Google Tag Manager and Google Analytics Configuration and Reporting for the Digital Collections of Colorado, a DSpace Digital Repository
Greg Vogl
Colorado State University
2016-03-10

Google Tag Manager

Containers
The two instances of the digital repository, DSpace and DSpaceTest, each have a corresponding container (GTM-WJMNZL for DSpace and GTM-5C48CR for DSpaceTest).

Built-In Variables
Built-in variables were enabled for Pages, Clicks and Forms.
User-Defined Variables

Most of the user-defined variables are Custom JavaScript variables. Initially, DOM Element variables with ID or CSS selectors were used, but the logic needed to be refined using JavaScript to deal with special cases. Many elements cannot be uniquely selected using a simple class, ID or CSS selector, so additional JavaScript variables are often necessary.

DSpace already uses jQuery, and jQuery makes the code simpler and uniform across browsers, so jQuery was often used instead of plain JavaScript.
To test whether a GTM variable is being set as expected, select the Preview and Debug option from the Publish menu and navigate to a page on the repository site to see the GTM debugger at the bottom of the page. Click the Page Load message in the left sidebar and click the Variables tab at the top to view the values of each variable.
Constants
The Google Analytics property of each instance of the digital repository is stored in a constant named gaPropertyID, which is UA-1399660-26 for DSpace and UA-1399660-27 for DSpaceTest. This property can then be easily added to each tag by selecting it from the menu instead of typing the property ID. Also, a container can then be copied between test and production, requiring only a single change to the value of the gaPropertyID constant.

General DOM element Variables
To track clicks of various interface elements, and to help identify what part of the page was clicked, the Click Element, Click Classes, Click ID, Click URL and Click Text built-in variables were enabled, and the ParentID, ParentClasses, GrandParentID and GrandParentClasses user-defined variables were created.

In CSS and jQuery, a class selector starts with a period (for example, .breadcrumb) and an ID selector starts with a hash symbol (for example, #ds-search-form).

Right-clicking a page element in Google Chrome, Mozilla Firefox or Opera web browser and clicking Inspect Element displays a list of the element’s selectors.
Using the Web Developer plugin for Chrome, Firefox or Opera and selecting Information > Display Id & Class Details makes it easy to visually identify the relevant attributes of the many different parts of the page.

Collection Variables
In the current digital repository theme, the navigation breadcrumbs list at the top of each page (CSS selector 'ul.breadcrumb li') contain a full path from the home page to the current collection or item. The Breadcrumbs variable stores the entire collection hierarchy in a string of collections separated by slashes, for example “DSpace Home/Colorado State University, Fort Collins/CSU Archives and Special Collections/Agricultural and Natural Resources Archive/Rocky Mountain Farmers Union”. “DSpace Home” almost always appears as the first item in the breadcrumb list.

The first collection is the university or institution, which is stored in the Institution variable, for example “Colorado State University, Fort Collins”.

The next collection is the top-level collection for that institution, which is stored in the CollectionTopLevel variable, for example “CSU Archives and Special Collections”.

The last collection is stored in the Collection variable, for example “Rocky Mountain Farmers Union”.

The BreadcrumbCount and ParentSiblingIndex variables are used to track the depth in the collection hierarchy, which is 4 in the above example.
Dublin Core (DC) Meta Tag Variables

Each Digital Repository item has many DC fields which appear in meta tags in the page’s source code.

Some DC fields are usually single-valued, e.g. DC.title, DC.publisher, DC.date, DC.type and DC.abstract:

- `<meta name="DC.title" content="Model study of liquified natural gas vapor cloud dispersion with water spray curtains : final report, March 1982-March 1983"/>
- `<meta name="DC.publisher" content="Fluid Mechanics and Wind Engineering Program, Dept. of Civil Engineering, Colorado State University"/>
- `<meta name="DC.date" content="1983-03" scheme="DCTERMS.W3CDTF"/>
- `<meta name="DC.type" content="Report"/>

For these, simple DOM element variables can be used:

Type: DOM element
Selection Method: CSS selector
Element Selector: meta[name="DC.date"]
Attribute Name: content

Most DC fields can be multi-valued, e.g. DC.creator, DC.contributor, DC.subject, DC.description, DC.identifier and DC.language.

- `<meta name="DC.creator" content="Meroney, R. N."/>
- `<meta name="DC.creator" content="Kothari, K. M."/>

JavaScript is needed to combine multi-valued tags into a single string.

A JavaScript function named cssAttribute returns a function to join the specified attribute of all selected elements into a delimited string. This function can be used to get the contents of any DC meta tag.

For example, this JavaScript function calls the cssAttribute function to get the content attribute of all the DC.creator meta tags and joins them in a semicolon-delimited string, e.g. Meroney, R. N.; Kothari, K. M.
Search Variables
Because DSpace does not include search terms in the query string, Google Analytics cannot use the normal method of tracking search terms using Admin > View Settings > Site Search Settings and displaying search terms in Reporting > Behavior > Site Search.

In the current digital repository theme, three different search boxes can appear: in the right sidebar, at the top of the page, or at the top of the page after a search has been performed. The contents of these search boxes are stored in the SearchQuery, SearchQueryTop or SearchQueryTopSimple variables, respectively. The Form Element and Form ID built-in variables need to be enabled. The SearchScope variable is set to SearchCollection, or SearchAll if the entire repository is being searched.

Download Variables
To track clicks of downloadable files, the FileType variable retrieves the file extension from the URL of the download link. Popular file extensions in the digital repository include pdf, jpg, zip, mp3 and mp4.

Page Type Variables
To measure which types of pages are most frequently viewed, the PageType variable uses the URL and breadcrumbs of the currently viewed page to categorize the page by type. Page types include home (the digital repository home page), institution, collection, search, browse, item, metadata, contact, feedback, login, register, and other. These could be further grouped into navigation, content and interactive pages.
Triggers

Searches
All search triggers use the event type Form Submission, and are triggered by the unique ID attribute of the form. The Show Advanced Filters click tracks when the user displays advanced search options.

Downloads
Downloads are tracked by checking the extension of the clicked URL against a list of commonly downloaded file extensions. The Download Link Click trigger is set to fire on Click URL matches RegEx

```
^.*\.(7z|aac|arc|arj|asf|asm|asx|avi|bin|csv|doc|exe|flv|gif|gz|gzip|hqx|jar|pe|g|js|mp(2|3|4|e?g|m|mov(ie)?|msi|mso|pdf|phps|png|ppt|qt|m(?)?|sea|sit|tar|tgz|torrent|txt|wav|wma|wmv|wpd|x|ls|xml|z|zip)\(?.*\)$
```

Outbound links
Outbound links are tracked by checking if the clicked URL does not match one of the DSpace instances or hdl.handle.net which manages the digital repository’s permanent URLs.

The Outbound Links trigger is set to fire on:

- Click URL does not match Regex .*$dspace(test)?.library.colostate.edu+
- Click URL does not match .*$hdl.handle.net+
Other click events
The remaining triggers track clicks of various parts of the page to help assess usability of the navigation, including breadcrumbs, browsing by sub-collection or metadata fields, pagination, and RSS feeds.

Tags
All tags are of type Universal Analytics. Each tag uses a corresponding trigger.

Tag Firing
The Universal Analytics tag fires on loads of all pages and provides all the statistics that would ordinarily appear in Google Analytics using the standard tracking code.

The three search tags fire on form submission events.

The remaining tags fire on clicks.

Tag Category, Action and Label
The tag category matches the trigger name and describes what element or part of the page was used, such as Search, Breadcrumbs or sub-collection links.

The tag action provides more information about the action. For example, the tag action for downloads includes the file extension, the tag action for searches includes the search scope (SearchCollection or SearchAll), and the tag action for navigation facets includes the field (such as Author or Subject).

The tag label includes the URL of the clicked link, or the text of the user’s search query.
Custom Dimensions
Each tag uses the same custom dimensions created by the Custom JavaScript variables that are based on the breadcrumbs and URL of the current page: Breadcrumbs, BreadcrumbCount, Institution, Collection, CollectionTopLevel and PageType.

The Download Event tag and Universal Analytics page view tag also include the DC meta tag variables.

Each custom dimension must first be created in Google Analytics, and the numerical index of each custom dimension in Google Tag Manager must match the corresponding custom dimension index in Google Analytics.
Google Analytics

Properties

Each digital repository instance has its own property and set of custom dimensions which work with the corresponding Google Tag Manager container.

To collect information about user age, gender and interests, in the Property Settings under Advertising Features, Enable Demographics and Interest Reports was set to On.
Views

In addition to the default view for All Web Site Data, a view was created for each institution in the Digital Collections of Colorado, with a filter to include only views of pages where the Institution Custom JavaScript variable matches the institution of the view. This allows each institution to view analytics for only their collections, without needing to create a separate custom report or segment. A non-staff view was also created with filters to exclude library staff IP addresses.
Reports

Standard Reports

Many Google Analytics reports are useful without needing to use Google Tag Manager or customize Google Analytics.

Audience Behavior

If the Digital Repository were just a database of unconnected items, people might simply find an item in a Google search and immediately leave the site. A few standard reports can assess multiple interactions:

Audience > Behavior > New vs Returning shows the percentage of users who returned to the site.

Audience > Behavior > Frequency & Recency indicates whether many visitors come to the site frequently or returned recently.

Audience > Behavior > Engagement shows the time visitors spent on the site and number of pages they viewed.

Audience > Users Flow shows pages visitors viewed after the home page or any other selected page.

Content

Behavior > Site Content includes reports of All Pages, Landing Pages and Exit Pages, to determine pages that are most used, attract users to the site, and promote other sites. Select Primary or Secondary Dimension: Page Title or DC.title to display readable page titles instead of URLs, which are mostly digital repository handle numbers. Only items have DC.title, so you must use DC.title to show only item views.

The Content Drilldown report shows pages by page path. In the repository, the top level path is the handle prefix for each institution (CSU=10217). To list the items in each institution with the most page views, click handle, click the handle prefix for the institution, and select Secondary dimension: DC.title.
Events

Reporting > Behavior > Events > Overview displays the top categories of events, and shows the overall relative use of outbound links, downloads, searching, collection browsing and other navigation features.
To produce a report of search terms, select Behavior > Events > Top Events, click Search under Event Category, and click Primary Dimension: Event Label.

To view all search terms together, select Behavior > Events > Top Events, select Primary Dimension: Event Category, select Secondary Dimension: Event Label, click the advanced link, Include Event Category begins with Search, click Apply, and click the Total Events column to sort descending.
Goals and Conversions

To create a goal based on an event, click the Admin tab, select a view, click Goals, click New Goal, select Custom, click Continue, type a goal name that identifies the event (e.g. Downloads), select type Event, click Continue, and select the event (e.g. by specifying Category Equals to Downloads).

After data has accumulated, click the Reports tab, click Conversions, click Goals, and click Overview. Select the goal from the Goal Option menu, and select Completions or Conversion Rate from the Metric menu. A high conversion rate indicates a high percentage of users are accomplishing your goal.
Five goals were created to measure engagement, using download events, search events, user registrations, session duration over five minutes, and sessions with five or more page views.

Use of Custom Dimensions

A custom dimension such as institution or breadcrumbs can be used as a secondary dimension to group or filter many standard reports.

In the Behavior > Site Content > Content Drilldown report, by using an Advanced Filter of Page path level 2 equals /10217/ (CSU’s handle prefix), and selecting a custom dimension as a secondary dimension, the report can group page views by top-level collection, DC.type, page type, or other custom dimension.

For a report of page views of faculty member Rolston Holmes’s articles, select Behavior > Site Content > All Pages, select Secondary dimension: DC.creator, and create an Advanced Filter of Include DC.creator Containing Rolston, Holmes. Click Primary Dimension: Page Title to view by the articles page title.
Custom Reports

Custom reports allow using custom dimensions as primary dimensions, and a choice of metrics.

To create a custom report, click the Customization tab. Select metrics like Pageviews or Total Events. Select dimensions to group by, such as institution, top level collection, collection, breadcrumbs, page type, or DC metadata. If needed, add filters such as Event Category to count specific types of events such as outbound links, downloads, searches, and page element clicks.

**Page Views by Institution**
### Page Views by Page Type

<table>
<thead>
<tr>
<th>Page Type</th>
<th>Pageviews</th>
<th>Unique Pageviews</th>
<th>Avg. Time on Page</th>
<th>Entrances</th>
<th>Bounce Rate</th>
<th>% Exit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>1,543</td>
<td>934</td>
<td>00:01:05</td>
<td>328</td>
<td>48.05%</td>
<td>23.27%</td>
</tr>
<tr>
<td>Browse</td>
<td>542</td>
<td>295</td>
<td>00:00:22</td>
<td>40</td>
<td>46.34%</td>
<td>8.12%</td>
</tr>
<tr>
<td>Search</td>
<td>428</td>
<td>294</td>
<td>00:00:31</td>
<td>11</td>
<td>63.64%</td>
<td>5.61%</td>
</tr>
<tr>
<td>Collection</td>
<td>423</td>
<td>200</td>
<td>00:00:25</td>
<td>64</td>
<td>46.03%</td>
<td>13.48%</td>
</tr>
<tr>
<td>Other</td>
<td>347</td>
<td>197</td>
<td>00:00:30</td>
<td>16</td>
<td>81.25%</td>
<td>4.93%</td>
</tr>
<tr>
<td>Home</td>
<td>158</td>
<td>101</td>
<td>00:00:48</td>
<td>89</td>
<td>44.94%</td>
<td>30.38%</td>
</tr>
<tr>
<td>Metadata</td>
<td>127</td>
<td>107</td>
<td>00:00:51</td>
<td>15</td>
<td>80.00%</td>
<td>18.99%</td>
</tr>
<tr>
<td>Institution</td>
<td>103</td>
<td>56</td>
<td>00:00:42</td>
<td>5</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Login</td>
<td>72</td>
<td>56</td>
<td>00:01:09</td>
<td>20</td>
<td>70.00%</td>
<td>40.28%</td>
</tr>
<tr>
<td>Recent Submissions</td>
<td>25</td>
<td>15</td>
<td>00:00:08</td>
<td>0</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
<tr>
<td>Register</td>
<td>10</td>
<td>9</td>
<td>00:00:23</td>
<td>0</td>
<td>0.00%</td>
<td>0.00%</td>
</tr>
</tbody>
</table>

This report was generated on 2016 at 3:30:29 PM - Refresh Report
Sharing Report Data

**Shortcuts**
To add any standard or custom report to the Shortcuts menu for easy access, click the Shortcut link at the top of the report. Shortcuts cannot be shared with other Google Analytics users.

**Dashboards**
To add any standard or custom report to a shared dashboard accessible to all Google Analytics users with Read & Analyze permissions on the view, click the Add to Dashboard link at the top of the report, then select the shared dashboard. Dashboard items contain only a small subset of the information in reports, so they are primarily useful as a quick overview.

**Emails**
To share a standard or custom report or dashboard in a single or scheduled email, click the Email link at the top of the report, enter the recipient email addresses, and select the email frequency and duration. Emails must be renewed to be used longer than a year.