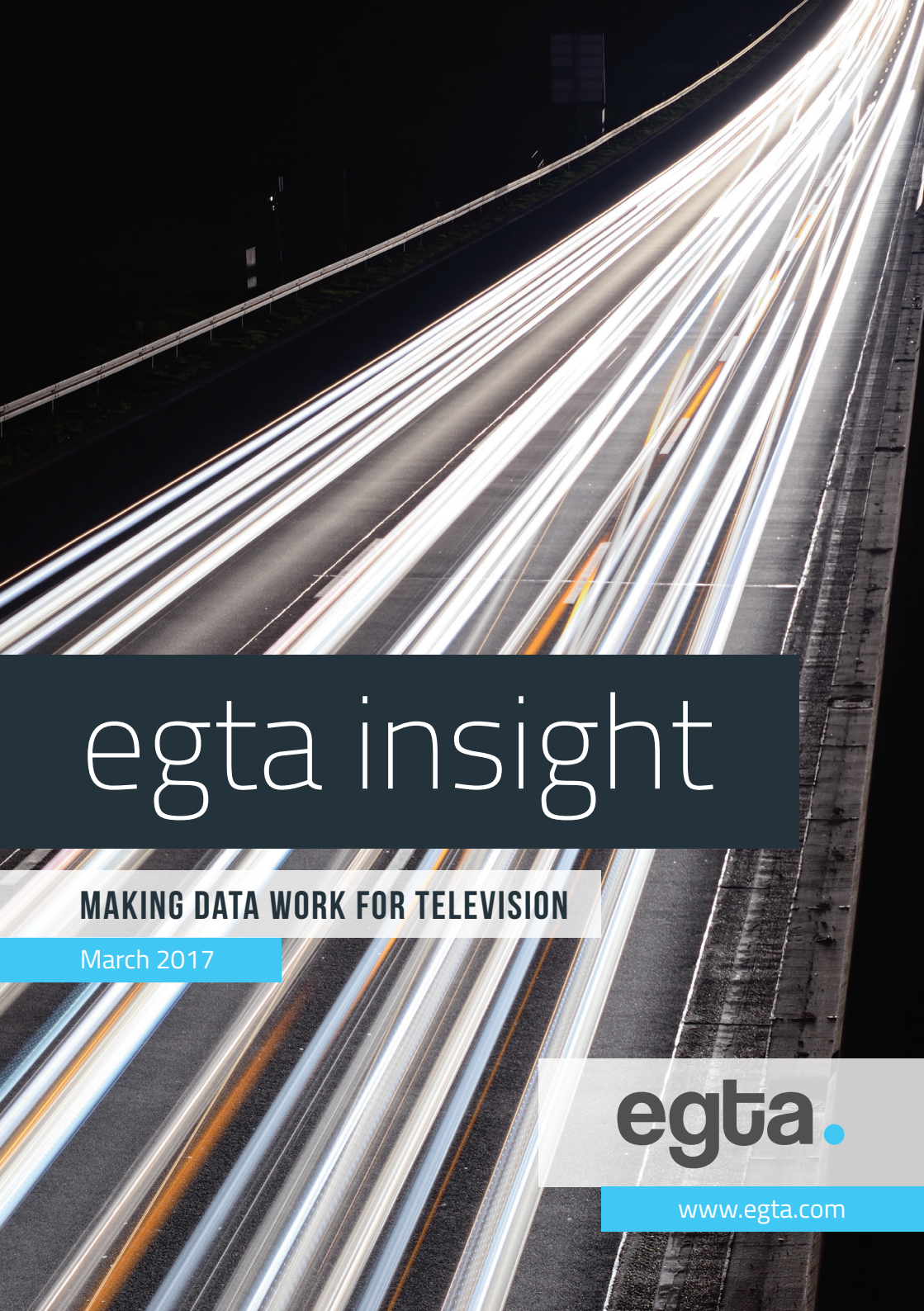


egta insight

MAKING DATA WORK FOR TELEVISION

March 2017

egta

www.egta.com


MAKE
DATA WORK
TURN
DATA
INTO
INSIGHT

TABLE OF CONTENTS

INTRODUCTION

Data, whether it be first-party data on the audience, second-party data from partnerships or third-party data from external sources, offers huge opportunities to reshape the way broadcasters interact with their viewers and to increase the value of their advertising inventories. However, harnessing these opportunities is no simple matter, often requiring large investments of resources, new ways of thinking about the business environment, new technologies and strategies. In addition, as broadcasters continue to embrace digital marketing, the challenges of fraud, viewability, ad blocking and data privacy associated with online advertising need to be addressed.

// The challenges of developing a data strategy: costs, time expectations, selection of partners and legal constraints

Media owners that have developed, or that are in the process of building, comprehensive new data strategies face significant challenges. Such projects are very costly and require time, specialised internal and external expertise and the reliance on new technology and data partnerships. Some systems can be developed in-house, in which case these costs are front-loaded, or they may be outsourced, which often involves ongoing operational costs. Data scientists with relevant technical and business skills are in high demand by all industrial sectors and are able to command compensation commensurate with their scarcity.

The traditional reliance on proxy measurement based on representative samples has served – and continues to serve – television very well for many years. However, panel measurement faces limitations in the granularity of information that it can deliver, meaning that media owners need to explore how they can collect and use their own first-party data and to consider leveraging third-party data sets to unlock targeted advertising.

Developing a new data strategy often involves a complete re-engineering of the relationship with the audience, particularly with regards to digital contacts and touchpoints. Media owners should not expect immediate results and return on their investments; for example, building the capability to acquire first-party data through a user login interface takes time, and it also requires a shift in viewers' behaviour, all of which needs to happen at scale before a significant boost in revenues through better targeted advertising can be achieved. However, as demonstrated by the companies featured in this report, the opportunities of data as the new "oil and soil" of broadcasting are now starting to positively impact advertising business models, unlocking new revenue streams and deepening the engagement between content producers, advertisers and audiences.

There are also many unanswered questions, particularly in relation to the possible impact of transitioning television advertising models closer to those seen in digital, which might, for example, account for the relatively slow adoption of programmatic TV technologies. Furthermore, selecting the right partners and separating real opportunities from optimistic promises is a daunting task.

One of the most significant considerations media owners face relates to the legal environment in which they operate. May 2018 will see the introduction of new European legislation in the form of the General Data Protection Regulation (GDPR). This will have important implications both in terms of organisations' structures, for example requiring many to hire a Data

Protection Officer, as well as the way they handle consumers' information. Preparing for GDPR now is of critical importance, as companies that fail to comply with the new rules will face fines of up to 4% of their global sales.

// The structure of this report

egta has designed this report to help its members as they consider the need to adopt new data strategies, as well as to better understand the types of data available, the tools to leverage data and the right partners to work with.

The first part of the report focuses on the questions egta members should be asking themselves when designing and implementing effective data strategies; the second part explores the data environment and the technical aspects of collecting, storing and activating data; the third part comprises of a series of case studies and interviews with broadcasters, sales houses, independent experts and technology providers, which examine approaches that have been taken and the learnings that can be derived from them.

RECOMMENDATIONS FOR MEDIA OWNERS

Based on the discussions held with industry experts in preparation of this report, egta offers its members the following series of questions to consider and recommendations that will help them define and implement a successful data strategy.

What will a good data strategy allow me to do?

- Stronger product development: programming and content management, personalisation, audience measurement and insights, etc.
- More effective marketing: retaining existing audiences, acquiring new viewers, promoting brands and products such as mobile apps, events, cross-brand initiatives, etc.
- Higher value advertising inventory: through targeting, retargeting, brand activations, direct marketing, second-party data partnerships, attribution, ROI studies, etc.
- Improved digital image and positioning towards data-savvy advertisers.

Recommendations:

» *Start with the end point: clearly define your goals and design your data strategy to achieve them. (P.56)*

» *Adopt the same language of data as your advertiser partners. (P.34)*

What steps do I need to take in order to ensure data security and compliance?

- EU ePrivacy Directive, General Data Protection Regulation (GDPR), anti-trust and other relevant legislation.

Recommendations:

» *Ensure compliance with existing data protection legislation. (P.28)*

» *Prepare for compliance with the GDPR rules set to come into force in May 2018.*

» *Have your data processes independently assessed and audited.*

What adtech/martech and infrastructure will I need to achieve the project's objectives?

- Taking into account: performance, availability, scalability, reliability, cost.

Recommendations:

» *Clearly define your objectives and engage technology partners that can deliver on your strategy.*

» *If you feel unqualified to assess a particular data source or technology, engage a specialist who is to advise on your strategy. (P.54)*

» *Define which devices you want to target.*

What kind of data sources do I already have; what additional data sources could I leverage?

- First-party data from media touchpoints: registration, website, mobile apps, video player, competitions, email and newsletters, social media presence and interactions, call centre, events, subscriptions, uploads, product sales, etc.
- Online and offline third-party data: data brokers, external DMPs, audience research, audience measurement, etc.
- Second-party data through partnerships with other publishers, social media networks, advertisers and media agencies.

Recommendations:

» *Consider registration models, including social login, to acquire rich first and second-party data. (P.29)*

» *Explore available third-party data sources to unlock better audience segmentation and targeted advertising.*

» *Consider the pros and cons of limiting yourself to your own data environment. (P.41)*

» *Ensure that your data is reliable and that it matches your needs.*

What skills and personnel will be required, and who should I involve in the data journey?

- For the data project design and implementation.
- For ongoing management of data within the organisation.

Recommendations:

» *Hire data specialists that can bridge the gap between science and business. (P.44)*

» *Ensure that involvement in your data strategy cuts across departments and specialisms and is not confined to IT; motivate colleagues throughout the company to work towards a coherent company-wide data strategy.*

» *Ensure that the data project has the support of the top levels of management.*

What opportunities may exist for developing a cooperative digital trading platform for the TV industry in my market?

- Cooperation with other broadcasters and publishers, cooperation with distributors such as telecoms operators

Recommendations:

» *Consider partnerships with other media owners and distributors on your market to unlock valuable data. (P.49)*


THE PRICE
OF LIGHT
IS LESS THAN
THE COST
OF DARKNESS
MacBook Pro

CHAPTER 01:
DATA IN THE BROADCAST TELEVISION ENVIRONMENT

Historically, television broadcasters, and the sales houses that commercialise their channels, have been able to acquire two main types of information about their audiences. The first of these, television audience measurement (TAM) data, is systematically collected by independent bodies, made available to all relevant parties on the market and used for the planning and transaction of advertising, among other things. This data, based on representative samples of the viewing population, has provided the only means for broadcasters to monetise their content through advertising.

The second source of data is through the many interactions television broadcasters have with their audience beyond the TV set. Competitions, votes, live shows and other events and so on are valuable touchpoints in their own right, but the data that they yield is sporadic and unsystematised. If collected at all, it has no commercial value beyond enhancing the connection between the broadcaster and its audience.

With the emergence of online video (OLV) in the form of broadcasters' video on-demand (VOD) in the mid 2000s, and the increasing household penetration of Internet-connected set-top boxes more recently, broadcasters have a third source of data, which can be generally classified as machine-generated return-path data. This has extended beyond interactions on desktop and laptops to mobile devices (smartphones and tablets), Internet-enabled TV devices such as Roku, AppleTV and Connected TV sets, and now includes other non-linear digital destinations such as channel websites and second-screen apps.

// New touchpoints lead to new monetisation opportunities

With the lion's share of television viewing – and indeed all video consumption – remaining live broadcast on the traditional TV screen, these new data sources have played a limited role to date, particularly when comparing Europe to the United States. However, the technologies developed and refined over the past few years are now becoming increasingly influential, with profound implications for the way television – and in particular its extensions on digital screens – is targeted and traded.

Broadly speaking, the data innovations have followed two main paths: on the one hand, data is being used to enrich the process of identifying the best audience for a given brand on linear TV, a concept referred to in this report as programmatic TV, and on the other hand data is being applied to OLV advertising inventory to harness the granular targeting capabilities of digital. The buying and selling of both broadcast TV inventory and OLV impressions can now also take advantage of the increased efficiency that can be delivered by automating workflows.

The evolution, while no less of a technical achievement, of programmatic TV is arguably less profound than the revolution in online video. Programmatic TV approaches typically retain audience measurement as the core datasets used for identifying an advertiser's required audience, layering in additional datasets for optimisation. OLV, however, can be traded on a completely different model, using the audience segmentation models and impression-based trading of other forms of digital inventory. However, both programmatic TV and OLV advertising share the objective of moving beyond age, gender and broad socio-

demographic targeting, and they use similar technologies to do so.

If, more precisely, the objective of programmatic linear TV is to identify the most appropriate audience for a particular campaign, in online video it is to identify the individual user. One of the most effective ways to achieve this is through registration and identity management systems, which may be used to partially or fully gate access to content, or as an additional identification layer alongside unregistered access.

This presents a number of challenges, including questions of whether a single currency should be used for linear TV and OLV, how to best manage cross-screen or cross-device campaigns, and how to find the balance between granular targeting and mass reach.

// The measurement challenge

TAM organisations in Europe and North America have been developing solutions to measure online video viewing – of both programming and advertising – on a comparable basis to broadcast television. Typically, the approach has been to complement television's panel-based measurement with census-level measurement of online video viewing across the whole population. As this census-level measurement only delivers information about the volume of consumption of OLV, and nothing about the identities or characteristics of the people behind the views, additional panel-based solutions are required to append demographic data to the census. The various approaches currently being developed are described in egta's January 2016 publication: Advances in hybrid television audience measurement[1].

The technology and data manipulation challenges are considerable, and progress cannot be described as rapid in most of those markets working on implementations, but the more intractable issue remains finding consensus on the most appropriate approach to currencies. That said, TAM organisations in France, Germany, the Netherlands, Sweden, the UK and elsewhere have started to deliver standardised, market-level reporting of online video consumption, representing a significant advance from the picture at the beginning of 2015.

CHAPTER 02: THE DATA ECOSYSTEM FOR TELEVISION COMPANIES

// Classifying the data available for television/video publishers

Data, audiences and analytics may be described and classified according to a number of criteria, including the source of the data, whether data is voluntarily given or inferred, whether it describes interests or intent and whether it serves to describe or predicts behaviour. The following section defines some of these characteristics.

// Classification based on ownership: first, second and third-party data

Data can be classified as first-party (publisher or advertiser)[2], second-party and third-party, and each offers different benefits within the advertising ecosystem. The determining feature is the relationship between the data collector and the individual about whom the information is being collected, which has implications for the value of the data and the uses to which it can be put.

First-party data

A "first party" is an entity that collects information from or about users and is the owner or controller of the website or service with which the user interacts directly[3]. There are broadly two types of first-party data: publisher first-party data and advertiser first-party data.

In the case of a television broadcaster, this may be names, email addresses, locations, ages and genders of people that have registered to access its online video platforms, data derived from its viewers' interactions on second screen apps or social media assets or purchase history information held in a CRM platform. Alternatively, it may take the form of machine-generated data derived from user activities across its digital platforms, such as statistics about videos viewed, articles accessed and so on.

First-party data is very valuable, as for the most part it is unique and accurate. In many cases, there are no operational costs associated with collecting the data, although this may not be the case with registration data. The owner of the data may use it solely for their own purposes (increasing the value of their advertising inventory in the case of a publisher or improving the effectiveness of their campaign in the case of an advertiser), or they may choose to make this data available to others under a negotiated deal of one form or another.

Second-party data

Second-party data can be described as someone else's first-party data. It has similar advantages, in that it is unique information about consumers, but it can add greater scale than first-party data on its own as well as the accuracy that is typically lacking from third-party data. Under an agreement between the two parties, an advertiser and a publisher can, for example, match their own data sets in order to identify attributes about their users that they did not know before. This can be used for better advertising targeting, for content personalisation or potentially as a source of revenue in the exchange of data itself.

Second-party data is a relatively recent phenomenon, and it is especially interesting as it requires actors in the advertising ecosystem to develop new, strategic partnerships, which can be very valuable in competitive marketing strategies. Second-party data is not usually commoditised, unlike third-party data, and it offers both audience extension and targeting. The data sharing may be negotiated between two or more publishers, between a publisher and an advertiser, or between two or more advertisers, providing the actors involved share some common marketing objectives. The data exchange may take place through a Data Management Platform (DMP) or elsewhere, and it may be negotiated on a simple transactional basis or under a more complex commercial deal.

One of the most useful sources of second-party data for broadcasters is that which can be collected using social login. When users register for access to a site, for example to watch live simulcast or archived VOD, using their social media profile, the broadcaster can – with the user's agreement – access certain information about them. The broadcaster can use this data, but it remains the property of the social media network in question.

Third-party data

Third-party data is aggregated and packaged information that is available for licence. In the online environment, third-party data is typically provided by DMPs or data aggregators, and these organisations do not have a direct relationship with the customer or user about whom the information is being collected. Third-party data is widely available on the open

FIGURE 01:
TYPES OF DATA THAT MAY BE COLLECTED AND USED FOR AUDIENCE SEGMENTATION

Data classification	Examples
Demographic/Socio-economic	Age Gender Marital status Income level/social grade Home ownership status Profession Education level
Psychographic	Spending habits Values Interests Personality traits Political preference
Behavioural	Website visits Video or audio streams Advertising interactions Web searches Brick and mortar visits (e.g. shops, showrooms) Purchases (online or offline) Interest (inferred from behaviour)
Geographic	Current location Places visited in the past Routes routinely travelled Planned trips
Contextual	Content metadata (descriptive information) Keywords Time of delivery / access Operating system or device platform

Source: egta - Automated & Programmatic Marketing (May 2015)

TYPES OF TARGETING THAT CAN BE APPLIED TO ADVERTISING CAMPAIGNS

Type of targeting	Examples
Targeting	Demographic Geographic Behavioural Contextual Time-based Semantic Emotional Lookalike
Retargeting	Site Search Creative Keyword

market, and while it is therefore not unique, it offers scale and targeting capabilities.

// Declared vs. inferred data

Declared data is personal or specific information that someone voluntarily shares. Examples include name, email addresses, age, gender, geographical or other information collected when a viewer creates a user profile to log in to a publisher's websites or signs up for a newsletter. It is generally considered to be of very high quality, as it is provided by the user directly, although it is not immune to deliberate or accidental inaccuracies.

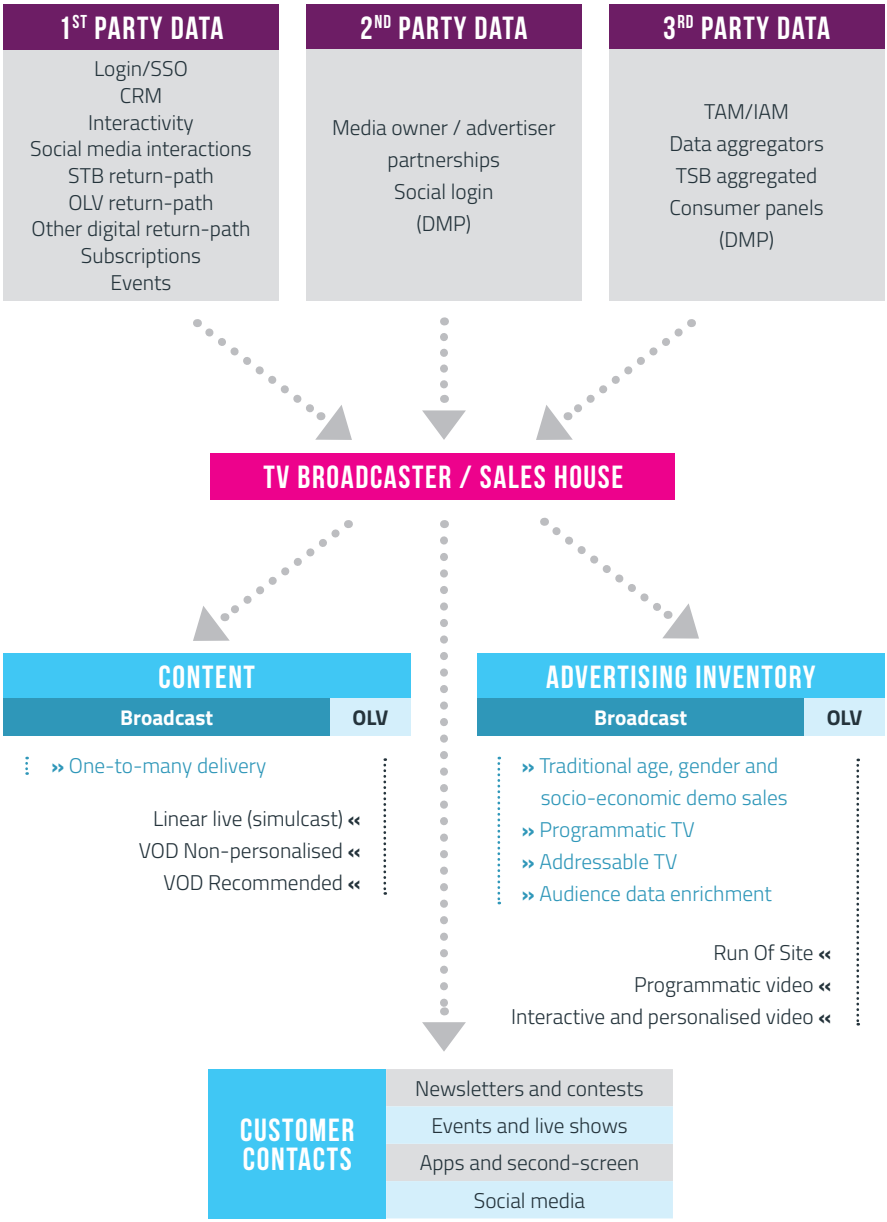
Inferred data segments are based on qualitative or algorithmic inferences, for instance by understanding a user's pattern of site visits or by more complex mathematical modelling of personality scoring[3].

// Interest vs. intent data

Interest data is information about what the user is interested in, and it may be collected by the type of content that individual consumes. It can be used to target display advertising, for example.

Intent data derives from a user's behaviour that suggests they have an intention to, for example, buy a particular product or service. This may be through searching for particular terms, comparing products or adding items to a shopping cart. This type of data is used in search engine marketing and for retargeting[4].

FIGURE 02:
AN OVERVIEW OF DATA SOURCES AND ACTIVATIONS



// Descriptive vs. predictive vs. prescriptive analytics

Descriptive analytics seeks to provide answers to the question: "what has happened?" Raw data is analysed in order to gain insights into the past, and in marketing this can be used to understand attributes about a particular audience segment and make inferences about how to target them with advertising.

Predictive analytics addresses the question: "what could happen?" This may be used, for example, to predict the future behaviour of particular groups of people or for finding new sales prospects.

Prescriptive analytics is used to offer advice on "what should we do?" These techniques are relatively complex to administer, mainly relevant to large organisations, and they leverage optimisation and simulation algorithms to predict the outcomes of ranges of possible courses of action and their likely outcomes[5].

// Sources of data and data collection tools

Depending on their digital and non-digital assets, marketing and programming activities and the sophistication of their technology infrastructures, broadcasters can extract data on their audiences (and in some cases link to the audiences of other media platforms) from a range of online and offline sources.

The broadcaster's own websites and OTT platforms

Television broadcasters possess high quality content that often cannot be accessed anywhere other than through their own platforms, a major differentiator from many digital publishers. This can take the form of short- and long-form video, articles about current affairs and editorials relating to their top programmes, games, competitions and other content.

Many broadcasters are choosing to place some or all of their video content behind registration, which allows them to collect valuable first-party data about their visitors. While registration may be enabled by a simple in-house solution, there are more sophisticated customer identification and access management (CIAM) tools available that offer better data privacy and security as well as considerable benefits to both the broadcaster and its users.

Even where no registration exists, or for those parts of a digital platform that can be freely accessed by any visitor, the broadcaster can collect a large amount of data about its users.

Web analytics

This is one of the longest established areas of online data collection and analytics. Using web analytics tools, such as Google Analytics and others, it is possible to extract information from cookies, server logs and tags (including third-party and social network tags). This can then be used to analyse the size of an audience and its dynamics, sources of referral traffic, the effectiveness of audience acquisition campaigns, user profiles and so on, as well as to connect with user IDs created from other sources. Web analytics data may also be enriched with third-party tracking data to match a publisher's visitors with wider audiences.

Video ad server

The video ad server acts as a central piece of technology in the decisioning, management, and delivery of spots to the video player, taking data in from various sources depending on the publisher's infrastructure and delivering data back to indicate whether an ad has been served, any interaction that took place and so on.

To facilitate the smooth integration of the

myriad different ad servers and video players available on the market, the IAB has developed – and continues to update – a set of standards around digital video. In simplified terms, the Video Ad Serving Template (VAST) standard instructs the video player on how to handle ads, while the Video Player-Ad Interface Definition (VPAID) protocol governs interaction between ad units and video players focused on "enabling a rich interactive in-stream ad experience"[6].

Viewer engagement, competitions, marketing and events

Television broadcasters have long had numerous touchpoints with their audiences beyond the TV screen, and these have been sharply increased with digital, mobile and social. The one-to-many nature of broadcasting is now complemented with the one-to-one communication enabled by viewer interactions, telephone and SMS voting, competitions, mobile second-screen apps, e-commerce, emailed newsletters and a host of other connections. Each such form of interaction is a valuable source of first-party data, as well as allowing for behavioural analysis of the audience.

This information is typically stored and managed using CRM databases and email services. These systems require maintenance, including the cleaning of data sets and, if applicable, integration within the publisher's wider data ecosystem.

Specialised engagement tools and competition engines are now available on the market to help publishers activate their audiences and collect additional first-party data, which can then be used in marketing or advertising activations.

Data collection on mobile apps

Mobile applications represent another important source of first-party data. Depending on settings and the opt-in status of users, this allows the generation and/or collection of the following (non-exhaustive) types of data: registration data; email addresses; mobile advertising IDs (Apple's IDFA, Google's Advertising ID, Facebook App User ID); geo-positioning; type of OS and device; activity time and usage patterns.

Registration tools

Self-registration is a cornerstone of user identification, and therefore a registration tool is a critical component of any login-based data ecosystem. Different approaches to registration mechanisms can be taken, and in many cases the registration tool will allow all of the following as options:

- Creation of a new account specifically for the publisher in question.

  Advantages: publisher retains control and ownership of all first-party data.
  Disadvantages: burden on user to create a separate new account and remember the login details; possibility for fake email addresses and/or false age and demographic information.

- Social login, in which a social media account is used as the registration.

  Advantages: easy to manage for the user; verified, accurate, granular and advanced data set from social network.
  Disadvantages: reliance on second- or third-party data; if user deactivates social account, publisher also loses access to this data.

- Single sign-on (SSO), a solution that allows to use a single registration to log in once to access a number of connected apps and platforms (for example, a single Google registration allows access to YouTube, Google Docs, Chrome, Gmail and other Google products).

  Advantages: publisher can acquire data

across a range of different touchpoints; easy access for the user.

Disadvantages: only accessible to publishers that own a portfolio of suitable assets; requires more sophisticated data processing and technology.

Registration is relatively complex and expensive to develop, and small to mid-sized publishers tend to use registration tool services from companies such as Gigya or Janrain. However, beyond a certain threshold of unique users, especially in cases where the publisher is leveraging numerous assets, it may be more cost effective in the long term to develop a solution in-house, even accounting for the additional length of development time this requires.

Social login data

While a key advantage of social login for users is simplicity, with no need to remember yet another account name and password, a major benefit for publishers is access to rich, verified and constantly updated insights about the individuals that visit its digital platforms. For example, a media owner that deploys a social login may be able to identify not only the age, gender and location of its users, but also information about their interests based on their interaction with content on the social media platform, on condition that the publisher has requested access to this data and the user has consented during the login process. Effectively, this is an automated ability for the publisher to take a look at the logged in user's entire history on Facebook, for example, see which musicians, bands and artists they have liked, or the locations they have been tagged in.

Taking the case of Facebook, as the social media platform with the largest number of users, this data can be accessed by the broadcaster and used for targeted advertising and building audience segments. However, when sharing data with app developers – in this case the television publisher – Facebook insists that the primary focus should be on improving the user experience[7].

Third-party data providers (data brokers)

A number of companies collect information about consumers that can be used by media buyers in targeted advertising campaigns. Data is not typically bought, in that no change of ownership takes place, but rather licenced for a particular use. The major players in this industry include Acxiom, Datalogix, Epsilon and Experian, and these large firms sit in a fragmented market estimated by the consultancy firm Gartner to include up to 5,000 data brokers worldwide[8].

Recent years have seen some convergence between data providers and data technology firms: for example, in 2014, Acxiom acquired the data onboarding provider LiveRamp, and Datalogix was acquired by Oracle, a matter of months after the company bought the DMP BlueKai.

DMPs, such as Lotame, also act as data market places, allowing companies to exchange their own first-party data sets, enabling them to leverage the benefits of second-party data.

This type of third-party data is useful for targeting both in programmatic (linear) TV and in online video advertising. It can also be used in addressable advertising through ad insertion into live programming, for example through Sky Media's Sky AdSmart technology (see pages 36-39).

FIGURE 03:
PERMISSIONS REFERENCE - FACEBOOK LOGIN

public_profile
user_friends
email
user_about_me
user_actions.books
user_actions.fitness
user_actions.music
user_actions.news
user_actions.video
user_actions:{app_namespace}
user_birthday
user_education_history
user_events
user_games_activity
user_hometown
user_likes
user_location
user_managed_groups
user_photos
user_posts
user_relationships
user_relationship_details
user_religion_politics
user_tagged_places
user_videos
user_website
user_work_history
read_custom_friendlists
read_insights
read_audience_network_insights
read_page_mailboxes
manage_pages
publish_pages
publish_actions
rsvp_event
pages_show_list
pages_manage_cta
pages_manage_instant_articles
ads_read
ads_management
business_management
pages_messaging
pages_messaging_subscriptions
pages_messaging_payments
pages_messaging_phone_number

public_profile

- id
- cover
- name
- first_name
- last_name
- age_range
- link
- gender
- locale
- picture
- timezone
- updated_time
- verified

Source: Facebook for Developers/Docs/Facebook Login

CHAPTER 03:
THE TOOLS AND TECHNIQUES TO ACQUIRE AND EXPLOIT DATA

Capturing, processing, storing and activating data involves a range of technologies, deployed by advertisers and media agencies at one end of the marketing chain (buy side) and media owners at the other (sell side). Some of these technologies, such as those allowing access to third-party datasets, serve both the buy and the sell sides.

For more information on the leading providers of advertising technology and data solutions to television and radio companies, please consult egta's interactive AdTech Navigator platform at adtech.egta.com.

// Data storage, processing and activation

There are several technologies that can be used to collect, organise, store and use data, and in many cases there is a degree of overlap in functionality depending on a given provider's offer. The following section looks at these processes through the lens of the Data Management Platform, but they are not all exclusively carried out by that particular component of the adtech stack.

// The data warehouse

A data warehouse allows an organisation to aggregate data from multiple sources so that it can be analysed and used to improve business intelligence. It may also be used as the basis on which to activate data for operations such as recommendation engines. A data warehouse typically handles incoming information at the enterprise level, in contrast with the data mart, which is typically oriented to a specific business line or department (sales, marketing, finance, etc.)[9]. Data warehouses may be built in-house, however there are a number of companies that offer the compatibility to analyse large data sets, some of the best known being Amazon RedShift, Vertica (Hewlett-Packard), and Teradata.

Data is moved from its source into a data warehouse using a trio of processes called ETL (Extract, Transform, Load). This ensures that data is properly formatted and normalised. The three processes are commonly carried out in parallel.

- Extract: data is collected from sources, validated and stored in a temporary database;
- Transform: data is processed to ensure all data conforms to common formats;
- Load: the transformed data is moved into the permanent, target database.

// The Data Management Platform

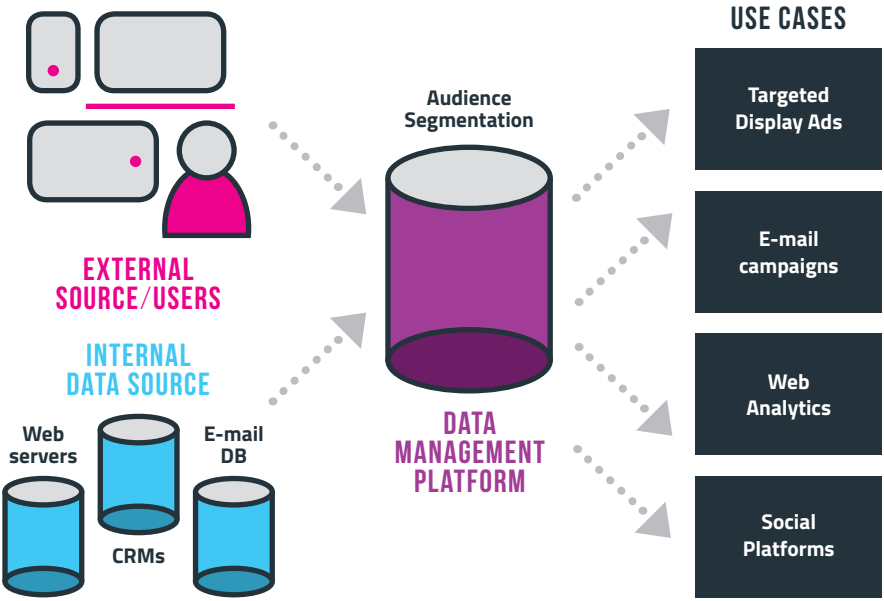
The Data Management Platform (DMP) is a component of digital marketing that can be used by publishers and advertisers. It has been described by [x+1], a pioneering DMP, as follows:

A DMP "… provides the requisite, yet somewhat unnoticed, function of data collection, translation, classification, indexing and storage. It's the 'plumbing' part of data-driven marketing online."

One way to understand the role of a DMP is to consider its four main functions[10]:

1. Import data from users and other data sources;
2. Group data into segments (audiences), which share common attributes;

FIGURE 04:
DATA MANAGEMENT PLATFORMS



3. Send instructions to place ads, usually via a Demand Side Platform (DSP). This is a combination of who to target, with what message and (in some cases) on what channel or device;

4. Measure the impact and improve the instructions.

In digital marketing, the DMP is used to collect, store and analyse data, generating audience segments that can be used to deliver targeted advertising to users. It is a smart data warehouse that can unify diverse data sets, aggregating information from a publisher's online, offline and mobile channels, organising it and making it actionable. The DMP can also be used to connect to third-party data sets or even to trade data between organisations in a second-party data relationship.

// Data collection

A DMP can collect data in three ways. It can receive data that is sent to it by the publisher (onboarding, see below), it can collect information from a publisher's digital properties using data collection tags, and it can use APIs to collect server-to-server data.

A publisher adds collection tags to its website to collect information about what visitors are doing and how often. Activity tags track the things users do on the site, while media tags are used to track media impressions and clicks. Activity within iOS and Android is tracked using mobile SDKs, which enter code into the app itself and deliver data to the DMP.

// Onboarding

This is the process through which offline data is transferred to the digital realm[11], and it offers a way to connect data between different silos. Customer records, for example from a publisher's CRM system, are uploaded, anonymised (personally identifiable information (PII) removed in a process known as hashing) and matched to online devices and digital IDs. Data segments can then be activated for use in marketing applications and media platforms.

One of the leading onboarding providers, LiveRamp, explains its service as resolving identity between offline CRM data and digital devices, but also integrating data from any consumer touchpoint that any stakeholder in the ecosystem might have. This allows a richer understanding of the consumer to be developed[12]:

1. Unifying everything a company knows about its users (resolving all first-party data together). This requires identity resolution.
2. Augmenting that understanding with what other organisations know about those users (bringing in second- and third-party data from different data owners).
3. Activating against that data, using it to create audiences and segments, reach users on whichever touchpoint is going to be most appropriate for them.
4. Taking the interaction data from users' engagements and using it to develop a better understanding of that customer.

This process has become much quicker in recent years; a process that could take days or even weeks can now be done in as little as an hour by some providers.

// Storage and organisation: IDs and attributes

A DMP uses IDs and attributes to identify information about people and categorise them as distinct user profiles that can then be activated. These categories are also known as data taxonomies. IDs may, for example, take the form of customer IDs, platform IDs, device IDs or DMP IDs, each of which usually represents an individual user. As one user may have many IDs, the DMP requires a key to map one ID to another and develop a coherent record associated with each individual.

Attributes describe what is known about these users, and these may be binary in form (male/female), categorical (high income/medium income/low income), or they may represent a narrow range of values (0 = aged <18, 1 = 19-25, etc.)[11].

// Cross-device identity

With the rapid increase in smartphone and tablet adoption, new technologies have been developed to identify users as they move between desktop and mobile environments. As the Internet of Things (IoT), wearables and connected TVs start to emerge as more significant platforms, the number of devices from which a user's data may be derived is steadily increasing.

The dominant tracking vehicle of the desktop, the cookie, is unreliable in the mobile web and ineffective in the app environment, as in-app cookies cannot be shared between apps. To overcome this limitation, two methods have been developed to track users across devices – deterministic and probabilistic matching.

Deterministic matching creates a cross-device

linkage when a user is logged onto the same service via a website on one device, for example, and an app on another. This use of personally identifiable information (PII) to make a match between devices is largely the preserve of companies with large, logged in user bases across desktop and mobile assets, such as Facebook, Twitter, Google and Apple. There are questions about the privacy issues related to this type of tracking, as well as the walled garden nature of the Internet giants' data universes. Deterministic data sets are likely to have higher accuracy than probabilistic sets, but it is much harder to achieve scale, and therefore advertising reach. Even Google and Facebook only count a percentage of the online population among its identifiable users, and in many cases people only log in to a service on one of their devices, making a deterministic match impossible.

Probabilistic cross-device matching relies on statistical modelling to create likely matches between devices, based on pseudonymous data points that arise from device usage. This observation data may include IP addresses, GPS data, browser cookies, mobile IDs, Wi-Fi networks accessed, operating system, time of day, and many other things beside. By avoiding PII, or in some cases using smaller, core PII data sets that be modelled to create larger audiences, probabilistic offers scale. It is also potentially less problematic from a privacy perspective, as the device matching occurs without the user ever being identified from PII[13].

// Audiences (segments) and targeting

Within the DMP, an audience is a group of users that share a common set of attributes. Belonging to that audience determines what actions will be taken with respect of that user: delivering targeted advertising; personalising websites and other digital assets; sharing with other systems (syndication).

The attributes used for a targeted advertising campaign may be first-party or third-party. First-party attributes are based on the publisher's own data about its users, for example activity on their website or some other forms of engagement. Third-party attributes are data acquired from data vendors, and could for example be a list of anonymised users that have recently visited a particular type of website.

A targeted campaign will also include other parameters, such as blacklisting certain websites, budget, objectives and so on, as well as pre-determined criteria for measuring the campaign's success. Success may be measured by Click-Through Rate (CTR), site landings or interaction, or whether a user has converted and purchased something.

// Targeting based on third-party data

For a television broadcaster (publisher) with a relatively limited number and variety of digital assets, it is possible to use a DMP to licence third-party data in order to build audience profiles for targeting and then to deliver campaigns against this data. In the absence of rich first-party registration data, and where the publisher is only collecting cookies from their websites, connection to a DMP is required to have access to third-party cookies and to widen the targeting opportunities.

One drawback of the use of third-party data is that its quality may be low, it may not be up-to-date and it may not be fully relevant. For example, third-party data might be used to target a campaign for a car maker to people by 'auto' attribute; however, this audience could include people who are interested in F1

racing, for example, which is not the equivalent of a 'people in market for a car' segment. This reduced relevance will lead to a drop in the performance of the targeting, and first-party data is therefore very important and valuable.

// Targeting and retargeting based on second-party data partnerships

Extending the example above, if a television publisher were to approach a car maker offering to enter into a second-party data partnership, further targeting opportunities would become available. By using a DMP to match users, the publisher could identify and target video spots to people already known to the car maker, for instance customers and others appearing in its CRM system. Alternatively, the publisher could re-target spots to any user on whom the car maker has first-party data, such as people who had previously been served a display ad for that brand. Furthermore, the car maker could use the publisher's first-party data to retarget any form of digital advertising to users that had been previously exposed to a video campaign on that publisher's online platforms.

// Lookalike audiences

The DMP can leverage third-party data and algorithmic modelling to build audience extensions (also described as audience expansion). An advertiser, for example, may use lookalike audiences to identify people that share similar characteristics or behaviour patterns with its best customers and target their advertising messages to them, as they have a higher probability of conversion.

The leading DMPs in the field use machine learning – a form of artificial intelligence – to create new lookalike audience segments that can be continuously refined and updated as the data sets grow[14].

// Delivering a personalised experience

The DMP can be used to make decisions about what content should be delivered to a particular user, based on the same principles as for targeted advertising. This may be content on a website or app, or it may be used to trigger other forms of communication such as mobile SMS messages or emails.

Content personalisation and recommendation can be used to ensure a publishers' visitors are presented with content that is more likely to fit their interests, based on the profiles developed by the DMP.

// Buying and selling inventory: SSP and DSP

While the DMP plays a critical role in ingesting, sorting and activating data, it requires connectivity to other marketing technologies to turn data into actions. Some DMPs remain as standalone platforms, and the companies that offer them promote the benefits of such an arrangement, while others have integrated DSP capabilities or vice-versa. Along with DMPs, DSPs and SSPs perform the core functions of automated advertising buying and selling, a central tenet of programmatic marketing.

DSPs are used by agencies and advertisers to buy digital inventory across search, display, video, and audio either on desktop or mobile devices (or both). A DSP may be a standalone product or integrated within a provider's wider technology offer, and it may offer access to several ad networks on an impartial basis or to just one major network, for example those of

Yahoo! or Google. In programmatic advertising, the DSP makes decisions on what impressions to buy on an automated basis according to the advertiser's pre-set requirements, and this automation extends to pricing in the case of RTB.

If the principle objective of the DSP is to allow advertisers to buy the impressions that best fit their requirements with the greatest cost efficiency, the SSP is designed to make a publisher's inventory accessible to buyers and to maximise the value that can be extracted from this inventory. An SSP connects a publisher's inventory to ad networks, ad exchanges and DSPs, and the publisher can set rules such as price floors and to determine which advertisers can buy their inventory.

As the technologies and commercial offers of providers have developed, it may be difficult to distinguish between an ad network, an SSP and a DSP[15]. Some of the leading providers in the video advertising space position themselves as serving both the buy and the sell sides – Adform, for example, offers both a DSP to serve agencies and advertisers and an SSP for publishers, while Videology also offers solutions to agencies, brands and media owners.

Companies that offer a video SSP as part of their technology stack also include WideOrbit, SpotX, FreeWheel, Ooyala and Improve Digital, all of which offer different suites of solutions to help broadcasters better monetise their video content. Further information on these and other technology companies can be found on egta's AdTech Navigator platform at adtech.egta.com.

// Programmatic TV

Defining programmatic TV is difficult, as it means different things to different people. Most commentators agree that it is a technology-automated and data-driven method of buying and delivering ads against TV content[16]. Alternatively, it may be seen as a solution that allows advertisers to buy and target TV advertising through software[17]. Where definitions tend to diverge is when it comes to the destinations those ads may be served and the degree to which addressability is a part of the package.

Programmatic TV technologies may be applied to linear television inventory served in the traditional way, without any addressability at the household level. In this *context-based audience targeting*[17], software capabilities and data are used to achieve the best mix of networks, dayparts, geographical areas and so on to reach an audience that over-indexes for the desired target audience. Where appropriate STB technology that allows TV ad insertion at the household level is in place, *addressable-based audience targeting* can leverage third-party data to deliver different ads to different households. Both of these options take targeting beyond traditional age, gender, socio-demographic, reach, and frequency components of targeting based on GRPs.

The definition of programmatic TV can be extended to combine linear TV with digital ads served across the web, mobile devices, set-top boxes and connected TVs, offering a centralised platform for brands to target their audiences across multiple screens.

It is worth noting that the television landscapes and the regulatory regimes regarding data use are markedly different in Europe when compared to the US, where many of the advances in programmatic TV have been made. In particular, US multichannel video programming distributors (MVPDs), such as Cablevision, Comcast and Dish Network, are able to sell a portion of local TV advertising (generally two minutes), and between them they can serve addressable ads to a significant proportion of the population.

This combination of technology, advertising sales models and access to third-party datasets does not currently exist at the same scale on European markets.

CHAPTER 04: CHALLENGES ASSOCIATED WITH THE USE OF NEW DATA SOURCES

// Ad fraud

While estimates of the level of ad fraud vary, the issue is widely recognised as one of the major challenges the digital industry faces. Bot Baseline, a high-profile research initiative led by the Association of National Advertisers (ANA) in the US estimated the annual cost of ad fraud to be $7.2 billion in 2016, roughly 5% of the total digital media market[18]. As recently as December 2016, a huge click-fraud scheme perpetrated by a group of Russian hackers, dubbed Methbot by White Ops, the security firm that uncovered it, was reported to be stealing between $3 and $5 million daily from top advertisers and publishers[19].

The Methbot fraud targeted high-value video advertising inventory, including that sold through private marketplaces, highlighting concerns that the high CPMs associated with this form of advertising make it particularly attractive to fraudsters. Indeed, in its comprehensive report on the subject, the World Federation of Advertisers (WFA) noted that almost any programmatic buy can be exposed to ad fraud, even including direct programmatic TV. The WFA advises to treat any claims to the contrary with caution[20].

// Ad blocking

Ad blocking affects publishers to different extents, depending on the type of content they offer, the technology they use to deliver advertising and the strategy – if any – they put in place to detect the presence of ad blocking software used by visitors to their online platforms. Some television broadcasters, such as Channel 4 and ITV in the UK, prevent users with an active ad blocker from watching content on their websites, while others deploy circumvention technologies, such as server-side ad insertion or video player plug-ins and SDKs, to ensure that advertising cannot be disrupted by ad blockers. A few broadcasters, such as TF1 in France, implement an obligation to switch off ad blockers for some of their flagship programming.

It should be noted that ad blocking software primarily affects access through desktop browsers, although it is a growing issue in the mobile operating system environment as well. These technologies do not affect the delivery of linear programmatic TV or household addressable television advertising.

For further details on this subject, please consult egta's June 2016 publication Ad blocking: strategies for television and radio companies[21].

// Strict new data protection rules to come into force in May 2018

A new General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679) comes into force from 25 May, 2018, replacing a set of rules that have been in place since 1995[22]. This legislation will apply to media companies and other organisations throughout the EU, and in broad terms it will give consumers more control over how their data is used and retained.

The potential liability for companies will also be extended. While under the existing rules, only the data controller is liable for data breaches,

the liability for any damages will extend to both the data controller and the data processor. As a simple example of this differentiation, if a media owner outsources the management of its customer databases to a third-party company, the media owner itself is the controller and the outsourcing firm the processor. There will therefore be a greater obligation for companies such as media owners to consider the privacy implications of their products and services, as well as to carry out privacy impact assessments. It will be mandatory for many organisations, including public companies, to have a Data Protection Officer (DPO).

The penalties for companies that breach the new rules include fines up to 4% of global sales.

People will have the right to have their own personal data (PD) corrected if it is inaccurate, and the regulation expands their right to remove irrelevant or outdated information, the so-called "right to be forgotten." Consumers will also have a right to stop marketing companies from building a data profile of them.

Some of the changes to EU data protection laws that will be relevant for publishers and media include:

- A right to be forgotten;
- "Clear and affirmative consent" to the processing of private data by the person concerned;
- A right for users to transfer their data to another service provider;
- The right for users to know when their data has been hacked;
- Ensuring that privacy policies are explained in clear and understandable language;
- Stronger enforcement and fines up to 4% of firms' total worldwide annual turnover, as a deterrent to breaking the rules.

Updating the ePrivacy legislation

In January 2017, the European Commission also proposed a revision to the existing ePrivacy legislation, which would convert the current Directive into a Regulation. In part, the objective is to correlate better with the GDPR when it comes into force.

Among other things, the existing ePrivacy Directive governs the rules for obtaining users' consent for the collection of data on their online activities, for example by the use of pop-up banners that inform visitors about the way a publisher collects data and offer an opt-out. Informally, the rules are known as the cookie directive for that reason. The proposed new rules will change the way publishers have to obtain consent and give consumers more control over the way their data is collected and used. Including the ability to easily make choices when users set up their browsers or change their settings is something the Commission refers to as "Privacy by Design"[23].

CASE STUDIES

This report contains a series of case studies, with summarised leanings, that were selected to offer the opinions, insights and recommendations of industry experts covering different perspectives, both from the media owner and technology provider sides.

PREPARING FOR THE GENERAL DATA PROTECTION REGULATION (GDPR)

Didier Angelo, *Legal advisor* – **TF1 Publicité**

egta: *When GDPR comes into force, what do you expect to be the main implications for media owners in terms of the ways they collect, store and use data about their viewers, listeners and visitors?*

Didier Angelo (DA): Media owners will need to give users back the control of their Personal Data (PD). This means giving users an easy access to policies and other documents relating to personal data protection, and making them understandable. Media owners will need to track their actions relating to users' requests to exercise their rights (right to access, right to know the purpose of the data processing, right to rectification, right to object, right to be forgotten, right to data portability).

They will need to make impact assessments for activities which are likely to jeopardise users' rights and individual freedoms and implement technical and organisational measures to secure users' PD.

egta: *How would you advise media owners to address their internal operations and strategies to be prepared for the GDPR?*

DA: They will need to increase the awareness of business teams on the regulatory changes and the increase of financial penalties, mobilise all departments of the company (Legal, Marketing, Information System, Product, Audit, …) that deal with PD, each one in its own area of expertise, in order to bring their activities into line with the GDPR.

They will also need to designate a competent person to hold the role of Data Protection Officer (DPO).

egta: *What type of legal questions should media owners be addressing to advertising technology and data providers as they further develop their digital presence?*

DA:

- When were you audited for the last time, in connection with PD protection (particularly data breaches)?
- How are you going to oversee the next PD transfers?
- Are you certified according to a mechanism approved under art. 42-43 of the GDPR?
- Have you signed a code of conduct approved under art. 40-41?

DEVELOPING A DATA STRATEGY BUILT AROUND LOGGED-IN VIEWERS

Martin Greenbank, *Head of Advertising Research and Development* – **Channel 4**

Channel 4 is an advertising-funded, public service broadcaster in the UK that has a long-standing reputation as an innovator, both in terms of the content it provides to its audience and the platforms through which it is delivered. Channel 4 was among the first European broadcasters to bring a VOD offer to the market, launching 4oD in November 2006.

Operating as a free A-VOD service almost from its inception, the platform has evolved over time, and it now offers access to a wide range of on-demand content, as well as live streaming simulcast of Channel 4's free-to-air channels, through a wide range of connected devices. 4oD was substantially rebuilt and relaunched as All 4 in March 2015. In addition to the web-based platform, All 4 is available on Android and iOS devices, game consoles and Connected TVs, and it is available on TV platforms such as Sky, Virgin Media, Freesat and YouView.

A new data strategy

Channel 4's current data strategy for its digital environment can be traced back to a speech made by David Abrahams, the broadcaster's then incoming CEO, in 2011: *"Central to our future ambitions will be to add a 4th dimension to our proposition in which viewers become a bigger part of the conversation than ever before. As we enter this decade of connected viewers, C4's mission will be to give people an even more personal and unique experience. We know Channel 4 already generates deep engagement with its viewers. Our new plans will exploit technology to bring them far closer to the programmes and people they love than was ever possible before."*

At that stage, the broadcaster had a successful digital platform, but no solid data strategy behind it, and there was no coherent way to collect data about visitors or use that information to improve the experience.

Right at the start of its new data strategy, Channel 4 made its *Viewer Promise* a central component. This explanation of why it collects data from users, what data it collects and what it does with it, as well as how users can control their own data held by Channel 4, was supported with a short video in very plain English that helped users understand what the broadcaster was doing with their personal information, and crucially this approach helped to build trust.

Mandatory registration allows the collection of first-party data

The archive of long-form content, comprising more than 9,000 hours of programming, and live streaming require mandatory registration, which is free. Users submit their name, email address, date of birth, gender and postcode, after which they can either access the platform using their user name and password or through social login. Users have access to a dashboard where they can see the personal information held and control their preferences over how their data is used. By the end of 2013, Channel 4 had reached 10 million registered users, including

half of all 16-24 year olds in the UK.

Having established the online platform and the data collection side of the operation, Channel 4 then focussed on developing a data-infused advertising offer. The broadcaster carried out what they refer to as a triple audit in 2012/2013 to verify the accuracy of its data and the accuracy of its methodology for targeting advertising based on first-party data. This involved trials with seven advertisers, verification against third-party data sets and an external audit by PricewaterhouseCoopers. This approach, which was unique in its robustness on the UK advertising market, provided the verification required to start selling targeted video advertising through 4oD (and later All 4) at a premium CPM.

A new way to buy VOD against demographic audiences

Channel 4 brought the buying of online video inventory in line with the way linear TV is traded. In addition to the possibility to buy run of site (ROS) or genre on Channel 4's digital platform, advertisers could now buy against the actual audience they wanted, for example a young male or female audience or social grade ABC1. The trials demonstrated significant uplifts in metrics such as brand awareness and intent to purchase, justifying a 30% price premium on online video inventory bought on demographics rather than ROS or genre. This capability is built on Channel 4's first-party registration data, which initially covered those platforms with the largest chunk of viewers (the web and mobile platforms, for example) and has now been rolled out onto other devices as well (such as Xbox game consoles). Data modelling can be used to predict audience characteristics where registration data is lacking.

Advertising can also be personalised to the individual user level. Channel 4 created the world's first personalised TV ad in summer 2014 for Coca-Cola's 'Share A Coke With' campaign. Delivered to 11 million 4oD users in the 16-34 target, each ad was adapted to include the user's name (as provided at registration) on a Coke bottle at the end of the spot. The campaign generated exceptionally high view-through rates, engagement, awareness, ad recall and purchase intent.

At the start of trading in 2014, Channel 4 was selling 15% of its VOD inventory as data-enabled (using first-party data), rising to 30% in 2015 and about 50% in 2016. In terms of its overall viewing, VOD represents about 5% of Channel 4's total hours, which is above the average for UK broadcasters. While all of the inventory is now technically addressable, in reality some clients will still want to buy ROS or genre, and registration has still to be rolled out to some of the platforms content can be accessed on. The business model has proved hugely successful, and having broken even on the investment Channel 4 saw a positive return in 2016.

Pioneering entry into programmatic video

Channel 4 has also developed a programmatic exchange for its VOD inventory. The Premium Video Ad Xchange (PVX) allows advertisers to buy audiences programmatically on All 4 across all devices and platforms, based on Channel 4's first-party data. PVX is an extension of Channel 4's earlier move into programmatic video, a private marketplace developed with FreeWheel as its technology partner in an exclusive deal and initially launched with Amplifi, the centralised media investment arm for Dentsu Aegis Network. Channel 4 built this programmatic infrastructure itself, as no off-the-shelf technologies existed with the capabilities they needed.

In addition to demographic targeting, Channel 4

offers interest-based audience segmentation, for example home, beauty and fashion enthusiasts, both through direct sales and the programmatic platform.

Dynamic Ad Insertion (DAI) has already rolled out on Channel 4's VOD platforms. The technological feat of DAI in linear broadcasts requires further advances in technology at broadcasters and in-home platforms, alongside developing new techniques for 'real time' audience measurement. R&D is currently hotting up in this area, and it is probably now years rather than decades away from reality.

egta spoke to Martin Greenbank, Head of Advertising Research & Development at Channel 4, about the broadcaster's data project.

egta: *Where did you start from in terms of data with your VOD platform?*

Martin Greenbank (MG): We had a platform, 4OD, but it didn't have a solid data strategy behind it, it wasn't cohesive in how it collected data about viewers or exploited that to make the experience better. It literally treated everyone who arrived at the site as an equal party.

And then after they left the platform, we couldn't piece together their viewing with that taking place through any future sessions. There was no login at that point, but there was some knowledge of some of our users. But it was literally using things like competition entries and other bits of information. We had some visibility of some of our viewers, but it wasn't anywhere near comprehensive or joined up at that moment. That was the moment the strategy came into place.

egta: *What was your experience with encouraging users to register?*

MG: It took us twenty-two months to go from zero to 10 million, which is a faster rate than Facebook and Twitter managed to achieve for their first 10 million registered users. Clearly, with Channel 4 we had a known entity, but there was something I think we were getting right with the Viewer Promise.

egta: *What was the rationale for building the capability for verifying your user data and targeting methodology in-house, rather than relying on a third-party verification service?*

MG: Part of it was the fact that we *could* do it ourselves. As David Abrahams stated in his speech, data is the new oil and soil of television; it's how you grow television's traction. We were awash with data in the TV industry, we were very data literate. The extension into data science was an investment that I think Channel 4 made because we're Channel 4: we're brave enough and we're set up in a way that we are there to challenge the way things are normally done. We are also acutely aware that a third-party company has their own commercial roadmap ahead of them, and we wanted to be in charge of our own destiny.

egta: *Is there an appetite for any cooperative digital trading platforms for the TV industry in line with initiatives like the Pangaea Alliance for press?*

MG: There is definitely potential for it. We have been concentrating first and foremost on getting our own product right, improving its performance and taking it to market. But with PVX we have created a platform which third party advertisers could in theory plug into, as the integrations with the buy side have already been created.

Having been on the buying side for many years, what is possible and therefore expected, often takes more time for broadcasters to develop and implement. When you're on the inside you realise the amount of machinery that has to change for new things to happen. And the market is not structured for quick changes,

the TV market especially. The way that TV is traded on an annualised basis isn't conducive to throwing in new things every quarter. It just doesn't work. You can try new things, but to get a scaled proposition you maybe have just one opportunity a year.

egta: *Where do you see data going in terms of linear TV?*

MG: There's a phrase that's being talked about much more than it has been in the past, which is addressable linear advertising. That means a number of things. People talk about linear programmatic, but it's not just that, it's really the ability to make smarter decisions and increase yield, efficiency and effectiveness of TV campaigns through the use of data.

As more and more TVs become connected, there is the opportunity to address them at a household level, or even an individual person level. As that opportunity grows, it unlocks more commercial opportunity. What we've done in the VOD space, we often term as 'the nursery' for what is likely to happen in the linear space. In terms of timescales, you'll have some people saying we might be only two or three years off, you have other people saying we're ten to fifteen years off. I'm probably somewhere in the middle there!

CASE LEARNINGS:

- Being transparent about data collection and usage creates trust with users
- Alongside a compelling content offer in exchange for the user's data, this motivates people to register for access
- First-party data allows audience buying on demographics and interest, not just run of site or genre
- If you cannot do DAI in linear online video, you lose potential impressions
- Programmatic video brings VOD into line with wider digital media buying, but it requires more than an off-the-shelf solution
- Getting the data strategy right is difficult: it has to be the number one priority
- People with data skills are a scarce resource and in high demand: having an attractive brand and interesting data helps bring the right personnel in
- Media markets are not structured for quick change, so improvements in the way television is traded are incremental
- What happens in the VOD space can act as a nursery for linear TV

IMPROVING TARGETING AND PROVING EFFICIENCY WITH ONE DATA

Nadège Verdurmen, *Marketing Insights, Data & Strategy Director* – **TF1 Publicité**

One Data is an initiative by the French sales house TF1 Publicité. One Data takes a new and innovative approach to analysing and using information about consumers and their TV viewing, allowing TF1 Publicité to harness the data revolution.

One Data was launched both to reinforce the message that TV can – and should – take its rightful place in the world of data as a precise and efficient medium, as well as to help TF1 Publicité talk the same language as today's increasingly data-focussed advertisers. Advertisers now use CRM, DMPs, social media and their own online properties to better understand their consumers, and it is therefore imperative that broadcasters and their sales houses step into their clients' world of data.

A three-step platform

One Data is composed of three parts:

- **Targeting** – using data to optimise campaign media planning and delivery;
- **Efficiency** – understanding the impact and ROI of TV advertising;
- **R&D** – identifying the data opportunities of the future.

1: Targeting

TF1 Publicité has developed a new approach to understanding and talking about consumers, moving beyond age and gender demographics to purchaser and decision-maker targets. The conversation that in the past has been about describing viewers by their basic characteristics therefore becomes one about *buyers*, and even one step further to *purchases*.

By merging data from two measurement panels – Médiamétrie for TV audience figures and Kantar's Worldpanel of consumer purchases – TF1 Publicité can identify not only who is watching a particular spot on all channels, but also what makes of car they own or what brands of shampoo they buy, for example. This allows the sales house to offer GRPs based on product sectors and market share data, rather than on simple demographics.

This technique has in the past been used to analyse post-campaign impact, but it has now been integrated as an option into the media planning tools, covering 48 sectors (44 FMCG sectors plus automotive, banking, insurance and telecoms).

The enhanced targeting capabilities of One Data also allow TF1 Publicité to use third-party profiling or behavioural data to improve the targeting of digital campaigns.

2: Efficiency

By using data to compare the purchasing behaviour of viewers who were exposed to a particular TV ad to those who were not, One Data offers compelling post campaign analysis.

TF1 Publicité worked with the data and statistics experts at Ekimetrics on a Big Data modelling project that identified the contribution and

impact of different marketing mix components on sales, using historical purchasing data from Nielsen representing more than 2,000 brands. The analysis proved that TV contributes 80% of incremental sales, giving a new and powerful narrative to take to advertisers when discussing their own brands.

One Data was also able to prove the effectiveness of TV in driving brands' website traffic, which it can increase by 34% (figures based on a new survey). This effect extends beyond the short-term impact of a spot, as it reinforces branding over the longer-term as well.

3: Research & Development

TF1 Publicité is engaging with a number of partners in an innovative search for tomorrow's data-driven marketing solutions. A partnership with the phone operator Orange and retailers, using smartphone and beacon tracking technologies, offers the potential to recreate the consumer's journey after exposure to a TV ad through the use of connected objects.

Ultimately, it is all about TV being the most efficient medium.

egta talked to Nadège Verdurmen, Marketing Insights, Data & Strategy Director at TF1 Publicité, about the One Data project.

egta: *What has been the reaction to One Data from advertisers and media agencies?*

Nadège Verdurmen (NV): The approach has been welcomed by our customers. They found it to be very interesting, because it fits with their expectations for a form of targeting that goes beyond the "simple" socio-demo. This has also allowed us to remind that the medium of television can also offer behavioural targeting.

egta: *How does TF1 Publicité take the message of One Data to the market?*

NV: We launched One Data at a major conference to which all of our clients were invited. Following this, we have held numerous one-to-one meetings to explain the approach and to look at how it could be applied to each individual client.

egta: *Has the One Data project evolved since it was launched?*

NV: The challenge today is to make One Data concrete and operational.

To achieve this, we will make new media planning metrics available, such as the proportion of purchases for a specific product category that have been reached by a campaign (described by TF1 Publicité as "couverture data", for example, the percentage of yogurt purchases reached at least one time by the TV campaign). In addition to this, we will also offer to some of our clients the possibility to monitor their campaigns based on metrics other than GRPs, such as website visits for example. This type of approach is 100% data-led, because it requires the matching of information from different sources and the analysis of a large volume of data.

Finally, we are developing a test and learn approach in terms of addressable TV. For instance, we are working on a proof-of-concept with Orange: how to improve the contribution of reach thanks to catch-up?

We are conducting a test on a panel of voluntary panellist households. This test allows targeting of catch-up viewers who have little or no exposure to an ad on linear TV.

CASE LEARNINGS:

- Third-party data can be used to improve targeting and efficiency in linear TV
- One Data merges data from the television audience measurement panel with data from Kantar's Worldpanel of consumer purchases
- This approach allows TF1 Publicité to speak the same language as today's increasingly data-focussed advertisers
- The discussion moves from being about the basic *demographic* descriptions of viewers to one about *buyers* and even *purchases*

BRINGING ADDRESSABILITY TO LINEAR TV WITH SKY ADSMART

Sky Media

Launched by Sky in the UK in January 2014, Sky AdSmart was the first platform to bring household-level addressability to broadcast TV in Europe. Sky AdSmart allows targeting based on the broadcaster's first-party data, third-party data from external aggregators and second-party data relationships with advertisers, offering brands the promise of increased efficiency and better effectiveness.

Sky holds a unique position in the European television media landscape, combining a broadcaster, Pay TV operator, media sales house, telecoms provider and OTT platform into a single entity with a high household penetration. For many years, the company has been consistently at the forefront of set-top box innovation, introducing the PVR and the possibility to digitally record and play back TV to British viewers at scale with the launch of the Sky+ box in 2001. This has been followed with the introduction of high definition services (HD and Ultra HD), Internet connectivity and the ability to watch different channels in multiple locations throughout the house. The company's latest set top box offer – Sky Q – comprises a set of services and a family of devices that enable a seamless connected television experience.

Data sources

Sky's addressable TV service is built on two unique attributes: its access to basic information about its subscribers, and its control over the technology used to deliver TV content to those households, including ownership of return-path data delivered back to the broadcaster. Sky AdSmart is available in over 7 million households in the UK and is set to launch in Ireland later this year, having already launched in Germany and Italy.

It should be noted that, unlike some cable operators in the US, Sky does not sell any advertising inventory on channels that do not form part of its commercial portfolio; therefore, Sky cannot offer addressable advertising on the UK's commercially funded free-to-air channels ITV or Channel 4, for example, when viewed through a Sky box.

Sky Media, the company's ad sales unit, sells advertising around the broadcaster's own channels, which offer premium entertainment and lifestyle programming, news, sports and movies, as well as a number of thematic Pay TV channels as a third-party sales house. Sky Media also commercialises the inventory of free-to-air Channel 5, part of a wider deal with the channel's parent company, Viacom, and the broadcaster became the first PSB to be included on Sky Adsmart in February 2017.

By using its own first-party data (postcode and household surname) and harnessing third-party data from providers such as Experian, Sky is able to offer a huge range of audience segments based on more than 1,000 attributes. This shifts the buying from *context* (channels, programmes and dayparts) to *audiences*; a Sky AdSmart client does not have forward visibility over when or

where their advertising will appear, on what channels or around which content, but rather they are guaranteed that their messages will be served only to households that match the criteria that correspond with their marketing objectives.

Advertisers can also use their own data to build custom audience segments, allowing them for example to target groups of existing customers with up-sell renewal messages, profile existing customers to understand their characteristics or exclude existing customers to target potential new customers.

Ad decisioning and delivery

To make Sky AdSmart work, Sky's STB acts as a dynamic ad server, storing both the household profile data and a set of relevant spots for that household, delivered via satellite. When a Sky AdSmart opportunity is identified during an ad break – when the household falls into the required audience segment – the box will select the most relevant stored ad to play and seamlessly integrate it into the video stream.

Other than receiving advertising that is more relevant to them, the viewer is unaware that the spot has been delivered as addressable advertising.

Sky AdSmart can deliver addressable advertising in both live and recorded content, when viewed at normal speed. Ads inserted into recorded content will be those stored on the box at the time of viewing, meaning that the ads will be the most recent available, regardless of how long before viewing the programme originally aired

Moving targeting, trading and reporting towards the digital model

Just as this audience-based approach moves Sky AdSmart closer to the targeting model more typically seen in digital buying than traditional television, the same can be said of the way Sky trades this part of its inventory on a cost-per-impression (CPM) basis, as opposed to the Station Average Price (a floating cost-per-GRP system) used for almost all TV ad trading in the UK. Clients are only billed for impressions where at least 75% of the spot has been watched at normal speed, as measured by return-path data from the STB, offering a further guarantee over viewability. Clients also have access to reporting tools that mirror the type of analytics dashboards clients would more likely use for their digital advertising campaigns.

Sky AdSmart Analytics

Sky's strategy has centred on adding value to the whole campaign lifecycle, moving the relationship with the advertiser beyond being simply transaction-based with client and provider, a scenario that inevitably sees power shift to the adtech provider or agency/trading desk. A key part of this is Sky's analytics capabilities. Sky AdSmart Analytics was launched in late 2015, with further tools being built out to both help advertisers understand their return on investment as well as to help them plan, optimise and execute their campaigns.

The advantages of addressable advertising in linear TV

The ability to target messages only to relevant audiences has obvious advantages to advertisers, and a premium is attached to spots delivered via Sky AdSmart. However, while individual impressions may be more expensive, the possibility to target much narrower audiences than can typically be achieved on national or regional TV channels, for example by post code area, lowers the threshold for advertisers that would not otherwise be able to afford television campaigns. Notably, figures released recently by Sky showed that of the nearly 1,500 advertisers to have used Sky

FIGURE 05:
SKY ADSMART AUDIENCE SELECTION

Attribute type	Details
Mosaic Attributes	Experian Mosaic Lifestyle
Finance related attributes	Experian Financial Strategy Affluence Senior Decision Makers
Location options	Regions Metropolitan Cities Postcode Area
Lifestyle related attributes	Newspaper Readership Early Tech Adopters Fashion Mobile Phone Provider Mobile Contract
Composition of household attributes	Lifestage Household Composition Expectant Families Age Of Baby & Kids Pet Ownership
Attributes relating to the house	Home Ownership Second Mortgage Home Movers House Type Home Insurance Renewal South Facing Garden
Vehicle related attributes	Age Of Vehicle Number Of Vehicles Car Insurance Renewal

Source: Sky Media

AdSmart since launch, 69% were new to TV or Sky Media.

Sky AdVance: connecting TV and online audiences

The latest advertising solution launched by Sky Media is Sky AdVance.

Sky AdVance connects Sky's ground breaking TV viewing data to the digital world, allowing advertisers for the first time to break down the barriers between broadcast and digital campaigns.

Advertisers can now take TV programme and ad viewing insights into the digital domain, reaching relevant audiences on and off the Sky estate (e.g. skysports.com or huffingtonpost.co.uk)

The award-winning technology can be used in multiple ways, whether an advertiser has TV presence or not. This new level of understanding gives advertisers creating a digital campaign the ability to:

- Extend or reinforce a TV advertising campaign;
- Further activate and enhance a TV sponsorship;
- Use TV programme viewing insights for an enhanced digital execution.

CASE LEARNINGS:

- Targeted advertising offers brands the promise of increased efficiency and better effectiveness
- Sky AdSmart's household-level addressable linear TV advertising is built on two unique attributes: Sky's access to basic subscriber information and control of the technology used to deliver TV content (set-top box)
- Buying is shifted from *context* to *audiences*
- Sky's STB selects the most relevant stored ad to deliver when a household falls into the required audience segment
- Sky AdSmart inventory is traded on CPM basis, rather than station average price used for most TV trading in the UK
- Clients are only billed for impressions where at least 75% of the spot has been watched at normal speed
- Sky Media's Sky AdVance connects TV and online audiences

CONSTRUCTING A NEW DATA STRATEGY TO DRIVE PERSONALISATION AND RECOMMENDATION

Dr. Pierre-Nicolas Schwab, *Big Data/CRM Manager* – **RTBF**

RTBF, the public service broadcaster serving Belgium's French-speaking population, has embarked on a new strategy that harnesses data and technology to deliver a more personalised and relevant experience to its audience. With people, especially in younger age groups, increasingly turning to social media as their primary – or indeed only – source of news and information, RTBF has recognised that its competitors are now the likes of Facebook and Twitter, rather than only traditional media as in the past. As a source of entertainment, its competitors now include Netflix and on-demand music streaming services, and all of these are currently far more advanced in personalisation and recommendation than legacy television and radio broadcasters.

egta spoke to Dr. Pierre-Nicolas Schwab, Big Data/CRM Manager at RTBF, who designed and directs the broadcaster's new data strategy, bringing together internal resources with external partner organisations.

The starting point

RTBF has a large audience and a great deal of unique, high quality content; however, it knows relatively little about the former and the characteristics of the latter cannot currently be identified automatically. Together, these two limitations prevent the broadcaster from recommending content to users or delivering a personalised experience.

As a broadcaster with television, radio and digital assets, RTBF has multiple touchpoints and interactions with its audience, both online and offline. However, very little information is systematically collected from these interactions, and while the broadcaster has about a million user ID's in its databases, for the most part this data is too limited to be actionable for personalisation.

"It can be when people call a radio station, send an email, attend an event or go to our websites, those are all sources of interaction", explains Schwab. "The problem was, for most of them we capture no data. When someone was calling a radio station, there was no track of it. For some of them we have databases. What is interesting is that the name of the database is the name of the supplier, and what you suddenly understand is that through the years, everything was very fragmented. We had to reconcile everything."

The second main challenge is the lack of robust metadata associated with RTBF's content. In theory, any piece of content, be it an article on the broadcaster's websites, a segment of audio or video, can be identified by a series of metadata points, such as its title, a description, associated tags and so on. In practice, this metadata is usually very patchy and, again, too incomplete to allow personalised content to be delivered to RTBF's audience. As noted by Schwab, "How can I do a recommendation based on metadata that is very poor? I have to recreate everything. I

could hire an army of people to write metadata, but that won't be scalable."

Building out the new data capabilities

At the start of the process, in 2015, the team at RTBF set about investigating the broadcaster's various audience touchpoints and identified more than 50 such sources of interaction. To be actionable, for example for targeted advertising, a person needs to be identifiable, for instance through an email address, a unique ID, or a combination of several other identifiers. However, for content recommendation, it is necessary to know something more about these people, what content they like to consume, their interests and their social environments.

SSO and social login: building richer user profiles

At the heart of the new data collection infrastructure is RTBF's single sign-on (SSO), launched in December 2016, which allows a larger proportion of the audience to be turned from anonymous visitors into known audience. Gigya was selected to deliver this service following a public tender. RTBF chose to outsource the implementation and ongoing management of the SSO, rather than build the technology in-house, as identity and access management (IAM) systems are complicated to develop and a number of providers already exist on the market. Furthermore, these providers take care of storing, processing and dashboarding the data collected through SSO, as well as managing the complex and often changing environment of logins using social media (social login).

RTBF offers users the possibility of registering either using their email address (first name, surname, email address, protected with a password) or through social login. "We recognised very early in the SSO process that we wanted to take advantage of Facebook Connect," explains Schwab. "Why? Because we want to understand what people like outside of the RTBF world. If I observe only what people do on my websites, it becomes tautological. So, I can recommend only what I am producing. And soon, I am trapped in my production bubble."

The broadcaster will request information about each user's network of Facebook friends, in addition to data about the user themselves, in order to build a more comprehensive understanding about its logged-in audience, which in turn allows better recommendation. At launch, RTBF chose to offer social login only through Facebook, as it delivers richer and more relevant data than alternatives such as Twitter and other social networks. However, as the system matures, it may be more appropriate to allow users to log in using alternative social media where consumers of particular types of content might have a better affinity with networks such as Instagram, for example.

The data from both social and non-social login is consolidated by Gigya, with which RTBF can also leverage progressive completion strategies to update users' profiles over time, for example by periodically asking them to add or confirm additional information. It should be noted that while the SSO has only been live for a few weeks at time of writing, registering with name and email is currently preferred by a majority of RTBF users over social login.

Connecting consumption with people

Complementing the Gigya database of people is a separate database to collect and store raw data about *consumption*, described by RTBF as a *datalake*. Additional infrastructure and technology of one of RTBF's partner organisations uses Big Data technologies to convert raw data on website interactions so that

it can be combined with the data held in Gigya to power the recommendation engine.

"We are thinking about non-relational databases to store transformed data and start visualising it and using it. When people visit the website, we drop a cookie. And when people register on Gigya, that cookie – or ID – is being replaced with a Gigya ID, which allows us to collect back all of the history of that person before he or she registered."

A reliance on first-party data

RTBF's recommendation strategy is dependent entirely on the broadcaster's own first-party data, either collected though the SSO or from anonymous digital visitors. "Very early, we decided that we will never buy data. We will not be selling data and we will not be buying data."

This reflects the broadcaster's priority, which is to develop a data strategy and architecture to enable targeted content to be delivered to users, rather than to deliver targeted advertising, which typically requires integration of third-party data sources, for example through a DMP.

Collecting data in the mobile environment

The collection of data on consumption through web browsers on desktop and laptops is enabled largely through the use of cookies, which are of more limited use in the mobile environment. In 2017, RTBF plans to launch a series of new mobile apps for its websites, Auvio video and audio platform and a new multimedia platform for young people currently going by the working title Média Z, all of which will be connected to the single sign-on.

The organisation is seeking to lead people to use the RTBF apps, and motivating them to log in through incentives, rather than to force all consumption to follow this path. This is a different strategy to some broadcasters, which now only allow their video – and in some cases online radio – to be consumed by logged-in listeners.

Preparing content for the recommendation engine

RTBF's recommendation engine is designed for its journalistic output, where metadata can be extracted based on what is *said*, rather than for music or television shows, for example, where metadata based on factors such as *emotions* can only be delivered by humans rather than algorithms.

"We have developed a prototype where you can inject a video or an audio file, which will go through a speech-to-text and a sequencing algorithm, then through an ontology algorithm that extracts semantic fields around what is said and attach them back to the content." Yet there are still some technical challenges to solve to make it a reliable product.

This prototype will be refined in 2017 to allow RTBF to build up a richer understanding of its own content using an automated, algorithmic process, which together with the user profiling described above will power the recommendation engine.

Two philosophies of recommendation and the conundrum of the filter bubble

Broadly speaking, recommendation engines can propose content following two models: *exploitation* and *exploration*. The former presents content that is similar to a user's past consumption, meaning that exposure to new ideas or concepts is limited. Amazon and Netflix are both examples of platforms that base their recommendations on exploitation. The latter, on the other hand, seeks to present content that – while it continues to be relevant – can expose the user to materials that they

might otherwise not have consumed. As part of its public service commitment, RTBF will design its recommendation engine to promote exploration while at the same time retaining the flexibility to adapt to meet the needs of users who may prefer to receive a narrower range of suggestions.

The nature of recommendation strategies has become increasingly relevant in discussions around the so-called *filter bubble*, in which people are exposed to information in the digital realm in a way that serves to reinforce their existing opinions and isolate them from alternative viewpoints. The subject of considerable debate in media, publishing, broadcasting and academic circles, there is as yet little consensus on how the filter bubble effect operates, whether it actually exists and how news publishers should respond to it. However, this effect is an important consideration when designing the operational parameters of a broadcaster's recommendation engine.

Implementation of the RTBF recommendation engine

Having established the data collection, storage and processing elements of its architecture, including the staging area and single sign-on, the next stage is to start recommending content to RTBF's users. The technology behind the recommendation engine is handled by the Liege part of the consortium, and it is linked to the RTBF's online platforms via API integrations.

RTBF refers to the first iteration of this system as its *similarity engine*, which will make the same recommendations to all people watching or viewing a piece of content, and this is due to launch in February 2017. "This content-based recommendation is not new, it is not innovative, but we had to rebuild the foundations in order to grow from there and have some more refinement and technological changes."

In December 2017, under specifications that are currently being developed, the system will become more sophisticated, allowing RTBF to expose people to content that they are not used to consuming. "The purpose being to expose people to the diversity of the world, to the diversity of opinions, to new pieces of content." However, if users choose not to receive this type of experience, "the system will just learn to not serve more of those unexpected types of content, using machine learning."

Data security and the GDPR

RTBF has overhauled its internal data privacy documentation and developed new communication materials about how and why it collects data for its users. Inspired in part by the Channel 4 Viewer Promise (see page 25), a set of straightforward texts and explanatory videos have been produced to increase transparency and help educate users on this topic. It is also important to recognise that recommendation itself works better when people understand the basis on which content is being selected and placed before them.

As a public organisation, RTBF will be required to have an in-house Data Protection Officer (DPO) when the new EU General Data Protection Regulations (GDPR, see page 22) come into force in May 2018. This raises a number of questions for the internal organisation of RTBF, such as where within the broadcaster's structure the DPO should be located in order to function properly.

Challenges and learnings

Almost 2 years into the RTBF's data project, Schwab reflects on the various challenges and learnings that have been made along the way. "There are two worlds. There are the people who believe in big data and correlation as being the sole predictor of future behaviour and the world of more sociology-oriented people who

think that you have to understand causation to make good recommendations. And I'm just in the middle!"

"It's so important that people do not believe that correlations between variables can predict the future of individuals. We need to understand *why* people are doing things. We must reflect the role of RTBF – also of broadcasting in general – in terms of what is the job that people hire us to do. When you are on public transport and you are watching the RTBF app, are you watching because you want to consume news or because you just want to spend time? These are two very different jobs.

"What are the consumption moments, why are people using RTBF? Hence also the SSO – I have to make the link between the user on their desktop, in their home, on their mobile or tablet. Yet the data I get will get will not be sufficient to infer *why* someone is doing what they are doing. And the problem with data scientists is that they just believe in *what*, they don't try to understand *why*.

"In terms of human resources, the challenge is currently between the business people here in the RTBF and the data scientists in the other organisations involved in the project. They don't always understand each other, and I have to bridge this gap. Making that link is the most crucial thing."

Looking back at the process so far, Schwab explained that "it is not a technological challenge, it is more of a human challenge. That's something that I did not believe two years ago, but now I believe it."

CASE LEARNINGS:

- As a public service broadcaster and news source, the RTBF's competitors now include the likes of Facebook and Twitter, Netflix and other on-demand video and music streaming services
- Challenge 1: RTBF historically had several touchpoints with its audience, but little data was systematically collected
- Challenge 2: the broadcaster had patchy and incomplete metadata about its content
- Personalisation and recommendation requires both of these challenges to be resolved
- User identification is now possible through a new single sign-on for all RTBF platforms, with social login also available (through Facebook Connect)
- The broadcaster is developing a recommendation engine based on *exploration*, rather than *exploitation*
- The General Data Protection Regulation requires RTBF to install a Data Protection Officer; this raises questions over where in the organisation this person will sit
- The biggest challenge of the project is human, rather than technical

THE CHALLENGES OF MEASUREMENT IN THE DIGITAL ENVIRONMENT

Charlene Weisler, *Founder & Consultant –* **Weisler Media**

Charlene Weisler is a market and media research consultant, columnist and blogger. She draws on her extensive experience in television, having held senior positions at AMC Networks, Cablevision Systems Corporation, A&E Television Networks, Discovery Networks and others.

Charlene shares insights on the latest direction of travel on media measurement and data in the US, noting the importance of achieving standardisation in cross-platform measurement, leveraging data from digital place-based media and the value of moving beyond traditional age and gender proxies towards behavioural audience segmentation.

egta: *To start with audience measurement, what do you see as the state of the industry at present?*

Charlene Weisler (CW): There are a number of challenges that need to be solved, and I think that there are a lot of efforts to that end. In my opinion, one of the biggest challenges is coming up with a standardisable message for cross-platform measurement. That would be capturing usage across devices and across platforms in a way that the industry can use the data to ascertain actual deliveries and help content providers maximise their revenue by capturing every single bit of usage. And it's easier said than done. Not just because we've yet to agree what the standard measurement might be, because there is a difference between digital and traditional television at this time. It's also having to take into account print, radio, outdoor and so on. There are a number of components, a number of media points that touch the consumer every day. And its growing, its morphing, its evolving, and it becomes almost a moving target.

egta: *Are we talking here about all touchpoints across all media, not just focussing on bringing traditional television together with video but rather the whole scope of how a marketer could reach a consumer?*

CW: Absolutely. I know that we're probably starting with the basics of television and video and digital, but there's a lot more. There is a lot of fascinating and very usable data being collected beyond the home in digital outdoor, and if you were a CNBC or an ESPN, those collections of data would be very important for you. So, yes there are the basic, traditional media points, but then there are also the new technologies and digital place-based media, which should be brought in as far as possible into the standard measurement.

egta: *You have companies, such as Nielsen in the US, that measure TV viewing across the whole population and more recent solutions, for example collecting return-path data, that are very good at capturing particular types of viewers or ways of receiving TV with high granularity. Is it possible to move towards a hybrid model where you can extrapolate some of that rich data onto the wider population that's not so well measured?*

CW: You have the whole area of OTT; it's not through the cable box but it's fragmented to a certain degree because it's being collected through different points. There's a lot of extrapolation, you've mentioned Nielsen, which is panel-based, and they extrapolate off of a panel sample of viewers, people in homes. So, extrapolation has been part of the industry for more than sixty years, and I think there will still be some of that. But I also think that with STB

data and with other ways of collecting other big data sets that can be merged with TV usage data, we'll have a larger base with which to make extrapolations.

egta: *With the companies that are now entering the space, with comScore going into TV with the acquisition of Rentrack for example, do you think that there could be a fundamental change to the way television itself is bought and sold, or is it more an icing-on-the-cake type of evolution?*

CW: I'm going to give my personal opinion on this. Age and gender is a proxy, and I think a lot of people would agree that not all woman aged 18-49 are alike. In fact, not all women aged 18 years old are alike. I think it's actually based on lifestyle and a variety of other factors. The age and gender proxies that we currently use to post against media buys have always been an easy way to ascertain delivery and that the contract was delivered to the guarantee. Will we move away from that? It would make sense that we should, because of the limitations of proxies, but I'm not sure we will do so fast enough!

Everybody's talking about segmentation, and it's a terrific idea to have behavioural segments, so that the advertiser as well as the content provider understands who is really consuming. Maybe age and gender doesn't matter if you have a certain product, you just want that brand's aficionados, no matter how old they are, no matter what gender they are. That would work better for everybody, because viewers would receive content that really speaks to them, no matter what their age or gender. Advertisers would reach people who are truly enthusiastic about the product they are advertising, and content providers could provide the best in entertainment for that special interest group.

It's just very hard to standardise that, and I think if we want to try and maintain some kind of traditional standard metrics by which the industry is measured, it's a pretty tall order to come up with a strict set of behavioural segments. If you look at one particular auto manufacturer, for example, different models will have different behavioural segments, so it's not just at the individual advertiser level, but at the individual product level.

egta: *You mentioned that we probably won't move fast enough, what would you say are the potential risks of not moving beyond proxies into audience segmentation and actually pushing that further forward?*

CW: I would think the biggest risk is that you're leaving money on the table and that you're not maximising the benefits of your budget or your inventory.

If you look at the recent claims about the role of traditional television, it still commands the majority of viewership. Whether that will erode over time remains to be seen, but I think more in the short term the risk is just not maximising what you already have. And, of course, in not really being able to fine tune to a specific behavioural segment. In a way one might think that that might accelerate viewership to other platforms yet to be developed or in development now, because they will target more effectively.

egta: *Would you say that the barriers we currently face are primarily technological limitations or legacy attitudes to doing business?*

CW: It's a little bit of both, I think. It's easy to stay within the traditional forms of measurement; your five-year plans are based on delivery of age and gender. And all of the systems that are being used by the agencies and the networks are built towards delivering along those age and gender targets. So, you have an immense job to transfer over in an automatic way to a new standard. It's a bit of the traditional mindset and it's also a bit of the technological, because you would have to modify the systems

to take into account behavioural segments in a standardised manner.

egta: *I understand there are companies like Simulmedia that are focussed on finding the value in inventory that could otherwise be seen as undervalued or even worthless. Do you think that there is quite a rich mine to be tapped there, or does this approach have its limitations?*

CW: I think it's incredibly rich. Part of the challenge with proxy measurements, with age and gender, is when you've crossed that threshold from aged 49 to aged 50 you are suddenly not very valuable to advertisers. And in fact, if you read the results of studies, the wealth is being concentrated in the 50+ cohort. And just because you've turned 50 doesn't mean you don't want to buy a car or buy a house, a vacation home, or furnish it, or take a holiday. So, I think that there is a lot more value that is not being mined for advertisers, and for content providers, and it's diminishing an audience that could actually prove to be very fertile in terms of variety of different purchasing behaviours.

It's not just the audience that might be undervalued, but there might also be inventory that is undervalued. Maybe you should be buying in overnight, as opposed to prime time. So, I think that there is a lot of value in looking beyond and using the technology and the data sets that are becoming available to recognise value where we may not ascribe it right now.

egta: *Are you seeing different types of skill sets coming into broadcasters and sales houses to be able to develop some of these new approaches?*

CW: I have always believed that there is an advantage in being creative, strategic and innovative, I would like to think more people are being hired with those skill sets. But I also think that they are very hard skill sets to find and to develop. I believe that the best media companies are those that encourage innovation and strategic thinking and will embrace challenges in order to overcome them. There are companies that are very traditional, and then there are those that want to expand beyond. And I think that those who hire to expand beyond the traditional will be the ones that will be around in the future and thrive.

egta: *It's sometimes said that the US might be more advanced in the way they are able to collect and use data; what advice would you give to European broadcasters in order to be able to maximise the opportunities of data in the next five to ten years?*

CW: It's interesting you should talk about US vs. Europe. I think that Europe has something I wish we had in the United States, which would be a JIC to help set policy. It's something that I think the time has come for in the US, so that we can all be heard at the same time and share ideas and solutions, and I think that's where Europe might have an advantage over the US. And my advice would be to work with the JIC that you have and start to develop protocols based on leanings internationally, and the data that you have on hand where you are.

COOPERATE OR COMPETE?

Thierry Tacheny, *CEO* – **Divedia**

Thierry Tacheny is CEO of the media consultancy Divedia and a former top TV executive. Thierry held a senior management position at IP, the sales house for RTL's Francophone television stations in Belgium, before co-founding the TV planning and yield management systems company Peaktime and later crossing the country's linguistic divide as CEO of the Flemish broadcaster SBS.

Thierry's advice to broadcasters and sales houses is to start thinking differently about the way they construct a data strategy, to look to the longer term as well as the short-term gain and to find ways to maintain control as an industry and not just as a single media owner.

egta: *What do you believe to be the opportunities for media owners working together, rather than solely focusing on single solutions?*

Thierry Tacheny (TT): While the opportunity exists, all media partners today are concentrating more on what data might provide to them, whether we are talking about internal knowledge to optimise their programming or to help the sales process. As a starting point, I'm not sure that all broadcasters have yet decided where they are going to locate their online VOD sales – are they going to do it through digital, and trade it as digital, or are they going to trade it as a part of television? Linear TV inventories are limited, due to the European regulations, while digital television inventories are unlimited.

Looking back to TV's history, and how data came together for television, broadcasters decided to define a common rating and the markets were willing to create Joint Industry Committees. And we don't see those popping up now, which I think might be a worrying item for the future, because what made television strong was that the data ecosystem was shared by everyone – data was the starting point of the efficient ecosystem of television.

There are some initiatives being taken here and there; for example, the newspapers in South Belgium are currently working together on an initiative that I would not quite describe as a common platform, although it might become that, but at least they are trying to have the same vocabulary. And I think this is something that television is missing – everyone is just playing in their own gardens, but no one is looking to the broader picture. It is probably much more difficult today, but this could harm television, because the technology stakeholders active today might take more of the market than we expect.

egta: *Does one of the difficulties lie in an environment where we are moving from trading on proxies, where everyone has to agree on definitions and values across the whole market, to digital impression-based trading, where all you then need is agreement between one buyer and one seller over the rules of the game?*

TT: You are right, but we also need to ask ourselves – what is television for and why does it perform so strongly? It is mainly because it is the only medium that brings such a high level of instant reach, and to asses that means you need a common data system that reports in terms of market share. If you move to an impression-

by-impression system, there is no question of market share, and that makes for a very different model.

Television is still very strong in itself of course, but also because it relies on commonly shared statistical figures delivered by peoplemeters. Nothing better than the peoplemeter systems has been created yet, and there is no answer today to bring something as effective as that to digital. Some tentative solutions are coming up here or there such as Mediametrie, for example, which is trying to measure what is watched on all devices. But the critical issue in such projects will derive to the way data are merged and put together. However, to protect its strength, television needs to develop a common data platform.

egta: *If we imagine a scenario in which broadcasters were willing to work together, would you say that there could be some benefits in extending the TV concept of a common data platform to the online part?*

TT: For sure, at least for the broadcasting area. If no one moves, the power will go to the technology companies. When you look at the history of how television was traded, the cost of sale was significant 30 years ago – there was 15% for the agency commission, there was also commission for the sales house, whether it's internal or external. Today these costs have sharply decreased. Whereas in digital, sales costs are very high, because (1) everyone needs to equip themselves with smart technology and (2) new technical entrants are taking a part of budget and charge commissions for the transaction process. That a third of the media budget goes to platforms and transaction fees does not bother advertisers as long as media cost do not increase. So the question is: "who is going to pay for these transactions?" And the answer – media.

To protect value, cost and efficiency, I think it would be wise for all broadcasters to impose something new, to develop common transaction platforms (even if these are run and managed by external and neutral technical partners). This was where television was clever in the past – television is efficient in itself, but that efficiency would be much lower without peoplemeters.

egta: *Are you aware of any moves for the TV and digital JICs to cooperate in Belgium?*

TT: It does happen in Belgium, but television is not ready to cede to digital the power to rule their data in the future. And they are probably right. There have been discussions over the past 4-5 years in Belgium trying to bring both the distributors and the broadcasters together in the same room, and I know that there are still conversations to try to come to the same standards in the data.

However, when it comes to platforms like Sky AdSmart, developed by Sky in the UK, one company may be smart and launch something new on the market, but the others may not be willing to join. The reason is clear – one cannot be the platform and the competitor at the same time.

The second issue is that with the new data protection rules, it is going to be even more difficult to gather information, because if you mix together the GDPR and the anti-trust obligations of TV, this is going to create a very difficult area for conversations. Having a common discussion about revenue or rate cards is strictly forbidden, so what about any discussion about data? Talking about data might speed up progress on the market, but at the same time the more you discuss, the more you take some risks. Media owners will not take the risk to pay fines of €20 million, or 4% of their revenue!

egta: *Platforms like Sky AdSmart offer those that develop them a way to differentiate themselves on the market: if you move towards market-level platforms, would media owners lose that edge over their competitors?*

TT: I'm not sure that platform is a key area for competition. What do advertisers want? They want to reach people, either instantly, or over the mid or long-terms. Instant reach is the mother of communication, and it is something that only television delivers. Having a better trading platform today than the competition is probably helpful, but the real value lies in being able to show what you can deliver and how quickly and strongly. If everyone is using the same data, you can make meaningful comparisons and then focus on the strength of your medium rather than on the strength of the technology.

If you look at the history of how a market develops, it always starts like this. One company is always cleverer than its competitors and creates a technical environment. And for a limited amount of time, they take more money and take a bigger piece of the pie. But this lasts only for a few years, and later on, what makes you strong is and remains your medium.

egta: *Changing from a situation where broadcasters' competitors were other broadcasters to today's reality where the competition is different media types – social, search, display, virtual reality – doesn't this make it difficult to talk just about the strength of the TV?*

TT: Yes it is, but the big question for me is that in order to manage protect and develop your business, you must know where you belong to and what you are good at. Television is strong because it is (still) the most efficient medium regarding instant reach. That's why advertisers need it and continue to use it.

In digital no single broadcaster alone will be strong enough to beat Google or Facebook. But together, broadcasters can and will.

That's why the industry should be thinking at duplicating the "peoplemeter ecosystem" in the digital world. Developing common data platforms and gathering their strengths in the digital world as they did it so well in the linear world. Short term, everyone needs to dig into data and to try to make the best for their own purpose. But something is missing, I think, it's more of a strategic vision – how to continue to protect television's strength.

DATA IN PROGRAMMATIC TELEVISION

Pete Doe, *Chief Research Officer* – **clypd**

clypd is an audience-based TV sales platform. Based in Boston, the company's clients in the US include Fox, Discovery and ESPN, and in 2015 clypd strengthened its position in Europe with investment from the RTL Group, which has itself taken majority stakes in the video SSP SpotXchange (now SpotX) and the video advertising technology company Smartclip. clypd opened an office in the UK in 2016, looking to start bringing its advanced audience buying technologies to Europe. clypd is one of a number of companies active in the programmatic TV space. In linear TV, the concept of programmatic differs somewhat from programmatic in digital advertising: in linear TV, programmatic refers principally to using data to optimise advertising placement (decisioning) and the automation of workflows, and it does not necessarily equate to the end-to-end automation seen in online programmatic trading, for example for digital display or video ad impressions.

In programmatic TV, audience targeting is typically broader than in digital, where impressions are targeted to the individual based on very granular data.

The clypd technology offers TV sellers two main services: workflow automation and data-driven decision-making. The clypd platform allows media owners to sell their inventory beyond traditional age and gender targets by layering in data from additional sources. In the US, for example, Nielsen's TV panel data can be fused with additional data sets to allow advanced audience decisioning. This may be set-top box data, third-party data, such as frequent shopper records, or panel data.

To take a European example, TAM data from the audience UK's measurement provider BARB can be fused with data from Kantar Media's TGI survey, which gathers information covering consumer attitudes, motivations, media habits and purchase behaviour from a panel of about 25,000 people. By merging these data sets, clypd's platform can identify TV inventory that over-indexes for a particular audience segment.

egta spoke to Pete Doe, Chief Research Officer at clypd, about enhanced data decisioning in linear TV.

egta: *How does data work within the Clypd platform?*

Pete Doe (PD): Think about it from our client's perspective. Imagine a client that is a cable network, or that owns a series of cable networks, and their role is to sell their linear TV inventory as efficiently as they can to agencies. Agencies in return are trying to get best value for their money and create campaigns that are as targeted as possible within the constraints of linear TV.

The platform allows both buyer and seller to do a better job than they currently are through what is called programmatic, although programmatic can be a problematic term because it means lots of different things to different people. The media owner puts its inventory into the clypd platform, with rates and forecasts for demographic impressions for their inventory, and we also have access to a lot of additional data sets for

what we call advanced targets, enabling buyer and seller to go beyond the traditional gender and age targets that are typically traded on.

In the US, we're adding in data that has been fused onto the Nielsen currency panel in most cases, and those data sets can be third party type data sets from companies like GfK MRI, they can be matched credit card transaction data or frequent shopper data, in some cases first party data. The platform takes the inventory, the rates, the additional data, and the forecasts for the impressions of those additional data, and creates an optimal deal for both buyer and seller within agreed constraints – time period, demographic impressions, constraints within networks.

So, it's all the normal things you care about when you're creating an ad campaign, but doing it in an optimised way using technology within the platform, optimisation routines to make the data work as efficiently as possible, both for buyer and for seller.

egta: *Do you get a sense that the objective of buyers in the US is to improve efficiency, or effectiveness, or is it to reduce media costs?*

PD: All of the above, I think. Any individual deal will have a difference in terms of the weight of the emphasis of those three. In some cases, a buyer may be really concerned about demonstrating effectiveness, because that's the brief they've been given, in other cases it may be more about trying to claw back some money that they've lost along the way somewhere.

egta: *What do you require as a minimum data set? Does the Nielsen data, or Rentrak or whoever it may be, form the base layer of data on which everything else sits?*

PD: It does, and we always say at clypd that we are data agnostic; it's not our job to tell buyer and seller what data sets they should be using to conduct transactions, it's our job to enable the agreed data sets.

In the US, Nielsen is still the primary currency vehicle for the vast majority of TV deals, so that's one of the most important and central data sets in our platform. We do have access to Rentrak data and have worked with it, but we don't force Rentrak data on anybody, in the same way that we wouldn't force Nielsen data on anybody. It's really the buyer and seller who decide.

So, those are the base currency data sets, and then the next discussion is about how we can enhance those databases with additional advanced target information to make this process more effective.

egta: *And what are the sources of that enhanced data?*

PD: It comes from various sources. The one data set that I'd say is most commonly available for these discussions and transactions is the Nielsen MRI Fusion dataset. MRI is produced by GfK in the US, it's actually the magazine currency measurement. MRI respondents answer a wide variety of questions on product usage, attitudes, psychographic questions, magazine readership, but also a lot of other media questions about TV shows watched, Internet usage, radio, cinema and so on.

As a standalone dataset, it's used by both media owners and agencies for planning and ad sales services, so you get a lot of insights into the media habits of particular brand consumers or people with certain attitudes.

What the Nielsen MRI Fusion does is take that big survey and fuse it onto the Nielsen National Peoplemeter currency TV panel, and that then connects directly to the currency. MRI on its own might tell you people who are international flyers, people who ride motorcycles, or people who buy particular products, and by fusing that

to the National Peoplemeter dataset, you can then look at the viewing habits and analyse that information as if the panel members themselves had all that information attached to them.

egta: *Do you also use other sources, such as credit card data?*

PD: Yes, there are Big Data solutions involving panel matches, where the National Peoplemeter panel is matched with aggregated credit card information, carried out by a company that prefers to remain anonymous. This company aggregates across a lot of banks, and the data covers 80% of all credit card transactions in the US and 20% of all debit card transactions.

That dataset is matched at the individual level with Nielsen panellists through a third-party anonymised matching process, which is a way of protecting privacy. The data gives spend at the retailer level.

Another source in the US is Nielsen Catalina Solutions, which matches millions of frequent shopper records data from supermarkets. And that gives you a lot more granular detail about brand purchasing that the credit card data lacks. However, it's more narrow in focus, because it's really in the FMCG space.

All these datasets have their pros and cons, of course, and one of the cons you could argue with these Big Data matches is that not everybody gets matched. When credit card data are matched, one problem is that some people still don't use credit cards, they just use cash, and that can actually bias things away from, for example, lower income and Spanish speakers in the US, so it's not a fully representative view of how the population is behaving.

egta: *Do you have to accept a certain amount of data fuzziness that perhaps you wouldn't have had to in the past?*

PD: I think that's right, and one aspect that we've addressed in the clypd approach is the fact that some of these data are not updated as frequently as the main panel information. The Nielsen panel gets updated every day, but something like the MRI Fusion is conducted monthly, and in some cases it may be two or three months out of date because of the cycle that the production has. And that's sort of OK if you're using it for planning and ad sales purposes, but if you want to use of for transaction purposes you need a way of addressing that latency, and so we've developed solutions for that.

I think more generally, one of the reasons gender and age has lasted so long as a means for transaction is that it's very clear cut and it's easy to agree on definitions. Although it may be a very broad and unfocussed way of looking at the world, it is something that permeates pretty much every database and is completely understood. But the moment you get into something a little bit fuzzier, like international travellers – what does that mean, exactly? What is international? If I fly from here to Hawaii, I'm classified as international traveller, even though Hawaii is part of the United States. Suddenly, you've got this discussion between buyer and seller – "what exactly are we talking about? Oh, I don't know, let's just do women 25-54!"

egta: *Coming back to European broadcasters and sales houses, what kind of strategies do you believe they need to be putting in place?*

PD: From my perspective, I always go back to the data. I'm sure that there are all sorts of technology issues that they need to address, but I certainly would not class myself in an expert in that area.

One thing I would say is that the opportunities in individual European countries are to an extent limited by the sample sizes of the current audience measurement panel. So, if media owners and agencies and buyers are focussing

on the linear TV capabilities and still wanting to transact through the standard currency panels, then they are dealing with sample sizes that are much smaller than we're blessed with in the US, where the Nielsen panel through its recent extension has gone up to about 35,000 households. I think the largest ones in Europe are about 5,000. So, that immediately means a more advanced target definition has to work within much tighter constraints.

egta: *Fraud is becoming a big problem in digital display – is it an issue for programmatic TV?*

PD: I think it's a non-issue compared with the digital issues that have been well publicised in the last few years. There's still a need for auditing, but I think the auditing of TV campaigns is well established. I don't think the development of programmatic and using advanced targets really changes that, probably where the challenge will be is demonstrating effectiveness. But that's certainly not a fraud question.

egta: *Linked to that, are there bad data sets or bad data sources, or sources that can be more questionable than others?*

PD: I would always say buyer beware! If you are presented with a dataset that purports to solve all your problems, find out about it. If you don't feel expert enough yourself, get somebody who you trust, who understands data, to look at it and review it and understand how the data are developed.

There is no perfect dataset, and that can in some cases result in people going too much the other way and rejecting anything that doesn't reach an incredibly high standard. You can be too lax one way and too tight the other way.

At the end of the day, you have to answer the question – will these data make our business better? And if the answer is "yes", and sufficiently better to justify the cost, then use the data. And having made that decision, review it in a year's time to see if that's still the case.

CASE LEARNINGS:

- Programmatic TV refers principally to the data-driven optimisation of advertising placement (decisioning) and workflow optimisation
- It does not necessarily equate to the end-to-end automation seen in other forms of programmatic advertising, such as digital display, video or search
- The clypd platform allows to trade beyond traditional age and gender targets
- The platform is data agnostic, and can use whatever datasets the buyer and seller agree on
- In the US, Nielsen TV peoplemeter data is one of the most important and central datasets
- Age and gender remain important metrics for transactions, because they are easily defined and understood; metrics based on interests or intent are harder to define
- If you don't feel qualified to assess a data source, get someone you trust and who understands data to review it

THE INNER WORKINGS OF THE DATA MANAGEMENT PLATFORM

Ian Curd, *Market Development EMEA* – **Lotame**

Ian Curd is responsible for Market Development EMEA at Lotame, a leading DMP and data exchange that enables publishers and advertisers to harness their own first-party data, improve targeting through third-party data and engage in second-party data partnerships.

egta: *Why would I as a publisher need a DMP, and what can it offer me?*

Ian Curd (IC): Let's start with how the DMP came to its current place. It's been much talked about in martech and adtech recently, but it's actually been around a long time. The DMP started out as a way for publishers to gain more intelligence about their audiences and to be able to package that as inventory that they could sell to trading desks and ad agencies with more intelligence behind it. The theory was, if you have better intelligence, you can sell for higher CPMs.

We – and any DMP that has existed as long as we have – all have histories in publisher use cases. Where that's grown over the last two years is towards the demand side. This trend has been happening probably since about 2014 in the US, and it is really starting to take shape here in Europe at the moment. But the real modus operandi of the DMP remains the same. It's the ability to pull data from multiple disparate sources, to intelligently organise that data in the middle, and to be able to activate and analyse that data after you've pushed it out to wherever you need it to go.

That's got a whole number of layers to it. It's got an advertising layer to it, so targeting new customers or speaking to existing customers, it's content management, in the corporate case it's feeding into business intelligence tools; it's all of those things.

But the point is where you start from, typically from a digital perspective at least: you have data sitting in a lot of different places. You have an ad server, you're collecting information about your users through your own tags or third-party analytics tools on your site, but there's nothing siting there in the middle bringing all of that together. That's what a DMP does. You often hear the analogy that the DMP is like the brain of the digital stack, or the central nervous system.

egta: *We're looking at this from two angles – the purely online space, but also bringing programmatic and data enrichment into linear advertising. If you need a DMP for the digital space or the linear space, is it a different animal?*

IC: It's not. Start from the fundamental building block of this brain. The way to think about it, whatever side of the corporate marketing or media angle that it sits on, the DMP is the place where the client can regain control over its own audience. You use third-party, but you base it off your own data. You see that third-party in relation to your own audience the whole time. Therefore, you are constantly building out control over your own data.

It doesn't start with third-party, it starts with first-party, even if first is small. We have a machine learning tool that can build lookalike

models using the power of Lotame's third-party data network, but off of a thousand uniques. We collect data through tags, at a very minute and unstructured data level, and we can apply machine learning algorithms to build models off it.

egta: *You use machine learning to develop audiences from smaller sets of data, are there other examples of how machine can be used in media sales?*

IC: We work with big publisher clients, and publisher-to-publisher some of them are doing some pretty interesting things. Our machine learning tool can be applied two ways: either to find an audience that looks or acts like the audience you can see, or to optimise campaigns.

egta: *How does the DMP help target consumers that you already know or a lookalike audience?*

IC: The first thing that we do with a new client is to put Javascript tags on everything in their digital environment, and we start pulling in all sorts of unstructured data. But within the platform that we build, those unstructured data points become building blocks. They become audience hierarchies, which are things against which you can target through DSPs. We've got about 80 DSPs server-to-server piped into Lotame.

So, this is the fundamental building block of first party data. But any kind of mature DMP will be able to connect to a list of third-party data sources on top of that to model. But in between, you should be able to use the DMP to actually trade your data either with another partner or directly, almost like you're a data provider as well.

egta: *In relation to that, would you say that second-party data could become considerably bigger by automating this process?*

IC: Absolutely. We say to our clients that we provide the technology and the plumbing for them to do that. The big hold back on this space of second-party data across the world has been the fear of trading data, a bank with a publisher for example, and the commercials around it. It's put it in a state of sedentary, but now it's happening. You see data partnerships happening.

egta: *If I'm a broadcaster looking at the possibilities of data, what kind of questions should I be asking myself?*

IC: One of the things we often hear is clients talking about the new data space and thinking they've got to get their new programmatic pipe set up, but we actually think it's the other way round. You should understand your data first. Because when you start pulling all your data together, you begin to realise things you didn't really know: your customer has certain tendencies to do this or that, you never really realised they had an affinity with a particular football club, for example. These things only come up when you start to master your own first-party environment. The next thing is what do you actually want to do with it? Do you want to extend the audience and target more people that are like that?

egta: *At the tech level, how do I build the link to get my data into the DMP?*

IC: Data from any source will typically come in or out of the platform through one of three ways. Either we build a server-to-server connection with that technology, we API integrate that technology or we can do batch file uploads.

egta: *Are there any pitfalls to avoid?*

IC: The difference between success and failure comes down to a client having a really strong idea of what it is that they want to do and having the right people running that. If you can imagine inside an organisation, you've got people

coming from a sales background, a marketing background, a technology background, and if all of them are fighting in their own corner with no one overseeing a coherent strategy, it can get really waylaid, messy and non-directional.

With publishers, for example, the temptation is sometimes to see the DMP as part of ad ops. It's not like a piece of programmatic tech – in the programmatic age there's sometimes been a tendency to commoditise the DMP and heap it in there as well. Where's the green button that makes me money? It doesn't work like that!

You need the right people and the right vendor that can understand the balancing point between achieving short term ROI goals for the platform and the longer-term strategy. If you just focus on short-term deals and short-term briefs in agencies, you're chasing after your tail with no defined strategy and you'll get blown around by the wind. To have a DMP in an organisation is about preparing for the future, because not having one could result in you being blown away by competitors that do.

egta: *What's the next step with regards to video and audio platforms?*

IC: The question often gets asked in meetings – what's the next big thing? Or can you guide us through your product roadmap?

Some clients are asking this before they have mastered a foundation strategy for their data. So in a sense we are often calling clients back to basics when it comes to their data sets. In other words, let's gain a thorough understanding of our first party data sets before we start using it in a complex way or chasing the next big thing that is 'en vogue'.

That said there are a couple of big waves that are gathering pace and will be top conversation in 2017. Firstly, in Europe GDPR and privacy is informing almost everything tech companies and media players are talking about. This despite the fact the implications haven't fully been realised or thought through by many businesses.

From a product perspective we are seeing big conversations around convergence between traditional broadcasting and digital. These have huge implications for both video and audio. 'TV DMP' is an emerging and big phrase in 2017.

Digital Audio is also reaching an exciting tipping point, but still lacks a genuine scale behind its ad inventory ability. DMPs and 3rd party DMP marketplaces are extremely well equipped to help vendors like AdsWizz and publishers like Spotify, Deezer and SoundCloud as well as traditional broadcasters take full advantage of this coming wave.

DATA IN ONLINE VIDEO ADVERTISING

Emmanuel Josserand, *Marketing Director* – **FreeWheel Europe**

FreeWheel offers a comprehensive suite of services for the management and monetisation of premium video inventory. The company works with many of the world's leading television broadcasters to solve the complexities of the new TV ecosystem.

egta: *Can you explain the different models for serving ads through the FreeWheel platform?*

Emmanuel Josserand (EJ): Back in May last year we acquired StickyADS.tv, a premium video Supply-Side Platform (SSP) solution, which we are integrating with our ad server to provide full-stack solutions, i.e. enabling a broadcaster or TV programmer to serve ads that are traded either from direct sold or programmatic sources.

egta: *In the direct sold scenario, what type of data is typically going into the platform?*

EJ: FreeWheel ad-serves against three parameters: *what* is being watched (the content); *where* it is being watched (the device) and *who* is watching (the user). The ad decisioning is made based on these parameters, and that is supplemented by the data layer available. That all feeds into the decision engine, which will deliver an ideal ad choreography, i.e. selecting the number of ads that should be served within a particular pod, how many times an ad should be repeated over the course of the entire piece of content that is being watched, the appropriate ads, etc.

Talking about the who, this can be first-party data the programmers obtain from registration/ authentication details or third-party data; we currently work with multiple data providers or DMPs which are already integrated in the platform. When enabled through IP-connected devices such as a STB, return-path data can also be fed into the ad server to ad-serve against this data. FreeWheel is an open platform and as such enables the ingest of any data as required by the broadcaster to improve on its campaign effectiveness.

Additionally, we've recently made some important developments to include audience viewability data. Working with, amongst others, comScore, Nielsen, BlueKai, Krux and Moat.

FreeWheel's Audience and Measurement suite is a unified solution for clients' evolving audience and measurement needs. It enables publishers and advertisers to transact on the most effective metrics available through one efficient workflow and provides a holistic solution to address audience targeting and measurement needs so they can accurately measure – and deliver on –audiences across platforms at scale.

egta: *For those broadcasters that do not have a registration or social login, what type of data can they bring to the platform? Can they integrate cookie data? And how does this work in the mobile environment, where you can't rely on persistent cookies?*

EJ: Yes, absolutely. They can use a DMP to add a layer of data for instance; but we are indeed using some cookie data to understand usage and define the three key parameters (what,

where and who) to make decisions on which ad to deliver.

In the mobile environment, the same principles apply, where we use our own cookies.

egta: *Do you have an idea of what percentage of the broadcasters you work with have implemented a registration system?*

EJ: This is now becoming more commonplace, today in Europe I would say it's probably about a 50-50 split. The companies we work with are often advanced on the digital side, and most of them have now implemented registration or authentication systems, so they can acquire and manage their own data.

egta: *What type of data does the FreeWheel platform send back to the broadcaster?*

EJ: The data that is being ingested in the platform is used to deliver greater campaign results according to programmers' strategic goals. What is being sent back from the platform is actual reporting data, typically the number of impressions, ad views and videos views, etc.

We produce a series of quarterly reports (Video Monetisation Report) looking at the trends and changing dynamics of how content owners and distributors are monetising premium digital video content. This is based on over 200 billion video views.

egta: *Are there differences between the various ad servers on the market today?*

EJ: Yes absolutely, though there are not that many on the market! Since its inception FreeWheel solely focused on the supply side of premium video and we champion the growth and unification of television. We are laser focussed, and well underway, to achieving true unification across IP, VOD and linear. FreeWheel is also unique in offering an integrated full-stack solution supplemented by our highly skilled and experienced advisory services team (product and market). FreeWheel made a strong commitment in leading advocacy efforts on behalf of the premium video industry through the FreeWheel Video Council.

egta: *When it comes to programmatic video, is the model that you simply integrate with technology to expose a broadcaster's inventory to the wider market?*

EJ: To some extent – yes, it's that simple!

Now, let's just go back to programmatic, one of the industry's most recent buzzwords but a poorly defined term resulting in misalignment of priorities and often friction between buyers and sellers. We're defining programmatic as based on *Data*, *Automation* and *Transaction Model*. When applied correctly, these 3 elements can be used in a variety of ways to target audiences, drive efficiency and greater monetisation.

Now, publishers need to first consider strategy and business needs, as certain transactional types will not fit the premium model. Compliance for instance will require different approaches. Long-form video with multiple ads per pod requires certain controls and quality assurances (e.g. creative quality, competitive separation, stream integrity). It is also difficult to execute within an open exchange, biddable marketplace, where there are multiple demand sources, none of whom are aware of what ad ran before them… These are some of the risks that differentiate premium video from the rest of the video ecosystem, and why transactional models must be considered as separate to the inventory and campaign in question.

There is a wide spectrum of execution options available, each with its own pros and cons for premium video so it is very important for buyers and sellers to align on objectives and develop mutually agreeable framework. One that considers overall business goals when

developing a premium strategy and includes value of employing data + automation via the right transaction model.

egta: *So, do you think that open marketplaces offer a potentially viable model for broadcasters when some of these issues have been resolved?*

EJ: Programmatic enables buying/selling of inventory through several distinct transaction models, which should be considered extensions of the technology itself. For instance, open exchanges enable opportunistic demand but come with significant risks, while private marketplaces offer more control, and automated guarantees create predictability for both sides.

Transaction model is too often put ahead of business strategies on both sides, at the expense of business rules and overall campaign goals.

Publishers and advertisers should embrace the value (e.g. Safety and Compliance) associated with premium video and collaborate on execution, which will result is increased ROI.

egta: *When you go into private or open marketplaces, you can apply some dynamic pricing models – are you starting to see that happen?*

EJ: We are working towards what we call a holistic ad management platform - any inventory delivered via any distribution platform to any screen – with full broadcast compliance. Through our SSP integration we've already enabled price comparisons between demand sources: direct-sold and programmatic. Which means that a publisher driven purely by CPMs could make a spot available to a direct-sold campaign or to a programmatic buyer, depending on who is willing to pay more.

We're now going one step further, making it possible to sell ads within a break either directly or programmatically, with holistic management of the inventory regardless of the demand source. Through the combination of holistic ad decisioning and holistic yield management we're enabling to mix programmatic and direct sold ads in the same break.

As an industry, we need to ensure we are being smart about how data and automation-fuelled deals are executed, getting this right will benefit brand marketers, publishers and users alike.

egta: *What do you expect to see in the coming couple of years?*

EJ:

1. More programmatic trading for premium video; while today's premium video ad inventory in Europe is mainly traded via direct sold deals (90%+) we expect programmatic to represent 35 to 40%+ in the next 2 to 3 years.
2. Further developments in addressability. This was one of the big topics at the Future TV Advertising Forum at the end of last year - addressability and the rise of telco operators. I believe their place in the ecosystem is going to get stronger, and we expect to see more collaboration between broadcasters and operators. In the next year or two we will see closer collaboration and the first deployments of advanced advertising based on operator set-top box data.

egta: *What advice would you offer to any broadcaster or TV sales house looking to improve their data strategy, particularly around first-party?*

EJ: The data strategy needs to complement the content and sales strategy. You have to be wary that too much data can kill data; i.e. going too deep in targeting can be counter-productive. So, a data strategy need to come as a complement to the sales strategy. You need to start small and grow with experience. Data should be an additional layer to help in achieving greater yield keeping in mind the advertiser requirements, the publisher revenue goals and the user experience.

CASE LEARNINGS:

- Ad decisioning is carried out on three parameters: *what* is being watched, *where* it is being watched, and *who* is watching
- The identity of viewers can be derived from first-party registration/authentifcation data or DMP third-party data
- About half of the publishers FreeWheel works with have implemented a registration or authentification system
- Programmatic is defined as based on *Data*, *Automation* and *Transaction model*
- Publishers need to first consider strategy and business needs, as certain transactional types will not fit the premium model.
- It is important that buyers and sellers align on objectives
- In programmatic, open exchanges enable opportunistic demand but come with significant risks
- Private marketplaces offer more control
- Automated guarantees create predictability for both sides
- A data data strategy needs to complement the content and sales strategy

GOING INTERNATIONAL – UNLOCKING THE POTENTIAL OF INTERNATIONAL BROADCASTING THROUGH NEW MEASUREMENT AND DATA APPROACHES

MTM

Jon Watts, Managing Partner at international research and strategy consultancy MTM, offered a point of view on the challenges of understanding TV audiences across international markets at egta's Market Intelligence Meeting (2 February 2017, Rome). The following article is an egta summary of his analysis and advice, bringing together findings from various projects and research initiatives undertaken by MTM.

Television in robust health, but challenges are growing

Television advertising remains competitively priced against other media in most markets, is proven to be effective in delivering high ROI, and viewing remains relatively robust, although consumption is becoming more widely distributed across different devices and windows. Some threats, such as Netflix and other OTT services, appear to have been overstated, at least in the short term, and media agencies remain positive about TV, as they face up to some of the challenges and limitations of online media.

However, competition from international Internet companies for ad investments is intensifying. Major players, such as Google and Facebook, offer excellent data and a single buying point for pan-regional campaigns, providing agencies with a simple, powerful proposition that broadcasters – national or international – struggle to match. Furthermore, these platforms are becoming more sophisticated, using data in increasingly complex ways to offer real-time advanced targeting and to drive programmatic advertising.

Audience measurement in a fragmented media landscape

Television audience measurement (TAM) is currently carried out using two methodological approaches – panels and census measurement – with data fusion models being deployed to provide a more holistic picture of viewing across different devices. A variety of tools are used for measurement in different markets, including peoplemeters, set-top boxes, software meters, router meters, portable meters and set meters.

While this adds complexity to audience measurement, the greater challenge lies in growing fragmentation and the blurring of media categories. Traditional media categories, such as TV, radio and print, are becoming more complex, as media owners transition into multi-media companies, investing in digital offerings that combine print, text, video and so on. As consumption becomes a cross-platform phenomenon, blind spots are appearing in television audience measurement, making holistic planning and reporting more challenging.

The TV industry responds

Several markets across Europe and elsewhere are responding to the challenges of cross-platform consumption by developing new total video measurement solutions, with innovations coming from research providers, national joint industry committees (JICs) and media owner committees (MOCs). The new statistical approaches being developed bring together panel and census data, using data modelling and other techniques to deliver a more comprehensive understanding of video consumption.

However, there are significant challenges associated with adapting television measurement towards a total video solution. Measuring new forms of video consumption is expensive, and with live TV remaining the dominant means of viewing – and the main revenue source for broadcasters – there is limited incentive to invest in upgraded measurement systems. Furthermore, it is difficult to provide integrated ratings for TV and online video content and advertising, given the granularity and fragmentation in online, and the complexity of the huge datasets involved leads to significant costs. There is a real risk of negatively affecting television's trusted currency while delivering questionable benefits.

A shifting focus for audience measurement

In traditional television, simply measuring viewer exposure forms a robust and trusted basis for transacting advertising, especially when combined with effectiveness and attribution research. However, as video is now consumed on a wide variety of different devices with screens large and small, and the context can be anything from live streamed TV to short, user-generated clips on social media platforms, one-size-fits-all measurement systems are becoming more problematic. For newer video formats, audience measurement needs to take account of more than just exposure, considering factors such as viewability, attention quality, environment and interactivity in order to deliver meaningful metrics that media agencies can use to plan and buy campaigns across different forms of video.

As digital advertising moves increasingly towards programmatic and real-time trading, there is a risk of a growing disconnect between media planning on the one hand and targeting optimisation on the other. Agencies are looking to deliver frequency capping and cross-media optimisation via targeting, rather than planning, but TV audience measurement will not easily support this.

An uneven pace of evolution of audience measurement from market to market

One of the challenges faced by international broadcasters stems from the different levels of audience measurement sophistication between countries. While some individual markets are moving quite quickly towards effective solutions for total video measurement, they are doing so in slightly different directions and at different speeds. A number of Western European countries, such as the UK, the Netherlands, Germany, France and the Scandinavian markets, for example, already have sophisticated audience measurement in place and are making progress towards four-screen solutions. A second tier of countries, for example Spain, Poland and Turkey, have relatively sophisticated audience measurement, but they currently do not have fusion models to bring TV and digital together. A third set of countries lack properly effective measurement.

The problem for international broadcasters is that the environment of measurement, planning

and schedule optimisation that they operate in is becoming increasingly fragmented and diverse. Rather than converging, different markets are actually diverging at the current time from a measurement perspective.

Strategies for international multichannel broadcasters to deal with measurement challenges

MTM's research led to seven recommendations for international broadcasters.

1. **Acceptance:** As there are no realistic prospects for a single approach to measuring TV and video throughout the EMEA region, broadcasters should accept that TV remains a medium that is largely measured and bought on a national basis.
2. **TAM tools:** There are common audience measurement tools, such as TechEdge, that broadcasters can use to better understand consumption of their content on a pan-regional basis.
3. **Index:** Broadcasters can develop a picture of how well their channels are performing in different markets by establishing performance indices that take into account how well they are rated by each country's TV panel.
4. **Data platforms:** Many major broadcasters now employ data scientists and leverage self-serve platforms to synthesise the large volumes of data they have access to.
5. **Strategic teams:** As markets become more fragmented, international broadcasters can put in place centralised teams to run programmes or projects that can deliver insights and lead to change.
6. **Research:** Following from the point above, centrally commissioned strategic research can be used to support local markets.
7. **Collaborate and invest:** Most importantly, international broadcasters need to make common cause with the rest of the industry and work together to develop better approaches to measuring total TV performance across different geographies. This process is critical if television is to compete with the major online players, which are competing aggressively for advertising investments.

REFERENCES

1. egta (January 2016). *Advances in hybrid television audience measurement*
2. Pike, L. (14 April, 2016). *"What is First Party Publisher Data?"* SpotX
3. Interactive Advertising Bureau (Updated January 2016). *Data Segments & Techniques Lexicon*
4. Creighton, M. (5 February, 2016). *"Intent Data Is Powerful Tool For Business Decisionmakers"*, Madison Logic
5. Halo (21 July, 2016). *"Descriptive, Predictive, and Prescriptive Analytics Explained"*, Halo Business Intelligence
6. IAB *Video Player-Ad Interface Definition (VPAID)*
7. Facebook for Developers: *Facebook Platform Policy*
8. Moore, S. (8 June, 2016). *"How to Choose a Data Broker"*, Gartner, Inc.
9. Levy, E. (21 September, 2016). "*Demystifying Data Warehouses, Data Lakes And Data Marts"*, Sisense, Inc.
10. Kihn, M. (30 July, 2015). *"How Does a DMP Really Work?"*, Gartner for Marketers
11. Advertising Age (24 June, 2014). *"What Is Data On-Boarding?"*, AdAge
12. Schuster, J. (16 February, 2016). *"How LiveRamp Data Onboarding Integrations Work"*, LiveRamp
13. Chang, H. (22 April, 2016). *"Deterministic vs Probabilistic Data"*, Adobe Systems, Inc.
14. Lotame (27 February, 2014). *"Lotame Optimizer Leverages Advanced Machine Learning"*, Lotame
15. Waterhouse, D. (3 March, 2015) *"Programmatic Video #WotW: What Is An SSP?"*, Unruly
16. Ng, R and Kattukaran, A. (March 2015). *"Evolution of TV: The Promise of Programmatic TV"*, Google
17. TubeMogul (April 2015). *A UK Advertiser's Guide to Programmatic TV*
18. ANA and White Ops (January 2016). *2015 Bot Baseline Report*
19. Ingram, M. (21 December, 2016). *"The Russian Methbot Scam Is Just the Tip of the Ad Fraud Iceberg"*, Fortune
20. World Federation of Advertisers (June 2016). *Compendium of Ad Fraud Knowledge for Media Investors*
21. egta (June 2016). *Ad blocking: strategies for television and radio publishers*
22. Sheftalovich, Z. and Meyer, D. (15 December, 2015). *"EU strikes deal on data protection rules"*, Politico Europe
23. European Commission - Press release (10 January, 2017). *"Commission proposes high level of privacy rules for all electronic communications and updates data protection rules for EU institutions"*, European Commission

GLOSSARY OF ACRONYMS

API:
Application programming interface

CPM:
Cost per mille (cost per thousand)

CTR:
Click-through rate

DMP:
Data management platform

DSP:
Demand-side platform

GDPR:
General Data Protection Regulation

OLV:
Online video

PII:
Personally identifiable information

ROS:
Run of site

RTB:
Real-time bidding

SDK:
Software development kit

SSO:
Single sign-on

SSP:
Supply-side platform

ACKNOWLEDGEMENTS

egta extends its thanks to the individuals and companies that provided information and expertise for this publication, as well as their colleagues who support egta in many ways.

This publication has been made possible through the contributions of:

Didier Angelo (TF1 Publicité), Caroline Beadle (Sky Media), Ian Curd (Lotame), Pete Doe (clypd), Jeff Eales (Sky Media), Martin Greenbank (Channel 4), Gonzalo Higueras (Gigya), Emmanuel Josserand (FreeWheel Europe), Thierry Tacheny (Divedia), Nadège Verdurmen (TF1 Publicité), Jon Watts (MTM), Charlene Weisler (Weisler Media).

LEGAL NOTICE

MORE ON THIS TOPIC

Editorial committee:

Matthew Carver
Head of Insight
E: matthew.carver@egta.com

Yuri Loburets
Director Radio
E: yuri.loburets@egta.com

Anne-Laure Dreyfus-Coutinho
Director TV
E: annelaure.dreyfus@egta.com

Caroline Brasseur
Director Research
E: caroline.brasseur@egta.com

Design:

Paulina Kott
Head of IT & Design
E: paulina.kott@egta.com

MORE ON EGTA

egta - association of television and radio sales houses

egta is the Brussels-based trade association of more than 130 television and radio advertising sales houses. egta's members are spread across 40 countries, mainly in Europe. Together, egta's TV members represent over 80% of the European television advertising market, whilst egta radio members collect 60% of radio advertising revenues in countries where they are active.

As sales houses of both public and private broadcasters, egta members commercialise the advertising space around audiovisual content available on platforms such as traditional television and radio sets, tablets, smartphones, PCs, Smart TVs and other Internet-connected devices.

22, Rue des Comédiens, boîte 4
1000 Brussels - Belgium
T: + 32 2 290 31 31
www.egta.com
@egtaconnect
www.egta.com

egta.

www.egta.com