The State Of Website Tag Security

An Analysis Of The Top 1,000 Consumer Goods & Services Websites

We asked the question, do companies in the Consumer Goods & Services industry have robust website tag security? Or are they at risk of leaking personal data through their website tags?

We found that almost 80% of the top 1,000 Consumer Goods & Services websites^ hosted 'piggyback' tags – tags that are neither deployed directly onto the site nor via a Tag Management System. While piggyback tags can be legitimate they can also pose risks of data leakage if the company is not aware those tags are present on their site.

Introduction

Companies are collecting ever-increasing amounts of data on their customers, to personalize user experience and to optimize marketing efforts. With this comes increasing responsibility to protect the data collected. Failing to do so can result in a significant breach of trust and brand damage, additional to legal consequences.

Most companies have tags on their websites that collect data about visitors. This data is often shared with third-party vendors to drive advertising, marketing and optimization tools. While companies seek to directly manage the tags on their site, data can be unknowingly shared with fourth or fifth-party tags that are not in the direct control of the company. We call these 'piggyback' tags. As much of the data collected on websites is personal or sensitive, these piggyback tags can pose risk.

At DataTrue, we monitor thousands of websites for data leakage by auditing and reporting on tags that may be collecting sensitive data. We used this capability to better understand the potential risk website tags present to data protection in the Consumer Goods & Services industry. To do this we analyzed the website tags on the Alexa Top 1,000 Consumer Goods and Services websites worldwide^.

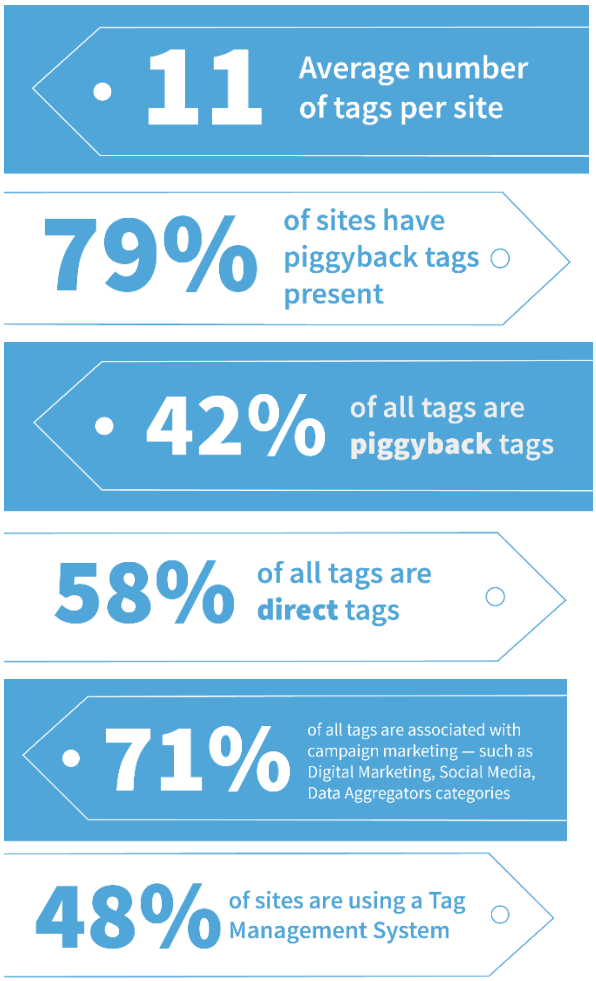
Summary Of Findings

We found that there was an average of 11 tags per site with a broad distribution. While over half of the websites had fewer than 10 tags, on one site the tag count exceeded 70.

More than 70% of all the tags related to campaign marketing activities – Digital Advertising, Social Media and Data Aggregators – indicating the influence that marketing teams and their suppliers have on website tags, and ultimately on data security.

Tags relating to Analytics and Tag Management Systems, the traditional territory of website tags, made up slightly over 20% of all the tags.

While 44% of all the tags were piggyback tags, we also found that 79% of the sites had at least one piggyback tag present. It is possible that the site owners are not aware that these piggyback tags are present on their site, or are not fully briefed on the data that these tags collect and share. As such they may not be in full control of their users' data. The clear majority of these piggyback tags related to Digital Advertising activities.



How Many Tags Are On The Websites?

While the average number of tags per site is 11, we found a broad distribution (chart 1). Almost 10% of sites had just one tag, and 58% of sites had fewer than ten tags. At the other end of the scale, 11% of the websites hosted more than 30 tags, with one site reporting over 70 tags. That's a lot of tags to keep track of!

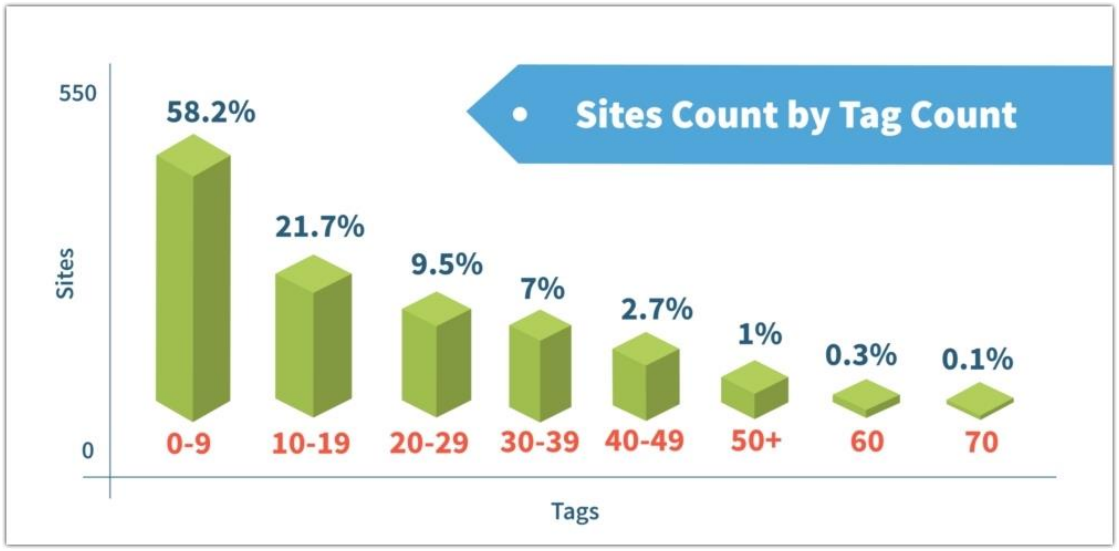


Chart 1

What Is The Purpose Of The Tags?

That's A Lot Of Tags. What Are They? And What Are They Used For?

Our analysis identified that just over 50% of all tags present were in the Digital Advertising category. Additional to this, Social Media tags accounted for over 16% of all tags and Data Aggregators tags tallied to just over 3%. These three tag categories, which are likely to be instigated by the marketing function in a company, accounted for 70.5% of all the tags on the websites analyzed.

While a lesser 20% of the tags deployed were for Tag Management Systems or Analytics purposes, the single most commonly used tag across all sites was Google Analytics, with Google Tag Manager the 6th most common tag.

In the Analytics and Tag Management Systems categories a small list of well-known tags is used, making knowledge and control of these tags far easier. Conversely in the Digital Analytics category, while there is a selection of predominant tags deployed across many sites, there is also a long tail of possibly lesser-known tags and vendors, all of which need to be researched and approved for data security according to individual companies' processes.

While Marketing teams have an increasing demand for and influence on website tags, typically the deployment of tags and security of websites is the remit of IT or Analyst teams. This evolution creates a growing intersect between Marketing, Analytics and IT functions. However corporate structures often house these roles and responsibilities in different specialist teams, often with different priorities. To deliver on evolving and fast-

moving Marketing data needs, while also maintaining website tag security, companies will need to embed governance protocols and effective ways of working between these functions.

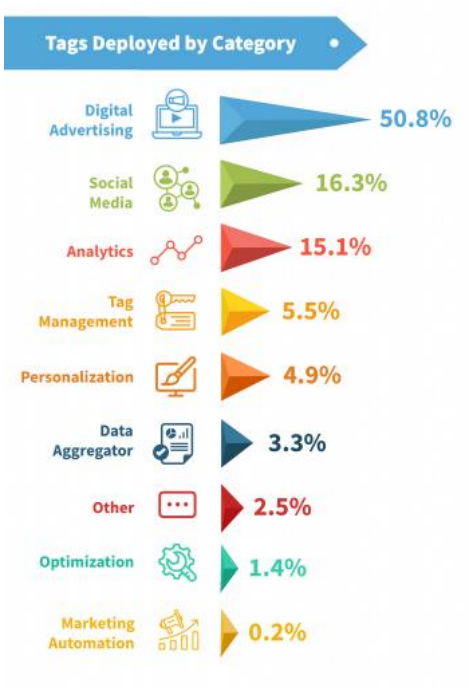


Chart 2

Direct And Piggyback Tags

What Are Piggyback Tags?

While companies seek to directly manage tags on their websites, data is often also on-shared with other tags that are not in the direct control of the company.

In the visual example of a tag hierarchy (image 1), we see Google Tag Manager and DART DFP tags directly on the site. We also see a number of tags that are at the second level deep, but deployed within the Google Tag Manager container, so also controlled by the company. We classify all of these as Direct tags, as they are clearly visible and within the control of the site owner.

Finally, we see one additional tag at the third level deep which is 'piggybacking' off the Google Analytics tag. We classify this as a piggyback tag, also known as an indirect tag. As anyone with administration permissions to Google Analytics can potentially enable this tag, those responsible for tag management and data security on this website may or may not be aware of this piggyback tag, what it collects, or where else it shares data.

While piggyback tagging can be legitimate, for example for identifying syncing, it can also pose risks if the company is not aware this piggyback tag is attached to the site, nor what data it is collecting and sharing.

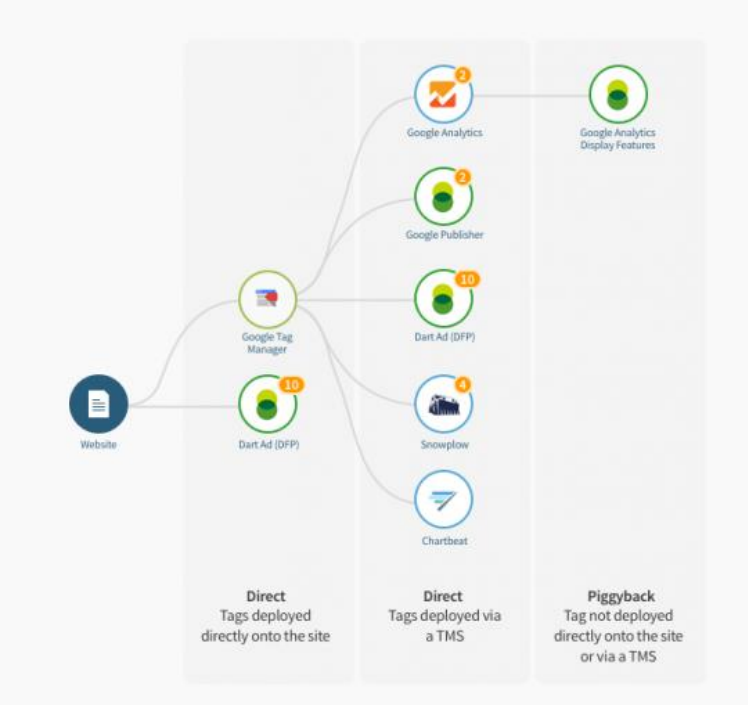


Image 1

Are Consumer Goods & Services Sites In Control Of The Data They Collect?

We evaluated how many tags were within the direct control of the top 1,000 Consumer Goods & Services sites. We classified as 'Direct' any tags that are coded directly onto the site or coded through a Tag Management System deployed on the site (indicating that the company has proactively deployed this tag and is aware of its presence). All other tags were classified as piggyback tags as they were not directly deployed.

Overall 58% of the tags were directly deployed and controlled by the website owner, with the remaining 42% piggyback tags. A significant 79% of the tagged sites we analyzed hosted at least one piggyback tag.

While not all piggyback tags are unknown to the site owner, there is a greater likelihood that the site owner is either not aware the tag exists or is not fully briefed on the data collected and shared via these tags.

We analysed this data by category (chart 3) and found that predictably 100% of Tag Management System tags were Direct. At the other extreme, only 31% of Digital Advertising tags were directly deployed onto the site or a Tag Management System, leaving 69% piggyback tags.

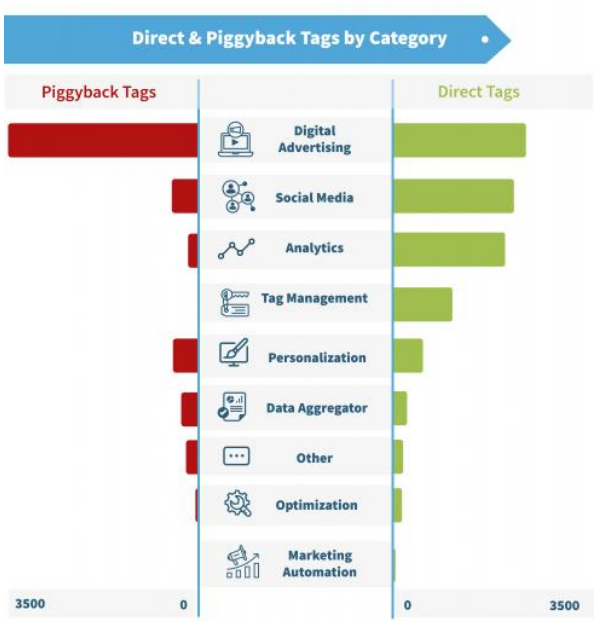


Chart 3

Breaking this down to an individual tag level (chart 4) we found that the Google Adwords Remarketing tag was indirectly deployed in 100% of cases. Other significant contributors to the piggyback tag count included: the GA Audiences tag, the Google Analytics Display Features tag and the Double Click Remarketing tag.

In many instances, these particular tags can be enabled through the administration interface of the parent tag, without a need for technical skills to deploy the tag. As such, anyone with the relevant permission level to that administration interface can enable the tag to be deployed. It may not even be apparent to the person enabling the feature that a tag is being deployed. As well as staff having access to these interfaces, suppliers are also sometimes given administrative permissions, meaning they could also enable the tags.

This demonstrates a scenario whereby piggyback tags could be deployed to a company's website without activating the usual internal governance required for tag deployment, and therefore those who are ultimately responsible for website tag security may be inadvertently excluded and unaware the tag is present.

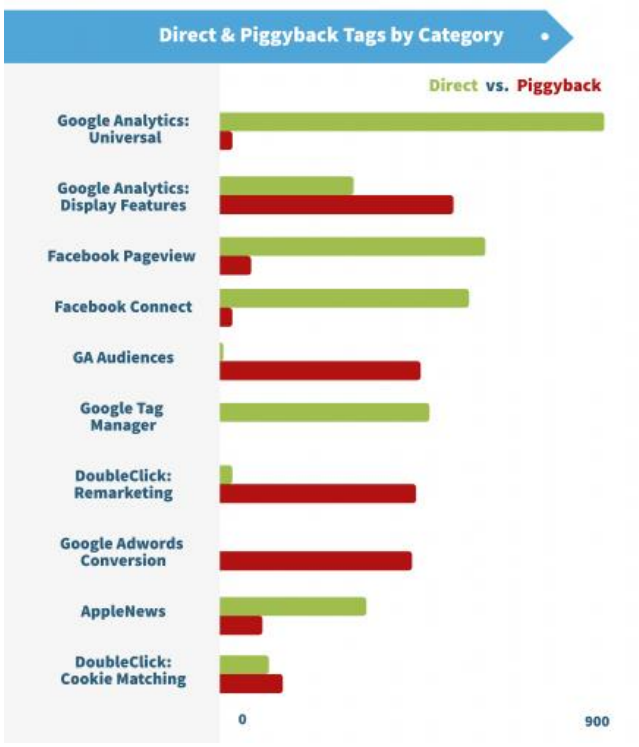


Chart 4

The Impact Of Tag Management Systems (TMS)

Our analysis found that 48% of the Top 1,000 Consumer Goods & Services sites use a Tag Management System (TMS). While one of the benefits of a TMS is that it enables a company to deploy and manage multiple tags through a central container and control point, we found that still 42% of tags on these sites were indirectly deployed tags – that is that were not deployed either within the TMS or directly onto the site. Again, the majority of these piggyback tags were in the Digital Advertising category. This is influenced by the capability to enable features and associated tags in the administrative interfaces of tools such as Google Analytics, rather than a traditional tag deployment process.

While there is not necessarily anything sinister about the tags deployed in this manner, this represents a point where the 'owner' of website tags and data security (often Analysts or IT team members) may not be included in the tag approval and deployment process and by extension may not be aware the tag is on the site.

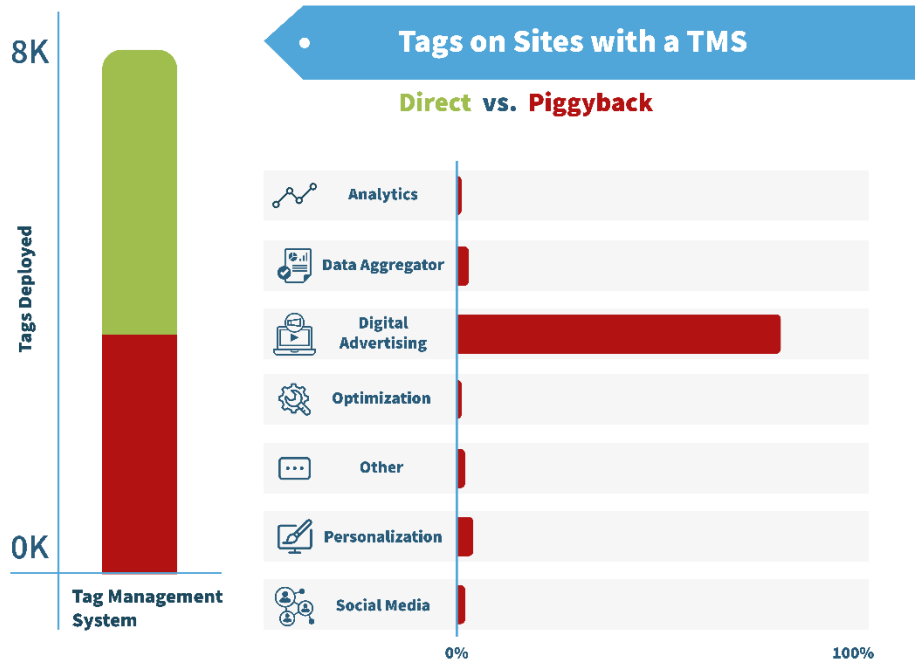


Chart 5

How Indirect Are The Tags?

We then peeled the onion back another level and analyzed the depth of the tags to show how removed these tags are from the direct control of the company (chart 6). For example a 4th level tag would be a tag that is piggybacking an indirect tag that itself is piggybacking a tag that is attached to a direct tag.

We evaluated the types of tags that are not under direct control of companies and found tags as many as six layers deep. While this occurred to an extent across most of the tag categories, the highest volume of tags at three or more layers deep related to Digital Advertising.

It should also be noted in the chart below (chart 6), that the presence of 'direct' tags at layer two, is a demonstration of a tag being deployed into a Tag Management System. As this is in the direct control of the company, we define these as direct tags.

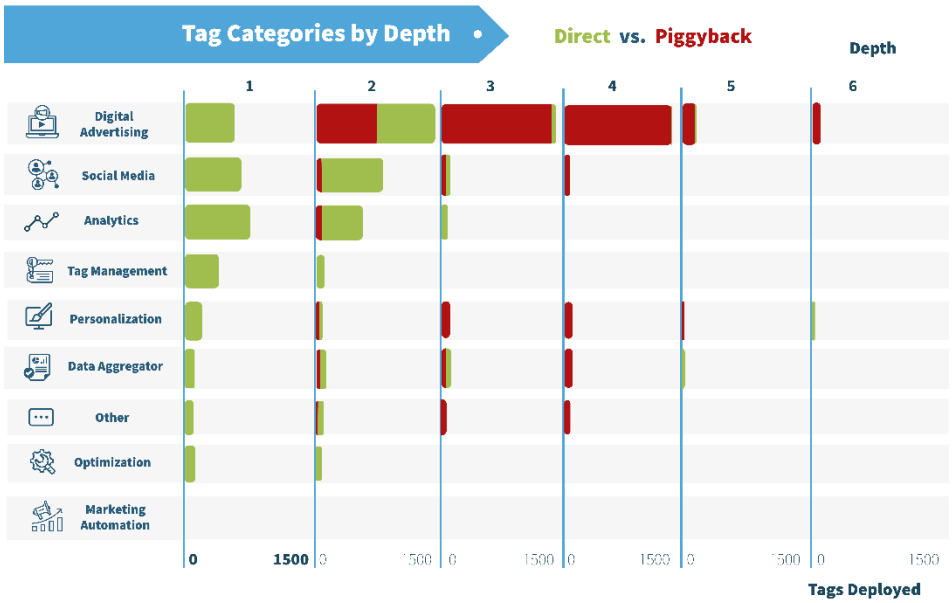


Chart 6

Conclusion

Our analysis shows us that 79% of the top 1,000 Consumer Goods & Service sites host at least one piggyback tag and that 44% of all the tags on these sites are piggybacking. From this, we draw the conclusion that the industry is presented with an ongoing risk of data leakage or non-compliance.

While it is necessary to share data via third-party tags to drive success for websites, companies need to manage the risk of unknown tags that are piggybacking off approved tags. It is important that companies understand all the tags – direct and piggyback – that are on their site and understand how each tag vendor manages and shares data collected from the site.

The rise of marketing campaign data collection and sharing is clear. From a marketing perspective, plenty of attention is usually given to data collection, use and permissions for databases, however marketing teams may be less aware of the risks posed by their website tags. Marketing is often not the team responsible for tag deployment and website data security, and campaign marketing staff may not be fully cognizant of company protocols relating to tag deployments, nor the potential or consequences of data leakage.

Internal corporate structures, specialist skill sets, trust placed in supplier agencies, and fast-paced environments can all transpire to allow leaky tags onto websites. But neither ignorance nor process is a defence for a breach of privacy, so it is the company's responsibility to understand all the tags on their site, including what data they are collecting and sharing, with whom.

The critical first step that lays the foundations for this is conducting a tag audit and creating a map of the relationships of all the current tags on your site. Once you have this visualized it is easy for you to then isolate and manage any tags that you are not in direct control of.

You can audit your tags by running a DataTrue coverage test. Find out more about DatatTrue, including a no-obligation 30-day free trial.

---

^ Data analysis source data: Alexa top sites for the Consumer Goods and Services category, 7 May 2018
https://www.alexa.com/topsites/category/Top/Business/Consumer_Goods_and_Services

---