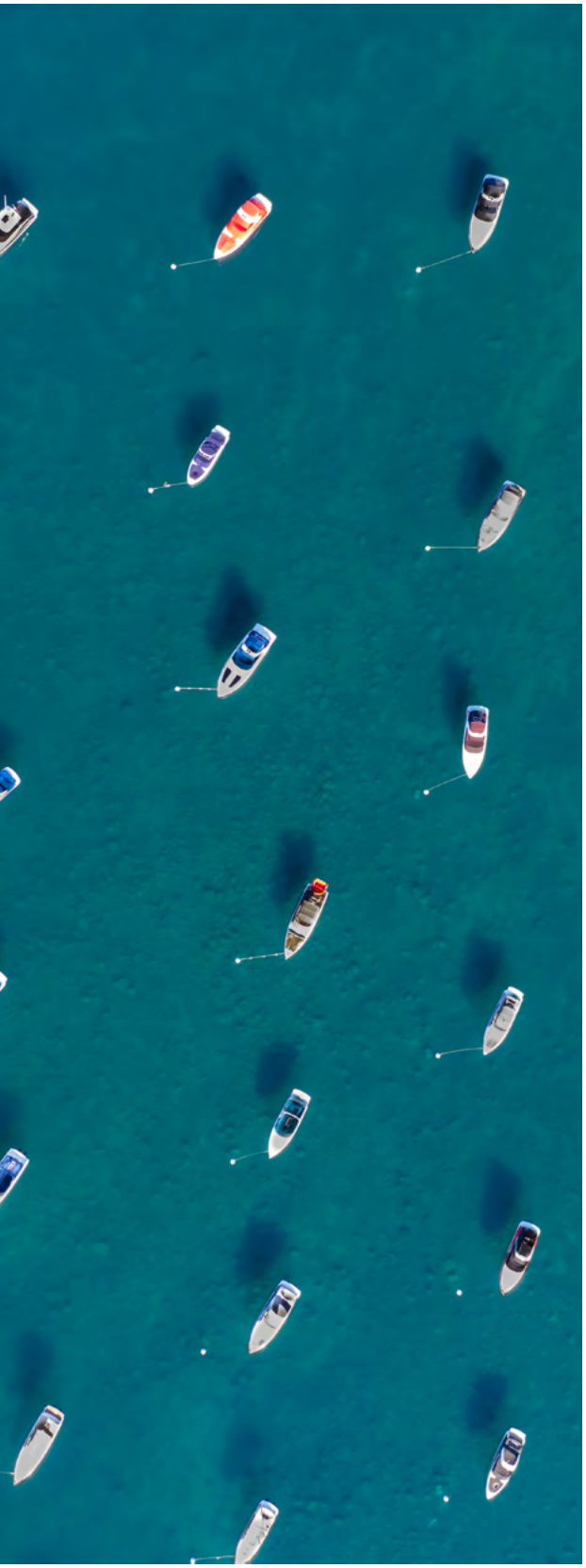
CardinalPath

Personally Identifiable Information (PII) in Google Analytics: Assessment, Mitigation and Hardening Against Risk



Outline

INTRODUCTION TO THE PII ISSUE

Sweeping changes to how we work with personal data

As marketers, we are experiencing a massive shift around the way we obtain and use personal data. The data privacy landscape has become increasingly complex, with disparate influences converging to shape what privacy means for your organization and to your customers.

Consumer trust is eroding. With large data breaches frequently in the headlines, customers are losing confidence in companies' ability to keep their private information safe and secure. Signals include the large swaths of consumers taking steps to opt out of data-sharing, whether it's simply not consenting to be tracked on their favorite websites or installing ad blocking software in their browser.

Legislative intervention is on the rise. Governments around the globe are becoming more active in regulating how organizations can utilize consumer data. In just the last two years, major statutory changes have gone into effect, drastically altering the privacy landscape. The most well known of these, the EU's General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA), may be a harbinger of additional governmental actions to come.

Tech industry leaders are listening. Google has announced that it will soon block third-party cookies in its Chrome web browser, joining Apple Safari's Intelligent Tracking Protection (ITP) and Mozilla Firefox's Enhanced Tracking Protection (ETP) feature in potentially fundamentally changing the way online tracking and privacy work. Meanwhile, new privacy-oriented alternatives to major digital players are coming to market, such as the Duck Duck Go search engine or Brave browser, and existing market leaders are updating their policies to address privacy concerns, following the lead of government.

With these coinciding influences, data privacy has become both highly complex and highly fluid as the ground continues to shift, forcing companies to adapt in real time. All of these elements are pushing organizations towards investing in and developing its first party data.

First party data is information a company has directly collected on a consumer. Sometimes that data might include elements that may be classified as "personal data" (GDPR), "personal information" (CCPA), or "personally identifiable information" (Google, among others), each of which are defined and regulated differently.

With companies collecting more first party data than ever, there are more opportunities to unintentionally collect consumer information that fits these definitions. For every company that has made headlines for customer data breaches, there are likely many others who have unknowingly collected individuals' personal data in their digital platforms, introducing risk to the organization.

PII and Google Analytics

This paper focuses on personally identifiable information, commonly known as "PII" and the problems that can be introduced by collecting it in Google Analytics.

Like legislative bodies around the world, Google has addressed the issue of privacy head-on by adding stipulations around data privacy to its terms of service for its core analytics product, Google Analytics (GA). Google Analytics policies clearly prohibit the collection of PII.[1]

When your GA account ends up containing personally identifiable information, it may have a significant impact on your organization, with challenges such as:

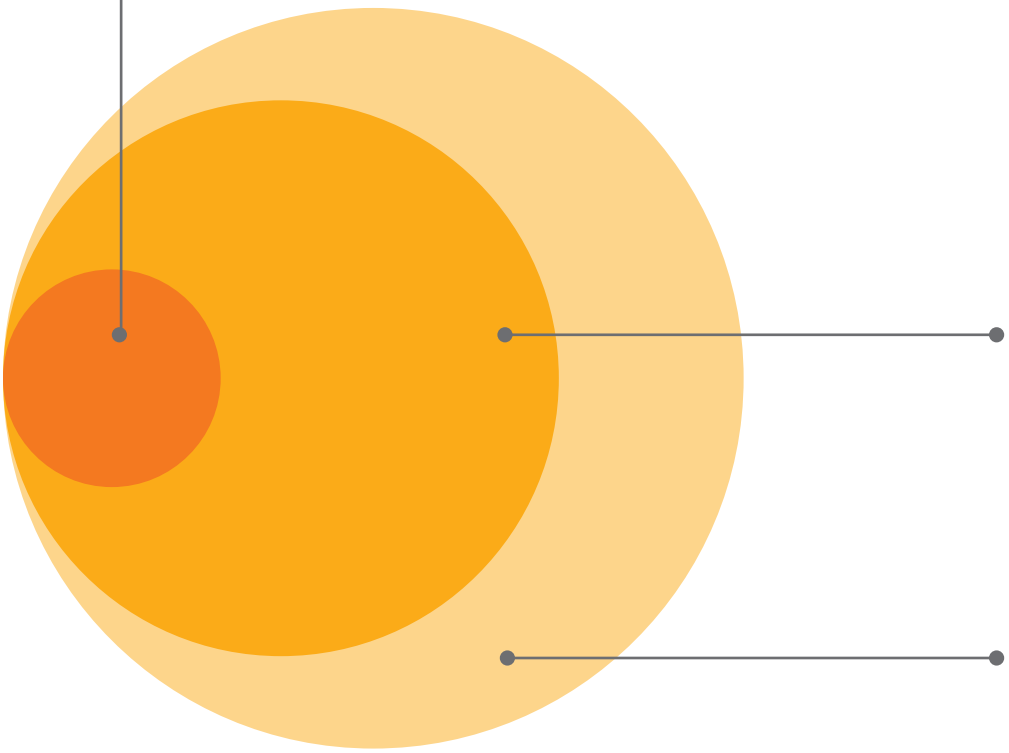
- increased probability that PII has infiltrated other data repositories, which may not be aligned with your organization's privacy policies and other obligations
- Triggering unexpected compliance obligations under privacy laws
- Increased level of effort and complexity in responding to personal information requests
- Broken trust with customers, prospects, and partners

Examples of PII as defined by Google commonly include email addresses, phone numbers, physical addresses and full names, among others. Examples of what may not be considered PII by Google include IP addresses or pseudonymous user identifiers such as a numeric ID that ties back to a platform outside of Google Analytics.[2]

Google's interpretation of PII today is "information that could be used on its own to directly identify, contact, or precisely locate an individual."[3]

The definition above is a good starting point for understanding the concept of PII, but it's important to involve your organization's legal team and complete a proper review of the contracts, terms of service, and policies that apply to platforms such as Google Analytics -- as well as your organization's own policies -- to determine what types of data may be considered personal in nature and may not be appropriate to collect in each case.

1. Terms of Service | Google Analytics. (n.d.). Retrieved August 25, 2020, from https://marketingplatform.google.com/about/analytics/terms/us/
2. Understanding PII in Google's contracts and policies - Analytics Help. (n.d.). Retrieved August 25, 2020, from https://support.google.com/analytics/answer/7686480?hl=en
3. Understanding PII in Google's contracts and policies - Analytics Help. (n.d.). Retrieved August 25, 2020, from https://support.google.com/analytics/answer/7686480?hl=en



Peronally Identifiable Information (PII)

- Username
- First Name
- Last Name
- Phone Number
- Email Address
- Physical Address
- Social Security Number

Personal Data (PD)

- PII +
- Political Opinions
- Pseudonymous IDs
- Religious Affiliation
- Ethnic Affiliation
- Biometric/Genetic Data

Personal Information (PI)

- Personal Data +
- Household/device data

PII can find a way into your Google Analytics account if proper safeguards are not applied within your technical layer and you are not proactively monitoring your data. As the account holder, you are ultimately responsible for what is captured and stored within your GA account.

While capturing PII in analytics data is often an error or oversight, Google's terms are clear: "Google policies mandate that no data be passed to Google that Google could use or recognize as personally identifiable information (PII)." Consequences of violating these policies potentially include Google Analytics account termination and data destruction , in addition to the larger organizational challenges noted above.

The longer the issue goes unnoticed, or worse, ignored, the more painful the impact may be.

4. Best practices to avoid sending Personally Identifiable Information (PII) - Analytics Help. (n.d.). Retrieved August 25, 2020, from https://support.google.com/analytics/answer/6366371
5. Universal Analytics usage guidelines - Analytics Help. (n.d.). Retrieved August 25, 2020, from https://support.google.com/analytics/answer/2795983?hl=en

WEBSITE VULNERABILITIES AND PII DATA COLLECTION

Often, PII Collection Starts With Your Website

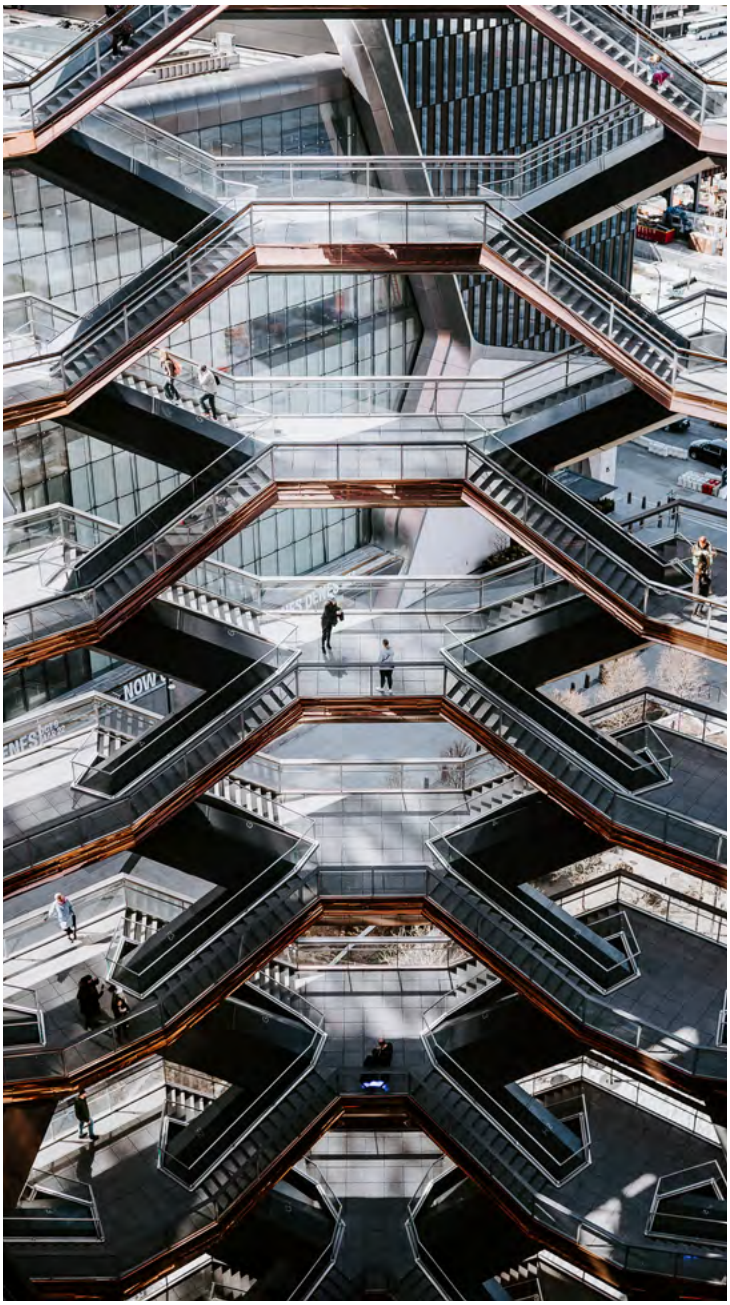
There are many ways PII ends up getting collected unintentionally in Google Analytics. Often, it's a result of the way a website was developed:

- In some environments, the site is set up to pass values from page to page using query string parameters (QSPs). These values live in the browser's address bar appended to the page address. If one of the appended values contains PII such as a user's email, it will be passed to GA when the tracking code collects the page URL.

- Custom tracking features such as Google Analytics events also allow for actions on a page to be tracked and values from that page to be collected. It is common practice to pull values directly from the page to populate the data in these events. If one of the dynamic values for the event is populated by a user's input, it can pose a risk. For example, when collecting the value of a blog comment and that user puts their email address in the comment field, it will be passed to GA in the collected event data.

- When a form is submitted using the GET method, it appends form data to the URL in pairs, and could expose data such as an email address or physical address.

In other cases, personally identifiable information in GA is the result data being sent to the site, such referrer information from another website or campaign tags.



Common Places PII is Collected in Google Analytics

Once the PII is exposed on the website, Google Analytics may then collect the data. Common locations Cardinal Path has found PII within GA web properties include:

- Page URLs, often via query string parameters
- Page titles
- Event dimensions
- Campaign tags
- Site search dimensions
- User IDs
- Custom dimensions, including those populated by data import

Some of these, like page URL and title, are default tracking configurations that come standard with Google Analytics, while others, like events or data import, are the result of customizations made during the platform's implementation.

With this baseline understanding in mind, you can run an assessment of your GA data, as well as look for possible weaknesses in your website development where PII leakage may occur.

ASSESSING YOUR ANALYTICS ACCOUNT

Conducting a Search of Your Historical Data

The first step in mitigating and preventing personally identifiable information from being collected in Google Analytics is assessing if you already have PII present within your account.

First, you'll need to conduct a search of your data. Below are a few sample regular expressions, commonly referred to as "regex", that match some common forms of PII: email address, phone number and problematic query string parameters. Regular expressions are a developer's tool used to identify patterns within strings of characters. They can be complicated, but to use them in the context provided, you don't have to understand what each symbol in the regular expression means, only that they provide one helpful method for searching Google Analytics for data patterns that often match PII values. While not every form or type of PII can be searched for using this method and the following search patterns, it is a good starting point for running a quick, initial assessment of your analytics data.

PII TYPE	REGEX
Email address	(.+)@([\da-zA-Z\.\-_]+)\.([a-zA-Z0-9\.\-_]{2,10})
Phone number (U.S. 10-digit format)	(\(?\d{3}(-|\.|\)|\)\+)\d{3}(-|\.)\d{4})
Query string parameters commonly associated with PII values	[?&](address|birthdate|dateofbirth|dob|email|firstname|fullname|lastname|password|phone|postalcode|pw|sin|ssn|ssnr|street|username|zip)=

In the next section, we'll show you how to apply these to your Google Analytics reports to search for PII.

Before you get started, remember: these regular expression patterns cover only a few common forms of PII, and will expose only some possible PII values. If you uncover personally identifiable information in your GA dataset using any of the above search patterns, you may have a larger PII issue that you will need to dig deeper to fully uncover. And even if you don't uncover anything with these regular expressions, it does not mean you do not have PII in your account; it could mean that you just have not uncovered it yet. Read on to learn how to apply these to your Analytics dataset and determine whether your assessment is complete.

Assessment Methods

There are several possible methods you can use to identify potential PII in Google Analytics: you can use the regular expressions listed above, as well as your own personal sleuthing skills to look for PII in the GA user interface, or if enabled, your GA dataset stored in Google BigQuery. Each has benefits and limitations.

For an initial assessment, we recommend a look back to the start of your implementation to the present date. You'll want to consider each web property separately, given that Google Analytics data collection happens per property, and choose a view in each that has both significant historical data and few customizations, such as filters that could obscure PII that was sent to Google from GA's reporting outputs. Most often, a raw data view established upon the initial creation of the property is the best candidate, if it exists.

Conducting a Search Using the Google Analytics User Interface

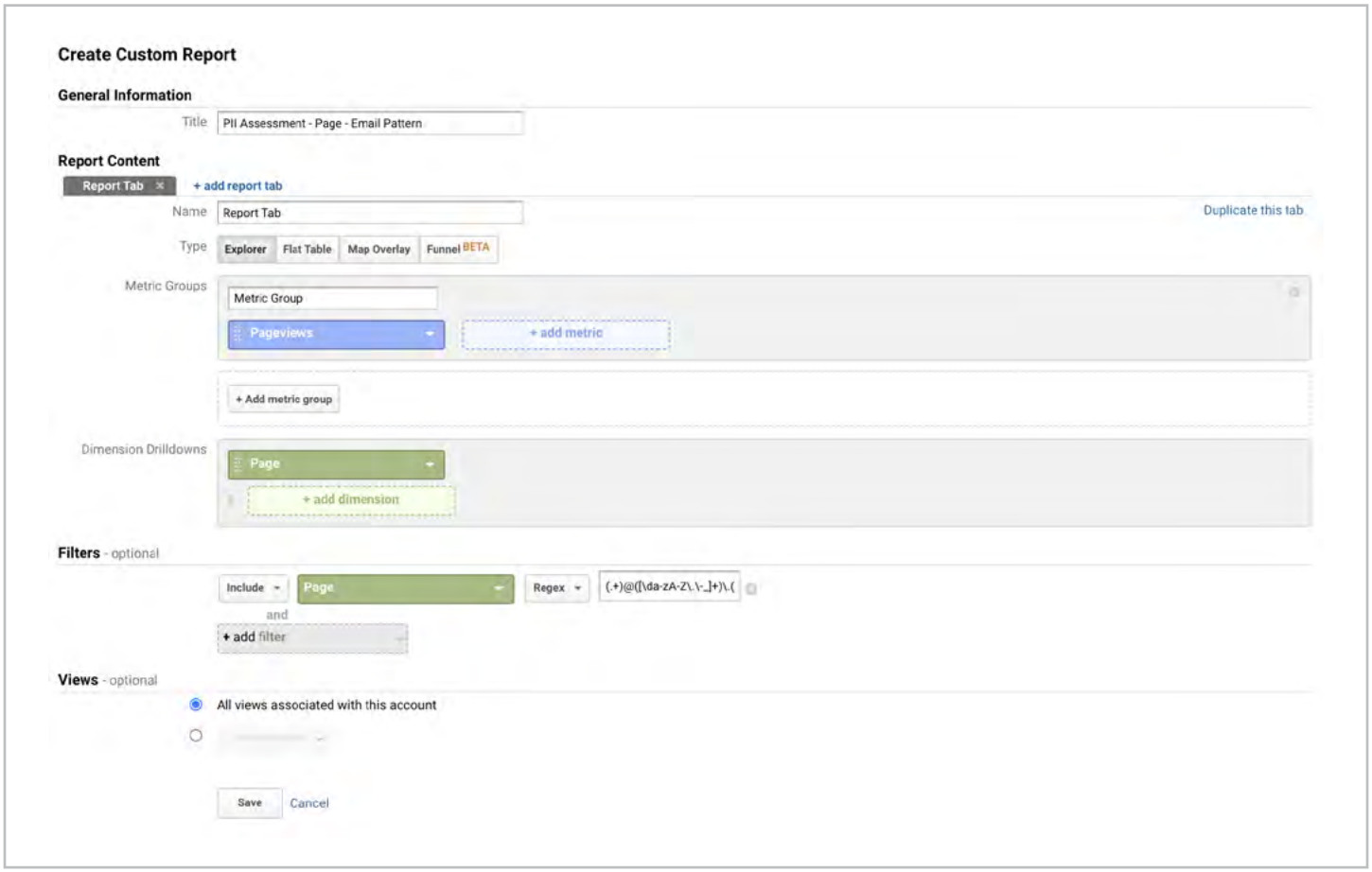
For GA properties that do not have a BigQuery connection enabled, or have collected a significant amount of data prior to the GA export for BigQuery being configured, you can use the Google Analytics user interface (UI) to run your PII assessment.

Benefits include easy access to the interface and relatively quick configuration of reports. Drawbacks include the possibility that you will encounter sampling and/or high-cardinality aggregation in your assessment's reports, both which may obscure the full scale of the PII your account has collected.

While there are multiple ways to construct a report in Google Analytics to assess whether the property may contain PII, one of the most straightforward is creating a custom report. A nice benefit of using the custom report method is that if sampling is encountered, it is possible to easily convert the report to an unsampled version for properties that are upgraded to the Google Analytics 360 (GA360) service level. This can help uncover more results that contain PII values that might otherwise not display. In addition, it's possible to export more rows of data from an unsampled report than a standard report with an inline advanced filter or segment applied, should you need to download your results.

Custom reports can be created for a chosen combination of multiple dimensions and metrics needed to meet specific reporting needs in Google Analytics, and can take several different forms. For a PII assessment report, the Explorer or Flat Table options may be the best option.

Custom reports can be created from the Customization menu or by clicking the Edit button at the top right of many standard reports.



You'll want to name the report, choose the report type, and pick your metric(s) and associated dimension(s), such as Pageviews combined with Page when searching URLs for PII. Finally, you will add an include filter that looks for matches of your chosen regular expression and save the report.

Your custom report will now be filtered to only show you data that includes possible PII values that match the regular expression you have chosen.

If you don't see any records, that is a good sign! It may mean that your data does not contain the type of PII you are searching for, but it's worth checking that the regular expression was entered and the report was configured correctly.

If you do see results, this means that your data might contain PII, false positives that match the search, or a combination of both, and you will need to confirm the makeup of the results and then address the issue.

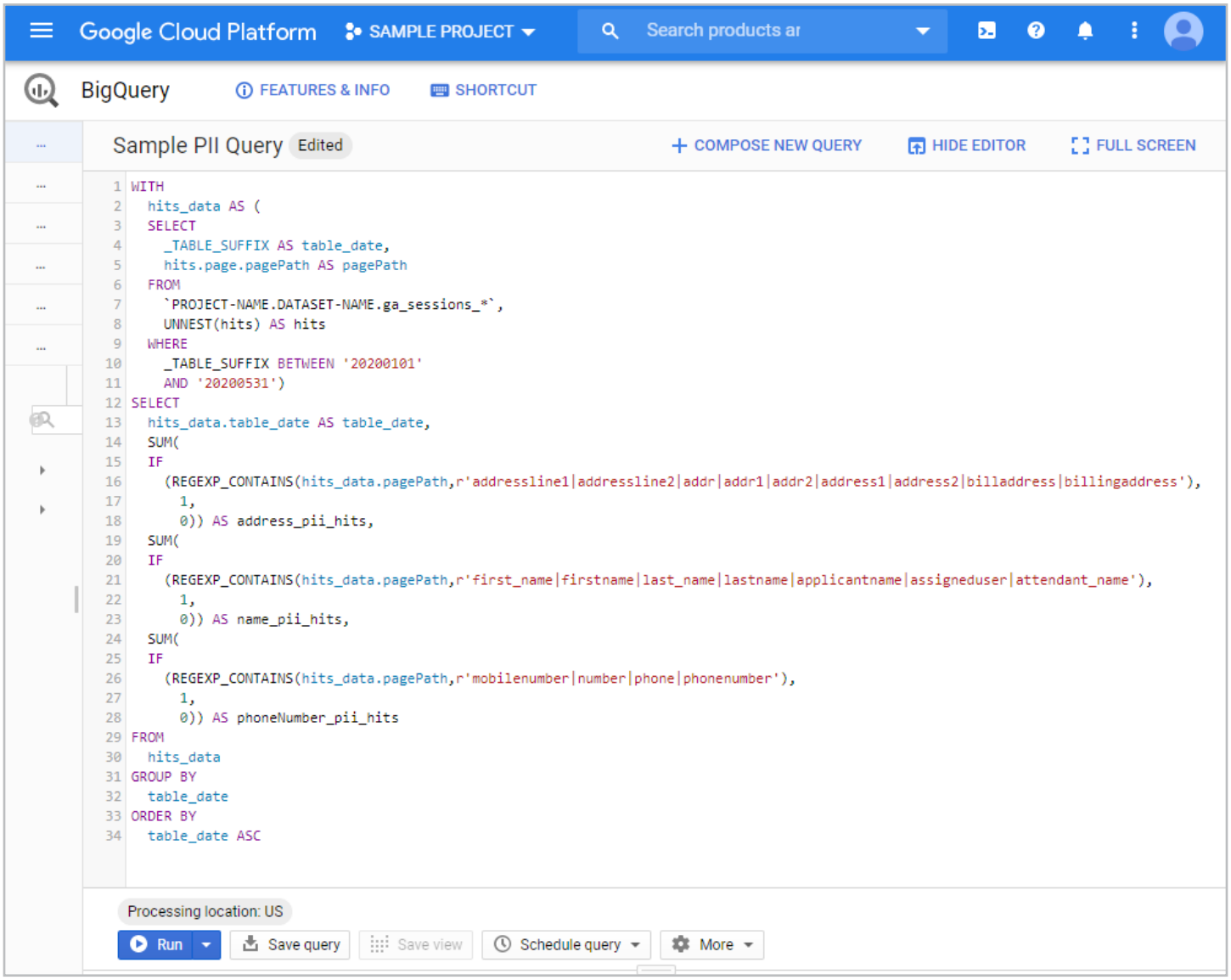
Conducting a Search Using Google BigQuery

If your Google Analytics property has the GA360 export to BigQuery enabled, or you are otherwise storing your GA data in BigQuery, you have another option for scanning your GA dataset for PII.

Benefits include access to the raw, hit-level Google Analytics data, which means you will not encounter sampling or high-cardinality aggregation in your assessment's reports, which may allow you to better reveal the full scale of the PII your account has collected and access each individual record. Drawbacks include needing to involve a BigQuery subject matter expert in organization to construct the reports, and possibly a higher level of effort to create them. You'll also want to consider if the GA360 view linked to BigQuery is appropriate for the analysis, such as a raw data view rather than a filtered view, and when the historical data export was created.

To run the search for PII:

1. **Locate your GA360 to BigQuery export dataset(s) within BigQuery.** The dataset will have the same name as the Google Analytics view ID it is linked to.
2. **Draft your query.** Consider how you'd like to group your regular expression matches. Separating into different groups can help you analyze how present a particular type of PII is within your data. Also consider what kind of insights you're looking for - are you wanting to understand PII as a percent representation of total hit volume? Are you wanting to generate a report to analyze the different manifestations of PII?
3. **Optimize your query.** Before you execute your query, take a final look. Scanning across large datasets for granular data can increase your GCP usage costs, especially if you are running your query multiple times during development. Some good gut-checks include: Am I limiting my query to only the required date range? Am I including only the fields that are necessary for this analysis? Is there a way for me to consolidate multiple queries to limit the number of requests needed for my analysis?
4. **Execute your query.** This is the fun part! Take note of the speed. Be sure to check out the execution summary to understand what aspects of your query had the most impact on query speed. Save your query and share it with others working within your GCP project if needed.
5. **Consider automation.** Is this the type of analysis that is going to feed an ongoing report? Consider creating a view or custom table with the data that you need. This way, the data can be easily visualized and shared with other team members using tools like Tableau, Looker, or Data Studio.



Above: Sample query. Results of this query would provide the number of hits by day where the page path dimension in the GA360 export included user address, name, or phone number based on basic URL parameter regex matches.

Bonus! If you really want to get fancy with your PII monitoring in Google Cloud Platform, turn to Google Cloud Data Loss Prevention (DLP) services. At Cardinal Path, this is a service we offer to clients who want to automate data identification and redaction. This service automatically discovers new forms of private or sensitive data, and handles them in the best manner for the data type identified. For example, it can find and obfuscate member IDs, while completely redacting Social Security numbers.

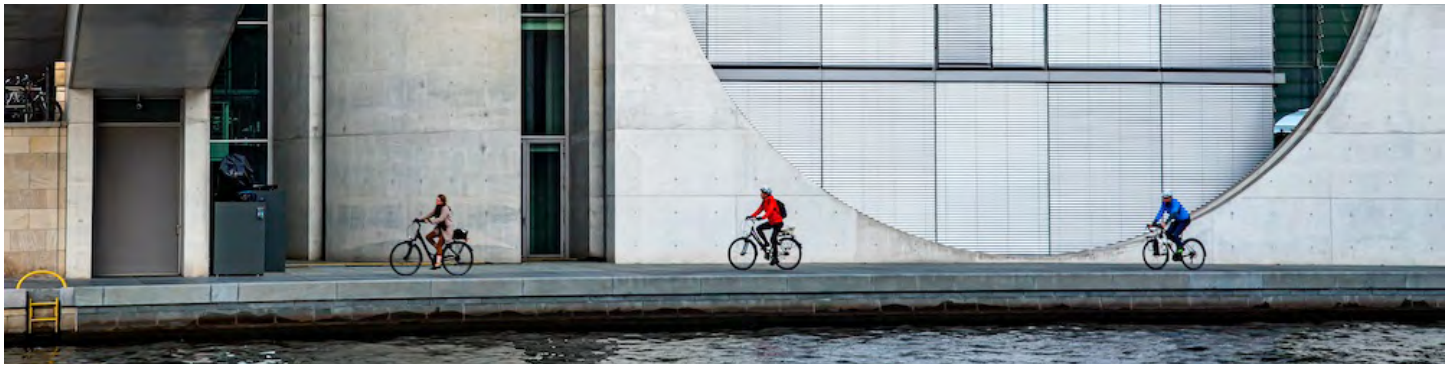
Analyzing the Results of Your Assessment

Once you've run your reports via your chosen method, it's time to analyze the data. Since the provided regular expressions don't cover all forms of PII or all possible PII values, you may need to continually refine your searches.

You can consider both broadening your searches to include more types of PII patterns and types, as well as more dimensions, and also narrowing your searches to reduce false positives in your results by making the queries more specific.

Because of the nature of this search methodology, you are unlikely to catch all possible instances of PII in your account. A good stopping point is when you reach a reasonable level of confidence that you have uncovered the bulk of the PII types, dimension locations, and dates -- or lack thereof -- contained in all of the properties in your Google Analytics account.

SOLUTIONS



If you have uncovered PII or sensitive data, or haven't but want to take proactive action to reduce the likelihood of collecting it in the future, it's now time to turn your attention toward hardening, remediation, and data management solutions.
This section addresses key solutions for the issues surrounding the collection of PII. This information is not meant as a legal advice and you may not rely on it as legal advice, nor as a recommendation of any particular legal understanding. Please consult your own attorneys, privacy and compliance experts to apply the law to your specific circumstances and advise you accordingly.

Hardening Solutions

"Hardening" means stopping the flow of personally identifiable information into Google Analytics and other repositories.

As noted earlier, Google's policies mandate that PII cannot be passed to Google, which means that implementing hardening measures in the technical layer are critical. Employing stop-gap solutions such as view filters inside the GA user interface to stop PII from displaying to users in Google Analytics are too little too late, as the data has already been collected by Google's servers. Instead, you'll want to focus on updating vulnerabilities on your website and your data collection methods.

Temporary Access Adjustments

For some organizations, uncovering PII in Analytics or other digital platforms leads to the need to take immediate action to lock down access to the collected data to prevent further leakage of the information outside of its storage location. In the case of GA, this may mean that you will want to change user permissions temporarily to ensure that current employees and vendors without a critical need to maintain access are not able to view or further distribute the PII while the hardening and remediation process is underway.

Website Updates

Once the root causes of the issue have been identified, site-side solutions are typically most effective for stopping PII leakage into GA. For example, if your assessment determined that PII was leaking into Google Analytics due to form submission data that used the GET method, you may want to update the site to use the POST method instead. With the POST method, the data sent from the browser to the server is not displayed in the URL or stored in the browser history or in web server logs, making it more secure than the GET method.

Hardening Code

In addition to updates made to your website, hardening solutions can be employed at the data collection layer.

If you use a tag management system (TMS) such as Google Tag Manager or Tealium to manage your code for Google Analytics and other digital platforms, you have the advantage of being able to more easily and quickly deploy solutions that can help reduce the likelihood that PII will continue to infiltrate those systems.

For Google Analytics specifically, custom JavaScript hardening code can be implemented within your TMS. This code intercepts the GA hit payload from the website and modifies the request to remove or redact suspected PII values before it is sent to Analytics, lessening the chances that PII will be collected by GA and its connected data repositories.

This PII hardening code can take many different forms, depending on your business requirements. Cardinal Path has developed custom hardening code that allows organizations to whitelist query parameters your company has deemed safe, blacklist query parameters it has deemed problematic, and scrub values that match common PII patterns, such as an email address or phone number.

Here's an example of what some key dimension values for a site search event hit might look like before versus after implementation of the hardening code:

GA DIMENSION	VALUE BEFORE HARDENING CODE	VALUE AFTER HARDENING CODE
Page	/search/?q=useremail@example.com	/search/?q=[redacted email]
Page Title	Search - YourSite.com	Search - YourSite.com
Event Category & Event Action	Site Search useremail@example.com	Site Search [redacted email]
Search Term	useremail@example.com	[redacted email]

Regardless of its exact form, the hardening code works in generally the same way: it intercepts the Google Analytics hit and transforms it to either remove or redact values that may contain PII, and then sends the modified hit to Analytics, resulting in cleaner data being collected by GA.

Remediation Solutions

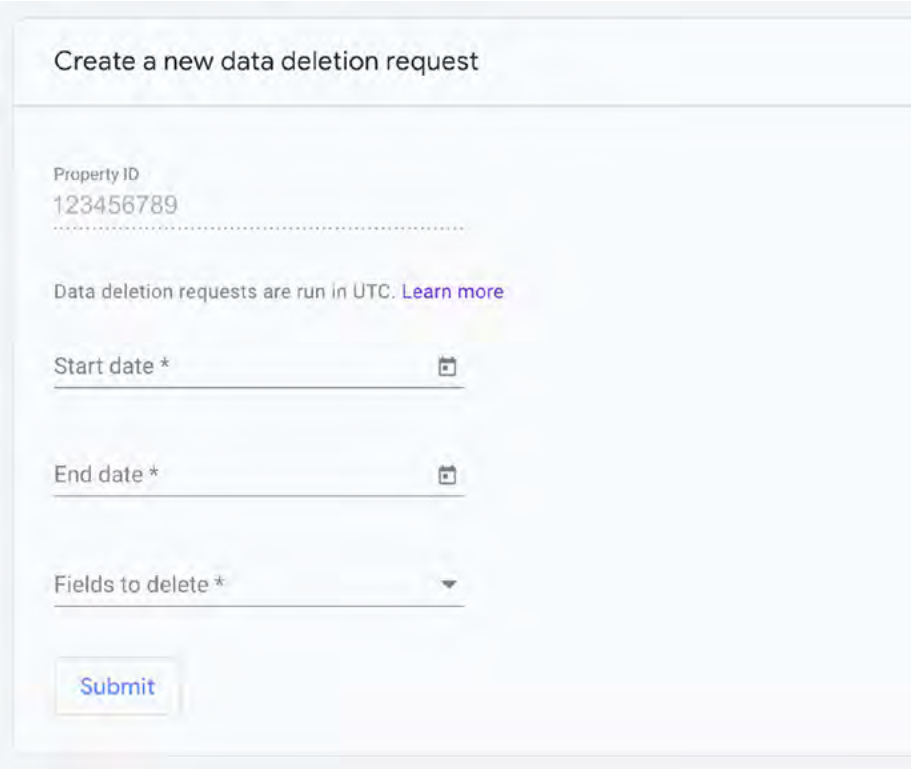
Once the flow of personally identifiable information from your website to your digital platforms like GA has been curtailed, it's necessary to clean up the PII that has already been collected.

Data Deletion in Analytics

For removal of existing PII in Analytics, Google currently provides a Data Deletion Request tool. This feature will delete historical data based on a specified date range and dimension combination that you specify.

For example, you can choose to delete the values of one or more select dimensions, such as URL or Event Category, or all data for the property, for as long as the length of your account history, or a smaller date range such as a week-long period in which PII was collected.

Before proceeding with data deletion in an Analytics property, it is critical to understand the possible implications, as deletion will impact all reporting for the property:



- **Data is deleted permanently and irrevocably** -- There's no going back once a deletion request has been commenced.
- **All values for the field(s) specified are deleted** -- You cannot delete only the offending PII values within a dimension; rather, all values for the entire dimension where data deletion is requested will have its values wiped. For example, if you request deletion of the URL field, you will no longer see any page URLs reported when looking back at the date range you deleted.
- **Date ranges are not precise** -- The deletion process uses UTC time rather than your account's time zone, and due to the way Google Analytics aggregates data, you may see data deleted up to three days before and/or after the beginning and end dates you select.[6] For example, if you choose to delete all of your Page dimension data for 1/1/20 to 1/7/20, Google may delete the data from as early as 12/29/19 through as late as 1/10/20.

As you can see, this feature currently takes a very broad stroke to PII deletion, and it can be very frustrating to lose years' worth of data because a small percentage of it is corrupted with PII. This reinforces the importance of ensuring your data is clean upon its initial collection through governance and hardening strategies, is backed up in BigQuery where a more controlled approach can be used for remediation, and is being monitored regularly to catch possible future issues early on.

6. Data deletion requests (Web) - Analytics Help. (n.d.). Retrieved August 25, 2020, from https://support.google.com/analytics/answer/9450800?hl=en

Data Deletion in BigQuery

Other platforms that house GA data containing PII, such as Google BigQuery, may have other custom solutions employed. While Google BigQuery's policies are different from those of Analytics and do not explicitly prohibit the collection of personally identifiable information,[7],[8],[9],[10] some organizations may prefer not to collect or maintain PII or sensitive data types in BigQuery.

Cardinal Path works with clients to create custom remediation approaches for data stored in BigQuery, including the Google Analytics dataset. One approach involves redacting PII by replacing the infringing values with a token to signify that the original value has been scrubbed.

To delete specific values from your GA360>BigQuery export, one option is to use the Data Manipulation Language (DML) supported by BigQuery. This language enables you to delete or modify values within tables or table partitions in BigQuery using Standard SQL queries. You can alternatively use the Cloud DLP platform for more advanced data redaction and obfuscation options.

When implementing either of these solutions for our clients, we consider:

- The client's specific privacy requirements
- The most optimized balance of cost and performance in solution design
- Ease of use for the client to manage moving forward

7. BigQuery is a cloud data warehouse that is one of over 75 components of Google Cloud Platform. The GCP terms apply to these various components.

8. "Google has been certified compliant with ISO 27018 for Google Cloud Platform products and G Suite.
ISO 27018 is an international standard of practice for protection of personally identifiable information (PII) in public cloud services." See Compliance | How Google complies with data protection laws. (n.d.). Retrieved August 27, 2020, from https://privacy.google.com/businesses/compliance/

9. Google Cloud Platform Terms of Service. (n.d.). Retrieved August 25, 2020, from https://cloud.google.com/terms

10. Data Processing and Security Terms (Customers). (n.d.). Retrieved August 25, 2020, from https://cloud.google.com/terms/data-processing-terms

Data Governance Solutions

Data governance is a collection of policies, practices, rules and regulations an organization follows relating to how it collects, stores and manages its data.

The collection and storage of PII in Google Analytics discussed thus far is just one part of data governance that your organization should be thinking about as you navigate the new reality that privacy is a major concern for consumers and organizations alike. Looking to the future, you should also be considering how your organization is handling data governance at a higher level, and implementing policies, procedures and technologies that embrace security and privacy.

Privacy Policies and Cookie Consent Management

Providing customers clear *notification* of how your company collects and uses their data is a cornerstone of data governance in the digital world. As a best practice, your organization should have a clear privacy policy that is easy to locate on your website and details your site's use of Google Analytics, cookies, mobile device IDs and similar data collection technologies you use. For Google Analytics account owners, this is another requirement of the Analytics terms of service, and is a common requirement for many other platforms such as advertising technologies.[11] This should be updated frequently as your data collection practices, policies and technologies change over time.

Taking it a step further, cookie consent management allows websites to automate the process of obtaining user *consent* to place cookies and use other means to collect their data during their visit. Now a standard part of privacy initiatives deployed by companies worldwide, cookie consent management platforms inform website visitors about what types of data is collected and how it is used, and allow users to opt-in or opt-out of data collection. A number of cookie consent management technologies are on the market today.

Making user privacy and consent a core part of your relationship with customers and prospects builds trust and makes compliance with applicable policies and regulations easier to manage.

As mentioned above, please consult your own attorneys, privacy and compliance experts to apply the law to your specific circumstances and advise you accordingly.

Data Minimization

Beyond notifying users of your practices and obtaining consent to track their activities, it's also worth assessing whether you are collecting the right data in the first place.

Data minimization involves limiting the data collected to include only information that is relevant or necessary to accomplish specific organizational purposes. Tracking only what you need can reduce your risk and the number of personal information requests you need to respond to.

Data minimization involves running an assessment of what data is being collected and how it is being joined to other data, and then determining whether each data point fulfills a necessary business requirement and is aligned with organization policies and applicable privacy regulations.

For data being collected that is no longer relevant to organizational goals, tracking of the element and joins to other data can be deprecated, or it can be changed to de-identify or pseudonymize the values. Historical data that is no longer needed can also be removed.

When it comes to data collection and storage in today's privacy-oriented environment, sometimes less is more.

11. Terms of Service | Google Analytics. (n.d.). Retrieved August 25, 2020, from https://marketingplatform.google.com/about/analytics/terms/us/

Data Mapping

To reduce the risk that Analytics data containing PII may leak into other data repositories and amplify your compliance problems, data mapping can be used to clearly identify connections between systems. Data mapping involves documenting what data your organization collects, what purpose it serves, where it lives, and ultimately, where else it flows.

If you know what data you're collecting and for what purpose, you can minimize it. If you know where that data is being stored and who it is shared with, assessing and remediating PII exposure and meeting other compliance obligations your organization has become easier.

Data Monitoring

Waiting for PII to be discovered by chance may delay detection, leading to more PII to find and remove once it is eventually identified. Data monitoring can help your organization get ahead of potential problems with automated solutions that scan and flag possible issues for investigation. This may save you from costly work stoppages, regulatory headaches and the types of time-consuming mitigation efforts discussed throughout this paper.

Today, there are a number of automated monitoring solutions that can be added to your solution suite to discover and classify sensitive data at scale, plus custom solutions that can be built to your needs.

As the axiom goes, an ounce of prevention is worth a pound of cure.

11. Terms of Service | Google Analytics. (n.d.). Retrieved August 25, 2020, from https://marketingplatform.google.com/about/analytics/terms/us/

CARDINAL PATH SUPPORT

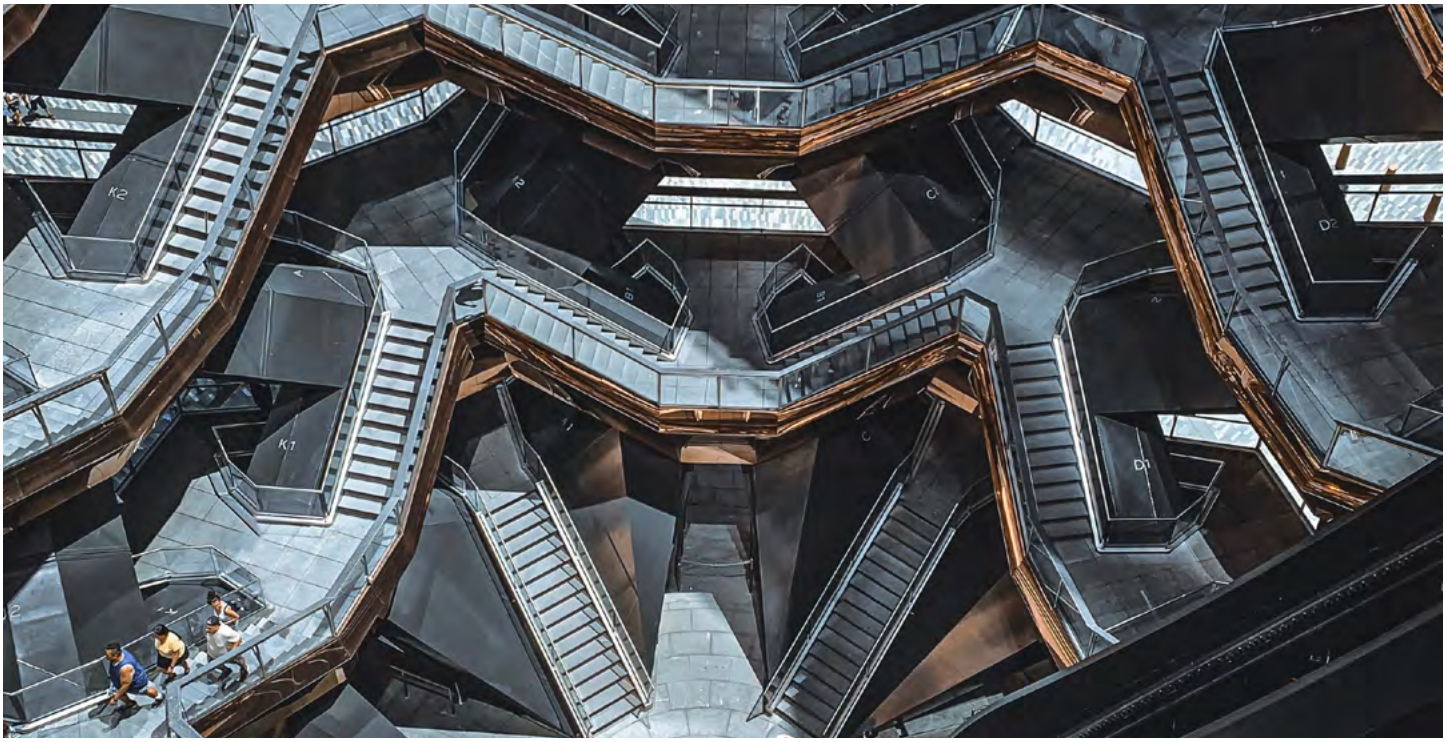
Cardinal Path's framework to address PII-related data issues is designed to understand the situation as quickly as possible, apply best practice solutions in the short term, and develop a custom solution based on your digital landscape for long-term governance. The process follows four phases:

- **Assessment:** We conduct a comprehensive review of your implementation and historical dataset to determine the depth and breadth of the issue, and its root causes. We'll clearly document what PII has been collected, where, and for how long and provide guidance on quickly stopping the flow of personally identifiable information.
- **Hardening:** The results of the assessment will determine the scope of mitigation actions, but this phase generally involves deployment of custom code to bolster your analytics data collection, plus consultation on other actions your organization can take to prevent recurrence of the issue and meet your privacy obligations to your company and customers.
- **Remediation:** Once we've stopped the flow of PII into your systems, we will help you scrub the problematic data from your affected platforms. This phase can move very quickly or entail a large project tailored to your organization's requirements.
- **Prevention:** To ensure that you're prepared to quickly identify and halt the emergence of new PII issues in the future, we will provide a prevention framework for ongoing monitoring and governance.

Addressing and managing the data captured within your Google Analytics could lead to a larger imperative to ensure proper oversight of your customer data. Capturing personally identifiable information is a risk not only to your analytics practices but also presents an unwanted liability . Most organizations benefit from having a team of analytics experts help you to assess your situation and develop a plan to mitigate this risk.

Analytics Experts in Your Corner

If you have questions or want to learn more, please reach out to the team that put this paper together.



AUTHORS

Ariana Wolf
Manager, Implementation
Cardinal Path

Tara Kincade
Director, Digital Intelligence
Cardinal Path

John Henson
Director, Digital Intelligence
Cardinal Path