Managed Data Layer in Ecommerce and Extending to BigData

Phani Bhooshan[1]; Dr. C.Umashankar[2]

[1]Research Scholar, Department of Computer Science and
Technology, Sri Krishnadevaraya University, Anantapur,
Andhra Pradesh, India
chbhooshanp@yahoo.com

[2]Retired Professor, Department of OR&SQC,
Rayalaseema University, Kurnool,
Andhra Pradesh, India
cumaor@rediffmail.com

Abstract- When talking about setting up analytical tracking or website monitoring, Data Layer, and Tag Management comes in play. In this Article author explained on MDL what do they do, and how do they work using a technic called Tag Management in an Ecommerce website.
Keywords- Metadata; Tags; MDL; Big Data

1. INTRODUCTION

Nowadays, everyone talks about Big Data. The reason being the most valuable asset for any firm in the modern world is DATA. One key aspect of the data is on how to store and retrieve data on need basis – that was an old school method. The latest technological development in software and hardware (specifically storage and processing capacity in terms of GPU) had multifold for past ten years. Most of the technological firms had developed AI, Machine learning algorithms to understand secret patterns within data to help businesses make decisions.

Given data coming in many shapes and forms (structured, unstructured, streamed, pictures, etc.) getting these rightly stored by using Managed Data Layer. In this article, the author had explored Tag management which fundamental concept for information collection for targeted customers.

2. A SCENARIO Explained

Let's explore the title by "when and why do you need a data layer?"

Let's take an example of a fictional website, DogCo. DogCo is a website dedicated to Dogs' pictures. It makes money by selling pictures frames of cats and has its goal to increase its user-base two folds in the coming 12 months. To do that, DogCo wants to increase its marketing reach and performance by integrating and pushing data to as many marketing platforms as possible and keeping a close look at the marketing's performance.
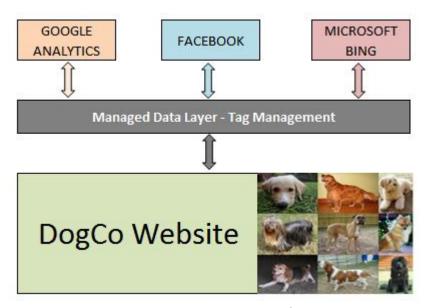
DogCo, however, is short staff in terms of engineers and cannot do by itself the integration with so many external systems. DogCo has a problem.

Here comes the data layer and tag management system. The data layer allows setting up a structure for the information needed to integrate with other websites.

A Managed Data Layer typically refers to the concept of data object available within a Web Browser for integration with third parties.

Let's have a look at DogCo; DogCo would like to integrate with social and service platforms like Google Analytics, Facebook, and Bing. It is not possible to directly integrate with them, however, combines with the help of a Tag Management System that enables information to be shared unanimously across these platforms:

Figure 1: Data Layer Integration with 3rd party websites

3. STRUCTURE AND CONSTRUCT OF MDL

When a user visits below DogCo web page, we can see that the page has a few characteristics:

Figure 2: Snowy Playing in Snow

•   This page expected to sell a particular Dog's picture frame

•   It has a title and specific URL address

•   One specific product is being displayed on this page

•   From this page, If a user would like to purchase the product he/she could be able to add to the cart

What DogCo would be interesting in knowing and tracking from that page would typically be collected by two distinct types of events, page Views and add to Carts. If we were to decompose the available data that could be pushed on that page for each of these events:

| PAGE VIEW | PAGE ATTRIBUTES | Page Category: PDP
Title: Snowy is Jumping
URL: /frame/SnowyJump.html |
| | PRODUCT ATTRIBUTES [IMPRESSION] | Product ID: 999
Product Name: Snowy is Jumping |

| ADD TO CART | PRODUCT ATTRIBUTES | Product ID: 999
Product Name: Snowy is Jumping |

Figure 3: Tags on Web page

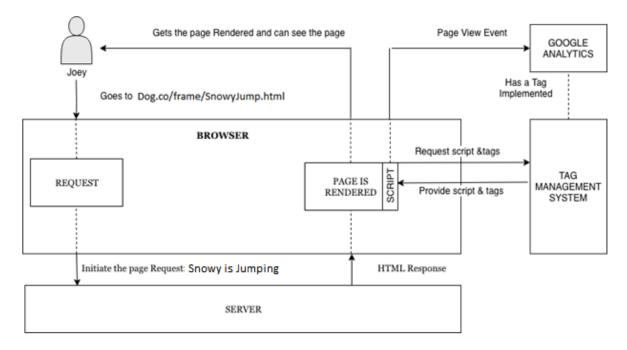Event flow upon user called Roja goes to this page on DogCo:

Figure 4: Event Flow Diagram on a Webpage

Roja goes onto frame/SnowyJump.html on her favorite browser, and this initiates a request to the server, which response with some HTML code back containing the code for the page "Snowy is Jumping." The HTML program contains within its definition of a JScript object, for instance, with Google tag manager it may look like:

```
dataLayer = [{
    'pageCategory': 'PDP',
    'pageTitle': 'Snowy is Jumping,
  }];
```

And 3rd party website's script, requesting some piece of JScript code from the tag management system.

```
<script src="mygenericmdl.com/mdlscript.js"></script>
```

As the page gets submitted and displayed to Roja, the program gets loaded and decides which tags need to be fired. The script also contains information related to the mapping of available data in the Data Layer subject to the bidden information in Google Analytics[4]. This translation is configured in the user interface of the tag management system itself. The script then fires the Google Analytics tag with the mapped data to trigger a page view event.

Different types of event from a page view would happen similarly, but the tag management script would be "listening" to a specific event happening rather than doing that when the page is rendered. If Roja decided to add Snowy's frame to her cart, for instance, there would need to be an event impacting the Data Layer for the tag management script to determine whether a tag needs to be updated. Google's Tag Manager does this through dataLayer.push action, and Tealium Tag Manager does it through utag.link action.
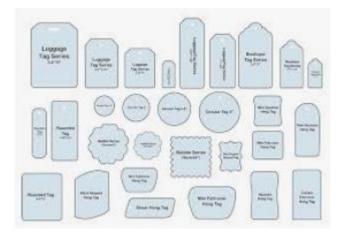
4. Usage and Benefits

Figure 5: Tags of Various Attributes

DataLayer provides a standard way of defining tags at front-end application with integration to external websites. Data within a DataLayer can be used to set up enhanced e-commerce with-in Google analytics, integrating with many more advertising partners out there also used for other purposed within the website such as when it is meaningful to show up a pop-up.

For Google Analytics website implementation, there is a need to provide quite a few events and information within its DataLayer:

- Add to/remove from/update cart
- Checkout
- Purchase
- Page View

Other information may be provided in other to have a better tracking of what is happening on your website. Besides using it for tracking purposes, the information provided in the data layer is useful to feed DMPs with specific target segments. Set up goals for digital marketing optimization, etc. [4].

5. EXTENDING TAG MANAGEMENT TO BIGDATA

While tag management exists and mostly inbuilt with Internet browser, this can be extended with uplift to BigData concept. We call this concept as ManagedDataLayer. When information is being collected across various platforms, it may not be practical or possible to know data about data. This opens up how better we understand the structure and relationship between the data that's been received/collected for storing, analyzing using BigData concepts.

Figure 6: Known and Unknown Sources of data into BigData

BigData concepts starts with data ingestion, before data can be ingested into HDFS, Physical structure and relationship has to be pre-defined. However, as we mentioned earlier, if data is being collected from known and unknown sources, it may pose hard to determine the structure that leads to spending much of resources (time, cost, human resources, etc.) to know about data. Also, this poses an additional risk of source people to define and own the structure of the data that's been provided — more constraints to the system.

By adding a layer called ManagedDataLayer, in short MDL to the BigData concept, we could overcome some of the challenges by knowing in advance data about data. This layer could be containerized to minimize its impact in terms of performance and change management perspective, as well as service-oriented to work with various open technologies out there.

6. CONCLUSION

In this journal publication, the author tried to understand the importance of Managed Data Layer using Tag Management. We applied the model in an Ecommerce Website for different type of information and usage. We have identified, Tag can provide valuable information that can be used to improve marketing and targeted clients. And can be extended to BigData concepts to evade some of the early challenges of adopting BigData Technology.

References

[1]   Dileep B. Desai, Dr. R.V.Kulkarni "A Review: Application of Data Mining Tools in CRM for Selected Banks," (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 4 (2), 2011, 199 – 201.

[2]   Deshpande, M.S.P., and D.V.M. Thakare, 2010. Data mining system and applications: A review. Int. J. Distrib. Parallel Syst., 1: 32-44.

[3]   S. Kotsiantis, D. Kanellopoulos, P. Pinellas, "Data Pre-processing for Supervised Learning," International Journal of Computer Science, 2006, Vol 1 N. 2, pp 111–117.

[4]   Julien Kervizic "Data Layer and Tag Management," Hacking Analytics, Oct 2018. https://medium.com/analytics-and-data

A Brief Author Biography

Phani Bhooshan – Pursuing Doctor of Philosophy. I had worked as a Teaching Assistant in the same department — currently a businessman. Interests include reading online articles and current affairs, world travel.

Dr.C.Umashankar – A Legend of the university with many statistical literature and journals published. Had awarded many Ph.D. under his guidance. He is being pursued as a Go-to person for any knowledge under the sun. He retired from his profession and currently on world tour. A die-hard cricket fan.