MIGHTYHIVE

Google News Initiative

Playbook

Google News Initiative Data Lab 2020

August 2020

This program was developed in partnership with the Google News Initiative and following partners: 55 the data company, Datalicious, e-Agency, MightyHive & Tatvic

55 the data company   datalicious   e-Agency

MIGHTYHIVE   Tatvic

Glossary

ARPU
Average Revenue Per User is a metric defined as the total revenue divided by the number of subscribers.

Brand Lover
A Data Lab audience segment defined as users with more than 30 visits, and 50 pageviews within a 30 day period and a visit within the last 3 days.

CCPA
The California Consumer Privacy Act is a state statute intended to enhance privacy rights and consumer protection for residents of California, United States.

Content Recommendation Engine (CRE)
An algorithm that uses data that describes the history of the user, the behavior of similar users and the article or content itself in order to suggest or 'recommend' another article to that user.

Deloitte
A multinational professional services network that provides audit & assurance, consulting, risk and financial advisory, risk management, tax, and related services to select clients.

Deloitte Study
The Deloitte digital transformation through data study: a guide for news and media companies to drive value with data. Link.

eCPM
Effective CPM, usually called CPM only, is the revenue a publisher can generate for every 1,000 impressions

Entity Extraction
An information extraction technique that seeks to locate and classify named entities mentioned in unstructured text into predefined categories (also called Entity Name Extraction or Named Entity Recognition)

GDPR
General Data Protection Regulation is a regulation in EU law on data protection and privacy in the European Union and the European Economic Area.

Google Analytics (360)
Also known as GA, is a free web analytics tool offered by Google to help you analyze your website traffic. Google Analytics 360 is the premium version which incurs a fee to use.

Google Ad Manager (360)
Also known as GAM, is an ad management platform for publishers who have significant direct sales.

Google Tag Manager
Also known as GTM, is a tag management system that allows you to quickly and easily update tracking codes and related code fragments collectively known as "tags" on your website or mobile app.

Glossary

Native advertising solution

An off the shelf native advertising system that a publisher can use to link third-party advertising and articles to pages of their site.

Natural Language Processing

A branch of artificial intelligence that helps computers understand, interpret and manipulate human language.

Recency, Frequency, Volume (RFV)

A method used by publishers to measure audience value based on the recency and frequency of their visits and the volume of content consumed.

Recirculation

The Recirculation rate is the percentage of users who visit another page of a publisher website after they finish reading their first article.

Revenue Percentile

A metric used to understand the distribution of revenue and eCPMs across inventory. Calculated by dividing users into 10 equal buckets based on the average impression revenue eCPM over 1000.

Share of house

A metric used to monitor non revenue driving inventory and allocation of this inventory to maximize potential for monetization. It is the number of impressions served as house line items divided by the number of impressions served in total.

Technologist

Google partner leading the development and delivery of a technical solution to one of the publishers in the program

Executive Summary

As the publishing industry discusses the state of its future, it's crucial publishers explore how sustainability and growth can be achieved through the use of data. Google News Initiative (hereafter, GNI) sponsored Deloitte's *Digital Transformation through Data: A Guide for News and Media Companies to Drive Value with Data (March 2019) study* (published by Deloitte with the support of GNI) to frame how the industry's current issues could be mitigated through data based solutions. Nine use cases were suggested as a means to achieve data maturity and future proof the publishing industry. Through the execution of these data focused projects, it's possible to affirm that maturity is essential for the success, growth and longevity of publishers. A publisher's investment in data orientated solutions is not only a practice in control, transparency and democratization of data, it's also the quickest way to grow digital maturity, realize better performance and drive revenue.

To ensure strong adoption rates and keep project momentum high, especially when COVID-19 hit, it was important all deadlines were met. To demonstrate practicality of the use cases, the implementation team sternly focused on delivering value in the quickest possible way. With this mentality and utilizing all available resources, the project was executed over a three month period whereby the impact could be observed. With some capability to work on these use cases in parallel, how much time would be required to master all nine of them and what should be the order of execution? There is no easy answer; although, based on the Data Lab cohort, content planning is the most obvious use case to begin the maturity journey followed by recirculation. Once a robust content planning strategy is implemented, experimenting with inventory yield management and audience advertising would be a natural progression as both are interconnected with pricing strategies.

When content, audiences and advertising data is harnessed correctly and thoroughly, a publisher can comfortably mature its business model to engage with a subscription model. The problem arises when many execute this project prematurely. Despite enjoying strong initial results (for example, higher conversion rates), often issues begin to arise (for example, regarding audience retention) that could have been avoided where data based solutions were adopted to aid in identifying and solving business challenges.

When it comes to lifetime value, the importance of this metric and related business outcome grows with maturity as a publisher's dependence on data starts to significantly influence strategic decisions. For this reason, this strategy shouldn't be implemented across nascent or developing organizations where such a structure is not developed nor sufficiently mature.

The GNI APAC Data Lab is a publisher-centric program that aims to help Asia-Pacific publishers transform from "developing" to "mature" by adopting data focused processes. Through collaboration, Google and MightyHive assisted publishers in becoming more data-driven through five Deloitte use cases. Publishers wishing to be assisted with the execution of these use cases were requested to self assess maturity. This self assessment was used to determine the appropriate publishers to become GNI's first APAC Data Lab cohort. Once selected, and crucial to the success of the program, MightyHive conducted in-depth publisher digital maturity evaluations. This consultative approach allowed for in-depth strategy and roadmaps to be developed for each participating publisher, outlining short and long term goals in addition to selecting focused use cases where the program could deliver direct impact of data maturity on the publisher's bottomline.

This program executed five use cases across a cohort of six publishers in six different Asia-Pacific markets: content planning, inventory yield management, recirculation, audience advertising and lifetime value. During the early stages of implementation, two main insights were immediately apparent:

As many publishers in the Asia-Pacific region ranked as "nascent", or low, in maturity, there is a wealth of opportunity available within the publishing industry for data-driven transformations.

Each publisher in the cohort already had a subscription model in place or was exploring the viability of introducing this model. To ensure the success of these subscription models, strong data infrastructures and strategies are necessary as an audience focused framework connected with content strategies enables a sustainable subscription service.

In the context of this playbook, maturity is constructed around the degree of audience segmentation and detailed analysis of accurate data available to lead financially successful business strategies.

Deloitte's "data maturity scale" defined four distinct levels of maturity, whereby a publisher's reliance on data, cross team capabilities, processes and decision making determines its readiness to implement a use case. For example, a 'developing' publisher has a basic level of readiness and functionality for the implementation of a use case, whilst a 'mature' publisher's infrastructure is far more detailed and ready to adopt more comprehensive data and audience driven business solutions.

Many publishers have an audience strategy but continue to rely heavily on third-party cookies, which will continue to be disrupted by external forces being rolled out or adopted globally (such as Intelligent Tracking Prevention (ITP), browser privacy updates and new privacy regulations (e.g. GDPR or CCPA)). As third party cookies are disappearing from the digital ecosystem, publishers need robust first party data strategies to maintain a competitive long term market position; imperatively, in markets where solutions like ITP are already impacting audience strategies. Whilst the Deloitte Study does not directly call out this urgency when describing audience advertising, there is a clear need for attention to be shifted towards a first party data strategy.

Coupled with this shift, is the need for publishers to adopt a strategy focused on growing its first-party cookie loyalty base by adapting content strategies to suit the need of a 'brand lover' rather than a casual reader. This control over data solutions allows for a reorientation and alignment in business strategy and cross functionality of teams. It creates a clear focus for editorial and content planning; sales teams can drive results through a comprehensive understanding of its audience; product and tech teams can tailor how the content is accessed to ensure ease, efficiency and an overall positive onsite experience for its loyal audience; and executive teams are able to forecast, strategize and undertake appropriate business planning around a focus on audience growth and retention strategies.

Whilst the implementation of agile data and technology solutions provide immediately positive results on a business, long term success is strongly dependent on far reaching organizational changes. The adoption of cross-functional teams aligned by mutual goals, education and appropriate processes (that are maintained) allows for true business growth to occur. Therefore, the use of a change management model is recommended to provide publishers with a framework to ensure the shift in roles, responsibilities and process is transparent and inclusive.

As COVID-19 lockdowns were in place when the final delivery of this project occurred, audience behavioral patterns impacted and limited quantitative results; despite this, many compelling observations were made regarding the immediate results of these use cases. Overall, a simple approach to segmentation and exposure to new metrics was very welcomed across publishers with many cross functional teams successfully adopting dashboards displaying these insights.

Some more specific highlights include, the share of highly engaged audiences increased between 20% to 40% on a weekly basis due to a new content planning strategy. Whilst with recirculation, the parity between a third-party recommendation engine and the bespoke solution click-through rate was achieved between a third-party recommendation engine and the bespoke solution. Importantly, a full transition from an existing recommendation system to a custom solution created an opportunity to monetize new (and revived) ad space. For inventory yield management, publishers who adopted a three pillar approach of performance, troubleshooting and forecasting, were able to significantly increase viewability (up to 50% increase). Further to this, some publishers observed a 5% increase in eCPMs.

Across all publishers and use cases, it was observed that a culture of experimentation and testing needed to be established. The program exposed many challenges related to how data is currently structured and managed across organizations with many available solutions not being easy to deploy; creating a low willingness by the publisher to address these issues

The purpose of this playbook is to empower readers to make a case for change, facilitated by data-driven processes. It outlines the execution of three use cases - content planning, inventory yield management and recirculation; based on the cohort experience, all provided the most immediately visible impact (i.e increase in revenue) on the business. The execution of these use cases follow an assumed order of implementation to support a look back loop. Based on the learnings from this program, these projects increase the likelihood of having successful subscription operations and more advanced revenue based business models.

Google News Initiative   MIGHTYHIVE

Cross-functional collaboration, implementation of agile solutions and overall organizational changes are easier as discussion points than actual implementation. As such, this playbook is valuable for organizations wanting to create and action change through cross functional teams engaged and aligned over mutual business goals. Whilst a step-by-step guide to implementation is not included, high-level technical details are explored. Such details include the strategic approach, high-level solution (including the technical solution), prerequisites to execute and change management features for each of the use cases. Each reflects the methods MightyHive, and partnering technologists, used in executing the use cases across multiple Asia-Pacific publishers as part of the GNI Data Lab. Importantly, hurdles and blockers coupled with successful outcomes have been shared to provide insights into the nuances involved with practical executions of the use cases.

All described use cases require a multidisciplinary approach, therefore it is advised to share this playbook across relevant departments in order to facilitate the preparations necessary for change.

Table of Contents

Google News Initiative   MIGHTYHIVE

Table
of Contents

Google News Initiative   MIGHTYHIVE

Table of Contents

Background

Background

Deloitte Digital Transformation Through Data Study

The GNI APAC Data Lab program was founded on *Digital Transformation through Data: A Guide for News and Media Companies to Drive Value with Data (March 2019)* study (published by Deloitte with the support of the GNI), the goal of which was to articulate how news and media companies are leveraging audience data to better engage readers and generate revenue.

The report into the 'data activation' practices of news media companies outlines nine use cases for using data to drive better reader insights and business outcomes, and is based on interviews with more than 80 individuals across more than 50 news and media companies from 16 countries. The use cases below, split by three key goals – to improve user engagement, to increase reader revenue and to grow advertising revenue, formed the scope of the Data Lab program.
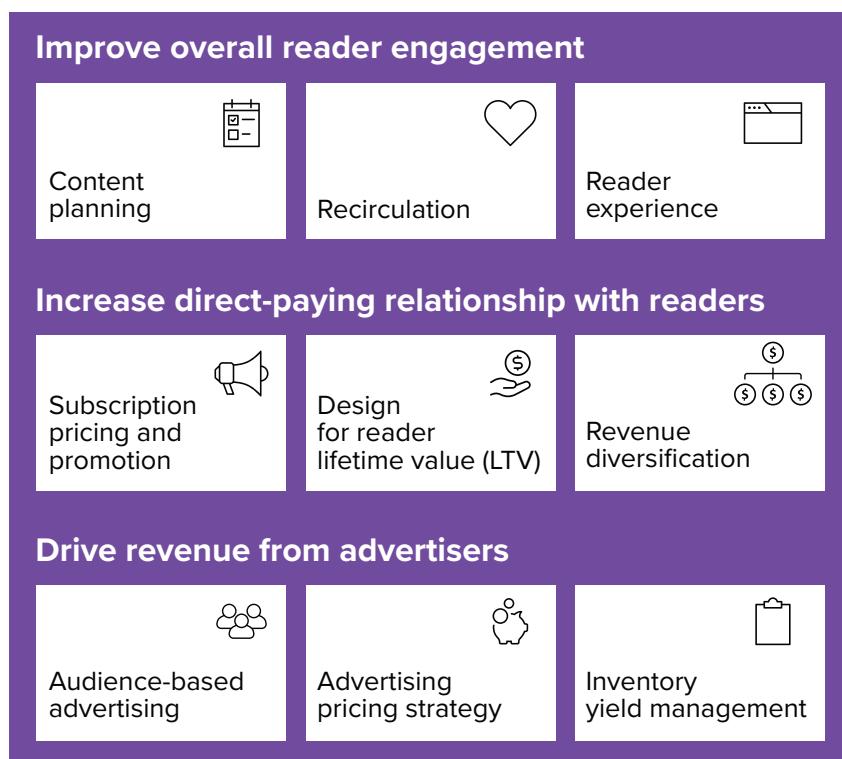
Improve overall reader engagement

Content planning	Recirculation	Reader experience

Increase direct-paying relationship with readers

Subscription pricing and promotion	Design for reader lifetime value (LTV)	Revenue diversification

Drive revenue from advertisers

Audience-based advertising	Advertising pricing strategy	Inventory yield management

Background

Based on an evaluation of all the publishers in the Deloitte Study, Deloitte created a data maturity scale that defines four distinct levels of maturity, articulating a specific criteria for each level by focusing on the degree of audience segmentation and detail of analysis available to end users:

- Level 1 *Nascent* - having a limited level of use case functionality
- Level 2 *Developing* - having a basic level of use case functionality
- Level 3 *Mature* - having a detailed level of use case functionality
- Level 4 *Leading* - having a dynamic level of use case functionality

The objective of the Data Lab is to move publishers from 'Developing' to 'Mature' for two use cases.

Nascent   ***Developing***   ***Mature***   *Leading*

Data Lab Program Methodology And Approach

In order to transform publishers from "developing" to "mature" data-driven organizations, MightyHive's approach was to:

- Assess the maturity of the organization for each of the nine use cases outlined in the Deloitte study using the framework laid out by it.
- Identify suitable use cases in which to develop the publisher's data capabilities which were aligned with the publisher's priorities and overall strategic direction.
- Identify execution gaps in culture and skillset (data foundations) that would lead to immediate bottom line impact
- Based on the above, identify the two best suited use cases outlined by the Deloitte Data
- Assist with the implementation of the two recommended use cases using a local technologist wherever possible.
- Execute a change management program to ensure adoption and ongoing usage of the new data tooling.

Google News Initiative Data Lab 2020
Background

Deloitte Digital Transformation Through Data Study | Data Lab Program Methodology And Approach |
Mightyhive: Selection Of Publishers And Scoping Of Strategy | Working With Technologists

MightyHive: Selection Of Publishers And Scoping Of Strategy

Publishers were selected based on their individual applications, maturity "self-assessment", tech readiness and commitment.

Once selected, interviews with key stakeholders in each organization were held – business leaders, marketers, content managers, data engineers, developers, analysts – in order to assess their maturity and capability to execute each of the nine use cases.

In order to make a case for change, a holistic understanding of exactly how successful an organization is, is required. This understanding can indicate what kind of impact a case would have on it. This, of course, is difficult as most employees work in siloed departments with their own objectives and bottom lines and viewing success is limited to the lense of that department. It is for this reason that some initiatives tend to get introduced to an organization but do not get adopted long term. Initiatives that cater to every department head's objectives can be difficult to identify and implement. The interviews across these different departments aim to find common bottom line goals so they can be linked and implemented as multifaceted initiatives that serve everyone. They captured the organizational status quo regarding data infrastructure, people, process and practices and identified strategic gaps and reporting needs across the organization.

Each selected publisher demonstrated a desire for data-driven change and innovation in their organizations, this research phase of the program was where we attempted to expose exactly what those needs were and how it could align with the Deloitte use cases.

Google News Initiative Data Lab 2020

Background

MightyHive, along with local APAC technologists were tasked with delivering bespoke use case solutions. The selected publishers were:

South China Morning Post

The South China Morning Post (SCMP) is a leading news media company in Hong Kong that has reported on China and Asia for more than a century with global impact. As well as an online news website, SCMP prints paper editions in Hong Kong.

Jagran

Jagran.com is the online version of the world's largest read Hindi daily newspaper - Dainik Jagran. They publish categories including Politics, National - International - Local news from Business, Auto and Technology segments. Jagran strives to provide meaningful content to consumers through data-driven journalism.

NZME — New Zealand Media and Entertainment

New Zealand Media and Entertainment (NZME) is a leading New Zealand media and entertainment business. The NZME brands include APN NZ, The Radio Network (TRN) and GrabOne. NZME has a portfolio of radio, digital, e-commerce and print brands that aims to produce excellent content for New Zealand audiences.

Mainichi 毎日新聞

The Mainichi Shimbun is one of the major newspapers in Japan, headquartered in Tokyo and published by The Mainichi Newspapers Co, Ltd. Mainichi also operates an English language news website called The Mainichi.

TheStar

Launched in 1995, as Malaysia's first news website, The Star Online has always strived to provide readers with up to date breaking news and compelling content. Their publication includes current news, business, sports, community, tech, lifestyle and world news, as well as expert analysis and dynamic videos.

聯合線上 udn.com

UDN (United Daily News) is a newspaper published in Taiwan, headquartered in Taipei, in Mandarin. UDN.com is its digital arm and its ultimate goal is to become the leading content service platform in the global Chinese Market for a Mandarin speaking audience.

Background

Working With Technologists

The technologist's role within the Data Lab program was to deliver a technical solution to each use case given the strategic direction derived from MightyHive's consulting of the publisher. Once publishers were scoped and onboarded, technologists were briefed on the use cases and worked with MightyHive to provide a workflow timeline, solution design, Proof of Concepts and Minimum Viable Products for use case execution. These different stages were to validate if the solution aligned the initial thought process, discover early problems, and gather feedback from the team. The selected technologists were:

- 55 the data company
- Datalicious
- eAgency
- MightyHive
- Tatvic Analytics

Google News Initiative

17

Background

Change Management - A Key Component

To maximize adoption of the new dashboard and measurement capabilities created by technologists, MightyHive conducted cross-functional workshops with each publisher. During these workshops, the new tools and the strategies behind them were introduced, practical activities were run to give participants the opportunity to use them, and co-design activities were completed to articulate how and when the new tools could be utilized in existing day to day workflows and processes. Recognizing that the success of any transformational program lies in the buy in and advocacy of influential individuals within the organization. MightyHive focused its time on creating these advocates who would be accountable for driving long term change and supporting them to form cross-functional teams to enable them to work together to achieve the quick wins necessary to persuade each publisher to roll out the tools at scale. These workshops were critical in providing teams with insight not only into "*what*" was created but "*how*" and "*why*", ensuring that all participants could become future "data champions" with the knowledge and skills to drive change long after the completion of the Data Lab themselves.

Use Case 1: Content Planning

Use Case 1: Content Planning

The findings of the content planning chapter were supported by the following technologists:

55 the data company, Datalicious, MightyHive, and Tatvic.

Business Problem

Most publisher newsroom's strive for a sustainable business model developed through the growth of its audience to drive revenue. The challenge for many in successfully executing this is due to a reliance, commonly by the editorial team, on performance metrics that lead to misinformed decision making. Focusing on baseline metrics such as "monthly users" or "article page views" to determine the success of an article does not provide the appropriate insight into how an engaged audience is responding to the publisher's content. Re-evaluating these metrics to focus and make decisions based on the performance of the article through an engaged audience, allows for a directed content strategy based on the response and reactions of this engaged and loyal audience.

Strategic Approach And Goals

Content Planning, the art and science of choosing what to write and how to write it, when and where to publish, and how to distribute and promote it, is the means by which a valuable, engaged and informed audience is created. Such a task is not easy in an era where the "when, where and how" of audience content consumption and the business model underpinning audience monetization has changed radically. Going back to basics and questioning key assumptions about the attributes of value and how to create it, is required.

Google News Initiative Data Lab 2020
Use Case 1: Content Planning

Business Problem | **Strategic Approach And Goals** | High Level Solution | What To Avoid | Prerequisites For Execution | Managing Change | Results And Impact

In today's competitive and constantly changing media environment, no newsroom can afford to take its audience or the value of that audience for granted. The objective of the Deloitte content planning use case is to create editorial analytics solutions that enable newsrooms to quickly, simply and frequently evaluate the performance of content in order to maximize engagement from audiences of greatest value. A robust data-informed content strategy, and the analytics tooling to support it, is therefore core to the data maturity of a publisher. Leading digital news organizations require bespoke editorial analytics, tailored to help them pursue particular goals, and align with specific editorial priorities and organizational strategy. These tools need to be oriented to inform both short and long term strategic development, and augmented with cultural change that drives greater familiarity to data within the newsroom.

The choice of metrics to evaluate performance is important, bearing in mind that industry standards like **pageviews and unique browsers or monthly active users were developed for advertising-based business model and are not necessarily the best mechanisms to support editorial decision-making on what and when to publish**, the best angle for a headline, ways to extend the lifetime and distributed reach of quality content or how to shape the habits of loyal and engaged audiences. For these tasks, editorial analytics solutions need to be created with the following best practice in mind.

1. **Report metrics which represent true measures of success.** A meaningful metric should correspond with a desired outcome or achievement of a goal. For example, whilst an article with high page views may, in the short term, appear to have driven a large monthly reach and an incremental increase in advertising revenue growth; when viewed holistically and through the lense of an engaged audience, the digital ad revenue was driven by the quality and characteristics of the audience and inventory available. Instead, reviewing a metric such as recirculation rate reflects user engagement with content and definitively shows a growth in engagement and value of an audience.

Google News Initiative Data Lab 2020
Use Case 1: Content Planning

Business Problem | **Strategic Approach And Goals** | High Level Solution | What To Avoid | Prerequisites For Execution | Managing Change | Results And Impact

2. **Make metrics actionable.** By showing metrics in context or in comparison to a benchmark, metrics become a source of action points. To act effectively, newsrooms must understand what the numbers, and subsequent insights, mean.

3. **Keep it simple and insightful. Less is more.** Journalists work under immense time pressure so fewer (meaningful) metrics mean quicker, better decision making at the point of need. Save everything else for periodic deep dives.

4. **Visible, accessible to and queryable by everyone, anytime.** Successful content strategies are not simply the purview of editors and management; decisions by journalists, producers, homepage editors and social media all contribute to the successful execution of a strategy over time. Interactive dashboards represent the sweet spot for ease of use and accessibility between static emailed reports and heavy weight analytics systems that require dedicated logins.

5. **Updated regularly.** Whilst real-time data isn't always possible or cost effective, there are many editorial analytics use cases that require fresh data. Yesterday's numbers are unable to help with the optimization or amplification of stories published today.

Aligning Metrics And Goals

Data is only valuable to an organization where it's used to help achieve goals. Ensuring a content planning strategy is utilized to its full capacity, an alignment across key metrics is necessary; whereby a clear understanding and joint narrative across all departments is established. For this to happen, the metrics chosen need to be both meaningful to and actionable by editorial teams.

Google News Initiative Data Lab 2020
Use Case 1: Content Planning

All publishers have a focus on revenue growth. A separation between editorial and commercial teams means, in many cases, editorial teams still express the goal of a newsroom is to "inform, educate and entertain" their audience. Within more commercially-oriented newsrooms, there too are many factors impacting revenue generation outside of the control of editorial teams (whereby responsibility sits with sales, ad operations or product and technology teams). To ensure data is meaningful and actionable, publishers need to choose metrics with a strong (causal) relationship to its business goal (revenue) and that editorial teams can influence (content).

The foundation for any successful publishing business model is a monetizable audience. Whereby, newsroom metrics must measure the success of its content to achieve its goal of creating *valuable, engaged and informed audiences* that can be effectively monetized, through the publisher's chosen revenue streams.

A Segmented View Of Audience

Macro analysis of basic content performance metrics does not allow for a true understanding of content engagement. To achieve a true understanding of how content is received, especially by a high-value audience, segmentation of the total audience by engagement (how often they visit and how much they read) and variables (to identify groups of most value to a publisher) is essential. Once created, these segments become key editorial metrics, for example completion rates by audience segmentation. This insight aligns editorial teams, newsroom goals and content success measures, as it enables evaluation of content on the basis of the audience quality (value).

Use Case 1: Content Planning

How To Determine A High-Value Audience

Recency, Frequency, Volume (RFV) is a standard method used for analyzing customer value in the retail sphere. More recently, it has been adapted by publishing business models to measure a reader's engagement and long-term value:

- **Recency** – when did they last visit?
- **Frequency** – how often do they visit?
- **Volume** – how many articles have they read?

By measuring these factors over a fixed period of time, it's possible to calculate the comparative engagement and value of different RFV audience cohorts. See Audience segments, below, for examples of engagement segments. Table 1, below, shows how the most engaged audiences contribute the lion's share of digital advertising revenues despite making up less than 5% of total audience. Brand Lovers, see below, make up on average 1% of total audience but contribute more than one-third of total ad revenue.

Example Audience Segments Used In The Data Lab

This approach to audience segments is inherited from a previous GNI project where they were originally established. Exact definitions may be localized, though segments should reflect outlined behavioral patterns;

- **Casual (Fly By):** Users who visited your site only once, with less than three pageviews in the last 30 days. This is to catch one off visits, from social media and news aggregator portals. This audience does not present much engagement, but it is high in quantity.

Use Case 1: Content Planning

- **Engaged:** Users who visit every few days although not necessarily evenly over a 30 day period. They are likely to read more than one piece of content per day. It is important to establish the maximum period of inactivity for these users, MightyHive proposes 2 weeks (and around 20 pageviews in the last 30 days). This audience may be divided into sub-audiences, if there is a meter implemented that forces users to register or pay. Where this occurs, one bucket should be defined as a group of people who may reach meter limits and those who registered.

- **Brand Lovers:** Highly engaged users who visit regularly and consume widely with few gaps between visits; to belong to this segment one would need to have more than 50 pageviews in the last 30 days

Once audience segments are defined and reported on, publishers are able to quickly identify which accrue the most revenue. Most commonly, publishers identify that a small percentage of readers drive the most revenue; with many, having a highly engaged and loyal (as defined by the rates of visitation and content consumption) audience that are likely to be paying subscribers.

As a result of audience segmentation, MightyHive was able to identify that across five publishers, the loyalty share (engaged and brand lover users) of each publisher's audience in each case represents under 5% of total monthly audience, but contributes 45-65% of all advertising revenue. By this measure **one brand lover is worth at least 50 casual users**.

Google News Initiative Data Lab 2020
Use Case 1: Content Planning

Business Problem | **Strategic Approach And Goals** | High Level Solution | What To Avoid | Prerequisites For Execution | Managing Change | Results And Impact

View Of Average Audience Segments Of 5 Publishers And Their Share Of Ad Revenue
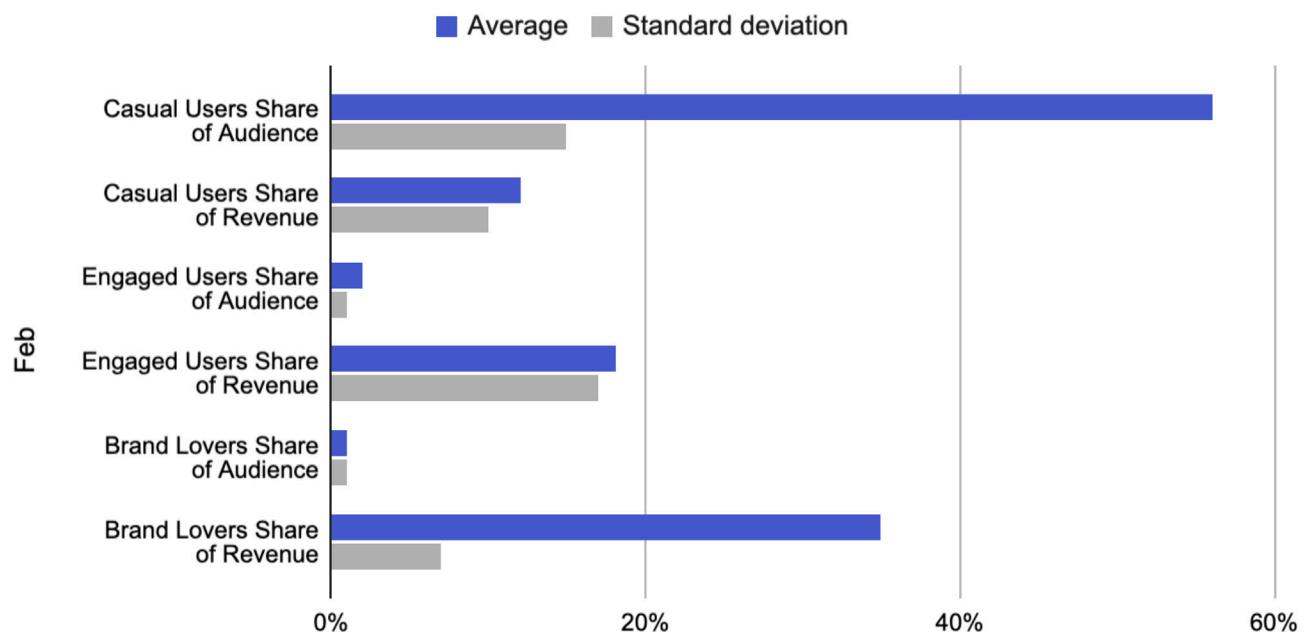
On average, the casual audience had a share of 56% (15% standard deviation) in the cohort, whereas this audience contributed only 12% of revenue (10% standard deviation); interestingly, for only one publisher this casual audience was delivering ~25% of the revenue. For all publishers, the engaged audience was in a realm of 0.5% to 3% achieving on average 18% of revenue (17% standard deviation). Only one publisher's engaged audience delivered around 40%. Whereas brand lovers numbers were between 0.5% to 1.65% share of voice with an average revenue share of 35% (7% standard deviation); with 3 publishers achieving shares around 40%.

Use Case 1: Content Planning

Once audience segmentation is established, analysis can go far deeper and become more intricate. More specifically, for mature data organizations, they can move away from manual rules-based audience segmentation which typically include subjective bias from their creator, and towards a data science audience segmentation model. The latter will help to reveal patterns of users characteristics and behaviors, so far gone unnoticed by the editorial team. This provides the publisher a detailed audience segmentation encompassing advanced characteristics and behavioral patterns that can be tracked and reported back on daily.

By monitoring key measures of engagement, such as recirculation rates for each segment, editorial teams can shape and evolve content and distribution strategies that increase user engagement and result in the growth of highly valuable readers.

Priority Editorial Metrics

Analytics is commonplace in most newsrooms. However, most editorial reporting relies on metrics to rate, rank and analyze news that weren't designed for the news and publishing industry. Additionally, the reporting designed by publisher data departments doesn't always appreciate the specific needs of the newsroom, i.e timeliness of data over completeness, meaningful labeling over technical detail or brevity over exhaustive data sets.

For deadline-driven, time poor journalists and editors, simplicity is key; the fewer the metrics, the more focus the editorial team will have. Establishing a few priority metrics that are meaningful and actionable is the first crucial step in building appropriate editorial analytics solutions for content planning purposes. The following priority metrics are foundational in forming a strong editorial analytics solution.

Use Case 1: Content Planning

Business Problem | **Strategic Approach And Goals** | High Level Solution | What To Avoid | Prerequisites For Execution | Managing Change | Results And Impact

How Can We Measure An Article's Success?

1. We measure what kind of audience the article attracted

Segment pageviews tell us how many article views are coming from priority segments.

2. We measure the reader satisfaction via the article completion rate

The **Completion Rate** is the proportion of readers who reach the bottom of the article

3. We measure the article's ability to generate another pageview using the recirculation rate

The **Recirculation Rate** is the proportion of readers who click through to another article

- **Priority Segment Users:** The number of users who meet the definition of a priority (highly engaged) audience segment. Targeting growth of the Brand Lover segment, say, represents a commercially meaningful goal for editorial teams which is under their direct influence.

- **Priority Segment Pageviews:** The number of pageviews generated by priority segment users (eg Brand Lovers). This metric enables teams to evaluate the performance of content based on the quality of its audience.

- **Article Recirculation Rate:** The proportion of an article's readers who go on to read another article. This metric enables teams to evaluate the performance of content based on its ability to increase its audience engagement and has a direct impact on growing priority segment users.

- **Article Completion Rate:** The proportion of readers who reach the bottom of the article. This measures how effectively the article engaged and satisfied its audience and newsroom efficiency.

Article-centric metrics provide us with one half of the required insight, the other is the size and habits of its reliable audience. To gain this insight, it is necessary to overlay the key metrics with audience segmentation; providing us with audience centric insights on content performance.

Use Case 1: Content Planning

The key audience (i.e primarily Brand Lovers) centric metrics that achieve growth

for publishers include:

- **Brand Lover Pageviews:** Number of pageviews for an article by Brand Lovers.

- **Brand Lover Recirculation:** The proportion of Brand Lovers who go on to read another article.

- **Brand Lover Completion Rate:** The proportion of Brand Lovers who reach the bottom of the article

- **Brand Lover Pages Per User:** The average number of pages viewed per Brand Lover over

   a defined period.

High Level Solution

Implementing a successful content planning strategy is achieved through a structured approach carried out by a cross functional team with defined mutual goals and KPIs. Upon determining key metrics and audience segmentation, the foundations have been laid and the strategy shifts to understanding how to execute the performance insights, the sharing of these insights and how a change in mindset will be established.

To reiterate the discussion undertaken earlier at "Strategic Approach and Goals", the foundation to any successful content planning strategy are the following,

1. **Introduce New Article And Content Strategy Performance Measures**

   Again, as explored earlier, defining the lens through which to view data is crucial. While the art of interpreting these metrics may be more nuanced, defining the metrics themselves should be straightforward. The pageviews, recirculation rate and completion rate of segmented audiences are the pillars of content planning strategy and should be focused on while executing this use case.

Use Case 1: Content Planning

2. **Segment Audience By Engagement**

   As explored earlier, once key performance metrics (i.e recirculation rate and completion rate) are established, a transition from a broad scope understanding of engagement insights across all readers to carving out defined audience segments that are weighted in value is necessary. Here, we see the creation of audience specific metrics that allow for cross audience comparison, for example comparing Brand Lovers recirculation rate to that of casual readers.

3. **Assign Segment Value Using Google Ad Manager Revenue Data Or LTV**

   Once audience segmentation is achieved, it's possible to go a step further with analysis by enhancing the segmentation with revenue data (found in Google Ads Manager or equivalent Ads management tools), thus creating revenue based audience segments. Viewing audience performance through a revenue lens highlights the commercial value of the audience. This is an achievable advanced enhancement of audience segmentation and helps drive a more sophisticated content planning strategy. This also enables the publisher to monitor the impact its initiatives have on the average LTV value of each audience segment: are your strategies effectively increasing the value of your key segments over time?

   Upon establishing a strong foundation of meaningful and actionable insights , a publisher can now develop a visualization tool to democratize the insights and provide actionable next steps for the relevant teams.

4. **Build New Dashboards**

   Once an understanding of what content is consumed by audiences is established, a dashboard allows for easy continued analysis. Visualising content performance is a scalable, unified and efficient solution to ensuring relevant business stakeholders can access and understand how content is performing; hence, a dashboard displaying this information is essential for providing

Use Case 1: Content Planning

Business Problem | Strategic Approach And Goals | **High Level Solution** | What To Avoid | Prerequisites For Execution | Managing Change | Results And Impact

actionable insights. The dashboards can be produced per desk, region or even author and should be supported by ongoing collaboration and reviewed with the data and analytics teams.

With such dashboards in place, publishers can begin iterating the content planning processes to utilize the new audience-based insight and reader behavior patterns to inform the content brief, format, channel, and long-term content calendar. Ultimately, the aim is to drive higher engagement from key audiences in order to grow the highest subscribing segments while increasing high value inventory generation.

Whilst the dashboard works as a visual tool, it must not be static; rather, be updated daily and be interactive (i.e application of custom filters and data drill downs). The ability to query data permits stakeholders to manipulate and isolate the data to understand the impact on or solve issues facing the content planning strategy.

The use of a dashboard provides publishers with a unified and visual understanding of content performance. Whilst it allows for tracking of audience segment growth, it also allows for,

- Comparison of engagement behaviors between segments
- Evaluation of article and topic performance by segment
- Identification of content to replicate, optimize, stop and amplify
- Aligning content scheduling and distribution with audience consumption peak
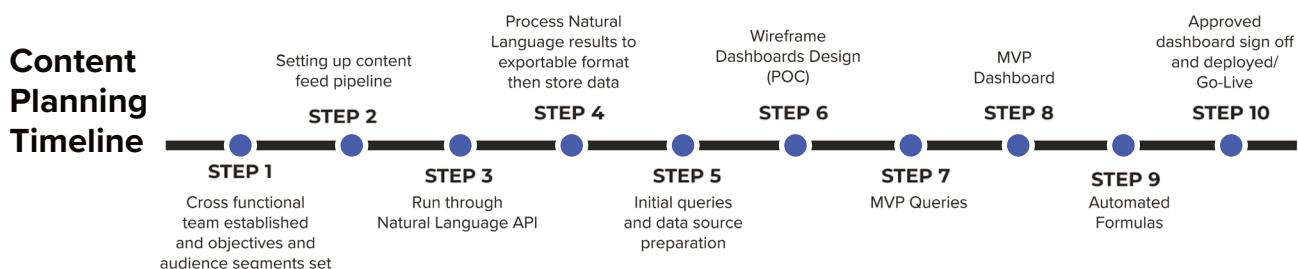
Content Planning Timeline

STEP 1 Cross functional team established and objectives and audience segments set	**STEP 2** Setting up content feed pipeline	**STEP 3** Run through Natural Language API
STEP 4 Process Natural Language results to exportable format then store data	**STEP 5** Initial queries and data source preparation	**STEP 6** Wireframe Dashboards Design (POC)
STEP 7 MVP Queries	**STEP 8** MVP Dashboard	**STEP 9** Automated Formulas
STEP 10 Approved dashboard sign off and deployed/ Go-Live		

Google News Initiative Data Lab 2020

Use Case 1: Content Planning

Technical Solution

The effective implementation of an editorial analytics solution for content planning, which utilizes a queryable dashboard to visualize data and insights, requires a number of technical prerequisites. It is crucial to have access to raw, hit level data from a reputable web analytics solution and a cloud data warehousing solution.
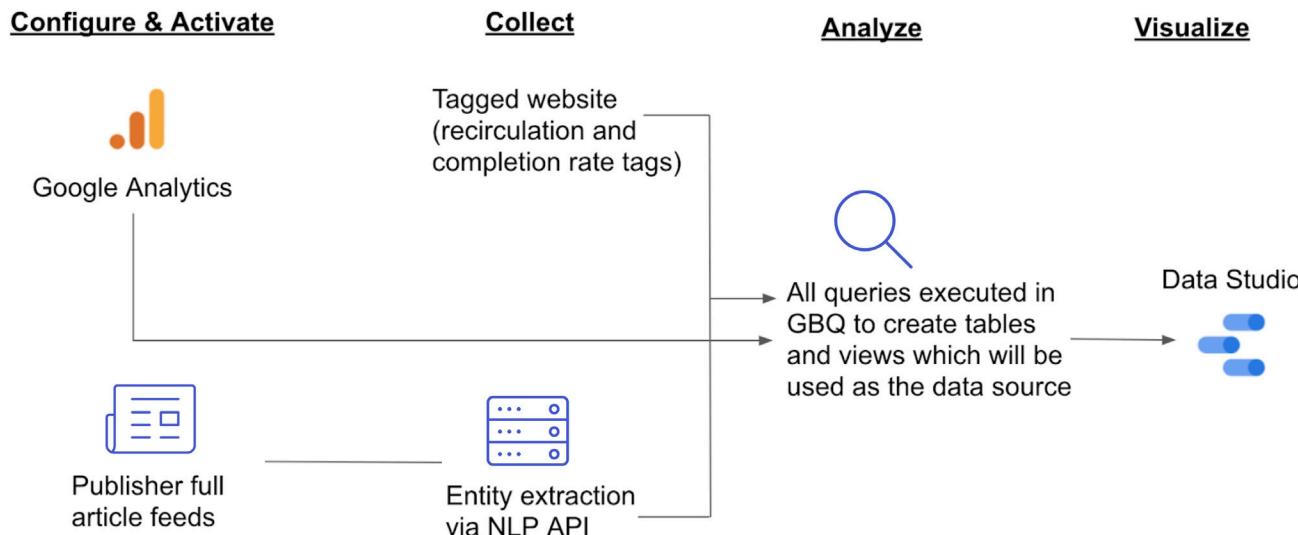
Tagging analysis should be completed upfront to assess feasibility and identify missing data points. Where missing data points are flagged through this process, it's possible to solve this issue through either;

- The implementation of specific Tag Management solutions that gather required information, or
- Joining article metadata (updated in a timely fashion through backend exports) to collect behavioral data in BigQuery.

Use Case 1: Content Planning

Business Problem | Strategic Approach And Goals | **High Level Solution** | What To Avoid | Prerequisites For Execution | Managing Change | Results And Impact

Dashboard Design

For a publisher to visualize, evaluate and help with editorial decisioning a dashboard, utilizing DataStudio was created. The intention of the dashboard is to outline current article performance, allow for comparison to previous periods and identify recommended changes to content to ensure user engagement is maintained (and grown).

Natural Language Processing

Due to slow, manual or outdated processes, many publishers struggle to effectively and reliably tag articles. Article tagging allows for content categorization and is another avenue to assess content performance (i.e by topic of popularity). To solve this, a solution utilizing Google Cloud Platform's (GCP) Natural Language Processing API (NLP API) was created (AWS Comprehend could be used as an alternative).

The NLP API offers an advanced and robust solution to article tagging. Machine learning is utilized to reveal the structure and meaning of an article's text and extract information about people, places, and events (also referred to as "entities"). This analysis of text was utilized across some publishers in the GNI cohort. To successfully execute this process, access to an article feed (either via API or RSS feeds where full article content (title and text) can be supplied) is necessary. The API reads these feeds to provide the extracted entities. There needs to be an article identifier available in the feed so extracted entities can be matched with data from other sources.

To successfully analyze and contextualize the extracted entities, a dashboard was developed to visualize this data in conjunction with the correlating article and its performance. This advanced data extraction and analysis method provides the editorial team with a deeper understanding of how specific entities perform.

Use Case 1: Content Planning

Business Problem | Strategic Approach And Goals | High Level Solution | **What To Avoid** | **Prerequisites For Execution** | Managing Change | Results And Impact

What To Avoid

When adopting new technology, systems or processes, publishers often experience a disconnect between the data and editorial teams. Often, data teams have explored or developed a sophisticated technical solution that doesn't align with the needs of an editorial team or comprehensive training hasn't been created to ensure adoption by all stakeholders. Inevitably, this leads to many solutions with low or incorrect adoption. To ensure adoption and continued use of the content planning dashboard, cooperation between the two teams is required.

It is important to prepare a deployment plan, i.e a pilot program, involving a small group of people. Due to new challenges (i.e around data processing or technology bugs etc), it usually takes three to four iterations to get to a solution usable at scale.

Unmonitored usage is another concern, as cloud solutions are billed in a manner of pay per use. It is important to understand how users are interacting with new tools to continuously optimize towards a reduction in billing costs. Through optimizations, it's possible to reduce dashboard usage costs by around 99% from the initial MVP versions.

Prerequisites For Execution

Editorial teams and newsrooms are the heart of any publishing business. The ability to create and sustain deep, habitual and valuable relationships with readers is the foundation for any successful publishing business model. For this reason, content planning is a use case that is broadly applicable to most publishers.

Google News Initiative   MIGHTYHIVE

Use Case 1: Content Planning

Business Problem | Strategic Approach And Goals | High Level Solution | What To Avoid | **Prerequisites For Execution** | Managing Change | Results And Impact

The assessment below has been designed to help determine whether content planning is appropriate for an organization. Agreeing with these statements is an indication the content planning use case for data investment may be of value to a publisher's newsroom.

In addition, the Data Maturity Scale and Diagnostic Tool developed by Google and Deloitte to help plot study participants on a defined spectrum can be used to assess an organization's data and analytics capabilities and determine how these compare to the competitive market.

Newsroom Assessment

☑ Our primary digital measure in the newsroom is monthly active users, total monthly unique users or total pageviews from all sources and we review audience data in aggregate e.g. total pageviews and unique visitors for the newsroom as a whole.

☑ Editorial meetings do not routinely incorporate data into story choices or other decision making.

☑ Audience segments and sources of traffic are not available within the editorial analytics solution most often used by the newsroom.

☑ The newsroom celebrates stories that "go viral" but rarely review stories that under perform.

☑ The newsroom is often torn between the desire to create quality journalism and the commercial imperative to deliver a big audience number.

☑ Not all newsroom staff have access to article-level performance data.

☑ It is hard to quantify to what degree a reporter or given desk has performed better over the last year in regard to audience development. It is rare for managers to compare numbers between authors or types of content.

☑ Many reporters don't know when their stories are published online and do not routinely check in to monitor their performance.

☑ Most of our stories follow a standard form: we don't often use different story forms to engage audiences in the content based on their needs and interests.

☑ Editorial teams do not have strong relationships with the data and analytics department, nor a regular cadence for performance deep dives to improve audience development.

Use Case 1: Content Planning

Technical Prerequisites

There are three primary technical prerequisites required to execute this use case. Without all three, this use case will not be suitable:

Off The Shelf Web Analytics Solution

Google Analytics 360 (GA360) was used throughout the Data Lab program, as such this playbook's solution utilizes the capabilities of this analytics platform. However, an interpretation using another analytics platform would also work (as long as there is access to raw data on the hit level associated with cookie IDs).

Data Warehouse

To successfully implement this content planning use case with GA360 data, Google BigQuery (GBQ) was utilized but any other database solution capable of working with Big Data would suit.

Dashboarding Solution

In this program, Google's Data Studio was the dashboarding tool of choice. Any dashboarding tool can be used as an alternative, provided it links with the data warehouse that is being used.

Required Skills

In order to successfully deliver the comprehensive solution discussed in this chapter, it's important to note that specific technical skills are required beyond the scope of most analytic teams. These include,

- An understanding of the web analytics solution and its limitations (sampling, approach to high cardinality etc), together with an ability to extract and work with raw data exports is needed.
- In terms of GA360, strong practitioners with GAIQ certification is a must.

Use Case 1: Content Planning

- Adopting solutions based on Google Cloud, especially around BigQuery requires analysts proficient with SQL, which is commonly used across various platforms. Currently GBQ is using a "standard" ANSI SQL, which means that individuals working with different databases will find it easy to adapt to using GBQ. It is important to ensure that all people included into the project are familiar with the GBQ pricing model, as it has a significant impact on implementation approaches. Automatic data transfers and scheduled queries remove ETL type of programming from the process.

- To build content feeds (for entity extraction), which may come at later stages of the project, typical software development skills are required. In order to use NLP API, the Google technology stack provides SDKs in many common programming languages, so developers have various options.

- As more complex data models are likely to be developed here, the above average knowledge of Google Data Studio (or any other desired dashboarding solution) will be required.

Managing Change

Many data initiatives fail due to a lack of engagement between the relevant stakeholders and departments. Adoption of new data tools, changes to existing processes and workflows starts with an understanding of what data is used day-to-day. However, in cases where new capabilities are being driven by centralized data teams, it is rare to find domain experts who also have a background in or detailed understanding of what it takes to run a newsroom. Whilst, editorial leaders, who understand these needs and context, rarely have the technical background to anticipate common issues with attribution, data matching, timely processing, and cost which can ruin any chance of success.

To solve for this misalignment and establish a solid foundation for adoption, early consultation between data teams and editorial leaders to share context and agree on the business problem to be solved is imperative to an initiative's success. Aligning on success metrics with the editorial team from the very

Google News Initiative Data Lab 2020
Use Case 1: Content Planning

Business Problem | Strategic Approach And Goals | High Level Solution | What To Avoid | Prerequisites For Execution | **Managing Change** | Results And Impact

start will be key, as well as getting their green light to perform several A/B tests to evaluate precisely how useful these new insights can be, once the dashboard is ready.

Data And Editorial Teams

Newsrooms are often fast paced, high pressure environments and there is little spare time outside of writing and publishing stories. While data teams may be looking for ways to make more data-driven decisions there can often be an oversaturation of dashboards or misaligned solutions that require high resource investments from the editorial team and in turn lead to low adoption rates.

Bearing in mind that data is key to editorial success and the growth of sustainable audiences, a strong relationship between these teams is valuable. To avoid a disconnect between these teams and alleviate foundational hurdles to the execution of the content planning strategy, it's important both teams view data through the same lens, i.e ensuring there are mutual definitions, objectives and KPIs. Coupled with this is a need for accountability through roles and responsibilities for each team in regard to use of the content planning strategy. Sharing tools and explanations through training and contextual background allows for and encourages adoption, as team motivation can develop by brainstorming and exploring possible outcomes from the new knowledge, skill set or solution provided.

Once both teams are on-boarded to the new content planning strategy, frequent communication is necessary to drive value and adoption. This can be achieved through data and individual desk biweekly meetings; joint team syncs to discuss interesting trends, findings, quirks, feedback or feature requests. Additionally, the dashboard insights should be leveraged during regular editorial meetings like morning conferences. These small tweaks to an existing workflow can begin to change the culture of content planning within a publisher organization. Whilst not essential, another way of assisting with adoption is assigning a dedicated data resource to the editorial team in order to moderate the usage of the dashboarding tools the newsroom is being asked by the data team to leverage.

Use Case 1: Content Planning

Business Problem | Strategic Approach And Goals | High Level Solution | What To Avoid | Prerequisites For Execution |
Managing Change | Results And Impact

The types of editorial decisions that should be made with the support of data (supplied by the data team) include:

Commissioning and creating

- › What to write
- › How much to write
- › How to write it
- › Whether to follow up

- **Improving performance**
  - › When to publish
  - › Where and how to promote it
  - › When to update or amend

- **Content strategy**
  - › How to improve audience engagement
  - › How to create audiences of value

Optimizations And Experiments

Dashboards with valuable data become crucial tools when immediate action points are clear for the editorial team. Whilst dashboards identifying insights are essential, amplifying this data to include actionable insights allows for opportunities to optimize and experiment.

There are various optimizations that can be made to a story with an average KPI performance, these include:

- **Recirculation**
  - › The stories that were linked
  - › The method they were linked e.g. Embeds or a Recommended module
  - › Identifying the strongest stories on the topic

- **Completion rate**
  - › The length of the story
  - › The story structure
  - › When the story was published/ promoted

- **Pageviews**
  - › How the article compare to other

Google News Initiative   MIGHTYHIVE

Use Case 1: Content Planning

Business Problem | Strategic Approach And Goals | High Level Solution | What To Avoid | Prerequisites For Execution | **Managing Change** | Results And Impact

While data can reveal what happened it doesn't reveal why, experiments and tests derived from analysis of data allows us to explore the effectiveness of current content strategies. By hypothesising reasons for particular results, experiments (i.e small or large tweaks to content as outlined above) can be conducted to understand what grows engagement rates and audience loyalty.

A recommended framework that allows for simple data actioning decisions is ROSA - Replicate, Optimize, Stop, Amplify.

Replicate

Identify stories achieving a high brand lover pageview count and above average completion rates. Consider the brand lover share and make plans to commission follow ups or create other stories that share similar characteristics.

Optimize

Identify stories that have a high brand lover pageview count but lower than average completion rate or lower than average recirculation rates. These articles have potential to drive greater value for the publisher if improved.

Stop

Identify stories that have low pageviews and low completion rates. These stories did not reach a material audience of any kind and did not sustain the attention of those who did read them. Set up watch lists for these types of articles. If the pattern is recurring, make plans to stop creating this kind of content.

Use Case 1: Content Planning

Business Problem | Strategic Approach And Goals | High Level Solution | What To Avoid | Prerequisites For Execution | **Managing Change** | Results And Impact

Amplify

Identify stories that have low brand lover pageviews and above average completion rates or recirculation rates. These stories sustain interest and drive article views but are not reaching a large audience. Consider promoting these stories in email newsletters, or on the homepage.

Once an editorial team member has evaluated an article's performance data, the ROSA framework provides a guide to identify next step actions to ensure the article's potential for the publisher's KPIs, reach and audience growth is maximized.

Example: Rosa Framework In Action

The performance data dashboards could reveal a celebrity COVID-19 story had high brand lover pageviews and high recirculation rates. Leaving the editorial team member with the following hypothesis: *"If we publish more celebrity COVID stories, we can grow total lifestyle pages per user and recirculation rates while maintaining overall lifestyle users, because these stories are a good source of quality external audience that recirculates effectively"*. This hypothesis, and subsequent experiment, would fall under the 'Replicate' stage of the framework.

Success for this experiment could be defined as any or all of the following metric changes:

1. Number of celebrity COVID-19 stories in the top stories table for the month
2. Number of celebrity COVID-19 stories in the AUDIENCE ACQUISITION BY ARTICLE table for the month
3. External audience recirculation rates
4. Lifestyle pages per user, recirculation rate

Use Case 1: Content Planning

Business Problem | Strategic Approach And Goals | High Level Solution | What To Avoid | Prerequisites For Execution | **Managing Change** | **Results And Impact**

Publishers that utilize data to formulate a hypothesis, experiment with various optimizations and analyze before and after results develop a mature and agile content planning strategy. Documenting these experiments allows for data to ingrain itself as an integral part in learning, driving change and assisting in decision making for the publisher. It's value is demonstrated through results, which in turn drives adoption of a strategy employing it.

Results And Impact

Please note, due to COVID-19 lockdowns being in place when the final delivery of this project occurred, audience behavioral patterns were affected; in some instances, to the level that impact analysis of the content planning strategy and dashboard was impossible. Therefore limited quantitative results are surfaced in this playbook.

As the Content Planning use case was executed for multiple publishers during the GNI Data Lab program, many nuanced lessons from different newsrooms were noted:

- A simple approach to segmentation and exposure to new metrics was very welcomed across publishers.
- Content Planning dashboards adoption continued to increase post workshops (across users and reports viewed per user in period of time). In the days after the workshop one publisher shared *"I have now heard the term 'brand lovers' being discussed in the newsroom in relation to story decisions at least 6 times in the last hour by people who were at the workshop."*
- Stress on actionability and the development of a "how to" framework to make decisions based on presented numbers is a new standard expectation across newsrooms (i.e a guide of how to action the figures in the dashboard).

Use Case 1: Content Planning

Business Problem | Strategic Approach And Goals | High Level Solution | What To Avoid | Prerequisites For Execution | Managing Change | **Results And Impact**

- It is observed that more efforts need to be put in place in order to create a culture of experimentation and testing. Publishers praised the use of the ROSA framework, post workshops because *"content creators own their artwork very dearly. Data helps neutralize emotions and help content teams become objective about the efficacy of content."*

- The program exposed many challenges related to how data is currently structured and managed across organizations. Here a challenge arose regarding many available solutions that can't be easily deployed and adopted; creating a low willingness by the publisher to address these issues. A publisher expressed that one of its data related vendors *"was unable to assure dashboard stability and data consistency. This further disengaged content teams and led to their questioning both availability and reliability of the little we did measure."* As such, the deployment of a content planning dashboard created a need to address these challenges (which had a negative impact on adoption short term, but it solved for these issues in the long term). The same publisher explained its *"content teams were focussing on efficiency of content production versus optimizing the effectiveness of content. With the GNI Data Lab's help, we see ourselves improving both short and long term visibility of our content's success in these result areas."*

- For some publishers it was observed the share of brand lovers increased between 20% to 40% (3%-5% points) on a weekly basis (e.g from 12% at the beginning of May, to 17% in early days of June).

- Thanks to Natural Language Processing, one publisher was able to analyze consumption patterns around COVID-19 related articles. These uncovered insights that drove a decrease in writing COVID-19 related articles and helped to steer efforts to maintain engagement in other categories.

Due to encouraging preliminary results, deployment of this use case should occur across many more publishers globally in a second edition of the Data Lab. Considering market dynamics and publishers exploring subscriptions models, content planning will play a crucial role in acquisition and, later, retention. Hence as Data Lab 2.0 is being discussed, content planning use cases will play a significant role in the next edition of the program.

Use Case 2: Recirculation

Google News Initiative Data Lab 2020
Use Case 2: Recirculation

Business Problem | **Strategic Approach And Goals** | High Level Solution | Prerequisites For Execution | A/B Testing Solution | Content Feed | Managing Change | Results And Impacts

The findings of the content planning chapter were supported by the following technologist:

datalicious

Business Problem

A publisher's ability to excel in audience growth is achieved where relevant content recommendations are provided for its reader, driving a successful and personalized continued onsite reading experience. The challenge for many publishers is understanding exactly what content audiences enjoy. Successfully utilizing audience insights, content planning and content classification with an appetite and culture for experimentation and testing is necessary for publishers wanting to adopt a mature, automated and bespoke recirculation strategy that drives audience growth.

Strategic Approach And Goals

The recirculation use case aims to increase reader engagement by serving relevant content recommendations to readers; increasing the likelihood of further on site and article engagement. For many publishers, recirculation strategies currently utilize solutions that link similar articles (either manually or in an automated way) to the piece of content being consumed. As automation and technical processes become more sophisticated, it seems natural the approach to this should be reviewed.

Recirculation algorithms that decide "the next best article" to present a user are often referred to as a "Content Recommendation Engine" (CRE). Historically, CRE's have been optimized solely toward the maximization of click through rates and advertising dollars.

Google News Initiative Data Lab 2020
Use Case 2: Recirculation

Business Problem | **Strategic Approach And Goals** | High Level Solution | Prerequisites For Execution |
A/B Testing Solution | Content Feed | Managing Change | Results And Impacts

When correctly configured, content recommendations play an important role in developing a publisher's business model; with the benefits including:

☑ Exposing the right content to the right user at the right time, increasing their likelihood to subscribe

☑ Increasing user satisfaction

☑ Reducing subscriber churn

☑ Better click through rate for low touch audiences

☑ Helps publishers understand the relevance between particular content and specific audience segments, and helps expose users to an 'optimal' mix of content to improve perceived value.

Unfortunately, these benefits are not guaranteed when using a generic or "plug and play" recirculation solution. Many of these "plug and play" solutions diminish the reader's experience and a publisher's audience retention and growth strategy, by sending readers to external websites or recommending content that doesn't correctly reflect the readers interests. From a revenue perspective, publishers may also suffer a loss in ad revenue due to ad slots being filled by the third-party solution instead of digital advertising opportunities. Further to this, many CREs that integrate with Content Management Systems, or other publisher related solutions, are often difficult or expensive to customize.

With a lower initial barrier to entry (i.e simpler implementation or lower price) a third-party generic solution can appear more attractive for a publisher wanting a quick solution. Here, publishers often discover a longer term impact to its business model with lower revenue generation and audience growth. As such, in order to gain control of the most appropriate recirculation strategy, that provides a long term solution of strong strategic value, a custom solution is recommended.

For a bespoke CRE to demonstrate its performance capabilities, say in comparison to a competitor generic or simpler solution, it's important publishers understand that optimizations and testing of the solution, whilst an ongoing exercise, allows for stronger long term success. It is advisable to manage

Google News Initiative Data Lab 2020
Use Case 2: Recirculation

Business Problem | **Strategic Approach And Goals** | High Level Solution | Prerequisites For Execution |
A/B Testing Solution | Content Feed | Managing Change | Results And Impacts

expectations around immediate commercial benefits that will flow from this solution. The beauty of a bespoke CRE, is the control a publisher has over achieving its goals. Custom algorithms achieve this control and success as an algorithm is generated to utilize only the publisher's data to drive results.

A successful model for any CRE is the adoption of a "recommendations model", as it allows for the development of an algorithm based on two key features: content similarity and user behavior. This algorithm can be further enhanced by including local conditions based on the publisher's local data, hereby providing a sophisticated model ingesting relevant data to ensure relevant content recommendations are provided. Here, recommendations are based on the semantic content of the articles extracted using Natural Language Processing (NLP) and the behavioral data of users.

This combination of content-based and collaborative filtering, taking both content similarity and reader signals into account, was determined as an effective solution to maximize the recirculation rate.

Article links produced by CRE widgets on publisher sites (to encourage recirculation) that accompany content are designed to speak to each reader on a personalized level. The links are one of many possible touchpoints in a deeper engagement journey. A CRE can then be used to push to other touchpoints, such as mobile app notifications, email newsletters, etc. These interactions all contribute to moving users down the funnel to subscribe, this makes it a valuable strategic asset as well as an important part of keeping readers engaged on a session by session basis. Furthermore, once publishers have established the CRE using NLP, other bespoke solutions can be built off these initial foundations such as automatic content tagging for consistent classification or contextual targeting for ad and article alignment.

Aligning Metrics And Goals

To ensure recommendations are aligned with a publisher's business model and strategic goals, it's important to set clear goals for recirculation as this allows algorithms to optimize towards a desired outcome. For example, if the primary goal is subscriptions (over advertising revenue), companies driving subscriptions might prioritize content that most often converts free users to paid users. It is important to identify the competing interest and overall impact achieving a primary goal may have over a secondary KPI, for example if the algorithm is aiming to increase the conversion to subscription rate, it may be at the cost of the click-through rate (CTR) and the ad revenue that goes with it.

The recirculation use case requires alignment across multiple teams, including editorial, product and sales with a unified position on reader experience. When common objectives are achieved, stakeholders subsequently become closer to achieving individual team KPIs, for example, the product team's goal might be subscriber growth and through recirculation increasing subs/retention/session length, this will have an positive impact on the growth of revenue management and inventory yield teams (e.g price structure, forecasting, etc.). Discussions to enable alignment and collaboration from all stakeholders to understand desired individual team outcomes and determine mutual goals should occur in the early stages of planning, pre-implementation.

Where a flexible rule-based approach is adopted, specific goals are set per audience segment. At a minimum, the adoption of rules similar to the following should be utilized:

- **Subscribers and non-subscribers goal:** Increased session length for readers (articles read per session per user)
- **Subscriber retention goal:** Premium (locked) articles per session per user was set for subscribing segments in order to reflect the publisher's desire to bolster the subscription value proposition.

Use Case 2: Recirculation

Business Problem | **Strategic Approach And Goals** | High Level Solution | Prerequisites For Execution |
A/B Testing Solution | Content Feed | Managing Change | Results And Impacts

- **Non-subscriber conversion goal:** To increase the conversion rate of non-subscribing users with a high propensity to subscribe. This goal would override the articles per session goal for this audience segment.

To adequately address these additional goals, a two phase approach is necessary,

- **Phase 1 - Recirculating all audiences**
  Both non-subscribers and subscribers are shown recommendations optimized for recirculation (click volume). This was quantified by measuring the click through rate with an aim to increase the rate.

- **Phase 2 - Fine tuning for conversion acquisition**
  Introduce a recommendation shown only to a subset of non-subscribers with a "high propensity to subscribe" flag. These recommendations are optimized for conversion (to subscription). The goal for this phase was to increase in average daily conversions (subscription acquisition).

The value of a bespoke approach to recirculation is in its alignment with specific goals and the nuances in content appreciation of each publisher's audience. As such, it is wise to expect ongoing fine tuning, testing and experimentation to find the right balance of content for each audience.

Example experiments a publisher could run include:

- Increasing page views to understand if this would increase ad impressions and revenue derived from non subscribed users each day,
- Exposing more premium content to non subscribed users with a high propensity to subscribe, thereby increasing subscription conversions (while maintaining a balance of exposure between free and premium content to ensure the page views per session metric increases),

Use Case 2: Recirculation

- Exposing more premium content to subscribers to understand if it increases subscriber engagement and reduces sessions where only a single article is read,
- Increasing yield and revenue by ceasing third-party native ad systems filling available inventory.

Metrics To Reconsider

With the focus on how to use data to create better recommendations with higher click through rates, it's easy to forget that CREs are not only an ingester of data assets but also a creator. A common criticism of using data to evaluate story performance surrounds pageviews simply reflecting the effort made to promote and distribute the content, *"the best performing stories are always the ones that are well promoted on the homepage",* is a common refrain within newsrooms.

To overcome this "distribution bias", there are a number of key factors to be considered and utilized:

- Recommendation CTRs, utilized as a measure of success for article performance. The CTR represents an unbiased measure of the article's (or headline's) appeal because it's represented as a proportion of the users who are presented with the recommendation. This may be useful both as a way to handle objections from editorial staff and as a helpful means to uncover high performing stories with great potential to drive readership, should they be promoted more broadly.
- When reviewing the CTR of different recommended articles, viewability provides a more detailed understanding of a user's interest in the recommendation. This is particularly important in A/B testing, whereby viewability ensures tests are created equally and the results are being reported accurately. The content of the recommendation (headline text and/or image) is all that should be different within the test, to minimize variables as much as possible.

Use Case 2: Recirculation

Business Problem | **Strategic Approach And Goals** | **High Level Solution** | Prerequisites For Execution |

A/B Testing Solution | Content Feed | Managing Change | Results And Impacts

- Ad revenue also needs to be considered when evaluating the content recommendation engine's return on investment. With a bespoke recommendation engine, the uplift in direct ad revenue can be recognized through newly available ad slots. Whether the ad slots are occupied by external advertisers or promoting internal domains, the impact of replacing third-party placements should be immediately realized.

High Level Solution

Executing a recirculation strategy with a bespoke CRE is an investment in reader engagement, it requires a number of prerequisites and steps to implement including content classification, user path analysis and audience segmentation, recommendation models and delivery architecture.
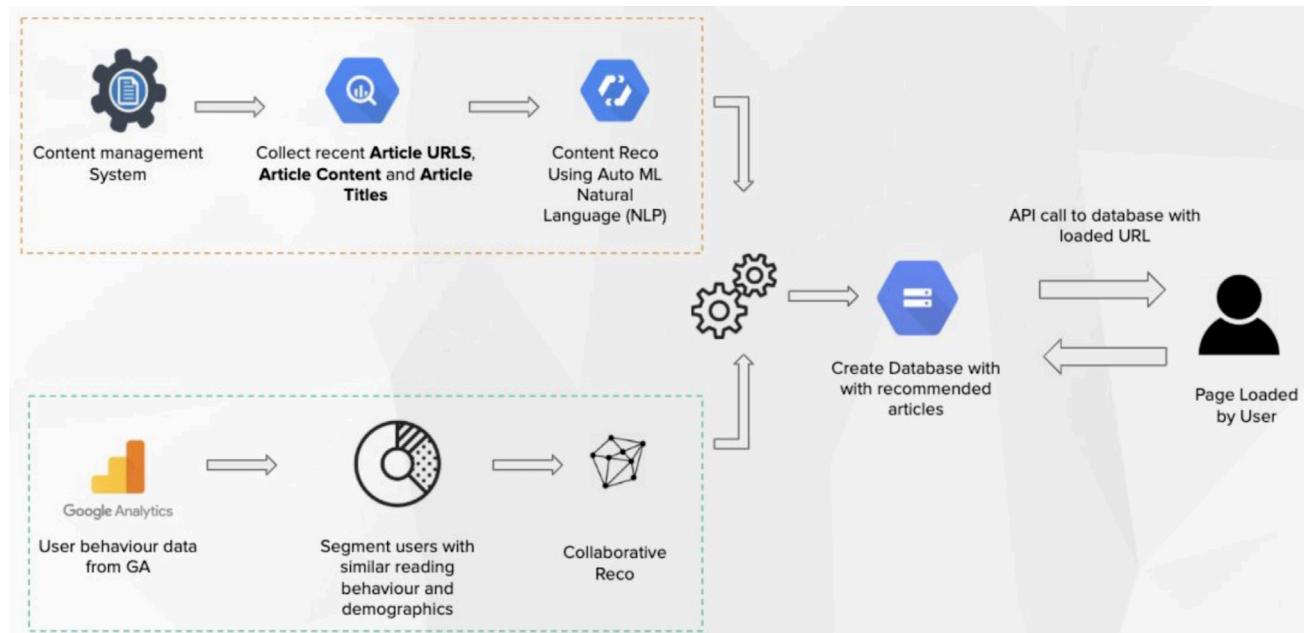

Image supplied by Datalicious

Use Case 2: Recirculation

1. **Content Classification**

   The process of content classification is essential for successful recommendations. It can be accomplished through manual tagging as part of the content creation process or via the more efficient and automated method of using NLP. A sophisticated recommendation model that relies on NLP to extract entities from raw article content allows for a semantically interpreted subject matter to determine similar articles and users with similar reading histories. Please note that some cleanup on results extracted through NLP may very likely be necessary as it may fail to identify the right context in ambiguous situations, for example, if the city Melbourne is mentioned in an article, is it supposed to be about Melbourne in Australia or Melbourne in Florida, USA?

2. **User Path Analysis And Audience Segmentation**

   Audience specific data, including user ID, article URL, topic, interaction timestamp and user type, should be utilized from a publisher's analytics platform (e.g GA360) and stored in a data warehouse platform (e.g BigQuery). Utilizing two-weeks worth of retrieved data allows for user segmentation. Whereby users that do not meet a prescribed article and session threshold (i.e are below the minimum or above the maximum) are removed. The data can then be queried to identify the strongest link between entities and behaviors of different types of users (e.g. location and subscription entitlements) to define the behavioral patterns to base article recommendations.

3. **Recommendations Models**

   As discussed earlier in the chapter, a recommendation model allows for a publisher to mature and take ownership of how and what content is suggested. A sophisticated recommendation system encompasses two elements: a content recommendation system - based on article and topic similarity and a user-based recommendation system - based on user behavior and reading history. Recent article URLs, content and titles are then collected to generate content recommendations using the auto machine learning NLP. Alongside this, Google Analytics user behavior data is

segmented by reading behavior and demographics. A "k means" algorithm is used to group users that behave similarly, allowing for the most popular articles to be served for each cluster; this is necessary for users with a low browsing history, which makes it difficult to build appropriate recommendations. From here, a rule based system can be used to define the conditions under which to recommend types of content to types of users in order to optimize recommendations for agreed outcomes.

4. **Delivery Architecture**

Recommendations must be delivered via a low-latency API linked to live content feeds and ensure they do not include content already viewed by the user. This ensures the user experience is not impacted and the content served to users maintained editorial standards. Further to this, there is a need for a low latency database that returns results quickly. HTTP endpoints to provide API responses can be deployed as cloud functions or a dedicated app engine instance (or in various ways). This use case depends on real time data in terms of responses but also needs to be near real time in terms of indexing new articles and behavioral data (e.g. to exclude articles which were already read by the user).

This architecture can be achieved utilising various Google Cloud Project functionalities, such as DataStore or BigTable. As different models may be applied from a machine learning perspective, different approaches may be considered. It is suggested to choose solutions that can be easily automated and run at scale (e.g. python based libraries).

Google News Initiative Data Lab 2020
Use Case 2: Recirculation

Business Problem | Strategic Approach And Goals | **High Level Solution** | Prerequisites For Execution | A/B Testing Solution | Content Feed | Managing Change | Results And Impacts

Recirculation Timeline
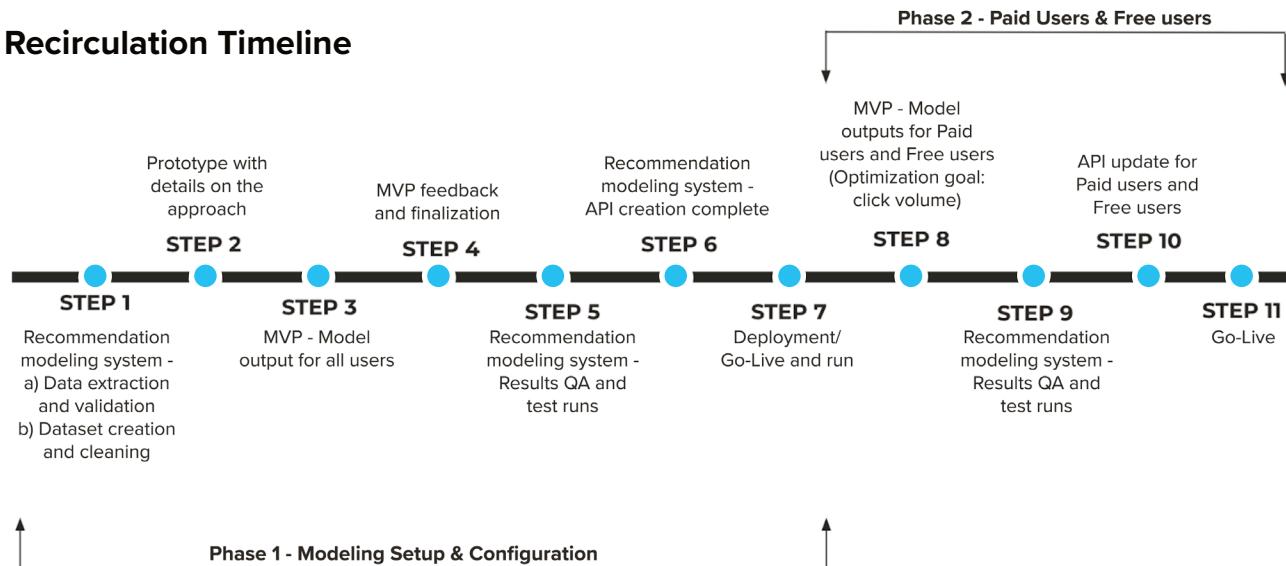
What To Avoid

There are a number of optimizations that can significantly improve the CTR on recommendations and reduce the overall number of errors when recommendations were not delivered at all. These include, but are not limited to:

Long Periods Between Updates To Content Feeds

Most publishers require a robust take down process for all content published; developing a mechanism to have specific stories removed from recommendations quickly, in case of legal breaches for example, is essential. Be mindful that feeds used for ingesting content may not be timely enough for certain scenarios, such as legal takedowns or contractual takedowns with third-party content providers.

Google News Initiative Data Lab 2020
Use Case 2: Recirculation

Business Problem | Strategic Approach And Goals | **High Level Solution** | Prerequisites For Execution | A/B Testing Solution | Content Feed | Managing Change | Results And Impacts

No Recommendations

If the user ID does not match any in the database and no recommendations can be pulled for that user, it's important to establish a rule for what article to display. This can be as simple as the most recent article published within the same category as the article currently being read. If a "fall back" recommendation rule is not established the recommendation engine will fail and a poor user experience will result.

Low Numbers Of Recommendations

If insufficient numbers of recommendations are available through the API, it is possible to "run out" of recommendations to show to a particular user. This can be avoided by monitoring and adjusting the algorithm to suit the patterns of consumption of the audience in question.

Bias In A/B Test Groups

When running A/B tests it's important to look at the audience attributes and ensure audiences are evenly split between test groups and control groups. e.g. Are the geographical locations evenly distributed between the two groups?

Google News Initiative Data Lab 2020
Use Case 2: Recirculation

Business Problem | Strategic Approach And Goals | High Level Solution | **Prerequisites For Execution** |
A/B Testing Solution | Content Feed | Managing Change | Results And Impacts

Prerequisites For Execution

Assessment

Agreeing with any or all of the statements below is an indication the case for change for recirculation is required in an organization:

- ☑ There are no processes supporting recirculation from an editorial perspective (e.g. tagging of articles).
- ☑ The value of the recirculation rate and its importance to reader engagement and function within the wider content planning strategy is either not understood nor valued by the editorial team and wider business.
- ☑ Native advertising tools, or 3rd party solutions, operating as a "blackbox", are in use to increase recirculation.
- ☑ A system is being used with an algorithm that does not take behavioral history into account to personalize recommendations.

In order to execute the recirculation use case described in this playbook, (including a CRE widget) a publisher will require the following:

A/B Testing Solution

If a publisher is implementing its first recirculation system there is nothing to test against, but should a bespoke CRE be developed to replace an existing solution, it is recommended the deployment of a new solution utilize A/B testing, with the existing solution being tracked alongside the new. This ensures the delta between the two can be monitored as optimizations are made to improve the new CRE. As a custom CRE requires an iterative approach to testing, it is important not to deploy 100% of the strategy straight away. The sample audience size allocated to the new solution can be gradually

increased as the CTR performance reaches parity with the existing solution. Once 100% deployment is reached the existing solution can be retired.

Content Feed

Full article content (text) needs to be pulled from freshly published articles into a feed from the Content Management System and passed to NLP to extract entities.

Database Of Historical Articles For Content Classification And Model Training

For NLP and entity extraction to be conducted and training models run, access to historical articles should be made available in a structured format (such as JSON). A data scientist or analyst should be consulted to determine how much historical data is appropriate to utilize; this could differ depending on the number of articles published within a period of time and traffic patterns.

Sophisticated segmentation

In order to deliver customized recommendations to audience segments, it's necessary for publishers to quickly recognize specific types of users (e.g. non-subscribing users with a high propensity to subscribe). A web analytics solution can be used in this scenario to automate this process through predictive analysis and audience segmentation.

Required Skills

In order to successfully deliver a comprehensive CRE solution, specific technical skills are required beyond the scope of most in house analytics teams. From a CRE perspective, besides advanced data analysis, additional factors need to be considered. Solutions need to work at scale and return recommendations in a fraction of second; meaning software engineering and networking capabilities are required for successful delivery.

Google News Initiative Data Lab 2020
Use Case 2: Recirculation

As with the content planning use case outlined earlier in this playbook, the following required skills are necessary:

- The recirculation use case needs to be resourced by personnel proficient with SQL for the Google Cloud and BigQuery (or other popular cloud environments such as AWS) solutions. Currently Google Big Query (GBQ) is using a "standard" ANSI SQL, meaning individuals working with different databases will find it easy to adapt to GBQ. As BigQuery does not provide low latency querying capabilities, it's likely other solutions will be used; therefore, experience with nosql databases may be handy, as this type of solution tends to provide the fastest response times (Google Datastore, Google BigTable).

- Google Cloud Platform offers various machine learning tools that assist in developing the required solutions without the need for heavy data science involvement (e.g the NLP API); this is unlike other activities that require solid data science skills to create feasible recommendation models commonly based on a customized code wrapped with additional business rules. Alternatively, Amazon's Comprehend tool could also be used for NLP.

- It is also important publishers have access to full stack developer resources for the development of the widget script and server application to receive and return content recommendation requests.

- Finally, ensuring those included in the planning and execution of the project are familiar with cloud pricing models is important, as it has a significant impact on implementation approaches. Machine learning solutions or infrastructure working at scale to provide timely responses may be expensive. It is advised to think about architecture upfront, as economics of scale may later make the project unprofitable.

Whilst some publishers may have an in house capability, many will use Google Partners that are analytics and cloud certified to assist in the creation and implementation of this strategy.

Managing Change

Achieving recirculation goals requires cross-functional collaboration

1 Our Goal	What are we trying to achieve?		Who sets targets? Who's accountable for them?
2 The Problem (Assumption)	What is preventing us from achieving our goal? What is the evidence to support this?		Who has the skills to analyze and define the problem?
3 Proposed Solution	How we can address the problem?		Who can deliver the proposed solutions?
4 Success Metrics	How will we know if we've solved the problem?		Who owns the experiment and analysis of results?
5 Documentation + Sharing	Who esle would benefit from this information?		How will we share insight to maximize impact?

Introducing new data capabilities in the form of a new recommendation engine represents not just a significant change to the user's experience of content, but new processes, roles and responsibilities within a publisher's data and analytics, product, sales and editorial teams. New responsibilities and workflows are required around managing the new product, it's ongoing development and maintenance and responsibilities for its results and revenue. The nature of the product and the skills required means cross team collaboration, of an unfamiliar nature within the publishing industry.

Clay Shirky's declaration, *"When the old model is broken, what will work in its place? Nothing will work, but everything might. Now is the time for experiments. Lots and lots of experiments,"* gives a true characterization of the task at hand; cross functional collaboration with strong alignment of shared goals is essential to run effective experiments that drive growth and the development of the CRE.

Use Case 2: Recirculation

Making Cross-Functional Working Work

Culture	Alignment	Empowered	Oversight
What's normal	Shared goals and priorities	Trust	Ongoing prioritisation, alignment and escalation
Encourage norms that support positive business outcomes	Align members across functions using shared goals, KPIs and incentives.	Trust teams with the resources they need to achieve their goals.	Establish portfolio governance to provide a clear view of all programs.

To maximize the value that can be derived from a CRE, it's vital for the cross functional team to develop a shared understanding of how recommendations are generated; to silo this information (e.g. to only the data or engineering team) would have a negative impact on the adoption and success of the project. To ensure responsibility and accountability is maintained, joint ownership of the CRE is recommended. Commonly, this would sit with product and data representatives. To ensure a broader range of commercial and editorial considerations are taken into account when goal setting, monetization and optimization, a cross functional working group is extremely valuable to act as a steering committee.

With a cross functional team, shared goals (KPIs for the CRE) and forums for collaboration in place the team should map out and identify required changes to existing A/B testing workflows and responsibilities. It is also valuable to consider the data assets created by the CRE, recommendation click through rates, or the relevance of content matched to a user's region or location. **These data assets can be used to derive insights which can be fed back into the newsroom and short and long term editorial decision making. During the Data Lab, these insights were used to identify content with potential to amplify to broader audiences, and to inform staffing decisions in regional areas.**

Use Case 2: Recirculation

Results And Impacts

As with the content planning use case, many nuanced learnings from different newsrooms were noted when implementing the recirculation use case:

- There are many benefits (and flexibility) to building a CRE from the ground up using first-party data, but running it at an optimal state requires investment from the necessary stakeholders. Roadblocks were encountered throughout the process of building and optimizing the CRE. With each barrier, an approach utilizing an iterative cycle of testing, review, improvement or elimination until the CTR reached parity and a full transition was made. This process should accompany a tight feedback loop to ensure alignment with business stakeholders.

- A full transition from an existing recommendation system to a custom CRE requires ample time to minimize volatility. An iterative development cycle should be followed with reported testing. Several rounds of changes and refinement should be planned for with appropriate time and team resources allocated.

- When it came to optimization efforts, a key takeaway was regarding the segmentation of results. By looking at the data differently e.g. devices (mobile vs desktop) or visitor (international vs domestic) break out, it's possible to expose individual areas of concern and success. Experimenting with these results is best practice to gain deeper insights on where the system could be improved.

Use Case 2: Recirculation

Business Problem | Strategic Approach And Goals | High Level Solution | Prerequisites For Execution | A/B Testing Solution | Content Feed | Managing Change | **Results And Impacts**

- Daily stand ups between the internal and external teams involved to execute the project allow for a more agile response to each optimization/change. Typically, optimization results were only considered accurate a few days after being published. When catch ups were restricted to a once per week cadence, progress was stifled. As such, stand ups should be used to:
    › discuss and review the performance (measuring actual results against expectations);
    › hypothesis avenues to improve results; and,
    › determine ways to test and plan for the next iteration.

Notable results from a custom CRE implementation include:

- Parity between third-party recommendation engine and bespoke solution CTR was achieved.
- For desktop devices, the new solution delivers more clicks than previously used systems. Further enhancements are planned for the new solution, driven by confidence that mobile traffic will also outperform previous systems.
- The performance (CTR) of the GNI Data Lab widget improved as the user pool increased to meet that of the third-party recommendation engine. This performance improvement demonstrates how the bespoke solution can be scaled.
- Due to changes in the supply side platform utilized, publishers were able to secure additional ad slots that were successfully monetized, providing enough revenue to create positive ROI.

Recirculation as a use case was definitely the most challenging exercise from a technical point of view, engaging the biggest variety of tools and techniques. It also required the most extensive iterations in order to achieve expected results. As machine learning solutions become commoditized, it is foreseen this kind of solution will become a more popular solution in the publishing market and the increased level of sophistication expected.

Use Case 3:
Inventory Yield Management

Use Case 3: Inventory Yield Management

Business Problem | **Strategic Approach And Goals** | High Level Solution | Prerequisites For Execution | Managing Change | Results And Impacts

The findings of the content planning chapter were supported by the following technologist:

MIGHTYHIVE

Business Problem

One of the biggest challenges faced by publishers today is the limited, often non-streamlined, approach towards managing digital advertising revenue. Traditionally, publishers operate a more siloed approach to business growth whereby each department is in charge of achieving a singular goal (increase sales, successful execution of a client's advertising campaign or an increase in content). Whilst this method provides a short-term solution, it isn't a transparent long term solution to achieve growth and maturity. For publishers to drive results, there must be a focus on cross-functional teams with a data-driven approach to commercializing inventory coupled with new processes, solutions and visibility.

Strategic Approach And Goals

The objective of the inventory yield management use case is to maximize advertising revenue for the available inventory on a publisher's website.

An effective framework to achieve higher yield is through the three pillars of optimization: performance, troubleshooting and forecasting. This framework addresses the thesis that an integrated approach to performance, troubleshooting and forecasting, via the utilization of a cross-functional team, will have a greater impact on overall yield. To ensure success, **it's important to develop a focus that unifies internal departments on achieving the same goal**, which, for example, may be improving revenue performance by increasing eCPMs without decreasing fill rates.

Google News Initiative   **MIGHTYHIVE**

Use Case 3: Inventory Yield Management

Business Problem | **Strategic Approach And Goals** | High Level Solution | Prerequisites For Execution | Managing Change | Results And Impacts

Each pillar of the yield management framework is built to provide a detailed and custom view for varying stakeholders - starting from the executive level, through to sales, product and ad operations teams. This structure works in tandem with the aim of driving the importance of cross-functional teams to support and direct optimization.

Let's focus on each pillar to understand why it is important and the value it provides to driving success for publishers.

Performance

Publishers need a holistic view of all media activity and an ability to easily identify significant changes. Stakeholders can utilize the performance management pillar as a framework to align parties with the same goals and identify processes to be improved, or for senior management to understand the effectiveness of different ad revenue streams or the progress of large client campaigns. Ideally, in the mid-term, possible changes can be identified that inform and introduce new processes that help achieve long term goals.

To improve yield performance, publishers strive to demonstrate how its high quality inventory will return a strong return on investment for advertisers. Therefore, discussions with advertisers whereby the inventory price is positioned against an actionable KPI, such as viewability, provides a tangible market differentiator for publishers. The relationship between viewability and pricing allows for highly viewable ad slots to be considered premium, as such, driving a higher eCPM. This also allows the actual impression viewability percentage on site to become an interest point for stakeholders beyond product and ad operations.

Use Case 3: Inventory Yield Management

Business Problem | **Strategic Approach And Goals** | High Level Solution | Prerequisites For Execution | Managing Change | Results And Impacts

From programmatic perspective, eCPM is function of 3 values:

1) **Viewability** - A metric that tracks which ad impressions can be seen by users and can be controlled by the product team through site design
2) **Cookie** - The presence of various types of cookies affects monetization, as every cookie has a different value. An AUNZ user visiting a webpage of an Indian site, will get AUNZ ads, as that AUNZ cookie will attract higher CPMs than local Indian advertisers
3) **Content** - As contextual targeting becomes more relevant in light of changing privacy regulations and cookie deprecation, certain publisher categories (finance) may increase in value over others (horoscopes)

Yield is a function of a few variables; the value of presented content (from a contextual targeting perspective), value of user behavior including viewability and CTR, browsing history (cookies), platform and the creative format (text, display or video). There is a need for publishers to understand the best way to handle this relationship, in particular from an organizational perspective, for success.

The operational responsibility for viewability largely falls with the ad operations and product teams, making it possible for this to be managed independently. This allows publishers to focus on the value of the cookie and the content quality. Managing content quality, requires involvement from the editorial teams and an understanding of what content drives higher eCPMs. Importantly, refreshing how an organization analyzes and understands these factors interrupts the long standing status quo of many publisher organizations by introducing editorial teams to the impact of content on revenue. As such, the value of the cookie becomes a link to cross-functional team work (especially around the alignment of content and acquisition strategies).

Use Case 3: Inventory Yield Management

Business Problem | **Strategic Approach And Goals** | High Level Solution | Prerequisites For Execution | Managing Change | Results And Impacts

The type of cookies publishers attract, and retain, may have a significant impact on eCPMs achieved on exchanges. Performance analysis, utilizing exchange insights, also allows for investigation and monitoring of high spending advertisers. If Google Adx's viewability is managed correctly and the inventory price reflects "quality", the case for increasing eCPMs is even stronger. This approach is also proven to be a good source of leads for business development, as sales teams can reach out to high spending advertisers to develop a partnership.

Troubleshooting

One of the easiest ways to increase revenue for many publishers in the Asia Pacfic region is through the swift addressing of technical challenges. Growing complexity and usage of various tools by publisher's makes identifying issues challenging. A recommended practice to help alleviate this issue is through tracking tools, allowing for issues to be immediately identified and swiftly rectified. Whilst tracking tools provide visibility into an issue or the source of it, the bigger challenge comes with how quickly the issue can be fixed. Timely fixes are often difficult to execute when a publisher has little to no established processes regarding the completion of tasks requiring significant input across departments.

A timely and thorough resolution of technical problems requires clear processes to be developed to ensure a workflow of accountability and solution is established. Two important actions to be included in this workflow are impact analysis and ability to reproduce the case. Impact analysis allows for a focus to be maintained on issues that have a significant impact on revenue; whilst the ability to reproduce the case allows for teams to explore and experiment to determine the best solution or develop practices to avoid the return of a similar issue. Establishing and codifying the most common technical issues faced by a publisher whilst also allowing cross-team functionalities (i.e accountability and assisting in developing and execution solutions) allows for publishers to future-proof the business against similar issues.

Google News Initiative Data Lab 2020
Use Case 3: Inventory Yield Management

Business Problem | **Strategic Approach And Goals** | High Level Solution | Prerequisites For Execution | Managing Change | Results And Impacts

Forecasting

The final aspect of effective yield management is the capability to predict market trends. Seasonal and cultural events should influence auction dynamics. Strategic changes such as adjustments to eCPM, packaging and ways to generate a sense of scarcity are some ways to capitalize on these calendar events. Forecasting should focus on managing pricing so that it is dynamic and following as many factors that could influence it as possible as static rates are not likely to reach yield potential. Whilst there are always going to be events without precedence that could not be predicted (e.g. COVID-19), there are overall market conditions that simple time series algorithms will help paint a picture of upcoming trends.

The goal of the forecasting pillar is to give editorial, sales and ad operations teams a chance to plan for or pre-empt changes before they happen and realize opportunities that otherwise could go unseen.

Aligning Metrics And Goals

Before yield management insights can be drawn, it's important for a publisher to prepare the business to establish or refine practices around the collection of meaningful and complete data. Actioning insights can only be effective where processes are established; meaning many existing workflows may need to be improved, updated or created. One workflow that commonly requires review is the data input for direct deals; it's a common practice for this data to be incomplete when input into a publisher's Ad Manager platform (especially in terms of pricing, which affects further reporting). To ensure a true data set is being utilized, complete direct deal data, at a minimum, is required.

Across the three aforementioned pillars (performance, troubleshooting and forecasting) it's necessary for core metrics to be defined with the available data to ensure there are no obstacles in reporting them. These core metrics include: impressions, fill rate, eCPM and revenue. These metrics

Google News Initiative Data Lab 2020
Use Case 3: Inventory Yield Management

Business Problem | **Strategic Approach And Goals** | High Level Solution | Prerequisites For Execution | Managing Change | Results And Impacts

are paramount in assessing a publishers current position and performance and assist in forecasting growth. Impressions and eCPM are imperative to understand the levels of traffic and the corresponding value of the traffic.

Other important metrics a publisher should define and utilize correctly include,

- **Ad Request eCPM:** Publishers often gauge the performance of their inventory by looking at eCPM, however eCPM doesn't give you a full picture as it only looks at the value derived from impressions (revenue / impressions) and doesn't take into consideration the opportunity cost of unfilled impressions. The ad request eCPM on the other hand takes into consideration all requests - both filled and unfilled, in the calculation of the metric (revenue / ad requests).

- **Viewability:** to understand the quality of a publisher's inventory to then develop appropriate pricing strategies.

- **Reach:** This metric is particularly useful for brand performance campaigns and helps publishers tell the story of how many unique users they are able to reach. More advanced

- **Revenue Percentile:** to understand the distribution of revenue and eCPMs across inventory

- **Unfilled Impressions:** for investigation into why inventory which is not being filled and therefore wasted

- **Share of house:** to monitor non-revenue driving inventory and allocation of this inventory to maximize potential for monetization

- **Ad Speed:** for a detailed understanding of user experience on site

In addition to the metrics mentioned above, pricing rules (e.g. floor pricing) must be monitored as fluctuations here may have a significant impact on revenue.

Use Case 3: Inventory Yield Management

Business Problem | **Strategic Approach And Goals** | **High Level Solution** | Prerequisites For Execution | Managing Change | Results And Impacts

Metrics To Reconsider

Some Asia-Pacific markets focus heavily on CTR as a metric of success and whilst this may provide an insight, not necessarily meaningful nor truly actionable, into content engagement, publishers do not have significant control over this; there are various factors that have an impact (usually in a negative way) on the movement of this metric. If CTR is an important metric for a publisher, perhaps due to advertiser pressure, it's important to build ad operation processes that audit and monitor CTR to help reduce its fluctuation. Example best practices here include, removal of poorly performing creatives or ad slots or ensuring CPMs do not dip to uncharacteristically low prices.

Across the publishing industry, there is an increasing trend towards the adoption of CPA models, whereby publisher revenue is based on advertiser conversions driven (often on 3rd party assets). This model must be carefully considered, as publishers have minimal control over measurement. The management of conversion collection tools (i.e pixels) on 3rd party assets is work intensive and is difficult to control the quality of implementation, which in most cases causes significant under reporting.

High Level Solution

To effectively manage and influence yield in the long term, utilizing a dashboard to visualize and analyze current performance and compare it with historical performance is imperative for success. In order to achieve a solution that drives continual improvement and involvement of cross functional teams, a number of technical prerequisites are required. It's essential a cloud data warehousing solution is available; this is necessary to successfully merge data from a publisher's ad manager platforms (e.g from Google Ad Manager) and on-site analytics data (e.g from Google Analytics) to create a holistic view.

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | **High Level Solution** | Prerequisites For Execution | Managing Change | Results And Impacts

If a publisher is using Google Ad Manager, its process and account set up should be audited and updated during the initial phase of project execution to ensure accessibility to all relevant data points and to address the collection of any missing inputs. In particular, addressing the process for direct deal activity input must be completed to ensure this data can be analyzed along-side programmatic activity; this complete picture is essential for managing yield.

The execution of the inventory yield management use case can broadly be broken down into the following steps:

1. **Consolidate Data**

    In order to ascertain a detailed and overarching understanding of a publisher's yield position, data from multiple sources must be integrated. It is important to undertake an extensive research and development stage to review all data available to a publisher and test varying processes of data blending. Successfully blending data from an ad manager platform, on site analytics data and, where available, Data Transfer Files is integral to the inventory yield management solution.

    All data should be housed in a data warehouse where it can be automated and advanced data transformation can be executed; allowing for all necessary data to then be used in a dashboard. At a minimum, publishers should ensure the following data is available for inventory yield management:
    - › Ad performance data (revenue + impressions + ad requests)
    - › Audience behavioral data
    - › First-party data
    - › Content variables
    - › Deal type
    - › Industry/client spend

Google News Initiative   MIGHTYHIVE

Use Case 3: Inventory Yield Management

Where a solution utilizing Google products is built, the GAM API allows for a publisher to access the majority of data integral to the development of an inventory yield management dashboard. The API provides an extended list of dimensions and metrics and the ability for customized datasets to be created on an ad hoc basis. BigQuery can be used for advanced data transformations. Whilst the GAM API allows for advanced and scheduled reports, one disadvantage is the inability to access live data. As such, dashboards need to be developed to permit for a 24 hour delay.

For publishers to gain deeper and more actionable insights (especially around distribution and nature of eCPM), Data Transfer files are essential. Event-level data is key in identifying high value users and the corresponding value across the publisher's website. Analysis from DT files allows publishers to understand what content is being consumed by high value users (those driving high eCPMs). Identifying this activity is a stepping stone for collaboration between the editorial and operations teams, leading to editorial pricing strategies focusing on maximizing yield by protecting high value content and inventory for premium advertising.

Further to this, user information, extracted from DT files, can be grouped into high value user percentiles, creating an opportunity for publishers to upload cookie encrypted audiences to GAM. This provides an opportunity for publishers to create and target new, viable first-party audiences that are used to create sophisticated rules to drive higher eCPMs.

Consolidation of SSPs is an important consideration for effectively managing yield. Disparity between data significantly impacts the ability to pull together a detailed and thorough understanding of a publisher's yield position. For publishers with many active SSPs, there are extensive and advanced technical requirements required to pull together all necessary data into a consolidated view and, please note in some instances, this consolidation is not possible. Therefore, it's recommended that where possible, SSPs are streamlined and in an ideal state, one SSP is utilized in order to manage all inventory.

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | **High Level Solution** | Prerequisites For Execution | Managing Change | Results And Impacts
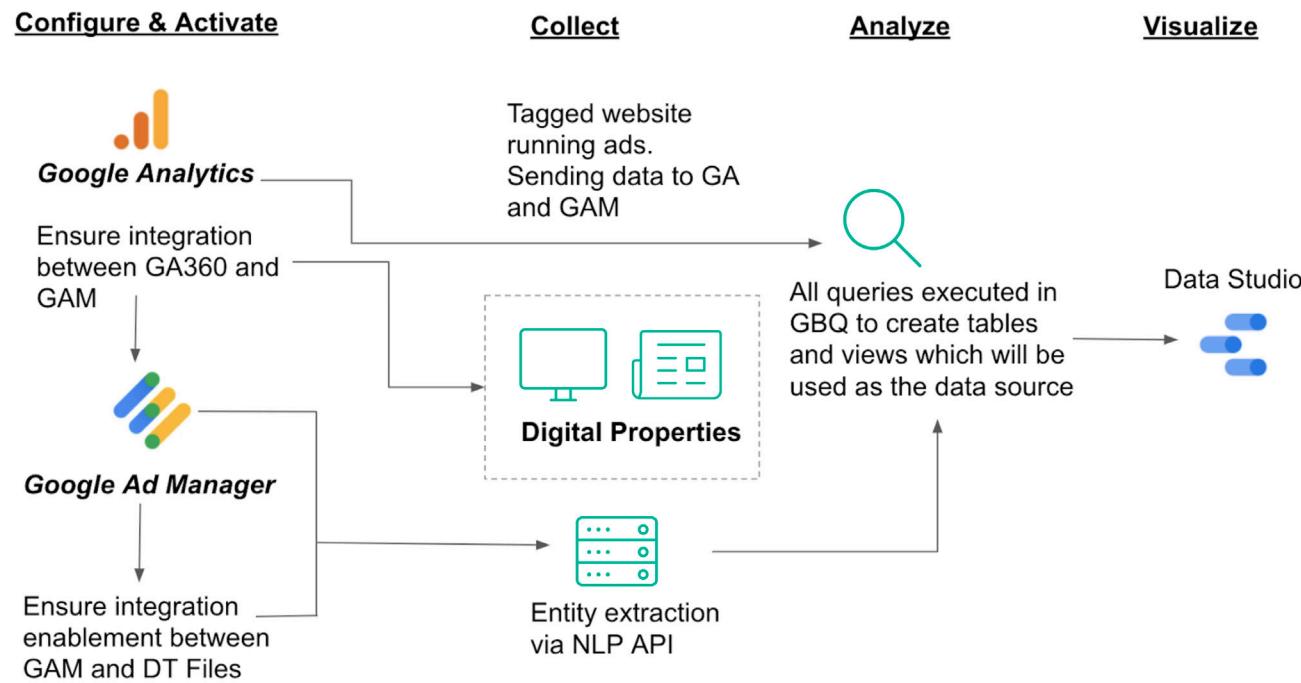
2. **Create New Dashboards**

   Implementing dashboards across the three strategic pillars of performance management, troubleshooting and forecasting is necessary to support cross-functional teamwork by reiterating the interdependence of successful yield management through stakeholder aligned goals. Whilst success metrics vary from publisher to publisher, core metrics of revenue, eCPM and viewability should always be included for analysis.

Performance Management Dashboard

Performance dashboards allow stakeholders to utilize performance insights to inform and align cross functional teams on shared goals, identify processes to be improved, or for senior management or sales, to understand the effectiveness of different ad revenue streams or the progress of large client campaigns.

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | **High Level Solution** | Prerequisites For Execution | Managing Change | Results And Impacts

Dashboard For: Executive Team

These dashboards are designed for the executive team to easily identify revenue streams split by advertiser and platform and also compare month on month and year on year performance. Further to this, these dashboards assist with identifying areas of growth and decline, performance against monthly revenue targets and analysis of the pricing structure as set by the sales team.

Example dashboard includes:

- Performance Management Summary
- Revenue By Direct Advertisers/ Ad Exchange Advertisers / Network Partners
- Revenue - Platform

Dashboard For: Sales Team

As the sales team is required to understand fluctuations in advertiser revenue in greater depth, performance management dashboards provide a granular insight that allow sales teams to identify new opportunities within the market. Further to this, the dashboards allow these teams to identify premium inventory and assess best performing content to be packaged into direct deals. Key metrics need to address the relationship between price, volume of impressions and achieved viewability. To ensure actionable insights, users should be presented with year on year comparisons to understand the fluctuations in advertiser spending. These insights can be used to understand if advertisers have increased or decreased spending with the advertiser, and where the later allow the sales team to explore methods to grow commercial partnerships or defend previous budgets spent by such advertisers.

Example dashboard includes:

- Sales Team Performance Management (this focuses on advertiser revenue month on month and year on year, as well as ad unit and order revenue breakdowns)

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | **High Level Solution** | Prerequisites For Execution | Managing Change | Results And Impacts

Dashboard For: Ad Operations Team And/Or Programmatic Team

Campaign management, monitoring performance and optimization at scale are key to the functioning of the in-platforms teams (meaning adops and programmatic). As such, the performance management dashboard for the operation team should focus on automation of reporting and providing granular data to quickly identify changes in performance to make actionable changes to mitigate under-delivery.

Example dashboard includes:

- Ad Operations Performance Management (this focuses on individual performance metrics like impressions, clicks, revenue, viewability,average eCPM for live ad units and CTR broken down by various dimensions)

Dashboard For: Cross Functional Teams

It is recommended that additional pages of the performance management dashboard be created to enable cross-functional team review; allowing for assessment and identification of content that is driving consumption by high value users and demand by high value advertisers. These dashboards act as conversation starters between the product, editorial, sales and ad operations teams to review varying approaches of maximising yield across high value content and inventory.

Example dashboard include:

- User CPM Percentile By Category
- Premium Inventory Analysis
- Viewability By Ad Unit By URL

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | **High Level Solution** | Prerequisites For Execution | Managing Change | Results And Impacts

Troubleshooting Dashboard

The troubleshooting dashboard allows relevant publisher stakeholders to easily identify issues, areas of concern or properties with low revenue. The focus here is enabling these teams to quickly influence and implement improvements. The troubleshooting dashboard is essential for identifying either product or technical errors in order to set up and quickly find solutions that mitigate the impact on revenue.

As such, a series of charts should be developed that surface key metrics and dimensions, which assist in drawing attention to any technical issues. Possible metrics include unfilled impressions, non monetized pageviews or perhaps a set of adspeed reports (currently in beta). The dashboard can also be used to identify abnormally low eCPMs compared to the rest of inventory and VAST errors for video publishers (particularly errors that lead to unfilled requests).

While some of these metrics are obtainable from GAM reporting it is worth calling out the nuanced benefit of having a troubleshooting dashboard over regular static reporting. Visual elements of the dashboard including prefiltered charts and graphs provide users with important information quickly enabling them to surface insights more efficiently than if it were in a tabular format. Many of the charts are also interactive with filtering, this facilitates easily drilling down on a specific dimension. Additionally, there are many metrics used in the dashboards that are extracted as a result of connecting different data sources (GA360, GAM API or DT files). These are not accessible purely through GAM reporting.

Business Problem | Strategic Approach And Goals | **High Level Solution** | Prerequisites For Execution | Managing Change | Results And Impacts

Dashboard For: Ad Operations Team And/Or Programmatic Team

These dashboards allow the ad operations team to quickly identify and rectify any errors made with its SSP set up. In addition, dashboards focusing on revenue percentiles assist the ad operations team to gain a deeper understanding of eCPM distribution across line items and allow for modifications or management of floor pricing rules.

Example dashboards include:

- Troubleshooting Low CPM Ad Units
- Troubleshooting High CPM Ad Units
- Troubleshooting Direct Deal CPM
- Troubleshooting on Pricing Rules
- Troubleshooting Revenue Percentile
- Troubleshooting VAST Errors

Dashboard For: Product Team

To ensure the growth of a publisher's offering, it's essential for product teams to access data that identifies possible changes necessary for site performance, for example the removal or alteration of ad slots that are causing negative user experience or indicate potential technical issues on site.

The troubleshooting dashboard significantly relies on the co-operation and co-analysis of the ad operations and product teams. It requires collaboration and varying skill sets between the teams to be able to quickly identify and respond to set up or technical issues in order to avoid negatively impacting yield.

Example dashboards include:

- Troubleshooting on Ad Speed
- Troubleshooting Unfilled Impressions

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | **High Level Solution** | Prerequisites For Execution | Managing Change | Results And Impacts

Forecasting Dashboard

A forecasting dashboard predicts the fluctuations and changes in impression and revenue volume, based on historic trends. This dashboard provides the executive and sales teams with an ability to prepare for changes, implement strategies to mitigate negative business impact or manage inventory to achieve business targets.

Forecast dashboards can be built off a predictive modeling tool using an root mean squared error (RMSE) and mean absolute error (MAE), measuring the error between the real and predicted number. One good practice of forecasting is to ensure campaigns are booked in GAM so that the native forecasting mechanism knows what has and has not been reserved. This can mitigate GAM overcast when publishers reserve inventory on an offline spreadsheet and don't transfer it to GAM until much closer to the launch day.

Dashboard For: Executive Team

The executive team can utilize the dashboard as a forecasting total for expected revenue on a monthly basis. Multiple filters can be included to ensure a more granular approach to forecasting is implemented. For detailed insights, a thorough forecasting dashboard will provide inside on expected revenue and impressions by deal type, inventory type, device, country, ad unit and placement. This insight enables strategy creation and planning to tackle impending revenue fluctuations.

Dashboard For: Sales Team

Forecasting allows the sales team to understand historical fluctuations in impressions and revenue to help influence and strategize sales tactics, through the lens of the current sales pipeline, for the months (and year) ahead.

Google News Initiative Data Lab 2020
Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | **High Level Solution** | Prerequisites For Execution |
Managing Change | Results And Impacts

3. Training

Once the dashboards are created, training workshops are necessary for cross-functional team adoption with the focus on building a case for change, using GAM and the framework. Adoption of the dashboards ultimately depend on the education of use (via training), communication of dashboard best practices and revenue benefits.

More on training can be found in the Managing Change chapter.

Inventory Yield Management Timeline

Initial Dashboard Setup — **STEP 2**

Wireframe Dashboards Design (POC) — **STEP 4**

MVP Dashboard — **STEP 6**

Documentation, Trainings and Walkthrough + Support — **STEP 8**

STEP 1
Cross functional team established and KPIs developed

STEP 3
Wireframe Dashboard Design (POC)

STEP 5
MVP Formulas & Charts

STEP 7
Deploy Dashboards/ Go-Live
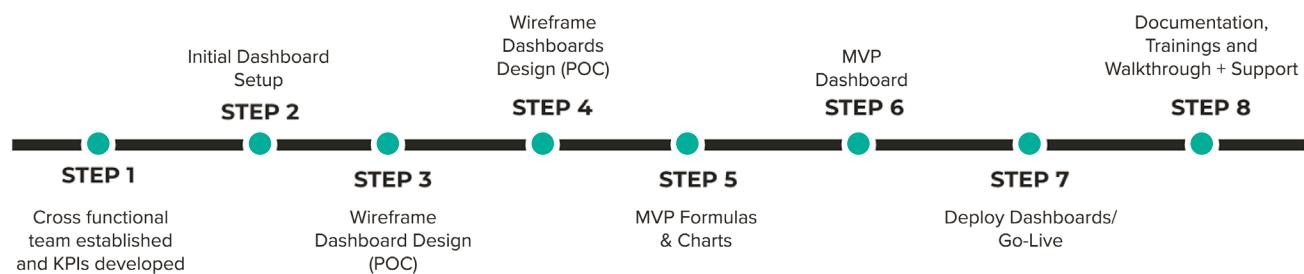
What To Avoid

A comprehensive setup within a publisher's SSP is required to maximize its ability to manage revenue and minimize waste. Many publishers lack consistency and a strong process for naming conventions across ad units, line item units and orders, which results in duplications of data and disparate reporting. A common example of this are variations when identifying elements in ad units such as app inventory being labelled as "App", "In-App", "in app", "app" or "In App". This can also be found in advertiser naming conventions, where single advertisers have been set up under multiple names within the SSP causing significant issues in ascertaining a clear picture of advertisers and activity.

Google News Initiative Data Lab 2020
# Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | **High Level Solution** | **Prerequisites For Execution** | Managing Change | Results And Impacts

In the same vein, a record of all CPMs for direct deals need to be recorded to effectively compare and analyze direct activity alongside programmatic activity. Without this detailed setup, publishers are unable to grasp a full and detailed understanding of revenue streams; impacting the ability to successfully manage and improve yield.

A workable, automated and streamlined pricing rule structure is necessary for the growth and maturity of a publisher. To develop this, it's recommended that publishers frequently review pricing rule structures as many are often disorganized or outdated, making it difficult to determine what pricing rules are effective or not. When determining the appropriate pricing structure, publishers need to adopt a testing and experiment regime; allowing for the right balance between protecting quality and maximising revenue to be established. Once a pricing strategy is adopted, publishers then need to explore the automation of its existing manual processes as most legacy processes are often redundant, time consuming or have a high propensity for human error.

Prerequisites For Execution

Assessment

Agreeing with any or all of the statements below is an indication the case for change for inventory yield management is required in an organization:

☑ There is focus on metrics which they have minimal control over such as click-through rates.

☑ The primary digital measure for ad operations is impressions, eCPM and revenue.

☑ Data is disparate making the development of a holistic view of performance difficult.

☑ Data is not reported on in a uniformed manner and as such, comparing previous periods is difficult (manual reporting processes, copy/pasting between different tools etc).

☑ Legacy processes for set up are time-consuming and prone to human error.

Use Case 3: Inventory Yield Management

☑ Strategy around pricing and pricing structure is siloed to the sales and executive teams, ad operations and product teams do not have as much insight or influence.

☑ No relationship between content and revenue is established.

☑ A lack of cooperation between product team and adops (lack of established KPI'ed premium inventory)

☑ Errors within SSP set up are not addressed (e.g. duplicate advertisements and inconsistent naming conventions) leading to inconsistencies and incorrect data.

☑ Ad operation teams do not have a strong relationship with the data and analytics department nor a regular cadence for performance deep dives to improve yield.

Technical Prerequisites

There are several prerequisites for executing this IYM use case successfully, these include:

Off The Shelf Web Analytics Solution

Google Analytics 360 (GA360) was used throughout the Data Lab program, as such this playbook will outline the implementation using GA360, some features were developed based on integration between GA360 and GAM.

Ad Server + Supply Side Platform With Minimum Api Access

Google Ad Manager 360 (GAM) was used throughout the Data Lab program, as such this playbook will outline the implementation using Google Ad Manager 360. As some of the reports could be built based on more granular data, Data Transfer files are particularly useful. In absence of Data Transfer files raw data from GA360 was helpful. Although, please note that due to various discrepancies in the GA360 + GAM reporting integration and limited data model DT files seems to be a better solution.

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | **Prerequisites For Execution** | Managing Change | Results And Impacts

Data Warehouse

All of the publishers that executed IYM as part of the Data Lab program used Google BigQuery (GBQ). The role of GBQ in this use case is to house data from GA360, GAM, and any Data Transfer Files that would then be queried in the dashboard.

Dashboarding Solution

In this program, Google's Data Studio was the dashboarding tool that was used for all publishers dashboards. Any dashboarding tool can be used as an alternative, provided it links with the data warehouse that is being used.

This use case can also be executed using non-Google analytics, data warehousing, dashboarding and SSP providers.

Required skills

In order to successfully deliver the inventory yield management dashboards outlined, there are specific technical skills required. These include,

- Detailed and experienced knowledge of the SSP and ad server utilized is necessary to ensure an understanding of the data being analyzed. Further to this, a detailed understanding regarding the SSPs main concepts and programmatic capabilities is crucial, including pricing rules, troubleshooting deals, deal setup/optimization/management and forecasting of inventory. The most common barrier to success is a publisher's lack of knowledge on the core platforms used in the IYM use case.

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | **Prerequisites For Execution** |
Managing Change | Results And Impacts

- To create a superior yield management solution it is a requirement to understand the analytics solution and its limitations, together with raw data exports. If the analytics solution is GA360 as was used within the above solution, it is a necessity for practitioners to be GAIQ certified.

- As with the other use cases outlined earlier in this playbook, this use case needs to be resourced by personnel proficient with SQL for the Google Cloud and BigQuery solutions. Currently Google Big Query (GBQ) is using a "standard" ANSI SQL, meaning individuals working with different databases will find it easy to adapt to GBQ. As BigQuery does not provide low latency querying capabilities, it's likely other solutions will be used; therefore, experience with nosql databases may be handy, as this type of solution tends to provide the fastest response times (Google Datastore, Google BigTable). It is also important to ensure that all people included in the project are familiar with the GBQ pricing model, as it has a significant impact on implementation approaches. Automatic data transfers and scheduled queries removes ETL type of programming from the process.

- As more complex data models are likely to be developed here, a strong knowledge of Google Data Studio (or any other desired dashboarding solution) will be required.

- Where data is pulled directly through the GAM API into BigQuery, some coding skills are required. For example, a script using the Google API python library could be developed to pull reporting data from GAM. The reporting data, in CSV format, can temporarily be uploaded to a Google Cloud Storage bucket and inserted into a BigQuery table. A script of this nature would be deployed as a Google Cloud Function utilizing Cloud Scheduler to run daily.

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | Prerequisites For Execution | **Managing Change** | Results And Impacts

Managing Change

In order to achieve maturity in this space publishers need to instill adoption of the best practices and procedures in this playbook long-term. A structured approach to change management will give team leaders the best chance to achieve success The case for change for IYM can be viewed in the following overarching segments:

Cross-Functional Team

Often there are limited interactions between the data team and those teams responsible for the publisher's commercial strategy (i.e sales, product and executive teams). Therefore, there is a great opportunity to use data to improve overall performance, quality (viewability or CTR) and monetization strategies through cross-functional teams driven by data-oriented decisions.

It is important to have a cross-functional team that includes representatives from different departments with the purpose of achieving a mutual goal. Representatives should include members from all teams directly and indirectly involved in a publisher's yield strategy, including sales, analytics/data, product, ad operations and content teams. Strong adoption of new technology or a change in process is achieved through empowerment by education. As such, training on how to best use the performance, troubleshooting and forecasting dashboards is essential. Once the team has a strong foundation of understanding, it's necessary to embark on a fixed time period pilot whereby new processes and the inventory yield management dashboard can be tried and tested. Success of this pilot is marked by the team's adoption of new technology workflows to better track inventory, review pricing rules and viewability of inventory in conjunction with a consistent uptick in the use of the inventory management dashboard for performance insights, troubleshooting and forecasting.

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | Prerequisites For Execution | **Managing Change** | Results And Impacts

Training Workshops

All members of the relevant teams required to adopt new yield management strategies should attend training regarding new quality assurance standards and a troubleshooting framework to guide future problem resolution. Most importantly, thorough GAM training is required to ensure continued success and maturity of yield management. This training must include implementation best practices allowing for more advanced and technical features to be adopted in the future without needing to reset or reimplement the GAM account structure.

For training to be relevant to its audience, e.g. executives or ad operations, it should be tailored to meet the differing needs and use cases per department and levels of maturity.

Documents For Roles And Trackers

To monitor adoption rate and keep a record of optimizations or experiments run by various stakeholders from different teams, the use of trackers is crucial. When managing new processes, centralizing roles, monitoring reporting or ensuring accountability for actions taken, trackers (or logs) provide a transparent and centralized location to store this information.

For the performance dashboard, it's best practice to adopt the following trackers:

- **Action Tracker**

  When teams meet to investigate significant performance changes or execute an optimization or experiment, an action tracker ensures the roles and responsibilities for specific tasks are assigned and maintained; allowing for accountability by team members.

- **Sales Tracker**

  Maintaining a tracker of current sales negotiations, assists the sales team in understanding the expected pipeline, forecast revenue and allows for capacity planning across the ad operations team.

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | Prerequisites For Execution |
Managing Change | Results And Impacts

- **History Record**

  Be sure to create a history log once the root cause of significant performance changes has been established. Once created, this becomes an important source of information that can be shared with the executive team, provides insight into fluctuations in performance and can be used for comparison (e.g. of technical challenges or performance) in future months.

- **Optimization Tracker**

  Where various stakeholders are making changes or wanting insight into an experiment or live campaign, keeping track of the changes made and by whom allows for accountability. It is also a powerful source of information as it demonstrates the impact certain changes can have on a campaign or insight (either in a negative or positive manner).

For the troubleshooting dashboard, an incident log should be kept up to date at all times. The purpose of the tracker is to maintain a detailed record of the issue and the solution implemented.
The forecasting dashboard should be accompanied by a forecasting tracker. This tracker outlines the line Item types, quarterly totals, estimates from the forecasting dashboard and details investigations undertaken into significant discrepancies.

Raci Team Matrix

To simply and clearly delineate roles and responsibilities of the new cross functional team, the project management practice of a RACI matrix can be used. This ensures visibility of each of the team members tasks and alleviates confusion around new operational roles.

The RACI matrix is an acronym derived from the four key responsibilities most typically used: responsible, accountable, consulted, and informed. The RACI matrix is usually displayed in a table, becoming a visual representation for all key stakeholders involved. Once roles are assigned, a letter (R, A, C or I) that corresponds with the level of responsibility each person has within the process

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | Prerequisites For Execution |
Managing Change | Results And Impacts

is listed on the chart. RACI also provides a guide regarding what level of engagement different stakeholders will have in any future decisions made in regard to that process.

Below is an example RACI matrix:

Task	Sales	Ad Operations	Editorial	Product
Creation of Trackers	Responsible, Accountable	Consulted	Informed	Informed
Training	Informed	Responsible, Accountable	Informed	Consulted
Site Optimization	Informed	Consulted	Informed	Responsible, Accountable

Operations

Ensuring the adoption of a new technical solution requires structured guidance around its application in daily operations. Development and implementation of the dashboards into every team's daily management processes will create frequent opportunities for the teams to identify successful inventory and activity set up to allow for amplification and replication.

The dashboards should be tailored to varying stakeholders within the organization, ensuring a holistic view but providing varying levels of detail depending on the users requirements. There should be an underlying focus on identifying areas for improvement to tweak product set up and pricing strategies. As such, an ongoing cadence and collaboration between teams to review and analyze the dashboards as a collective is a necessary tool for success.

In terms of practical application and usage of the dashboard and trackers, the interplay between dashboard and tracker usage is outlined below:

Google News Initiative Data Lab 2020
Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | Prerequisites For Execution |
Managing Change | Results And Impacts

Performance

Prioritizing performance metrics in the following order:

1. Ad Performance (rev, impressions)

2. Product Performance (viewability, monetized pageviews)

3. Audience Performance ("Brand Lovers" audience size, ARPU)

4. Content Performance (avg eCPM for categories, and share of pageviews).

From a usage perspective, the performance dashboard should be reviewed and analyzed:

• Daily by the ad operations team

• Twice a week by the sales team

• Weekly by the executive team

• On an ad hoc basis by the cross functional team when significant changes are flagged to collaborate and investigate root causes.

In regard to trackers,

• The action tracker would be updated, by the team or personnel responsible, once an action related to performance management is undertaken.

• When performance changes are recognized, teams would update a performance history record to keep a historical log of fluctuations.

• In addition to this, the sales team should keep a sales tracker to provide detail on the sales pipeline.

• For action and sales trackers all actionable items should tag or alert relevant team members who are required to action a task.

• Upon completion of a task, the tracker is updated to reflection actions taken, observations and any important call outs/notes.

• Once the root cause of the performance change has been found, this should be updated in the Performance History Record and updated to the executive team.

Google News Initiative   MIGHTYHIVE   88

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | Prerequisites For Execution | **Managing Change** | Results And Impacts

This process of task documentation, assignment and status is designed to streamline the workflow and responsibilities for all those involved in the performance side of the use case.

Troubleshooting

Prioritizing troubleshooting metrics in the following order:

1. Ad Performance (ad speed, fulfill rate)

2. Product Performance (viewability, monetized page views)

3. Content Performance (GA page load data, js errors)

4. Audience performance (increase / decrease in "Engaged" / "Brand Lovers")

From a usage perspective, the troubleshooting dashboard should be reviewed and analyzed by the ad operations team in the morning every day

In regard to trackers,

- As soon as any issues are flagged the ad operations and product teams should organize meetings to determine appropriate next steps to solve the problem.

- An incident log should be updated every time an issue is flagged. Fortnightly reviews of the log is necessary to assess if there are recurring issues. By having the incident log updated regularly, there is a record of issues that can be referenced in future.

- Where there are recurring issues, a supervisor of the appropriate team that owns this process or product must be notified. For example, if URLs frequently have slow load times, the head of the product team should be notified.

Forecasting

From a usage perspective, the Forecasting dashboard should be reviewed and analyzed:

- Weekly by the sales team

- Prior to all sales meetings, ensuring that all orders have been updated in the platform.

- Every quarter to review the current and next quarter.

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | Prerequisites For Execution |
Managing Change | Results And Impacts

In regard to best practices to maintain a thorough data set for future forecasting,

- Once actualized, the forecast tracker must be updated with final revenue amounts; allowing for both comparison against forecasts to assess accuracy and to assist with forecasting for future time periods.

- The sales team should identify and investigate areas of significant discrepancy between forecasts and actuals, and report these back to the executive team, by tagging or alerting them within the tracker.

Integration Of Dashboards Into Key Processes

The power of the data provided through the dashboards extends further when embedded into the appropriate processes; reliance on it to assist in decision making by various stakeholders is crucial to the success of long-term culture shifts. From the sales processes right through to post delivery, reporting from the three dashboards should be consulted regularly in place of previous measures. The forecasting dashboard is a powerful tool that seperate to providing revenue forecasting can also be used during the sales brief process to check inventory availability. The performance dashboard can also aid in monitoring deal creation to ensure no setup errors have been made, an example would be wrong CPMs being entered or the incorrect advertiser being used.

Once the deal is delivering, monitoring processes should incorporate the performance and troubleshooting dashboards to make optimizations and rectify errors respectively. This can replace traditional performance reporting rather than accompanying it.

For post campaign reporting, the performance dashboard can be used to look at previous campaigns and performance, influencing deals created in the future.

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | Prerequisites For Execution | Managing Change | **Results And Impacts**

Results And Impact

Please note, due to COVID-19 lockdowns being in place when the final delivery of this project occurred, audience behavioral patterns were affected as were advertising budgets. Despite the unprecedented environment at the time of execution, many nuanced lessons from different publishers were noted:

- In some countries, the ad request eCPM dropped significantly therefore our reporting and attribution became very difficult.
- Publishers who adopted the three pillars were able to significantly increase viewability (up to 50% increase). This is a positive solid metric that was not affected by the economy.
- For one of the publishers, an increase in ad request eCPM (by around 20%) was observed before COVID-19 lockdowns were introduced. Post lockdown in this country, further results were skewed.
- In another case publisher observed around 5% increase in eCPMs after a month from deployment.

Importantly, despite strong economical impact there was a strong adoption of provided solutions; all publishers use the inventory yield management dashboards on a daily basis. The training workshops provided were successful, as each outlined example use cases for stakeholders on how to best utilize the dashboards. Further to this, many publisher's updated processes across the organization to ensure the dashboards became the source of truth and central to discussions around sales and product. A shift in culture towards a new reliance on data to maintain higher standards in terms of taxonomy (ad units) and orders (eCPMs for direct deals) was also observed. One publisher explained we *"are able to identify each factor of performance separately such as ad unit performance, conversion rate and viewability and draw a separate cause and correction."*

Use Case 3: Inventory Yield Management

Business Problem | Strategic Approach And Goals | High Level Solution | Prerequisites For Execution | Managing Change | **Results And Impacts**

Finally, encouraging feedback surfaced post deployment when sales departments were able to structure sales materials to position the publisher better and see more satisfying results. One publisher stated "with this new program, we have added several new layers of reasoning, optimization, efficiency and other value parameters. This has helped direct sales teams design stronger and more logical pitches in addition to vanilla metrics like reach and market positioning."

Sponsors

Rohan Tiwary

Head of Media, News & Entertainment

Google

Sheena Bhalla

News & Publishing Lead

Google

Mathew Pretel

Partner Technical Lead

Google

datalab-apac@google.com

Contributors

Jakub Otrzasek

Senior Director of Analytics in Data APAC

MightyHive

Jess Ross

Change Management Lead

Samantha Vickers

Lead Project Manager

MightyHive

Isabella Spragg

Technologist Project Manager

(Content Planning)

MightyHive

Pip Hughes

Technologist Project Manager

(Inventory Yield Management)

MightyHive

Jayden Clayton

Publisher Project Manager (Recirculation)

MightyHive

Google News Initiative   MIGHTYHIVE

Technologist Partners

55 the data company

contact@fifty-five.com

Datalicious

solutions@datalicious.com

eAgency

di@ml.e-agency.co.jp

MightyHive

aunzsales@mightyhive.com

Tatvic

hello@tatvic.com