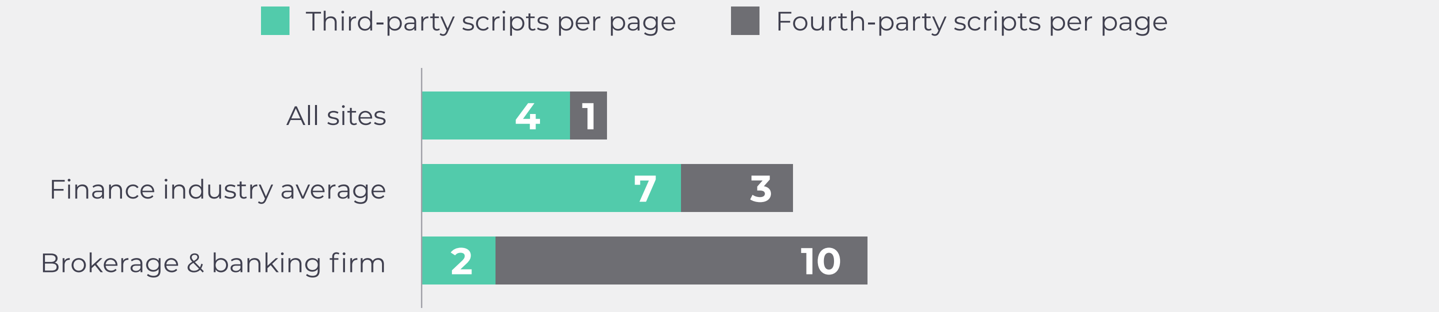

Figure 9: Scripts per page: all sites, finance industry, and example brokerage firm

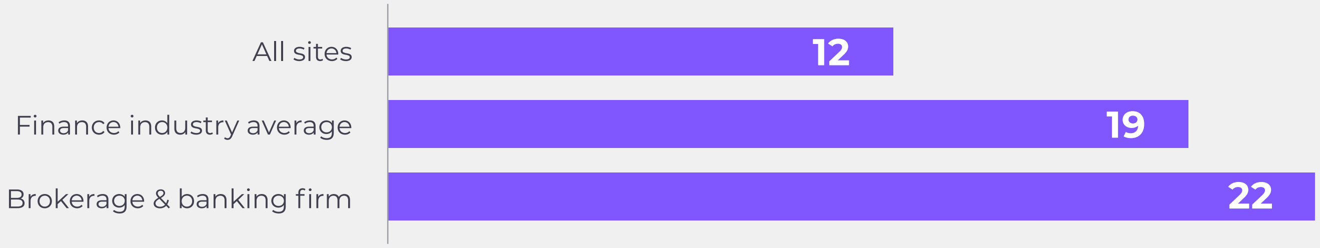

Figure 10: Scripts on sensitive pages: all sites, finance industry, and example brokerage firm

Compared with both the overall and finance industry averages, this firm uses significantly fewer third-party scripts, both on the site as a whole and on each page. That might suggest that third-party scripts should not be a major area of concern for its security team.

However, the firm's website includes many more fourth-party scripts. The total number (16) is five times the overall average (3), and more than double the finance industry average (6). The average page on its site has 10, which is an order of magnitude more than the overall average, and more than three times the finance industry average.

Finance websites tend to access more fourth-party scripts than those of most other industries. Typically, they use third-party scripts that include a lot of data feeds from fourth-party sources related to securities and commodity prices, news feeds, and client financial data. However, this particular firm far exceeds the industry norm. And as we discussed earlier, fourth-party scripts involve more risk factors than third-party scripts.

The data also shows more scripts on sensitive pages (22) than either the overall average (12) or the finance industry average (19). This also indicates a higher than typical exposure to risk.

We can drill down even further and see the specific scripts on this company's website (Figure 11). This shows seven different scripts used for advertising, four for analytics, and three each for social media and tag management.

What does this data suggest? Well, perhaps there are sound reasons why this organization's website includes so many scripts, including so many fourth-party scripts. But the security team should do some investigation to understand:

- Why are there so many fourth-party scripts?
- Could the number of scripts be reduced (perhaps by consolidating to fewer scripts for advertising, analytics, or tag management)?
- Could some scripts be removed from the pages that handle sensitive information?
- Does the firm need to invest more in tools or people to assess and monitor these scripts?
- Better yet, are there steps the security team could take to prevent client-side attacks from occurring, which would reduce risk and the costs of monitoring third- and fourth-party scripts?

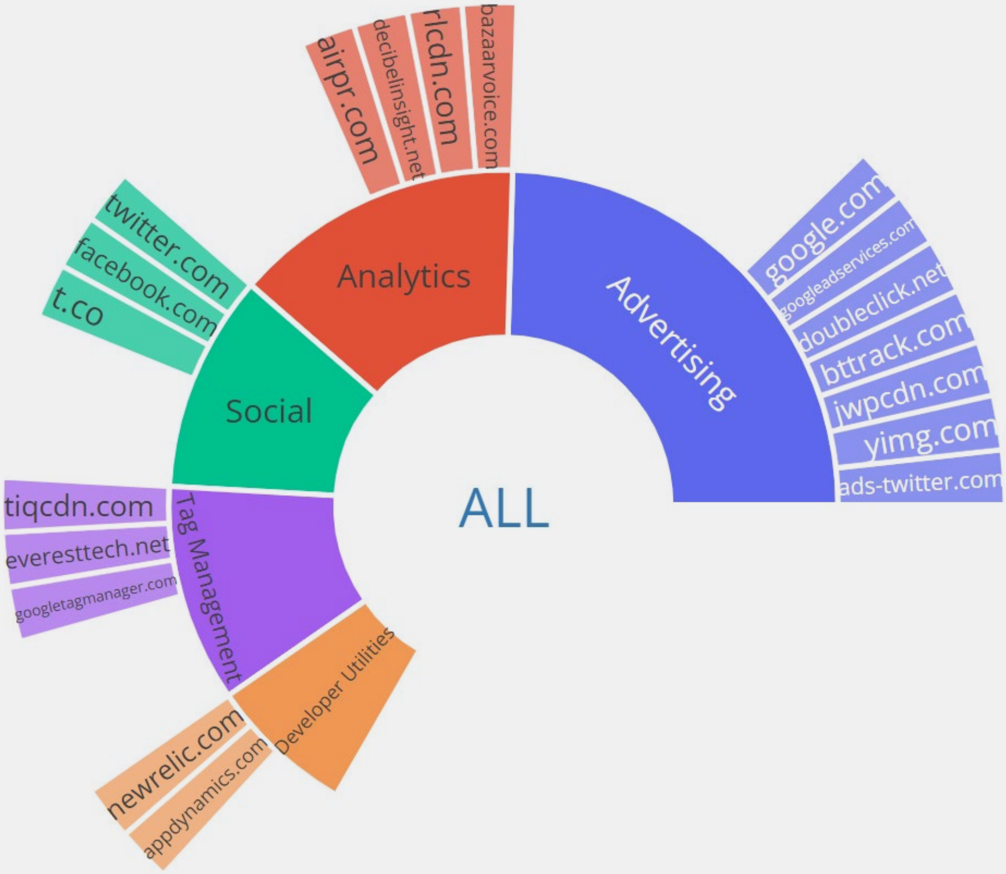

Figure 11: Breakdown of scripts on the financial firm's website

Example #2: The Hotels and Hospitality Company

Let's look now at how an example hotel and hospitality company could use data about scripts to strengthen its security posture.

Figures 12 and 13 show that this company's web footprint involves roughly three times the number of third- and fourth-party scripts as typical organizations in the sample and in the travel industry. Because this corporation runs several businesses, above-average numbers might be expected. Its websites support a wide array of services, including hotel and resort reservations, meeting and conference planning, guest reward programs, additional travel-related offerings, and even credit cards. However, the abundance of scripts indicates a very high level of risk. The fact that 34 scripts appear on sensitive pages (Figure 14) is another conspicuous red flag.

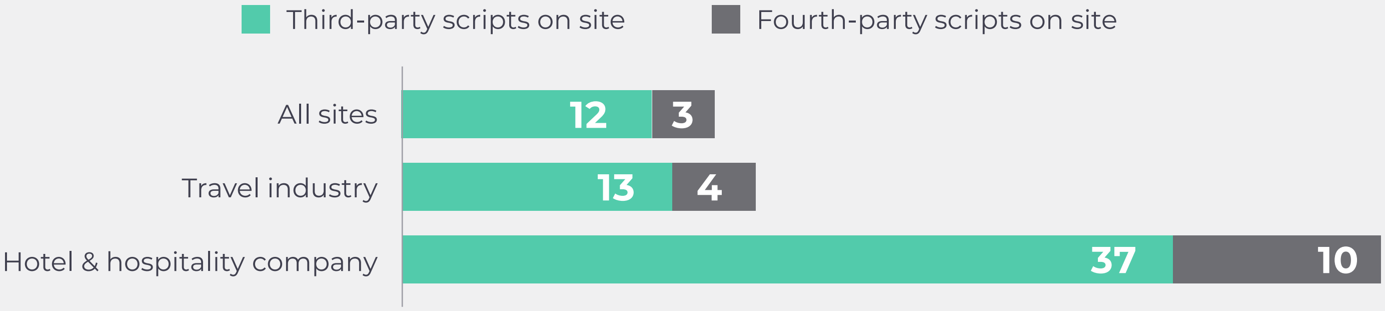

Figure 12: Scripts on the website: all sites, travel industry, and example hotel company

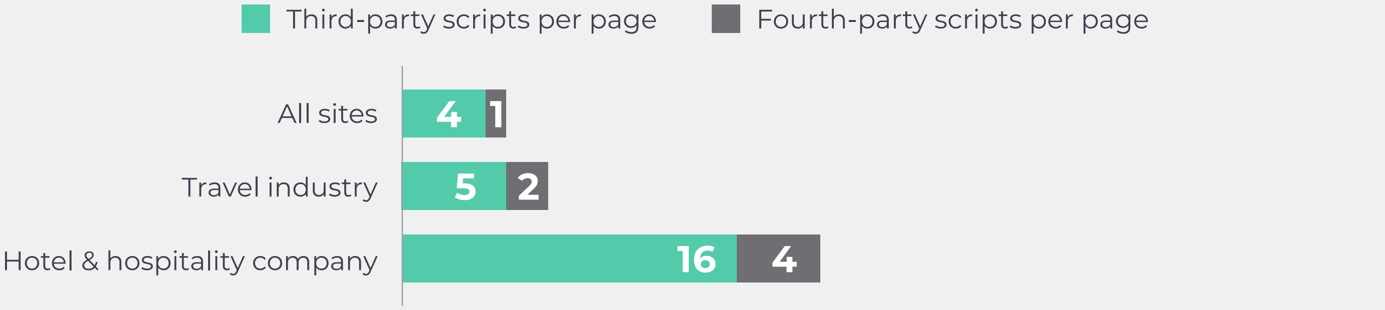

Figure 13: Scripts per page: all sites, travel industry, and example hotel company

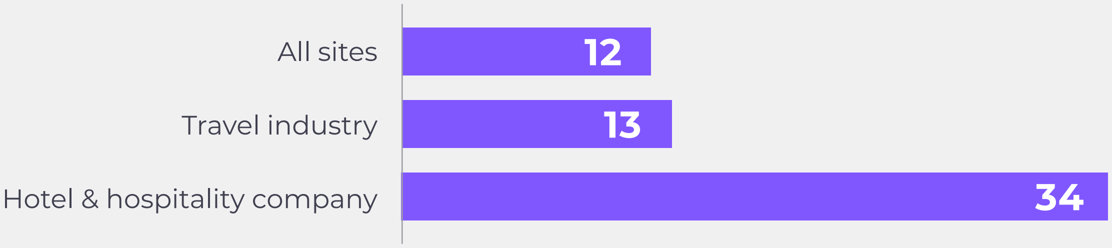

Figure 14: Scripts on sensitive pages: all sites, travel industry, and example hotel company

The data in Figure 15 shows multiple third- and fourth-party scripts in several categories. The analysis found:

- 15 scripts related to advertising
- 7 for analytics
- 7 involving CDNs
- 6 related to social media
- 5 for developer utilities

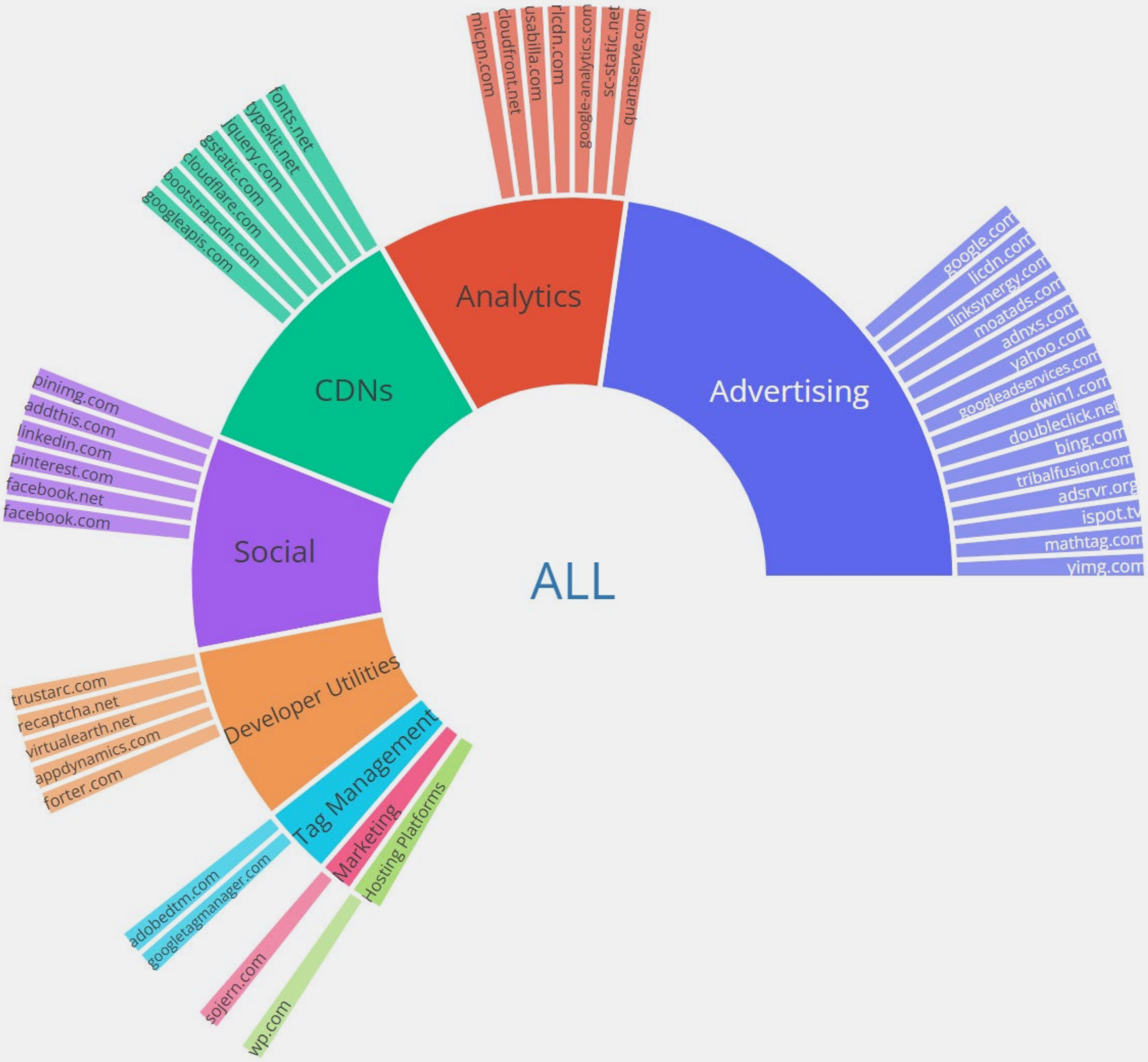

Figure 15: Breakdown of scripts on the hotel company's website

This proliferation suggests that multiple web development teams are operating independently of each other, each selecting scripts on their own. We suspect that no individual or group has an overall perspective on the number of scripts involved or the risks they create.

The company's security team, along with IT management, can use the data in this analysis as to explore question such as:

- Do we have a good grasp of the risk created by third- and fourth-party scripts?
- Can we bring our siloed web development groups together to agree on fewer scripts, and perhaps fewer partners, in areas like advertising, analytics, CDNs, and developer utilities? (Besides strengthening security, consolidation might reduce costs and improve operations.)
- Why do we have 34 third- and fourth-party scripts on pages that handle sensitive information, when other companies in our industry average 13?
- Can we possibly assess and monitor 47 scripts with the resources available to us?
- Are there preventative solutions that would allow us to minimize risks from third- and fourth-party scripts quickly, at a low cost?

These are all important questions. But organizations can also implement a systematic program to manage client-side application risks. We will now describe the key activities in such a program.

Developing a Program to Manage Client-Side Application Risks

To help assess your website, you can request a custom report from Source Defense, which takes less than a week.

For quick wins with a website client-side security solution, consider Source Defense VICE, which uses behavioral analysis and advanced machine learning to automatically detect vulnerable scripts, suspicious PII access, and data leakage from browsers. It provides real-time visibility into script interactions, protects websites against digital skimming, formjacking, and Magecart attacks, prevents data breaches, and reduces the risk of non-compliance with regulations and industry standards. Because VICE can be deployed simply by pasting two lines of code into the website's page headers, it can reduce risks associated with shadow code immediately, giving web and security teams breathing room to rationalize the management of third-party scripts on their websites.

Assess your website

Acquire the types of data described here: total scripts on the site and average scripts per page, scripts on sensitive pages, and code on scripts that potentially provides unauthorized access to form data. Compare this data with averages for your industry to provide insights on where to concentrate your efforts.

Educate management

Use the data from the assessment to educate IT and corporate management on the risks inherent in third-party scripts. The additional resources listed below also can help. Get a commitment to create a program to reduce and manage client-side application risks, or to expand an existing third-party risk management program to include this function.

Provide quick wins with a website client-side security solution

Technology solutions are available to detect and isolate third-, fourth-, and Nth-party scripts on your website and control their behavior. The right approach can be implemented quickly, without affecting the performance of websites or networks or adding to the burden on web or operations staffs, producing quick wins for the security team with minimal effort.

Categorize and consolidate scripts

Categorize the scripts based on their purpose. Determine if categories can be consolidated down to fewer entries without losing value. Besides reducing the size of your attack surface, that will simplify website administration.

Find ways to reduce exposure and compliance risks

Assess whether all the scripts are really adding value. Can any be eliminated? Can some be removed from sensitive pages? Determine whether compromised scripts might result in compliance violations or regulatory fines. Assess whether your organization's efforts to manage this risk represent reasonable efforts based on the standards in your industry.

Additional Resources

Blog posts

The Top 3 Things You Need to Know About Client-Side Web Application Attacks

Web App Client-Side Protection – Little Effort for a Big Win

Beyond the Server: Why CISOs Must Boost Their Defense of Client-Side Attacks

White papers

The Hidden Risk in Your Digital Supply Chain

Out of Compliance and Out of Sight: The Client-Side Web Security Gap That's Putting Your Business at Major Risk

Protecting eCommerce & Retail Sites from Client Side Attacks

Articles, research reports, and other resources

Article: British Airways breach caused by credit card skimming malware

Article: Macy's breach is a game-changing Magecart attack

Research report: Website Trust & Client-side Security Report 2021

Industry analyst report: Source Defense looks to close web app security gaps

Background FAQ: Source Defense FAQ

Video: Cyber Academy | Magecart 101

Offers from Source Defense:

- **Request a custom risk report**
- **Request a personal demo**
- **Subscribe to the Source Defense blog**

source DEFENSE