**Introduction**

The care and handling of personal information is a top concern for consumers and governments alike. Unlike many issues, which gain public attention and struggle to keep it, an endless stream of publicized data breaches serves to keep data privacy in the public eye. As a result, we’re seeing increasingly onerous regulation coming into effect to improve the data management practices of organizations and protect the confidential information of citizens. Major internet players also weigh in in an effort to make the user experience more secure. The remedial work for organizations can be painful, but the result will be a more secure digital environment to support their digital transformation efforts.

While there are many aspects of the lifecycle management of customer data, it all begins with collection. For organizations with a large digital presence, identifying all the places that personal information, or in the case of the new General Data Protection Regulation (GDPR) put in place by the European Union (EU), personally identifiable information (PII), is collected can be a daunting task. The GDPR is unique in that its PII collection requirements apply to any business that collects data about EU citizens—including those who have no physical presence in an EU nation and are instead headquartered in other countries. This greatly broadens the scope of organizations that must comply with the GDPR. It leads to many questions, including, can we identify and monitor all websites collecting PII on behalf of our company? Are those collection points secure? Are they accompanied by compliance statements and controls? Research carried out by RiskIQ highlights that there is much left to do in this area.

**A Closer Look**

Though the new GDPR applies to companies outside the EU who do business with EU countries, we set up a research project to look at the public-facing websites of the top-30 UK companies (FTSE 30 or FT 30 as it is also known). Overall, our research identified nearly 100,000 live websites belonging to FTSE 30 organizations, of which 13,000 pages are collecting PII—an average of 400 pages per organization. One-third of these pages are still collecting information insecurely, either through lack of encryption or as a result of using vulnerable, obsolete encryption algorithms.

![Fig-1 As the data indicates, insecure web collection of personal information is still a major concern]
According to a recent survey by IDG-Connect, 70% of organizations have no to modest confidence in their ability to manage and reduce their attack surface, which is made up of their external assets, such as web pages, web components, or web apps. This poor visibility results in unknown assets and exploitable attack vectors that may compromise customer data.

Insecure collection of personal information can affect consumers through loss and fraudulent use of their data, and organizations through loss of revenue, brand reputation, and damages.

Under the new EU GDPR which comes into effect in May 2018, fines can be considerable if collected data is compromised. The GDPR applies to EU companies, as well as any non-EU companies who collect and store data about EU citizens. In addition to secure collection, GDPR has many additional requirements to ensure that citizens know how their information will be used and give their consent.

Commenting on this research, Bob Tarzey, Analyst and Director, Quocirca Ltd., said, “While this RiskIQ research is focused on large UK companies, the findings will be representative of all organizations. Many will already have the data security basics in place to comply with the regulations that precede GDPR. However, GDPR has many additional requirements, especially around the way data is captured and processed. These include obtaining explicit opt-in from data subjects. Before an organization can address GDPR, it needs to fully understand the extent of its online data-gathering activities. With enforcement of GDPR less than a year away, the time to act is now.”

Outside of regulation, we see a general view across the industry that insecure data collection is a hangover from the early days of the internet and is no longer an acceptable practice given today’s threat landscape. Google is proactively addressing this issue by alerting browser users when they are entering data into insecure forms. Chrome 56 displayed the “Not Secure” label for certain types of data entry (login and payment card data) and this will be extended to cover any data input into HTTP: pages with Chrome 62 due in October 2018. These expanded warnings for HTTP pages will reflect badly on organizations and cause concern for visitors to those pages.

Before an organization can address the insecure capture of personal information, it needs to fully understand the extent of its online data gathering activities. At RiskIQ, we have extended our Digital Footprint™ solution capabilities to help organizations address the PII discovery and assessment problem. Our solution uses comprehensive internet reconnaissance to identify an organization’s sites and pages collecting personal and personally identifiable information, regardless of language or layout. It highlights instances where that information is currently being collected insecurely, and in the case of GDPR compliance, where usage notices are not present or active opt-in mechanisms
are missing. Once a baseline inventory is established, any new collection points that arise in the future will be brought to the attention of compliance teams.

The PII/GDPR analytics feature to accelerate GDPR readiness and assessment processes is now available to U.S. customers as part of our Digital Footprint, Enterprise Edition product—providing comprehensive external attack surface discovery and monitoring, along with a quick and cost-effective way to help ensure compliant PII collection. Users of the solution benefit through significant time and resource savings in the discovery process, a more comprehensive discovery of their internet-facing assets, website collection points, and active identification of new data collection points in the future.

To find out more about RiskIQ’s PII Discovery, contact us at info@riskiq.com.