



# IJIRSET

International Journal of Innovative Research in  
**SCIENCE | ENGINEERING | TECHNOLOGY**

e-ISSN: 2319-8753 | p-ISSN: 2320-6710



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 10, Issue 7, July 2021

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

Impact Factor: 7.569



9940 572 462



6381 907 438



ijirset@gmail.com



www.ijirset.com

# Creation, Enrichment and measurement of First Party Data

Rohan Rajkumar<sup>1</sup>, P.K Manjesh Ganapathy<sup>2</sup>, Dr. Shobha N.S<sup>3</sup>

U.G. Student, Department of Industrial Engineering and Management, R.V College of Engineering, Bangalore, India<sup>1</sup>

U.G. Student, Department of Industrial Engineering and Management, R.V College of Engineering, Bangalore, India<sup>2</sup>

Assistant Professor, Department of Industrial Engineering and Management, R.V College of Engineering,  
Bangalore, India<sup>3</sup>

**ABSTRACT:** The project is based on Creation, Enriching and Measurement of First Party data, carried out at Wavemaker India Pvt.Ltd. The problem at hand is to realize the unfeasibility of **Third-party data**, which has always been the go-to method for companies to collect data on users and help with advertising. With the new data policies, the use of third-party data has to stop and there needs to be a complete shift to **first-party data**. The primary objective is to establish a First Party Data Architecture for clients through [m]insights integration, using Data from **DMP** (Data Management Platforms/sources). Use of First party data to create **DCO** (Dynamic Creative Optimization) strategy-based ads and Insights. The analytical and programming tools used are: Python, excel and [m]insights. The motivation for the project is fuelled by the rules put out by the Global Data Regulations such as GDPR (Global Data Protection Regulation) which highlights the importance of First Party data and pushes for a complete First Party Data Platform and Strategy by the year 2025. The process starts with Collection of Data, Unification and ID resolution, enrich the data and Activation.

**KEYWORDS:** 1PD – 1<sup>st</sup> Party Data, 2DP – 2<sup>nd</sup> Party Data, 3PD – 3<sup>rd</sup> Party data, DSP – Demand side Platform, DCO – Dynamic creative optimization, DMP – Data Management Platform, IAB – Interactive Advertising Bureau, [m] insights

## I. INTRODUCTION

Media Agency is a term that is traditionally used to describe an agency that strategically plans where and when the advertising and marketing message should appear. Media agencies use a combination of Software, Data analytics and Algorithms to deploy and make sure marketing message appeals to consumers, appears in the right place, at the right time and that the advertiser pays the best possible price.

The term “Media Agency was initially coined to portray an organization that creates, makes and deliberately designs where and when promotion and advertising messages are to be displayed. That is, it helps purchase the advertisement/media space and helps portray the message on those spaces. The media agencies are not only responsible for the transaction of advertisement spaces, but also managed advancement in communication technologies

With advancement and progress in the digital space, media agencies have numerous efficiencies as buyers who change the way the digital space plans the strategies and techniques. They have adapted to numerous new and unused channels to handle their problems. These agencies today provide key strategic counselling and consultation to help best use digital for a brand's communication. How to optimize advanced media touchpoints, create vital engagement plans across various medium, help find unused and new digital spaces e.g. Social networking, mobile and Web 3.0 for advanced and accurate integration.

New guidelines by various Government Data regulatory organizations have made it very clear that Cookies should be phased out completely citing privacy concerns. This has resulted in companies shifting to adoption of First-party data in order to create insights and strategies in order to target their consumers on digital spaces.

The project involves solving this problem by coming up with a method of using First-party data instead of third-party data, performing analysis on this data and, enriching it by using Target segmentation and creating Strategy and Insights to reach audience based on Identity-first approach leading to higher accuracy and reach.

## II. RELATED WORK

Several Journals were studied to understand the value of First Party, Second Party and Third-Party data and how to handle them for the future.

### [1] Timothy Morey, Theodore “Theo” Forbath, and Allison Schoop [2014]

A study of 900 individuals across five countries, which took a gander at perspectives about information protection and security. It analyzed what individuals thought about the data trails they leave on the web, which associations they did—and didn't—trust with their information, and which information they esteemed the most.

### [2] Kathryn Montgomery, Jeff Chester, Laura Nixon, Lillian Levy and Lori Dorfman [2014]

One recognizes six key highlights of this new Big Data food promoting framework. Investigate how they leave from customary types of publicizing, showcasing, and retail tasks, and present ideas for research procedures and general wellbeing intercessions.

### [3] Matthew J. Schneider, Sharan Jagpal, Sachin Gupta, Shaobo Li, Yan Yu [2015]

Ideal information assurance methodology relies upon a danger return tradeoff dependent on the probabilities of misclassification of clients into fragments. The chance expenses of incorrectly allotting section participation, and the expected expense of an information break.

### [4] Andrei Hagiu, Daniel Spulber [2015]

Shows the essential utilization of first-party content by two-sided stages is driven by two key factors: The idea of purchaser and merchant assumptions (ideal versus ominous) and the idea of the connection between first-party content and outsider substance (supplements or substitutes).

### [5] Oliver Gertz and Deirdre McGlashan [2016]

Each automatic arrangement consolidates a DSP, First and Third-Party Data, Media Inventory, Brand-Safety approaches, Adserver and to wrap things up specialists that drive the entire framework. Promoters have increasingly more decision on arranging their own arrangement which moves the conversation away from "office versus in-house" to "what is the right mix of in-house and rethought for necessities."

### [6] Jonathan R. Mayer and John C. Mitchell [2016]

It is seen that this paper overviews the current approach banter encompassing outsider web following and clarifies the applicable innovation. It additionally presents the Fourth Party web estimation stage and studies that they have led with it.

### [7] Gian M. Fulgoni [2016]

The next paper is about the understanding about the data barriers Where The on the web and the disconnected world act like various biological systems, lessening the utilization of online media to advertisers. Going ahead, advertisers can either run showcasing blend models that reject the causal segments for which information are inaccessible, or they can utilize causal information that are collected across clients.

[8] Geoffrey Precourt [2017]

The next part involves the study of the real-world example of a company incorporating a First Party Data Strategy Web based business request has flooded on stages claimed by Anheuser-Busch InBev, Budweiser's parent organization, just as outsider conveyance administrations. First-party information is a basic asset to Budweiser, and can assist it with enhancing pack organizations, recommendations and other basic components of the essential blend for web-based business.

### III. METHODOLOGY

The entire methodology of the project is divided into three main steps, which are Data collection phase, Data enrichment phase and the Data activation phase. Data Collection: Based on research, observation, and analysis of customer requirements, and with use of quantitative and methodologies, the type of Data and the source for it is defined. Enrichment of Data: In this phase certain measures and practices are performed to enhance and improve the Data that is collected using certain mathematical methods to help make the data more relevant and improve accuracy. Data Activation: In this phase of the project a pre-check is conducted on the data that has been enriched, by looking at the size of the data, the relevancy and ensure if it is fit for use. Excel is used to perform basic analysis to develop Insights and help activate a target segment for customers/clients.

**1. COLLECTION:**

Ensuring data is collected at scale across a market to provide accurate audience insight. This gives an overall market view which is a valuable starting point for the client.

**2. UNIFICATION:**

1PD 2PD 3PD are combined to create a holistic view of the consumer. The [m]PLATFORM identifier [mp]ID combines all data points on an individual to give this detailed profile. Thus, data management was done in order to aggregate all relevant data gateways to reaching a unified structure, set for further processing.

**3. CLASSIFICATION:**

[m]insights provide consumer and media insight tools to understand which consumers are relevant and how these relevant consumers behave in a digital ecosystem.

This digital ecosystem consists of Consumer intelligence, media insights and planning intelligence. Thus, classifying the aggregated data into meaningful subdivisions for further analysis

**4. AMPLIFICATION:**

[m]PLATFORM has over 40 data scientists globally who have built automated machine learning algorithms to accurately scale data collected by each advertiser in the platform.

Thus, creating an audience studio, i.e., a lookalike model of the target audience for the client which is a valuable resource for their understanding of what a target population would look like.

**5. ACTIVATION:**

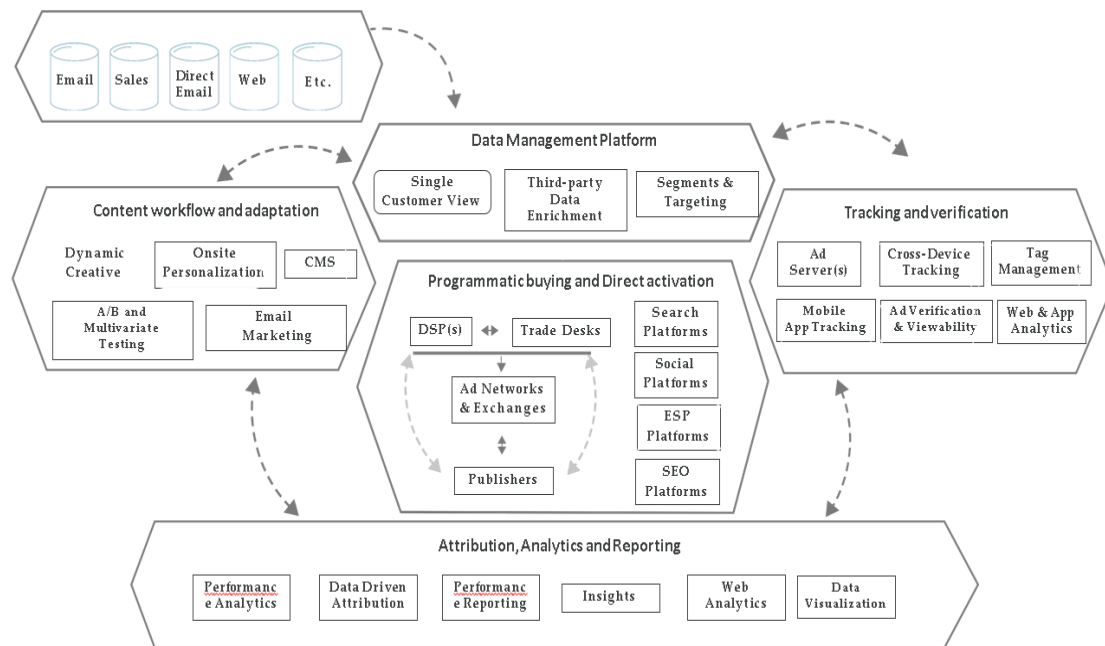
All data is propagated in real time to DSPs for digital activation ensuring if a consumer can be found to be relevant, they can then be targeted programmatically.

### The First Party deployment architecture

First-party information is profoundly significant in light of its quality. Since it is gathered straightforwardly from the source, it is known it's precise, and in light of the fact that it comes directly from the crowd, know it's applicable to the business. Another advantage of first gathering information is that security concerns encompassing it stay insignificant in light of the fact that one knows precisely where it came from and, as the advertiser who gathered it from the crowd, own it. As a result of the great nature of first gathering information, there are numerous alternatives for how advertisers can utilize first gathering information. The construction is delineated where the first party information's precision and



importance permit to anticipate future examples, like crowd conduct, with certainty. The information permits to pick promotions that bid more to the crowd which help utilize first gathering information to customize the substance or promotions that show to a specific client, as displayed in the sending design. The information is gathered on a guest to the site will give further developed bits of knowledge into their inclinations and necessities, permitting to serve content to them that feels customized. Say, for example, one has different advanced recordings on various themes that need to impart to likely clients as a component of an inbound showcasing procedure. By taking a gander at information about the substance the crowd has seen previously, one can decide clients' inclinations and send them recordings about science, sports, music, design, nature, wellbeing or whatever regions their inclinations lie in.



**Figure 1. IPD deployment architecture**

### Look-alike Modelling

This is the way toward distinguishing new clients that look and act like the current crowd. By taking a seed crowd and characterizing key attributes which separate this load of different qualities. [m]insights stage has been displaying and different cycles that will assist with distinguishing another bigger, crowd that is like the current shoppers.

## IV. RESULTS AND CONCLUSIONS

The Target Segmentation process generates a DSP report after the Seed attributes gets populated for 24 hours. The DSP report is released in a Microsoft Excel format (.xls). This report comprises a dataset based on the custom audience that is created, the audience is segmented based on Audience type, audience name, audience size, relevancy score and overlap percentage.

The main outcome of this project was to develop a new approach to creating a complete First-Party data strategy to generate audience Insights. The platform used to perform this is Wavemaker's own Analytic software tool called [m]insights which is easily customizable to work with different Data types. The project's goal was to develop a new approach to audience targeting with the use of the clients First-party data only that reduced data processing time, had a very large-scale Audience footprint while also being extremely accurate.

## Project Outcomes

The outcomes of this project were as follows:

Generated consumer insights with the help of [m]insights tool and Excel by only using First-party data from the client. Using this data, a highly accurate target audience was developed by collecting Portrait data (2nd party data) and Cookies (3rd party data) based on Seed attributes and integrating it along with the client's data by incorporating analytical modelling methods like K-means clustering and Look-alike clusters.

Demonstrated how powerful the [m]insights platform is and showcased its breadth of ability by building a custom audience from the basics using first-party data.

Developed a K-means clustering model from scratch to showcase the understanding of the how [m]insights work and created clusters for an audience size of 57.3 million by accurately replicating the tool.

## V. FUTURE SCOPE

With Third-party cookies being essentially dead now that Google has announced their intention to phase out cross-site tracking (3rd party cookies) in browsers by 2022, there shall be a wide variety of ways software vendors, companies and marketers try to address these new challenges, including “device graphs,” “Privacy Box”, non-cookie-based ID matching technologies, publisher alliances, and other experimental means to target Consumers. The methods used in order to generate audience insights and strategies using First-party data only can be used in other applications as well. This change in reliance on Data will result in many large organizations realizing that they are moving to a cookie less future and that it's already time to take total control of their first-party data, by making it as comprehensive and usable as possible. This will result in new approaches companies will have to take in order to b Future-proof, some methods are:

**User-Registration:** Most companies now will be forced to make users register before they access their websites as they will need to be appended to users' data records in order to deliver truly personalized experiences throughout the entire digital marketing stack. The companies will have to come up with creative ways to ensure user registration, this will also result in companies adopting higher data-protection methods in order to keep their data and users' data safe thereby increasing privacy [21].

**Progressive Consumer profiling:** Since companies can't use third-party cookies to track users in the background, they will need to be open about asking users for consent and permission to track and collect information on them, this will thereby ensure that user has control of his/her data and can control what they're shown. By-product of this is that ads will be less cluttered and content will get even more personalised now, thereby improving engagement rates and consumer satisfaction [20].

**Tracking based on Event only:** This will ensure that advertisers can track user's data only during certain events i.e.- surfing a particular section of a page, using fitness apps, using specific apps etc. This will ensure that data gets tracked only for a specific event and won't intrude unnecessarily, thereby enriching First-party data and ensuring consumer has data privacy

## REFERENCES

1. Customer Data: Designing for Transparency and Trust, Timothy Morey, Theodore “Theo” Forbath and Allison Schoop. [2014]
2. Protecting customer privacy when marketing with second-party data, Matthew J. Schneider, Sharan Jagpal, Sachin Gupta, Shaobo Li, Yan Yu, PII: S0167-8116(17)30008-3. [2014]
3. First-Party Content and Coordination in Two-Sided Markets, Andrei Hagiu Harvard Business School, Harvard University, Boston, Massachusetts 02163, Daniel Spulber Kellogg School of Management, Northwestern University, Evanston, Illinois 60208. [2015]
4. Consumer-Centric Programmatic Advertising, Oliver Gertz and Deirdre McGlashan. [2015]
5. Third-Party Web Tracking: Policy and Technology, Jonathan R. Mayer and John C. Mitchell, Stanford University Stanford, CA. [ 2016]



6. Why data barriers are preventing marketers from optimizing marketing spend, Gian M. Fulgoni, Journal of Advertising Research, Vol. 58, No. 4, 2018. [2016]
7. Budweiser builds an e-commerce business with first party data and smarter advertising, Geoffrey Precourt. [ 2017]
8. Seed audiences – “From not knowing you to tapping into millions like you”: Leveraging hypertargeting to maximise media ROI. Mario Dughi, Utsav Mamoria, Abhinav Mohan, Souvik Roy, ESOMAR, Insights Festival, 2020. [2017]
9. Predictive programmatic re-targeting to improve website conversion rates, Abhiram Prasad et al 2021 J. Phys.: Conf. Ser. 1714012027. [2017]
10. An Analysis of First-Party Cookie Exfiltration due to CNAME Redirections, Tongwei Ren\_, Alexander Wittmany, Lorenzo De Carli, Drew Davidsony  
Worcester Polytechnic Institute, University of Kansas. [m.2017]



**Impact Factor:  
7.569**



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 **9940 572 462**  **6381 907 438**  **ijirset@gmail.com**



[www.ijirset.com](http://www.ijirset.com)

Scan to save the contact details