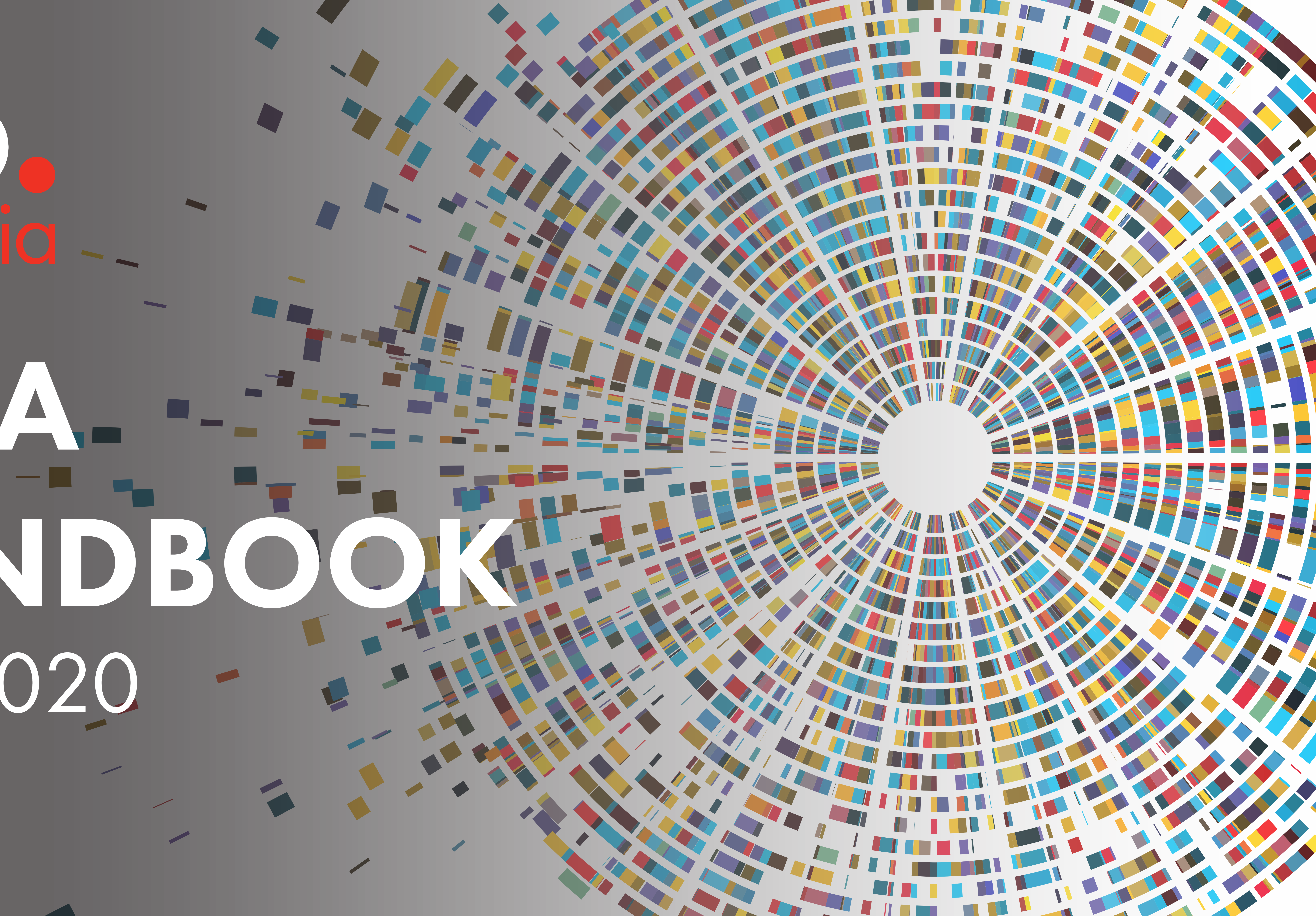
iab.
australia

DATA HANDBOOK

JULY 2020



CONTENTS

IAB AUSTRALIA'S DATA COUNCIL

This paper was produced by the following members of the IAB Australia Data Council:

Sian Whitnall
Chief Digital Officer
OMD Australia

Iman Behzadian
Senior Manager, Data Scientist
Woolworths

Isabelle Dunn
Chief Digital Officer
Hearts & Science

Dan Richardson
Head of Data, ANZ
Verizon Media

Jonas Jaanimagi
Technology Lead
IAB Australia

INTRODUCTION

Data is a very popular term these days and one that is constantly referred to in online advertising. Hence the IAB Australia Data Council recently have been very keen to provide some guidance on the various types and definitions – along with some best practices on its collection, management and usage. This document is the resulting output and intends to build upon the efforts of the Data Handbook released in 2017, so as to bring us more up-to-speed.

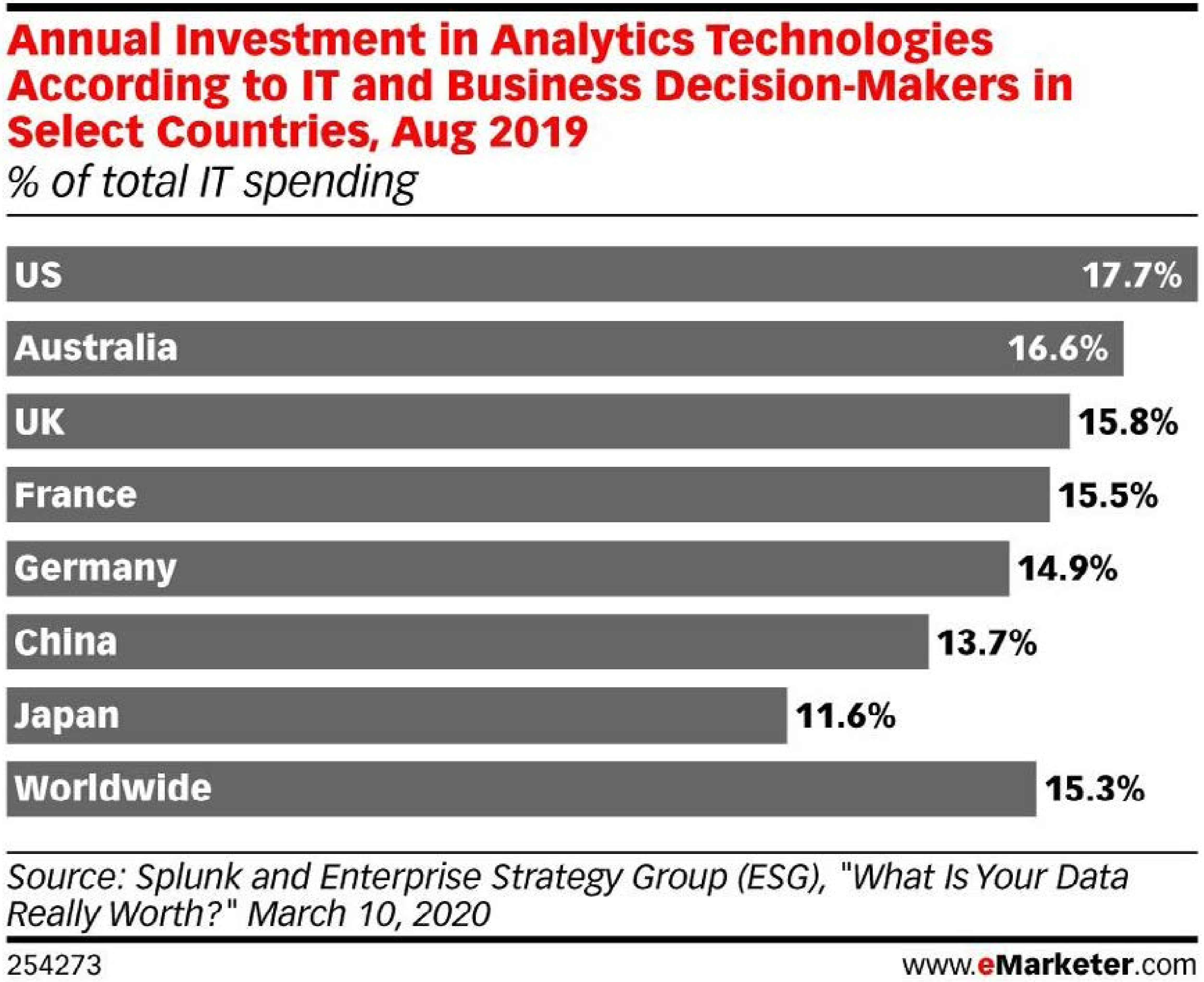
In this document we'll be focusing specifically on the definitions and usage in relation to digital marketing and our intent is to provide you with a solid set of useful considerations. Additional content can also be found in our other relative workbooks and sites, a list of links to which you can find at the end of this document for further reference.

Investment in data technology, talent and solutions has been growing aggressively for many years now here in Australia – and the trend here follows what we have been seeing in other markets globally for some time, but Australia still remains slightly above the global average.

Therefore our Data Council ensuring that the content and outputs in this area for digital advertising is fit-for-purpose and regularly updated will remain a priority. There are plenty of forthcoming changes that we are aware of, such as the recent Apple announcements and the impending deadline for third-party-cookies, but there will be many others to come that we cannot easily predict. Hence ensuring that we have a clear set of definitions, recommendations and best practices to help support our members for both the knowns and the unknowns to come felt timely, with the core intent being to enable our members to feel confident in how they choose to invest in data - in terms of strategy, people and technology.

Please feel free to feedback on this document, we are always open to constructive inputs – and also look out for updates in relation to data via our weekly newsletter, online articles or though our regular industry events.

- Jonas Jaanimagi
Technology Lead
IAB Australia



Annual Investment in Analytics Technologies According to IT and Business Decision-Makers in Select Countries, Aug 2019

% of total IT spending

Source: Splunk and Enterprise Strategy Group (ESG), "What Is Your Data Really Worth?" March 10, 2020

254273 www.eMarketer.com

THE DIFFERING TYPES OF DATA

In this section we'll provide some simple definitions of the different types of data related to online advertising that we are keen to provide then provide further guidelines for.

The basic definitions are below:

- Known consumer data - directly attributed to consumers via CRM data or eCommerce.
- Anonymous behavioural data - generated via website analytics or linked to identifiers such as cookie or a cross-device ID.
- Financial data - from sales, investments or other commercial activities.
- Research data - consumer insights generated by online studies, panels or consumer segmentation.
- Campaign data - analytics measuring impressions, clicks, engagements tracked by ad servers or website analytics.

For the requirements of advertising, all of this data tends to be categorised as being either first, second or third party.

FIRST-PARTY DATA

WHAT IS IT?

First-party data collected from your assets. In short, it's your data. It might be collected from your customers' in-store purchasing habits and stored in a CRM or it might be behaviours people exhibit on your website. Either way, it is a unique data set to your business that you and you alone own. It's often argued that this is the most powerful form of data as it is a true indication of how people are interacting with your brand and what they really want from you.

WHAT ARE ITS MAIN USES?

First-party data has a number of different uses, but it's primarily used in retargeting campaigns and customer marketing campaigns.

HOW DO I COLLECT IT?

First-party data can be collected from any asset that you own, with the most popular sources being your website or app (through a pixel or SDK) or at the point of sale (either through an email or customer loyalty card number).

WHAT ARE THE ADVANTAGES AND DISADVANTAGES?

Firstly, first-party data belongs to your brand and this gives it some strong advantages. There's no dispute on ownership, this is your data and you can do what you want with it (within the law). It's a completely unique data set to your company so this data will give you the best results for your company. Be aware though it's incredibly difficult to scale your first-party data. It is taken from your assets so unless more people take actions with your brand (i.e. visit your website or subscribe to your database) it's very difficult to grow your first-party data.

SECOND-PARTY DATA

WHAT IS IT?

Second-party data is first-party data that you're getting directly from the source. It's a relatively new type of data and most people think about it as data that isn't first-party or third-party. It is generally a unique data set (similar to first-party data), however, that data set isn't unique to your brand. A great local example is an airline (like Qantas) might team up with a global accommodation supplier (like Airbnb) to share their first-party data sets with each other.

WHAT ARE ITS MAIN USES?

Second-party data is a great way to enrich your first-party data. By integrating first-party and second-party data, brands are able to scale their first-party data, find new customers and learn more about the behaviours of their current customers.

HOW DO I COLLECT IT?

The majority of second-party data relationships happen through private deals and direct relationships with other brands. Data is collected in the same way as first-party data so a combination of pixels, tags and subscriber information. Most brands, however, choose to employ a Data Management Platform (DMP) to allow them more flexibility when it comes to sharing and integrating different data sources.

WHAT ARE THE ADVANTAGES AND DISADVANTAGES?

Second-party data has the advantage of adding scale to your first-party data, however, it's a lot more work to make its collection privacy compliant. People need to know what you're using their data for and how it's going to be shared. Brands also need to make sure that the data they are collecting, and sharing is being done so in a secure manner. Second-party data can also be hard to come by. As most data partnership deals are done privately, they can bring several complications and take a long time to put in place.

THIRD-PARTY DATA

WHAT IS IT?

Third-party data is probably the most common type of data that marketers and brands are used to talking about and working with. Third-party data is collected from an external source that doesn't have a direct relationship with the people it's collecting data about. For example, a third-party data company might pay publishers to put their pixel on their site and then use that information to piece together online profiles.

WHAT ARE ITS MAIN USES?

Third-party data is a useful tool for finding new customers. If you know what your customer profile looks like, then you can purchase third-party data to match that profile and then integrate it into your marketing efforts.

HOW DO I COLLECT IT?

To put it quite simply, you buy it or lease it.

WHAT ARE THE ADVANTAGES AND DISADVANTAGES?

Third-party data is readily available through a number of different companies like Eyeota, Oracle and Equifax. It's fantastic for helping businesses grow their customer base quickly and easily. The quality of third-party data can vary widely, however. While some companies do a good job of matching profiles, there is always some form of wastage. Brands take a generalisation of who their customer is and try to match that with other profiles and that won't always give you a high success rate.

Third-party data is also subject to different regulations. While Australia isn't yet as strict as Europe, marketers need to be mindful of what type of data they're using (usually health or financial data will attract tighter restrictions), how they're using that data, how they are storing that data and how they are informing their customers about the data.

The challenge is often how to make all these various data assets often siloed within entities manageable and actionable by integrating data sources and then successfully building an analytics layer. The aim being to enable the owners and users that have access to fully implement the technological capabilities required.

PII (PERSONALLY IDENTIFIABLE INFORMATION)

This refers to information used or intended to be used to identify a particular individual, including name, address, telephone number, email address, financial account number, and government-issued identifier. For legal purposes the Australian Privacy Act defines personal information as:

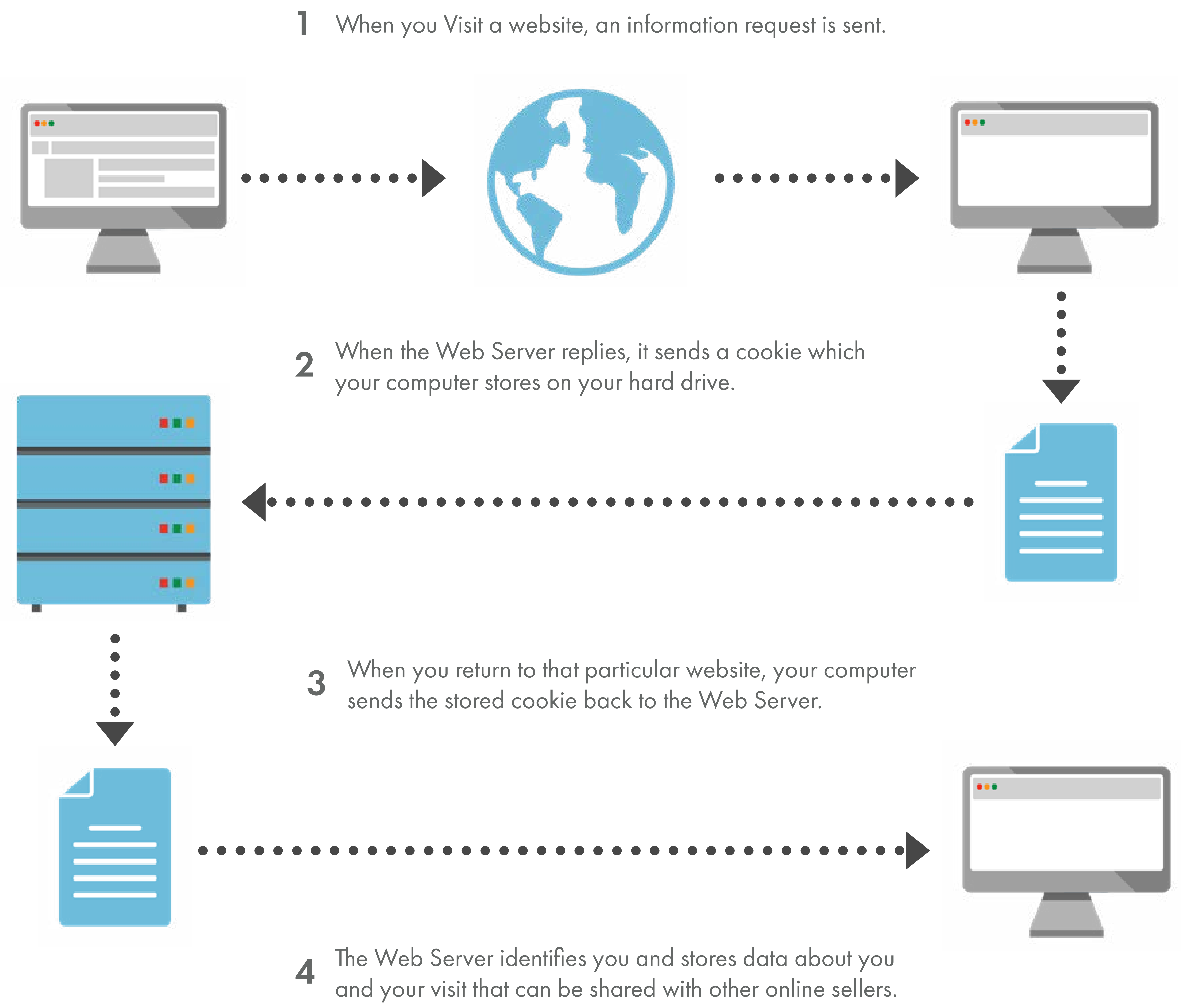
Information or an opinion about an identified individual, or an individual who is reasonably identifiable:

- Whether the information or opinion is true or not; and
- Whether the information or opinion is recorded in a material form or not

COOKIES

Also known as an HTTP cookie, web cookie, or browser cookie, this is a string of text sent from a web server to a user's browser that the browser is expected to send back to the web server in subsequent interactions. A cookie has a few core attributes: the cookie value, the domain and path within which it is valid, and the cookie expiry. There are other attributes as well that limit the cookie to https-

THE COOKIE PROCESS



1 When you Visit a website, an information request is sent.

2 When the Web Server replies, it sends a cookie which your computer stores on your hard drive.

3 When you return to that particular website, your computer sends the stored cookie back to the Web Server.

4 The Web Server identifies you and stores data about you and your visit that can be shared with other online sellers.

only transactions, or hide it from JavaScript.

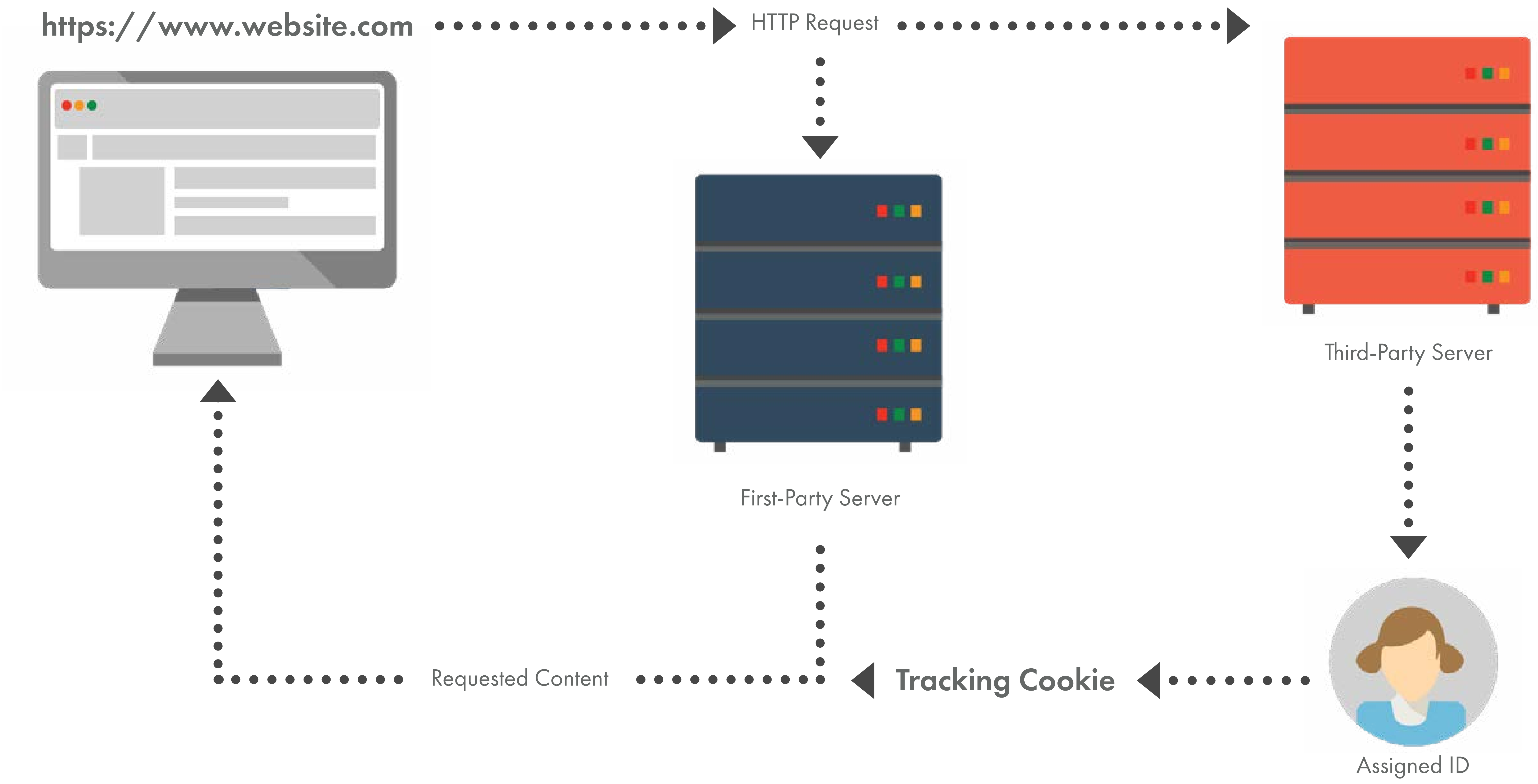
The domain and path define the scope of the cookie – they tell the browser that cookies should only be sent back to the server for the given domain and path. Cookies that do not have a specific expiration date and time are automatically deleted when the web browser is next closed. Cookies with a set expiry time are considered persistent cookies, while cookies without set expiry times are considered session cookies.

In online advertising, cookies generally store a unique identifier, and may contain information like what ads were recently seen (for frequency capping), when the cookie was created (to discover short duration identities), and other simple attributes.

Often cookies are placed on your site that are managed externally and sent to external servers. In this instance, the domain associated with the cookie does not match your website's domain. When the domains differ, these cookies are considered cross-site or third-party cookies.

Changes in browser tracking continue to roll out across the industry with Google announcing that Chrome will no longer support third party cookies within two years. This change follows similar moves from Mozilla's Firefox browser and Apple's Safari browser. These changes will have a major impact on much of digital marketing from analytics, targeting, measurement and attribution. The industry is now working on new privacy complaint approaches and there is more on this further on in the document.

CROSS-SITE COOKIES IN PRACTICE



GUIDANCE ON IDENTIFIERS AND IDENTITY

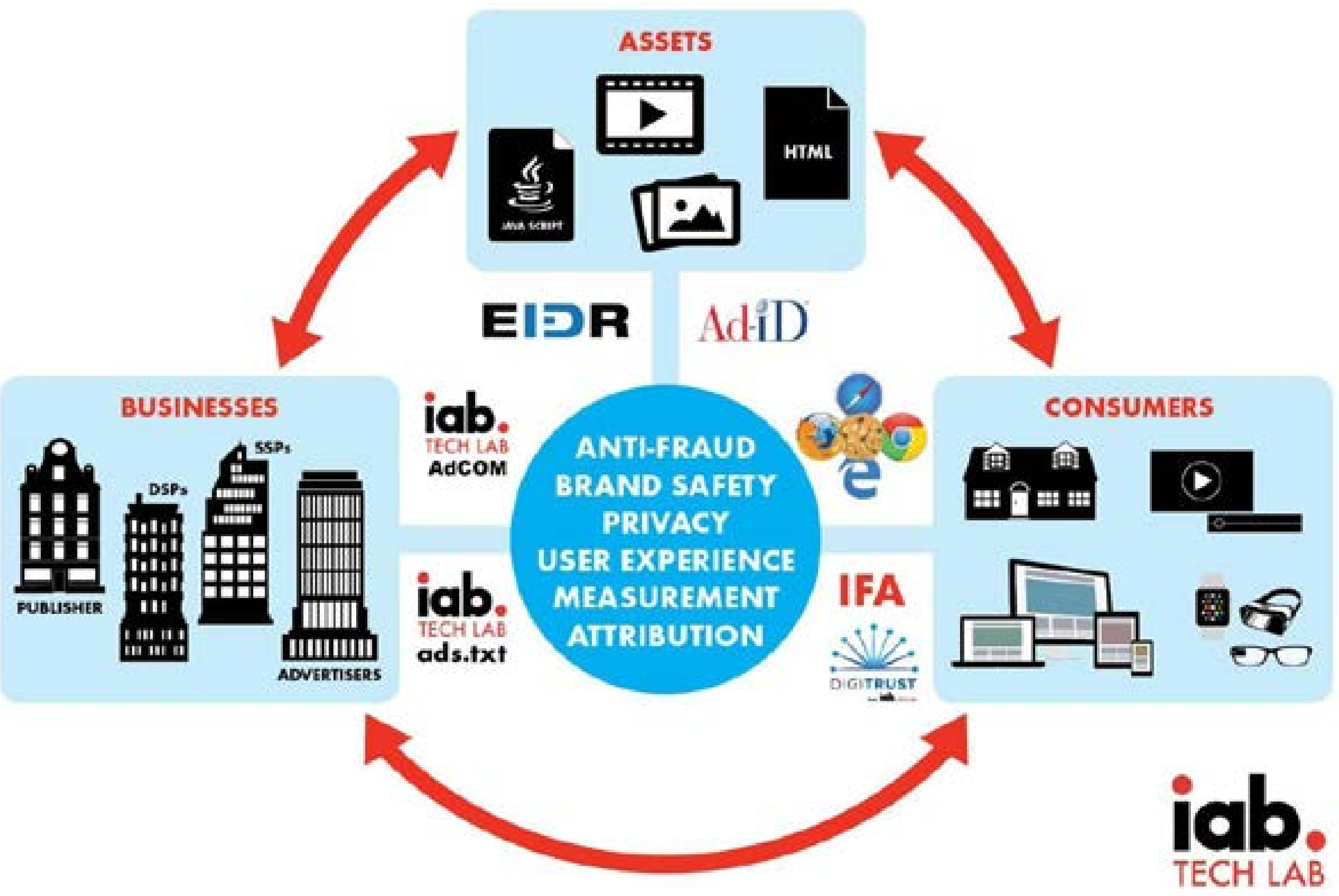
We should also define the meaning and details around ‘identity’ and ‘identifiers’, as these terms are often used liberally to refer to similar things.

IDENTIFIERS

We’ll start with identifiers, which are prolific and consumers can easily have hundreds or thousands of identifiers across all of their browsers and devices.

We align to the IAB Tech Lab view that identifiers come in three types: consumer, creative assets, and the businesses involved in the supply chain. These identifiers are the core building blocks that help fight fraud, improve brand safety, deliver a better experience to consumers, and support measurement and attribution.

IDENTIFIERS FOR EFFECTIVE ADVERTISING



CONSUMER IDS

try to identify individual users or a group of people within a household (all generally anonymously), but may ultimately be tied to devices or browsers, depending on the available data (such as logins) on various platforms. Examples are cookies, DeviceIDs or IFAs (Identifier For Advertising) on mobile and/or OTT (Over-The-Top video) devices. These are utilised for the purpose of understanding user behaviors and interests for targeting and personalisation, assessing where/when a person saw an ad (for measurement and attribution), and applying known privacy preferences consistently across sites, apps, and devices. This allows platforms to develop insights into users’ needs and deliver a better experience by providing more relevant ads.

ASSET IDS

Identify creative assets as they go through the advertising supply chain, to make it easier to understand what was or will be shown to a consumer, ensuring that the right content is delivered to the right individual (separating ads from competitors, age appropriate etc.), and enabling accurate measurement/tracking of which creatives were displayed where and who they were presented to. Asset IDs are also important to help with brand safety by tying the ID to metadata about the creatives.

BUSINESS IDS

Identify the various companies such as publishers, advertisers, and vendors that provide content and ads to consumers, and execute a range of other functions across the advertising supply chain. These IDs are used mainly to manage trust, reduce fraud, and improve transparency.

IDENTITY

As a consumer can have so many different identifiers across all their browsers and devices, the real effort is in being able to manage and leverage these into a genuinely useful version for marketers and publishers. Cookies, for instance, are useful short-term identifiers as they can be used (for now) and shared fairly widely, but often for the purposes of true identity management they are simply too inferred and depreciate too quickly.

However, through the unification of the various identifiers and appended data points a persistent Individual ID can be created, which can be utilised as being a more meaningful and shareable profile per individual. Ultimately all the various device-level identifiers are merely an enabler of advanced identity resolution solutions. This process is often referred to as Identity Resolution and the tools used to align the various identifiers and store them is known as an Identity Graph.

These persistent customer identities can also be supplemented with other data, including offline to build out a fuller and more accurate anonymised profiles. This ability to onboard offline customer data in order to further enrich managed identities and target online ads has proven very successful for platforms such as Google and Facebook.

This approach is also the most effective technically for consistently and cleanly managing a consumer’s digital privacy. An identity resolution solution can permanently keep track of any preferences and/or consent signals from identifier to identifier, as well at each data touchpoint. This allows for competent management but also a quality consumer experience as the updates and controls can be made more seamless.

GUIDANCE ON DATA COLLECTION, USAGE AND EXECUTION

This section focuses on the considerations related to the collection of data, it's unification, segmentation and ongoing management.

At a high level the image to the right captures the overall requirements to aggregate and normalise disparate data sets for delivery, advanced campaign analytics and reporting.

ADVERTISING TECHNOLOGIES

BUY-SIDE AD SERVER (THIRD-PARTY AD SERVER)

A web-based platform for buyers to host, manage and serve digital assets for advertising campaigns. Its primary goal is to centralise and standardise delivery and performance data with a uniform methodology for counting across key metrics. Some may also be able provide viewability capabilities (e.g. Google TrueView).

AUDIENCE MEASUREMENT

Best defined as the independent measurement of digital advertising audience delivery across all device types. Audience measurement can refer to the measurement of ads and content/media environments.

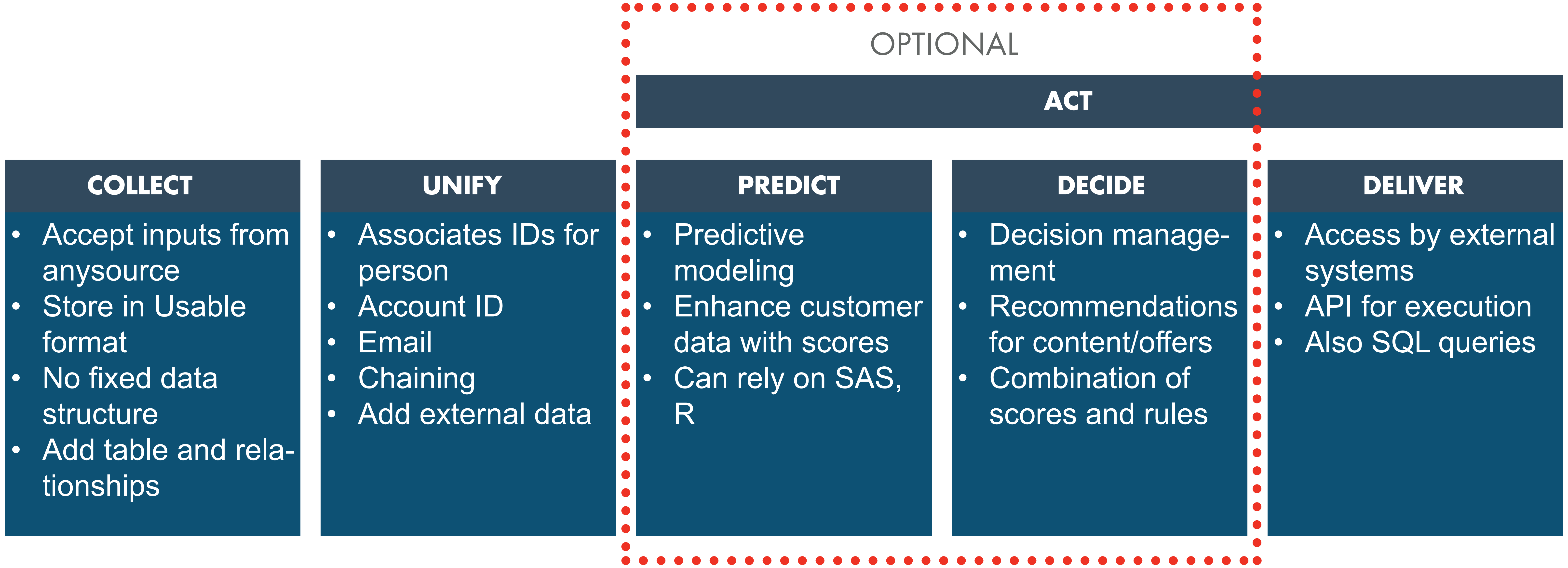
DMP

A Data Management Platform (DMP) is used to collect, store and manage online and offline data sets to gain insights before creating actionable segments to target digital campaigns.

CDP

A Customer Data Platform (CDP) is a technology capable of collecting data from multiple online and offline interactions and matching them to a single customer profile. One of the main features of CDPs is they can profile interactions from anonymous customers and retroactively tie that data to a customer once it is identified.

REQUIREMENTS TO AGGREGATE AND NORMALISE DATA SETS FOR DELIVERY, ANALYTICS AND REPORTING



OPTIONAL

ACT

COLLECT

- Accept inputs from anysource
- Store in Usable format
- No fixed data structure
- Add table and relationships

UNIFY

- Associates IDs for person
- Account ID
- Email
- Chaining
- Add external data

PREDICT

- Predictive modeling
- Enhance customer data with scores
- Can rely on SAS, R

DECIDE

- Decision management
- Recommendations for content/offers
- Combination of scores and rules

DELIVER

- Access by external systems
- API for execution
- Also SQL queries

DSP

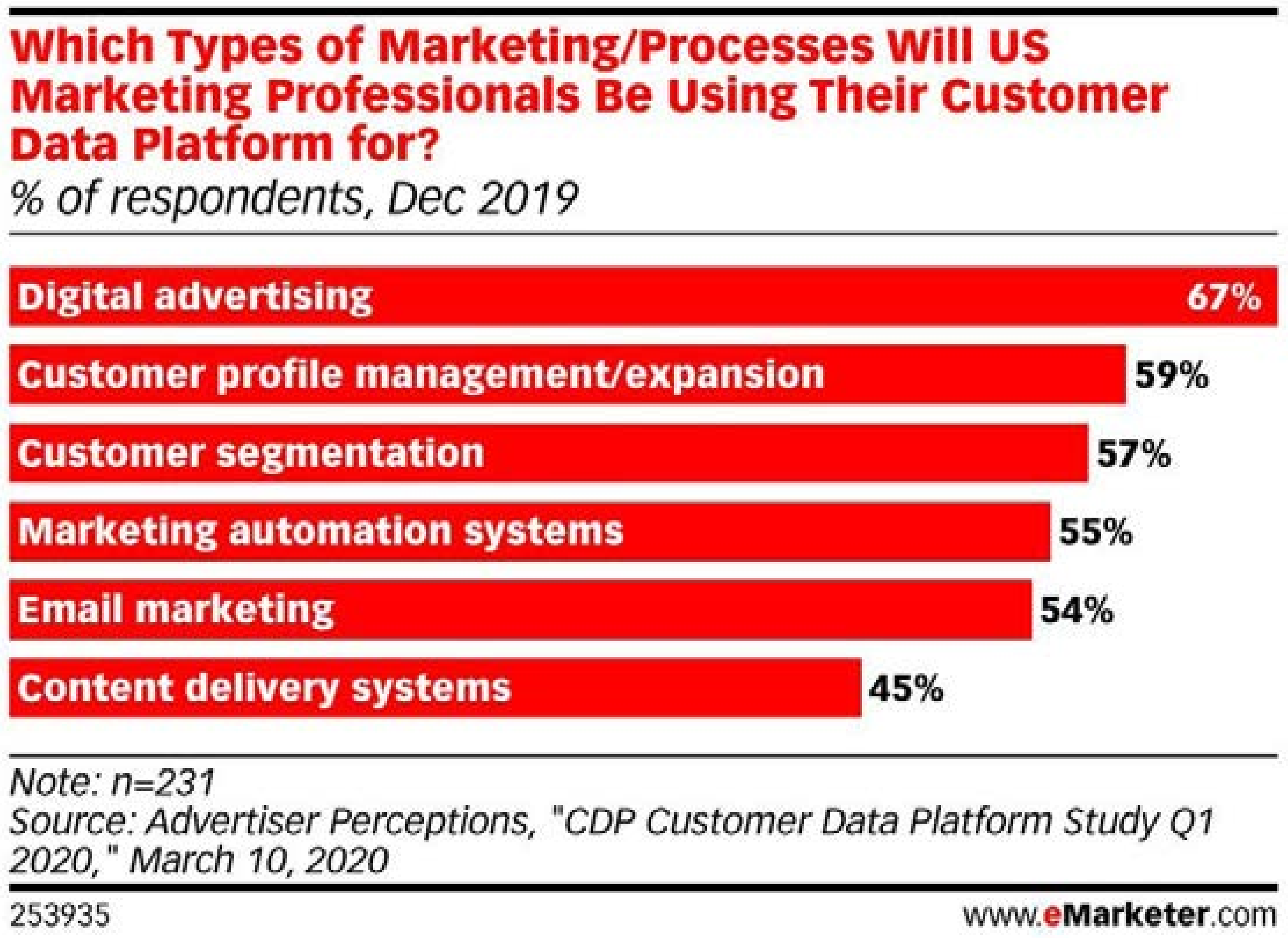
A Demand-Side Platform (DSP) handles automated media buying across multiple inventory sources using targeting, data and real-time optimisation. The DSP is designed to buy an audience versus an objective. Its goals are to pay the lowest possible price for inventory whilst fulfilling the buyer's campaign objectives. A buyer will use a DSP to access publisher inventory made available via an SSP.

SSP

A Sell-Side Platform (SSP) is used by sellers (publishers) to make digital inventory available for DSPs to bid upon. The SSP looks to maximise yields for the seller whilst meeting the buyer's key campaign objectives.

SELL-SIDE AD SERVER (PUBLISHER AD SERVER)

A Sell-Side Ad Server is a web-based platform for publishers to store, manage and serve digital assets across digital properties. Its primary functions are to manage the pacing and delivery of advertising campaigns and provide data on campaign performance. Some platforms may specialise specifically in mobile, in-app environments or video.



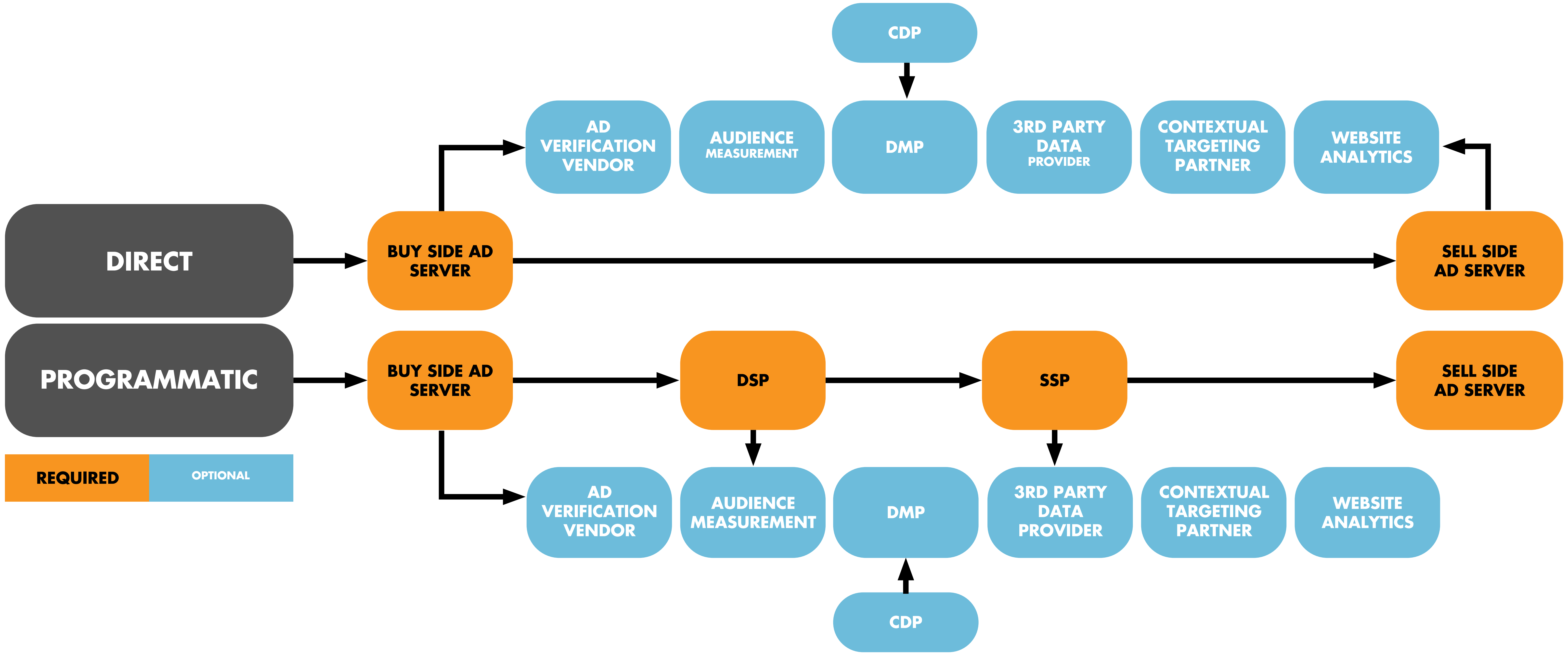
Which Types of Marketing/Processes Will US Marketing Professionals Be Using Their Customer Data Platform for?

% of respondents, Dec 2019

Note: n=231
Source: Advertiser Perceptions, "CDP Customer Data Platform Study Q1 2020," March 10, 2020

253935 www.eMarketer.com

ADVERTISING ECOSYSTEM TECHNOLOGIES



AD VERIFICATION VENDORS

Ad Verification Vendors offer technology that can give independent data on measurement metrics including viewability, fraud and brand safety.

THIRD-PARTY DATA PROVIDERS

Third-Party Data Providers are agents that build audience data sets through partnerships and integrations with external audience collection providers. This data can then be used to target audiences via a DSP or DMP.

CONTEXTUAL TARGETING PARTNERS

Contextual Targeting Partners offer technology that determines the content of a web page allowing programmatic buyers to deliver contextually targeted ads.

Website Analytics Platforms Website Analytics Platforms are platforms that track and report on online traffic. These platforms can optimise digital properties and analyse online user behaviour. While they are optional, they are essential to understanding website visitor behaviour.

DMP AND CDP TECHNOLOGIES

DMPs will work primarily with anonymous behavioural data such as cookies, device IDs, and IP addresses generated from pages of websites.

Meanwhile CDPs can store the same information as DMPs, but also very detailed deterministic information on people's profiles and behaviours aggregated from both online and offline sources. These are often generated from purchase transactions, Customer Relationship Management (CRM) database tools or filled forms and can contain data such as purchase transactions, postal addresses, email addresses and phone numbers as well as numerous web behaviours - and very often containing sensitive PII (personally identifiable information).

For the purposes of digital marketing CDPs will have to attempt to authenticate any users they have access to online and then also (often via DMPs) make those users addressable via cookie-based advertising platforms.

GUIDANCE ON GOVERNANCE & CONSUMER PRIVACY

Increasingly, with regards to data collection, management and utilisation, the topic of privacy and the best practices related to privacy compliance are front of mind.

DATA PRIVACY

For businesses handling consumer data, demonstrating the ability to secure and protect said data, both your own and that of your customers, is a business imperative that yields a competitive advantage.

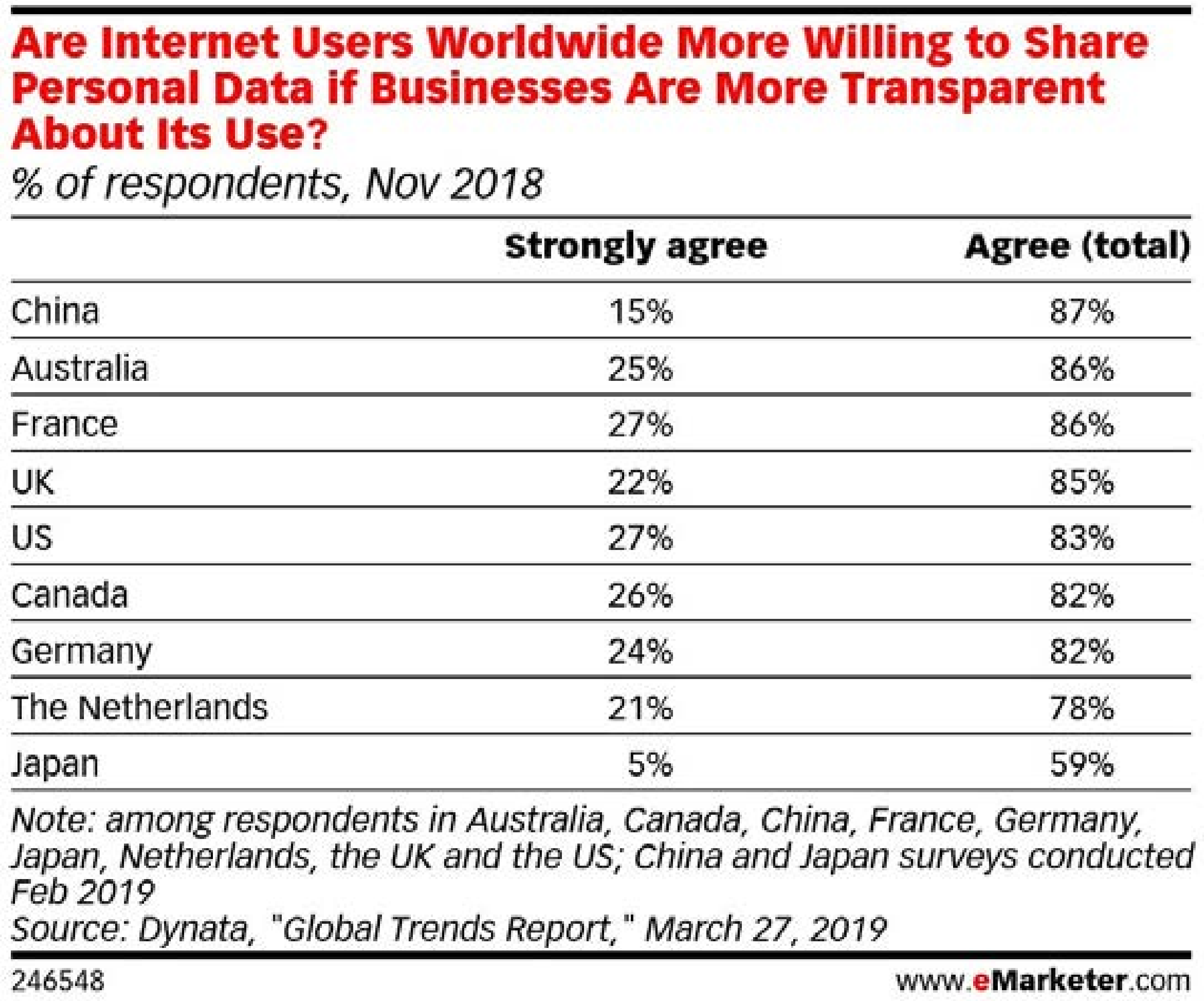
Every phase of the data life cycle (like collection, use, retention, storage, disposal or deletion) must be managed to guarantee compliance with the law, protect the brand and preserve customer confidence.

Organisations understand the need to innovate and safeguard the personal and confidential data of customers, employees and business partners. Maintaining best practices of privacy and security controls to comply with the law will help better manage post-breach incidents. Effective data protection will:

- Reduce the chance of reputational damage from a data or cyber breach.
- Strengthen the business by increasing consumer confidence.

The Australia Privacy Act regulates how the digital advertising ecosystem handles personal information, including sensitive information. The Act includes the Australian Privacy Principles (APPs) which set out standards, rights and obligations in relation to handling, holding, accessing and correcting personal information. Parts of the digital advertising ecosystem rely on data analytics which may include personal information.

Under Australian law, personal information is information or an opinion about an identified individual, or an individual who is reasonably identifiable.



Are Internet Users Worldwide More Willing to Share Personal Data if Businesses Are More Transparent About Its Use?

% of respondents, Nov 2018

	Strongly agree	Agree (total)
China	15%	87%
Australia	25%	86%
France	27%	86%
UK	22%	85%
US	27%	83%
Canada	26%	82%
Germany	24%	82%
The Netherlands	21%	78%
Japan	5%	59%

Note: among respondents in Australia, Canada, China, France, Germany, Japan, Netherlands, the UK and the US; China and Japan surveys conducted Feb 2019
Source: Dynata, "Global Trends Report," March 27, 2019

246548 www.eMarketer.com

DATA GOVERNANCE

MANAGE CONSUMER CONSENT AND CONTROL: CHAMPION THE USER EXPERIENCE

Determine the controls to give to customers when it comes to consent. Under Australian Law, personal information collected by an entity may only be used or disclosed for the primary purpose for which it was collected, unless an exception applies.

This means the way personal information is collected, and what the individual is told about the collection is important for data activities. Businesses should implement the Office of the Australian Privacy Commissioner's (OAIC) Privacy Management Framework which encourages the use of a Privacy Impact Assessment (PIA), to inform big-data activities. This includes mapping the information life cycle, identifying what information is collected, whether it might be utilised for big data and for what purpose. Undertaking these steps will inform what personal information should be collected, how it should be collected and what notice should be given.

PROACTIVELY MANAGING PRIVACY PROTECTIONS: BE FORWARD THINKING

The Australian Privacy Act requires Australian companies to implement data practices, procedures and systems to ensure ongoing compliance with Australian privacy law. This means compliance is a dynamic and ongoing process. When collecting data on a consumer, the collector becomes a custodian of that data. The Office of the Australian Privacy Commissioner requires the consideration of privacy and data protections throughout the data life cycle including when:

- Existing owned data is used for new purposes.
- Collaborating with vendors (such as CRM, marketers, or cloud IT providers) that involve data sharing.
- Staying up to date with newly introduced legal requirements (like those introduced in early 2018).
- Developing internal policies or strategies with privacy implications.

PRACTICE DATA MINIMISATION

Under Australian Privacy Law, companies are required to actively consider whether they are permitted to retain personal information. When a business no longer needs personal information for any purpose for which it may be used or disclosed, the company should destroy or de-identify that information. A business must take reasonable steps to destroy or de-identify personal information. Best practice guidance on de-identification has been compiled by the OAIC and can be found on the IAB Australia website.

BE IN COMPLIANCE WITH DATA BREACH LAWS

New data breach laws introduced to Australia in February 2018 are very similar in substance to new data breach requirements in place in Europe under the General Data Protection Regulation (GDPR). These data breach laws force any company in the digital advertising ecosystem (advertiser, agency or publisher) to notify all affected individuals and the Office of the Australian Privacy Commissioner if they experience a data breach which poses a risk of harm.

MINIMISE THE RISK OF A DATA BREACH THROUGH THE FOLLOWING TECHNIQUES:

Educate employees: The oft-used example of a misplaced company laptop can lead to a data
breach of hundreds of clients' data. Teach staff the most secure ways of data sharing and storing and how to identify and deal with data breaches.

- Evaluate technology: Check if existing software and hardware can adequately identify and deal with data breaches in real time.
- Minimise the amount of personal information held: This can be a tough one, especially when it
- comes to advertising databases, but where possible, try and decrease the personal data stored.
- Encrypt and anonymise personal data: Encrypt or anonymise personal information where possible.

Manage data protections in your advertising tech stack
Any vendor in the digital ecosystem that is processing customer data is liable for the protection of data used. Many marketing professionals rely on third-party vendors (such as CRMs, email service providers, cloud IT services) to interact with customers. The owner of the first-party data is responsible to speak with third-party partners processing customer data to ensure they are taking the proper steps to remain compliant with privacy law. This includes having the tools in place that will allow vendors to both retrieve and destroy data at the end of its life cycle.

ETHICALLY AND TRANSPARENTLY SOURCED DATA

A whole-of-cycle view of data with deep insight into how and why data is being used is essential for ensuring the right balance between compliance, privacy and innovation. The revelations around U.K. firm Cambridge Analytica in 2018 demonstrate that it is not good enough to think about data sharing in a linear way. Anticipate any risks that might be involved in the processes employed. Ethical decisions around the use of data will be expected of the company when handling customer data – but don't expect these requirements to be spelled out in every case.

LEARN MORE ABOUT THE AUSTRALIAN PRIVACY ACT & PRIVACY PRINCIPLES

LEARN MORE ABOUT GDPR

IAB Europe, in partnership with IAB Tech Lab, recently launched the second iteration of the Transparency and Consent Framework (TCF), which is an industry tool that supports companies within the digital advertising ecosystem as they manage their compliance obligations under the GDPR and ePrivacy Directive.

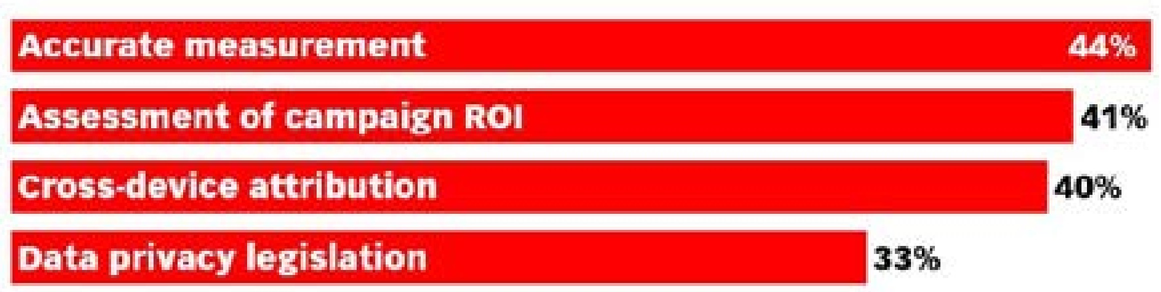
For more information please visit: https://iabtechlab.com/standards/gdpr-transparency-and-consent-framework/

GUIDANCE ON DIGITAL ATTRIBUTION

WHY IS IT IMPORTANT?

Because of the increasing popularity of alternative digital marketing channels such as social media and video sharing platforms, digital attribution models have attracted a lot of attention in recent years. William Hesketh Lever (1851-1925), founder of Unilever has a famous quote that highlights the importance of an accurate marketing ROI measurement tool: “I know that half the money I spend on advertising is wasted. My only problem is that I don’t know which half.”

Leading Digital Media Challenges Digital Media Professionals in Asia-Pacific* Expect Their Company to Face, Dec 2019
% of respondents



*Note: respondents selected up to 3; in the next 12 months; *Australia, Japan, New Zealand, Singapore*
Source: Integral Ad Science (IAS), "Asia Pacific: Industry Pulse Report," Jan 30, 2020
252692 www.eMarketer.com

To understand the complexity of the problem, imagine this scenario: A customer was exposed to five social media Ads and clicked on one of them, watched the video Ad on Youtube, had 10 impressions on digital displays and finally when you sent a marketing email the user converts and spends $100 on your website.

In this case, how much of the $100 sales value should be attributed to each channel? Do we take recency as the most influential factor hence email gets all or majority of the credit or do we consider frequency as the main factor and digital display gets more credit? How about the delivery metrics? How important is an impression in comparison with a click? These are the types of questions that require a Digital Attribution (DA) in place.

In the following sections it will be discussed that there are traditional methodologies that are rule-based and depend on the qualitative understanding of the marketing managers and there are data-driven methodologies based on machine learning approaches that quantitatively decide the answer of each of these questions.

DIGITAL ATTRIBUTION MODEL: OBJECTIVE AND DELIVERABLES

Digital Attribution is a model that determines how credit for the success KPI (sales value, number of conversions, website visitation, and etc.) is assigned to touchpoints in the conversion paths.

Based on this definition, it is clear that the main deliverable of a DA model is the attributed sales (or any other success KPI) to each digital channel. In the next step, by overlaying it with the marketing investment data we can calculate attributed Cost Per Acquisition (attributed CPA) per channel.

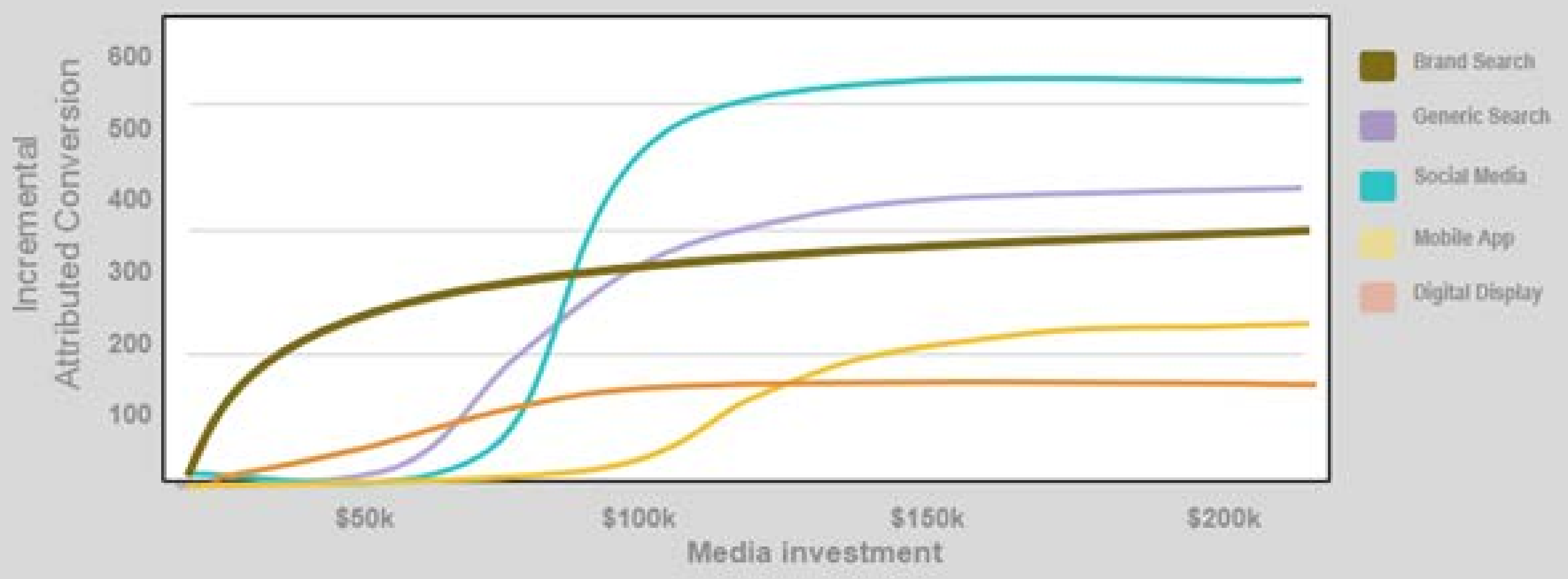
Finally, we can use historic data to calculate expected returns from different levels of investments in various channels and fit a logistic curve to the investment- attributed sales plot. This will give us the response curves. a mock-up response curve for monthly investment value on different marketing channels is visualised in the Attribution Conversion And Media Investment plot below.

These curves help us to identify some key actionable insights:

1. Minimum investment: Because of the S-shaped nature of the fitted logistic curve, the minimum level of the investment required for each channel can be estimated. For instance, in this example, the minimum required investment on mobile apps is more than social media.
2. Optimum investment point: The knee of each curve is the sweet spot of investment. It is simply because of the diminishing returns that will occur after this point and a minor incremental conversion is generated for more investment. For instance, in this plot, $90k is the optimum monthly investment for digital display.
3. Saturation level: These curves also highlight the saturation level or the maximum number of conversions that one can expect from a given channel. For instance, in the plot above while the optimum investment for social media is more than brand search, a higher number of conversions can be expected from it as well (500 vs 300).

In summary DA models on channel level provide two metrics: attributed sales as a measure of volume and attributed CPA as a measure of efficiency.

ATTRIBUTION CONVERSION AND MEDIA INVESTMENT



MMM OR DIGITAL ATTRIBUTION, THIS IS THE QUESTION

In the next section, a high-level roadmap is suggested to implement a DA model which can respond to your company business questions. But before starting your journey on the road towards attribution, there is one fundamental question that you need to ask yourself: Do I need a digital attribution or a marketing mix model as my measurement platform? To answer this question, two main differences between these two measurement approaches is highlighted here:

- Deep dive vs big picture: MMMs paint the big picture of BTL and ATL channels, health of brand, environmental variables such as exchange rate or seasonality, effects of competitors activities, and etc. and estimate the effect of each variable on sales. On the other hand, DA models just have digital channels in scope (digital display, social media, email, video, SEM, SEO, and etc.) in which the customer-level data is available.
- Sophistication level of the methodology: DA methodologies are all deterministic and are built on millions of actual customers journeys that are tracked. Deterministic methodologies are generally more accurate. Also, as it is explained in the following sections, they can be used for targeting customers as well. On the other hand, MMM uses probabilistic methodologies that try to explain the changes of sales based on changes on aggregated investment level on the channel probabilistically. This is mainly because ATL channels are in scope for MMMs and the customer level data is not available for them. Probabilistic models are more likely to get negatively affected by correlation and pick them as causation.

In summary, to decide whether you need an MMM or a DA model you need to consider the marketing investment mix. If it is heavily skewed towards ATL channels MMM is the answer since it has a broader scope but if majority of your marketing media is going towards Digital and BTL channels, digital attribution will be a more suitable and sophisticated approach.

STRATEGIC ROAD TOWARDS ATTRIBUTION

1. Define the measure of success: Discuss the KPI that measure success in the model with the stakeholders and scope the work. Some of the options are:
   - Website visitation
   - Click on “request a demo” or “call a sales person” button on the website
   - Number of conversions
   - Sales value

   Usually managers are more interested in explaining sales based on their marketing investment to define a marketing ROI and they jump to the last option on the list: sales value. However, it should be considered that the further we go down in this list the more external variables affect the customers decision to purchase. For instance, if you are a financial institution and you are running an acquisition campaign to get new customers for your personal loans, the value of their loan applications might not be the best KPI because it heavily depends on some variables that have nothing to do with your marketing channels (e.g. customers’ risk level and their approved loan value). Instead the number of applications as a KPI truly reflect the effectiveness of the marketing campaign on channel level.

2. Assess the limitations: Understand each of the data sources by investigating methods, opportunities and limitations for data collection.
   - Availability of channel data: for instance, Facebook impression data is not published for the client.
   - Granularity of the data: For instance, customer demographics is not available with cookie-based tracking techniques and unless you have a loyalty or a similar program, you can’t find the attributed sales for different demographics.
   - Lack of influential variables and ATL channels effect: As mentioned before, if your organisation invests heavily on ATL channels and you are not attributing any sales or conversion to those channels, chances are that you’ll get non-analytical results. For example, how do you know who has seen the Ads on TV and what percentage of the sales uplift is because of those untracked TV audiences.
   - Limitations in forming unique customer digital journey: This can be considered the main risk to the attribution model that you need to manage before starting the project. For example if a customer is exposed to ten digital touch points but you capture him as two separate customers one with three and one with seven interactions with the digital channels, the result of your DA model will not reflect reality and will be inaccurate.

   Three major reasons for this issue and some examples are listed below:
   - Cross Device Tracking: A customer use his mobile, tablet and desktop computer and each of them get a separate cookie
   - Walled Gardens: a cookied customer on Google Display Network is exposed to a few ads on Facebook as well.
   - New regulations: GDPR stops Google from publishing the cookieID that is essential to identify the same customer and form a customer journey

   In summary you need to make sure you have a true multisource DA model not two or three isolated ones. For example, if you use Facebook measurement platform where attribute sales to your marketing activities on FB, you are not taking into account that the converted customers have also been targeted on other channels (e.g. email, display, SEM) and end up converting on Facebook.

3. Review the tagging processes: Note the limitations of historical

data for each of the channels and improve methods for the future.

4. Automation: DA models can be designed on campaign level, or on time period (quarterly, annually and etc.). Regardless, the collection of the data will happen frequently. Therefore it is wise to invest in automating the processes of collecting, cleaning and processing the data and avoid any ad-hoc hacky solution.

5. Identify the appropriate methodology: Last but not the least is choosing the best methodology. Next section discusses different attribution methodologies and the consideration points in picking one in detail.

ATTRIBUTION METHODOLOGIES

There are two type of attribution methodologies:
- Traditional
- Data-Driven

Traditional approaches are all rule-based and are designed based on human intuition. The most famous ones are first click, last click, even and linear. As their names suggest - first & last click attribute all the credit to the first / last touch point, even distribute the credit among all the touch points on the digital journey and linear approach credit linearly more toward the latest touch points.

The main drawbacks on all these traditional approaches are listed below:
1. They all ignore the path to conversion hence no credit is given to supporting channels;
2. It does not enable us to optimise spend over the supporting channels and measure their effectiveness;
3. Marketers have low levels of visibility. For instance, they cannot determine which placements, creatives, strategies etc. actually work.

As you see the major limitation of these methods is that each of them assumes the most influential factor for success of a media channel is something different (e.g. Even approach find frequency more important and linear approach considers recency to be the main one). Data-driven models, on the other hand, are based on machine learning algorithms and attribute credit purely based on the data. The extensive description of this methods is out of scope of this work but in short, they select the importance coefficient of each influential variable (recency, frequency, ad id, placement id, time of the day and etc.) in a way that if a predictive model is built to estimate the chance of conversion of each customer, it has the minimum error.

The most famous data-driven approaches are:
- Logistic regression
- Random forest (and other tree-based models)
- Markov Chain
- Game theory (Shapley value)

While all the data-driven approaches are more insightful and more sophisticated that the traditional ones, they each have their own cons and pros that some of them are showcased here:
- Logistic regression can also report on Ad-stock rate of each Ad but it does not take into account all the possible interactions between marketing channels.
- Random forests are one of the most advanced techniques and also enables you ,in the next step, to calculate the propensity of conversion of each unfinished path but it is high maintenance and each time a data scientist needs to set its hyper parameters hence it might not be the best case of you are making a DA model for each of your hundreds of campaigns.
- Markov chain is very low maintenance and can be scaled easily. However, it does not give you the targeting power by calculating propensity scores.
- Game Theory is another very fast approach, but it is not as scalable as Markov Chain and it does not consider the recency of the touch point on the channel into account.

In summary, it is advised to use one of the machine learning models rather than traditional ones. But the question of “which data-driven methodology works best for my organisation? “ is a very technical question that needs to be answered by a data scientist that has gone through the road map and assessed the limitations, challenges and opportunities.

SUMMARY

In the last few sections we discussed the fundamentals of an attribution model. We reviewed the model objective and presented a sample output and actionable insights that can be extracted from it. In addition, a strategic roadmap for developing a DA model was presented and the most popular methodologies were explained briefly. Lastly, it was briefly mentioned in the previous section that one future enhancement to digital attribution models is using them as targeting tools to calculate the conversion probability of unfinished paths. After all, measuring success and targeting customers are two sides of the same coin.

If we have a model that can report on the importance of each feature (touchpoints) on the customers by observing who has and who hasn’t converted, why not consider using such a powerful model the other way around and estimate the chance of conversion of each customer given the calculated importance of each feature?

CONCLUSION

We hope that this brings up-to-date the work done back in 2017 with the original Data Handbook and provides some further guidance in terms of the direction in which data management, it’s usage and identity management is headed.

Two more topics require regular oversight and we will be keeping all of our members regularly updated throughout any of the related changes.

CURRENT IAB AUSTRALIA PROJECTS IN-PLAY FOR 2020

GLOBAL REGULATIONS

We've seen changes in Europe with the General Data Protection Regulation (GDPR) and more recently in the US state of California, with the California Consumer Privacy Act (CCPA).
The regulatory landscape is shifting quickly in relation to digital advertising. We expect the outcomes of the ACCC Digital platforms inquiry here in Australia to have an impact in terms of what's expected from us all in terms of legislation in relation to the management and usage of Consumer Data in advertising.

- **IAB Australia Statement on the ACCC Digital Platforms Inquiry Report**
-
- **Data, Privacy and the ACCC Report - Peter Leonard of Data Synergies**
-
- **IAB CCPA Compliance Framework**
- **IAB Europe Guide to the Post Third-Party Cookie Era**

BROWSER TECHNOLOGY CHANGES

Alongside the regulatory changes each of the major browsers have been gradually restricting the usage of third-party cookies and identifiers and with Google Chrome announcing in early 2020 that they will also stop supporting third-party cookies this essentially started the clock ticking on the time left for third-party cookies.

The IAB's response to this has been to launch 'Project Rearc' as a project for all global IAB members to collaboratively re-architecture a replacement for third-party cookies.

IAB TECH LAB - PROJECT REARC

With the impending changes to third-party cookies and other identifiers, Project Rearc is a global call-to-action for stakeholders across the digital supply chain to re-think and re-architect digital marketing to support core industry use cases, while balancing consumer privacy and personalization.

WHAT IS BEING PROPOSED?

With the loss of third-party cookies, and potentially mobile ad IDs thereafter, the default future state of digital media will be 100% anonymous, non-addressable to third-party vendors that support advertising-funded media and services today.

The feasibility of direct addressability going forward, for any advertising-related use case, rests on trusted relationships between consumers and first parties: brands and publishers. Addressability might only exist in the future state via a consumer-provided, consented identifier tied to privacy preferences. As an example, today many brands and publishers ask consumers for an email address.

Tech Lab proposes to develop rigorous technical standards and guidelines that inform how companies collect and use such an identifier so that:

- Consumers are in control of the use of the ID and any related data. Any privacy preferences attached to the identifier are strictly followed.
- The identifier is sufficiently encrypted so that it cannot be reverse-engineered to identify the person.
- Brands and publishers have auditable, technical assurances that third-party vendors cannot track consumers on this basis without explicit consent.
- Third-party vendors are able to execute on behalf of trusted first parties, without compromising any of the above objectives.
- Tech Lab also proposes the need for standardized consumer-facing messaging, and accountability mechanisms that ascertain responsible privacy practices.

WHAT ARE TECH LAB DOING?

- Acknowledging existing discussions or practices among first parties to utilize consumer-provided, consented identifiers for addressability.
- Proposing that the industry collaborate to ensure responsible use of consumer-provided identifiers with privacy, transparency, and control.
- Suggesting that technical standards and a compliance program will be critical to ensuring that a range of addressability practices – including some employed today – are much more tightly constrained, coupled with privacy and accountability.

WHAT TECH LAB ARE NOT DOING

- Tech Lab is not creating an identifier product/service.
- We are not advocating for the broad collection, use or sharing of email addresses or phone numbers as IDs across the ecosystem. We specifically proposed this should NOT happen.

More information on Project Rearc can be found here.

APPLE WWDC 2020 UPDATES AND CONSIDERATIONS FOR APP DEVELOPMENT & MONETISATION

Although not killing off IDFA, Apple at their Worldwide Developers Conference made some major announcements that will impact app development and app ad monetisation. Many of the changes place more control in consumers but will make it more challenging for ad funded apps that are made available for free to consumers. Key changes that the industry needs to be aware of:

- App developers will be required to report their privacy practices within the App Store so users can view prior to download,
- Consumers will be explicitly asked if they are willing to be tracked across apps and sites from other companies,
- Apple will alert people to the types of data that an app might collect,
- Users will have the option to only share their approximate location with an app, rather than their precise location and, on the web in Safari

In a market like Australia that has a very high iOS market share these changes need to be reviewed closely by all companies developing apps whether they be paid or ad funded.

Below is a summary of the key changes and considerations for the market.

CROSS-APP TRACKING

Summary: Cross-tracking is a tool advertisers and data brokers often use to glean more information about people, particularly in ad-supported apps. Code in these apps and their ads allow advertisers and data brokers to follow you as you jump between apps by assigning you a unique identifier. This identifier lets an advertiser build up a profile around you and target you across a range of apps that you use. Via these new app-tracking controls, users will now be able to see what apps they've granted permission to cross-track them and revoke that permission at any time. Users, for example, will see just-in-time notifications that alert them when an app wants permission to track them across sites and apps followed by two options: "Allow tracking" or "Ask app not to track". This applies to Apple's own apps as well – and all app developers will have to follow Apple's AppTrackingTransparency framework.

Impact: Apple is lifting Limit Ad Tracking (LAT) functionality from where it's currently buried within a phone's Settings menu and calling attention to it at the moment of use, which will likely lead to more people enabling it. While good for privacy-conscientious consumers, the new notifications will be detrimental for developers that rely on an ad-driven business model and will limit reach and scale for advertisers.

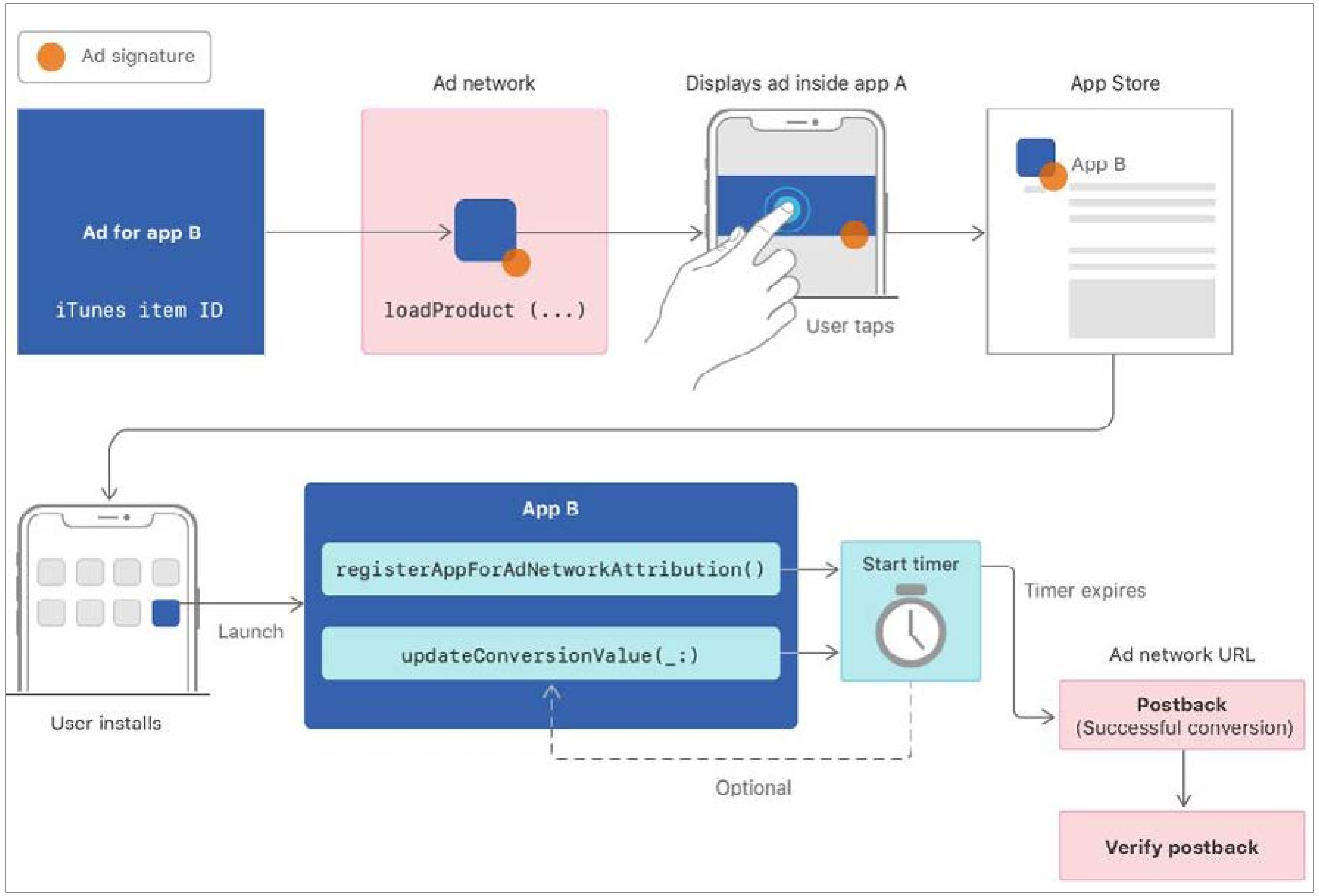


Image courtesy of Apple Inc.

Considerations: App developers will have to focus building out their own data-related capabilities for buyers and carefully consider the consumer value of any cross-app tracking functionalities. The consumer benefits and the value exchange for any related advertising experiences will have to be made very clear.

APPROXIMATE LOCATION

Summary: iOS 14 and iPadOS 14 will include optional "Approximate Location" controls, a system that prevents sharing an exact location of an iPhone or iPad. Many apps that use location to provide services, such as weather apps and local news apps, don't need to know your exact location. Instead, all they need is a general idea of where you are, which is often enough to provide the same level of service without intruding on users' privacy to the same degree.

To achieve the "approximate location" feature, Apple have divided the entire planet into regions roughly 15-20 square km in size. Each region has its own name and boundaries, and the area of the region is not based on a radius from the user as it's fixed. That means that an app can't extrapolate your precise location from approximate location data, because you aren't necessarily at the centre point of that approximate location boundary.

Impact: Increased limitations on location data for advertising.

Considerations: Developers ensuring that the value exchange for any related advertising experiences are at the forefront. Still the opportunities for stadiums, events and shopping centres can exist as long as the consumer benefits are made very clear.

APP STORE 'NUTRITION LABEL' FOR APPS

Summary: App Store listings for apps will include an easy-to-read list of privacy details so you know what data is collected before you download an app. Internally, Apple is referring to this as a 'nutrition label for apps'. It will include details on the user data an app wants to utilise across 31 categories. These labels will appear in every app's listing across all of Apple's various app stores, giving users clear upfront insights into each app's data practices – including the types of data the apps might collect, whether that data is shared with third parties, and the option for users to opt out. This won't be available in iOS 14 v1, but it is coming before Xmas 2020.

Impact: Increased focus on clear data transparency for app developers and clarity around consumer data practices. Consumer benefits and the value exchange for any related advertising experiences will have to be made very clear.

Considerations: Increased levels of consumers opting out of certain data collecting/utilisation unless it's of clear benefit to them.

DOUBLING-DOWN ON THE SKADNETWORK API

Summary: In 2018, Apple released an ad network API called SKAdNetwork (see image below) – which allows advertisers to know which ads resulted in desired actions without revealing which specific devices, or which specific people, took those desired actions. This API will now be improved to enable improved measurement KPI's for app downloads and re-downloads.

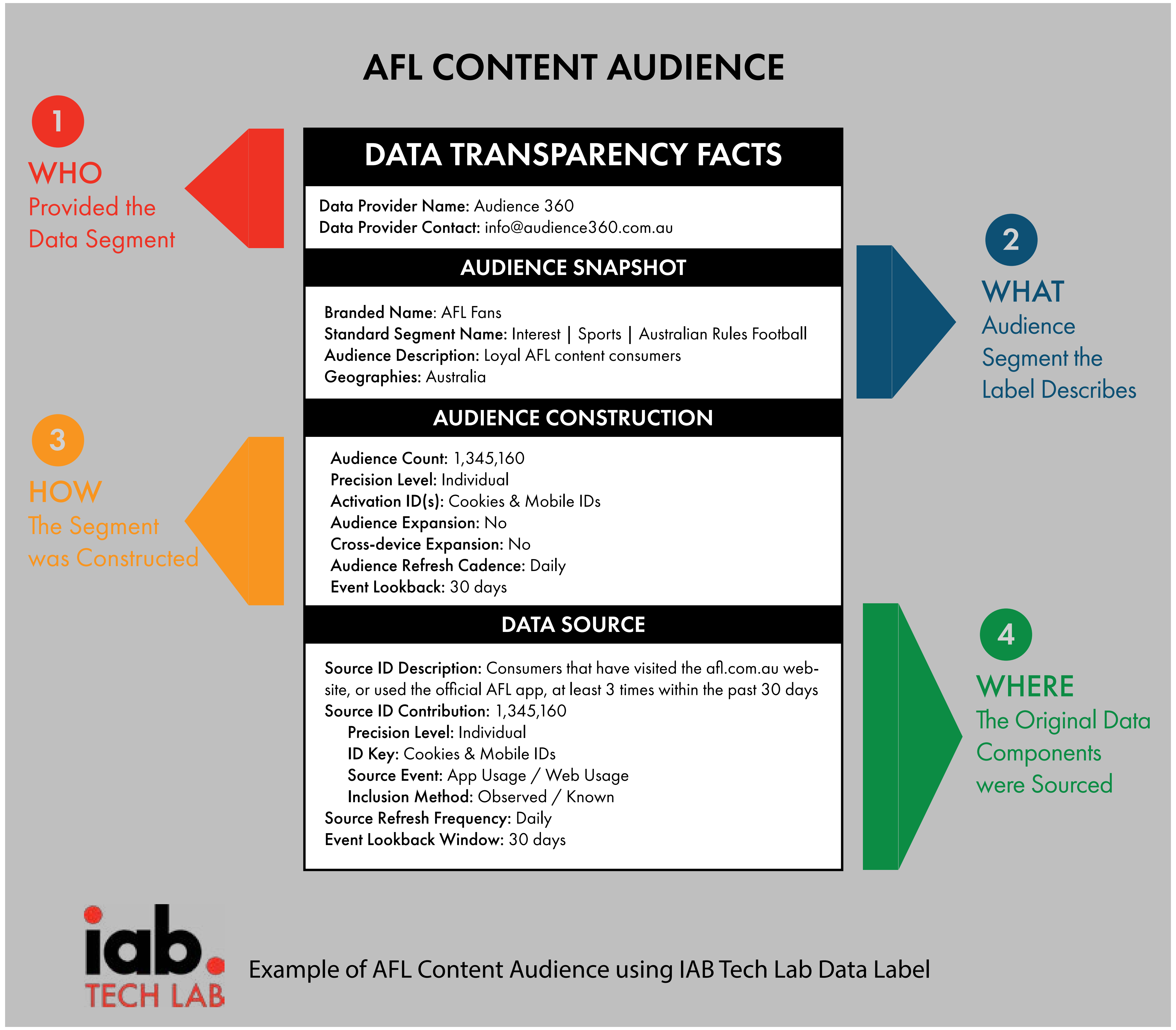
Impact: This will be attractive to marketers as this now adds extra info such as app-level attribution to measurement, but mobile attribution providers will suffer as Apple would keep all the information within its own proprietary environment – thus reinforcing Apple's dominance over the overall in-app eco-system here in Australia.

Considerations: Advertisers have until September 2020 to be ready to measure the results of their mobile ad campaigns using the SKAdNetwork API.

IAB TECH LAB'S DATA TRANSPARENCY STANDARDS AND DATA LABEL

The Data Label is based upon the IAB Tech Lab's Data Transparency Standards and is comprised of four descriptive sections designed to better inform buyers of each data set of the related critical details and displaying them as easy-to-understand ingredients:

1. Data Solution Provider and Distributor Information – Who provided the data segment, inclusive of contact information, for both data solution distributor and, where applicable, original data provider.
2. Audience Construction – How the segment was constructed, inclusive of details such as audience count, any applicable modelling or cross-device ID expansion that may have been applied, audience refresh rates, and event lookback window for inclusion.
3. Audience Snapshot – What audience segment the label describes, including both the provider's branded audience segment name as well as the most relevant segment name from a new standardized taxonomy, a top-line audience description and applicable geographic coverage.
4. Source Information – Where the original data components were sourced. Required for each significant data source, this component includes details on data provenance, data collection techniques, refresh frequency, and event lookback window.



AFL CONTENT AUDIENCE

DATA TRANSPARENCY FACTS

Data Provider Name: Audience 360
Data Provider Contact: info@audience360.com.au

AUDIENCE SNAPSHOT

Branded Name: AFL Fans
Standard Segment Name: Interest | Sports | Australian Rules Football
Audience Description: Loyal AFL content consumers
Geographies: Australia

AUDIENCE CONSTRUCTION

Audience Count: 1,345,160
Precision Level: Individual
Activation ID(s): Cookies & Mobile IDs
Audience Expansion: No
Cross-device Expansion: No
Audience Refresh Cadence: Daily
Event Lookback: 30 days

DATA SOURCE

Source ID Description: Consumers that have visited the afl.com.au website, or used the official AFL app, at least 3 times within the past 30 days
Source ID Contribution: 1,345,160
Precision Level: Individual
ID Key: Cookies & Mobile IDs
Source Event: App Usage / Web Usage
Inclusion Method: Observed / Known
Source Refresh Frequency: Daily
Event Lookback Window: 30 days

1
WHO
Provided the Data Segment
2
WHAT
Audience Segment the Label Describes
3
HOW
The Segment was Constructed
4
WHERE
The Original Data Components were Sourced
iab. TECH LAB

Example of AFL Content Audience using IAB Tech Lab Data Label

OTHER IAB TECH LAB PROJECTS IN-PLAY RELATED TO DATA

IAB TECH LAB'S CAMPAIGN PERFORMANCE METRICS DATA (OPENDATA)

This is a central standard for reporting campaign performance metrics, which provides publishers, agencies, and data management vendors a common language and mapping tool intended to improve workflow in day-to-day campaign analytics processes. Read more OpenData.

IAB TECH LAB'S CONTENT TAXONOMY (VERSION 2.1)

This allows sellers of data to more accurately and consistently describe their content and buyers can target and/or block certain content categories. Learn more here.

IAB TECH LAB'S AUDIENCE TAXONOMY (VERSION 1.1)

A standardized audience taxonomy / data segment naming convention. Learn more about Naming here.

FURTHER READING

- **MOBILE ADVERTISING AND LOCATION DATA**
  SEPTEMBER 2018
- **AUSTRALIAN DIGITAL ADVERTISING PRACTICES**
  JULY 2018
- **ADVERTISING TECHNOLOGY PURCHASE GUIDELINES**
  MARCH 2018
- **DIGITAL DATA BEST PRACTICE HANDBOOK**
  APRIL 2017