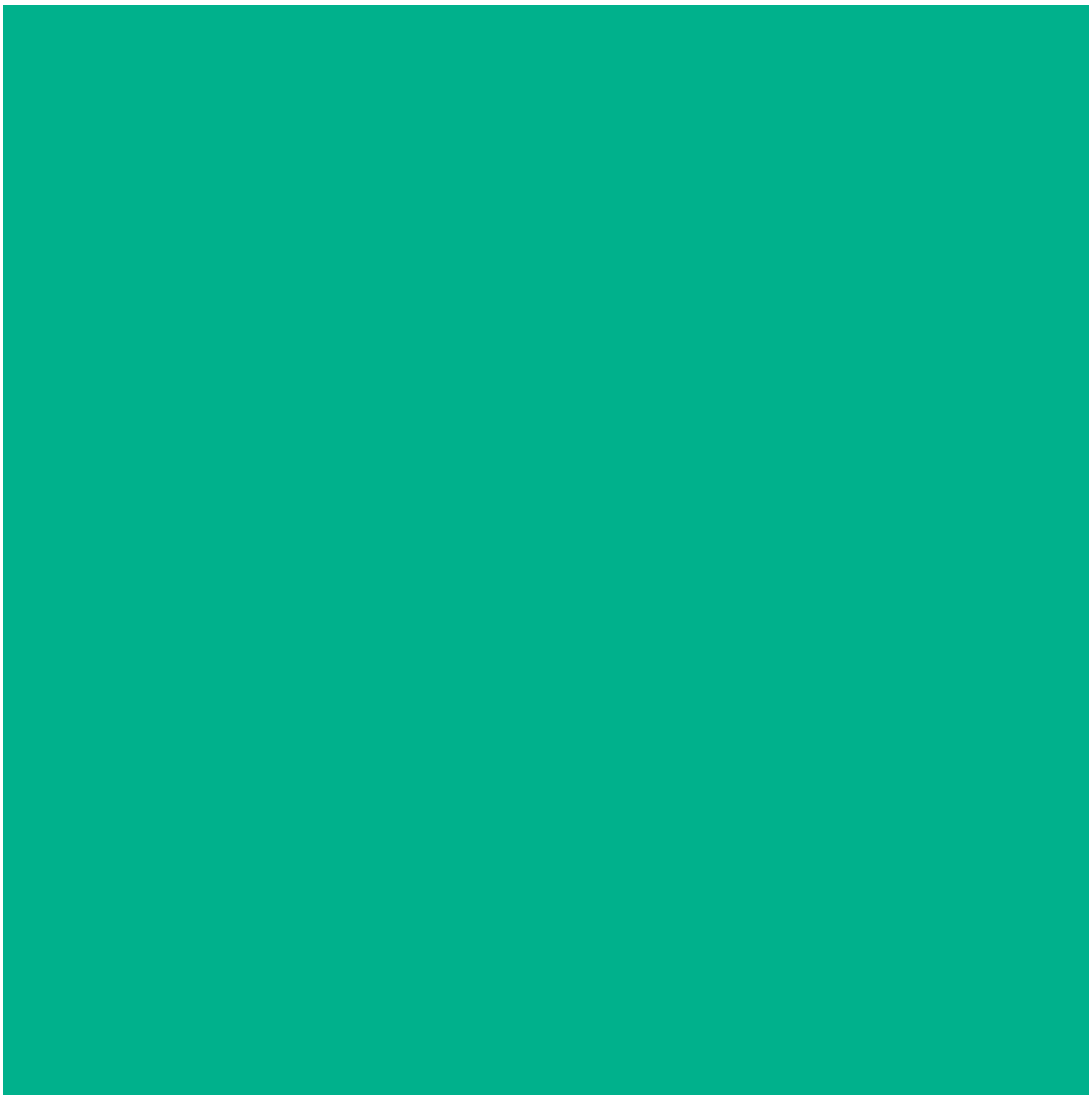
Department for Digital, Culture, Media & Sport

REVEALING REALITY

November 2021

How user data shapes the media sector



Contents

1. Introduction

Unlocking the value of data is a key priority for the Government's Digital Strategy, and the National Data Strategy has asserted the need to build a world-leading data economy—both for economic gain and the opportunity to deal with some of our biggest societal challenges.

Within the media sector, data and digitisation have dramatically changed the markets in which businesses now operate. Newspapers have experienced long-term decline in both print circulation and traditional advertising revenues as consumers have moved online, and press businesses have struggled to compete against dominant players in the digital advertising market.

Television has similarly experienced a decline with consumers, especially the young, increasingly turning to digital platforms and SVOD (subscription video on demand) platforms such as Netflix.

Radio is suffering similar challenges through changing listening trends and the growing prominence of smart speakers.

DCMS aims to support a healthy and world-leading UK media by ensuring sustainable and successful business models. It is recognised that the collection of better data about audiences and possessing the expertise to leverage this to improve personalisation, enhance services, build up subscriber bases and increase the value of advertising is one route to this. Following the Cairncross Review[1] into what a sustainable future for quality journalism looks like, better use of consumer data was a key recommendation: "news publishers must get better at differentiating their advertising space from other online sources, and in particular, at collecting more granular information on their readers to offer advertisers better targeting opportunities". However, this research has found that often, it's not simply a case of getting more, or becoming better at using, consumer data.

Overall, the research asked the question: **What are the barriers to media businesses and organisations making better use of data?**

Some of the more detailed objectives for the research were:

- Exploring how media businesses use consumer data and how this differs between different types of organisation
- Exploring *attitudes* towards consumer data in different businesses
- Exploring data and digital *skills* across businesses
- Exploring barriers and challenges to using data—practical / technical, strategic, legal, ethical

[1] The Cairncross Review: a sustainable future for journalism, 2019.

2. Executive summary

Data can be extremely valuable for media organisations

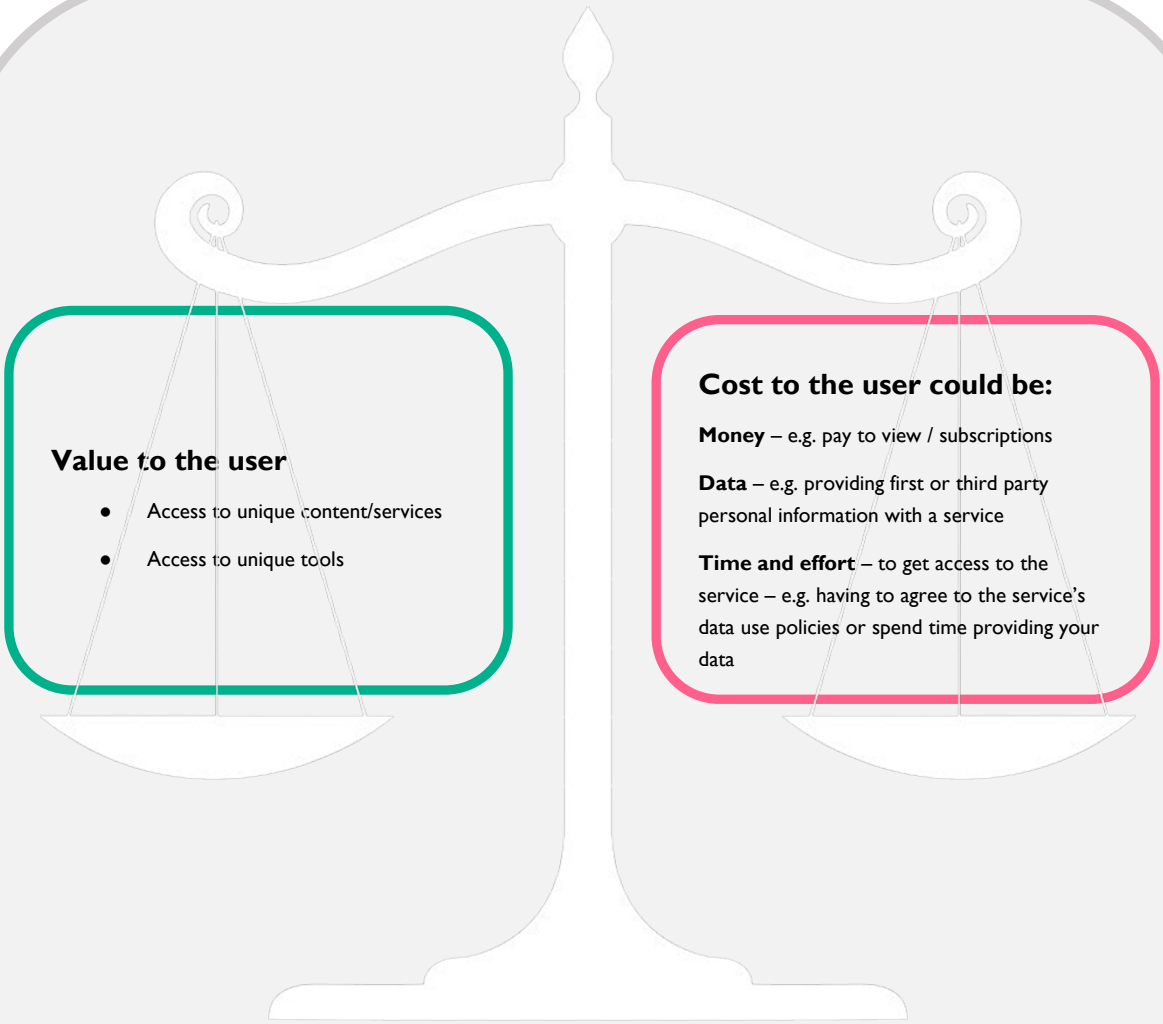
Across the press, TV and radio sectors, consumer data was seen as extremely valuable. It was being used for a range of purposes, broadly:

- **Generating revenue through advertising**. Selling how many users they have, and who the users are. Some were selling targeted advertising where they were able to get data on who individual users were
  **For more detail on how data is valuable for advertising, see the pull-out box at the end of this section*
- **Content / editorial decisions.** Understanding which content is doing well and not as well, to shape future content
- **Strategic decisions about service / product development**. Understanding how users are engaging with services and products, to shape future feature and product development. For example, creating an app due to lots of people accessing their service via mobile
- **Personalising content.** Personalising the content individual users see based on data the organisation has about them

It's not as simple as more data = better business

It is true that there is scope for businesses to generate more revenue from increasing their use of data. They have the potential to boost the value of advertising space, drive subscriptions, and improve their service through better audience insight or increased personalisation.

But there are also costs, and risks, associated with doing so. There are the obvious financial costs to a business of investing in the capability to collect and process data, as well as the challenge of navigating legal parameters. And for the consumer there is a **value exchange** associated with being asked to give up their personal data.



On the one side is – **value to the user**

- **Access to content or a tool they can't get elsewhere** – if sharing data is the only way to access a certain type of content or service that is in demand, users are far more likely to accept the value exchange
- **The promise of access to more tailored or exclusive content** – sometimes sharing data provides additional value to the user and comes with the promise of an improved experience

On the other side is – **cost to the user**

- **Organisations need to consider the different kinds of friction incurred on the consumer** – sign-in, paywalls (financial cost), data-use agreements. If consumers are not sufficiently motivated to access the content on offer, they may easily drop away or go elsewhere.
- **Many consumers have an instinctive aversion to businesses collecting their data for commercial gain** – this is reinforced by an increasingly heated public narrative surrounding privacy. If it is not clear to consumers that they will personally benefit from sharing data, brand perceptions can be harmed, and loyalty lost.

If there is an imbalance within the value exchange and the cost is too great, efforts to collect data could mean a significant drop in audience numbers for media organisations. Organisations have to carefully weigh up the two sides to the exchange before making any changes to their data strategy.

For many tech giants, who can often provide the user with access to huge amounts of in-demand content and keep the friction of data-sharing low, the benefits have historically clearly outweighed the costs. The tech giants are the current winners in the data economy.

For a range of reasons, many media businesses do not have access to the same potential benefits or are more at risk from the potential costs. They are not operating on a **level playing field**—thus the cost-benefit equation can easily sway in the opposite direction.

The barriers to a level playing field

The big players have a status quo advantage

Having relied on collecting customer data for many years, the big players (Apple, Google, Amazon, Facebook) often hold established long-term customer relationships. They have one-time consent agreements with customers, from which customers agree to sharing their data thereafter, and are not asked again. This data can then be shared across all the company's products and platforms.

> *"So when you sign in, although you think 'I'm signing in so I can get a Gmail account', you're also agreeing for Google to collect all your data on all their services"* – Large national newspaper

Some (e.g., Apple, Google) are making efforts to reinforce their dedication to protecting user privacy, reinforcing this status quo. For example, through Google's 'Privacy Sandbox' (explained later).

Huge volumes of content = more value to offer from personalisation

User generated content platforms (e.g. social media) or platforms that offer aggregated content (e.g. Google) can offer far more value from personalisation through data sharing, as the amount of content they serve means they can provide an entirely personalised experience. The content these platforms distribute often includes the content produced by media organisations (e.g., news articles).

Media organisations have historically relied on providing comparatively limited volumes of editorial or commissioned content, therefore the value proposition to customers through personalisation is much more limited.

> *"Some organisations don't have a massive pool of content in order to give everyone a personalised content page"* – Medium sized media group

Tech businesses sit upstream and can collect wider ranging data

Many of the larger tech businesses provide the digital devices, operating systems, interfaces or applications through which consumers access content from media organisations. Therefore, the data they collect often sits 'upstream' of media organisations, meaning they have an inevitable advantage in terms of the volume of data they can access.

> *"So they're all vertically integrated through the device, through the back end ecosystem and navigation tools and voice interfaces. Also, the consumer facing applications. So they've got a level of data there that is all encompassing"* – Large radio station attached to large media group

Tech businesses exert control over how data is used

As upstream providers of digital devices, OS, etc., several large tech companies are able to exert control over how data can be shared, accessed and used by other organisations, including media businesses.

For example, currently, many media businesses rely on third party cookies to gather data on user behaviour beyond their own website/app—which is hugely valuable to prospective advertisers. Google's announcements (and subsequent delays) of their intention to restrict use of third party cookies via their services is of great

concern to many media organisations, not least because one of the proposed alternatives, Google's 'Privacy Sandbox' will likely end up driving more business in Google's own direction:

> *"They [Google] under this ['Privacy Sandbox'] proposal, they'll make packages of data available to third parties like us. They call this FLOC, which stands for Federated Learning of Cohort's. But, this isn't the same as us having our own data and being able to go to other companies for data and being able to match data across a number of different sources, and it puts Google in complete control."* – Large, national newspaper

Media organisations hold themselves to different standards than the tech companies

Media organisations have to comply with a range of different rules and regulations. Some that are imposed on them by law, others which are more self-imposed. PSBs have to follow regulatory principles, radio stations have to follow certain rules to keep their licence and press organisations often self-regulate and are expected by the public to uphold certain journalistic standards.

For example, in the press sector, IPSO (Independent Press Standards Organisation) is an independent regulator with a role to uphold standards in the sector. There is another such body called IMPRESS.

Big tech companies do not necessarily hold themselves to the same standards when it comes to these kinds of self-imposed rules, and are benefitting because of it.

If these self-imposed standards are something which society itself values—there has been extensive research into the value of press and local journalism to society[2], for example—this shouldn't leave media organisations at a disadvantage in relation to the big tech companies.

Where does this leave media organisations?

These factors all shape the equation when a media organisation is setting out their strategy and vision for using data.

For many, especially in the press sector, an additional factor to consider was the potential impact of data use on their organisational vision and editorial principles. Using data to push increased personalisation or higher value tailored advertising was seen to be at odds with their core proposition and value to audiences.

Many of the organisations we spoke to talked about the *role of the editor*, which is regularly in tension with the idea of maximising engagement.

> *"We know that curation, and editorial role is valued by many. So we want to avoid this commoditisation of data."* – Large media organisation

While the impact of these pressures were felt differently by individual organisations, the decisions and trade-offs that they faced were often similar. Many businesses had opted for (or felt pressured to choose between) one of the following routes:

- Use data to increase the value of advertising, accepting the risk and/or embracing the pressure this could place on content production (e.g., producing more 'click bait'/lower quality content to drive ad revenue)
- Prioritise editorial quality and move towards a subscription funded or signed-in environment, using data to fuel subscription sales and target content / advertising, but to a potentially smaller audience
- Prioritise editorial quality and a universal access offer (e.g., no paywalls), retaining audiences and their company values but sacrificing the potential income and personalisation gains to be made from increased use of data

The following chapters outline in greater detail the landscape described by media organisations across the sectors of press, TV and radio—and the pressures and challenges faced by them in making use of data.

[2] See report 'The Social Value of Local Journalism', for example.

***A note on *why* data is valuable for advertising**

A key use of customer data is revenue generation via programmatic advertising and other advertising deals—but what is the actual value of data to advertisers?

CPM (cost per mile (1,000)) is the value of 1,000 ad impressions to an advertiser. Where impressions can be measured this is exact, where not, these are estimated. There is huge variation in the CPM.

Value to advertisers is driven by how much their adverts are seen by the right kinds of consumer. The more targeted the ad impressions, the fewer which are 'wasted':

- If you put an advert for men's shoes in front of 1000 people, the half that are placed in front of women are wasted (for the purposes of this example). Of the remaining ads that do go in front of men, the vast majority will be wasted too, because the ad is still not relevant: people of the wrong age for the product, with different styles and taste, not in the market for shoes at the time etc. The value of this 1000 adverts is limited.
- If you can put 1000 ads in front of only men in the target age range, with a certain taste in shoes and at least six months since they last bought a pair of shoes, there is likely to be a lot less wastage because they are more likely to be relevant—the value of that 1000 ads is much higher because there is a much greater chance of them turning into sales.

The data needed to drive this targeted advertising comes from the publishers, and can be collected in numerous ways as long as it is attributable to an individual. The most common way for press publishers to access this kind of data currently is via third party cookies.

One way to generate a similar revenue but from less targeted advertising, is through increasing the volume of ad impressions for the same cost—a lower CPM:

- If someone reads one article on your website you can serve them one advert.
- If someone reads two articles, you can serve them two ads.

So rather than having one article that takes five minutes to read, you can encourage them to jump from one one-minute article to the next. You've sold five impressions instead of one to compensate for the lower CPM.

Getting this much engagement with your site requires making editorial decisions that preference engagement over other factors (possibly including quality).

3. Method and sample

Methodology

The insights throughout this report were gathered from detailed interviews with representatives from the press, television and radio sectors. Interviews were carried out remotely on video conferencing platforms, and lasted approximately one hour.

Topics broadly covered:

- Background to the organisation and individual
- Current knowledge about their audiences
- Organisation's business model
- Overview of the role of data within their organisation
- The skills/teams the organisation holds in relation to data use
- Their data strategy
- Data sources
- Barriers the organisation faces to accessing and using data
- Their plans for future data use
- Knowledge of innovations or changes in the data landscape
- Support they feel is needed
- Deep dives into specific data use cases (dependent on organisation): advertising, customer insight, content decisions, selling alternative products, selling to third parties, social media analysis

Sample

We spoke to 25 media organisations—sampled based on size[3], type (radio, press, TV), format (e.g. online or offline), ownership (independent or part of a group), audience (national or local), location within the UK and proficiency with data—this ensured the research covered a range of experiences and perspectives on data collection, use and the potential barriers at play.

The organisations included:

Press (x12)
- x5 large traditional newspapers
- x3 publishing groups of local newspapers
- x2 independent locals
- x2 social media-first news platforms

TV (x8)
- x5 large Public Service Broadcasters
- x2 independent stations
- x1 trade body for tech organisations in the UK (including TV manufacturers)

Radio (x5)
- x3 large radio groups
- x1 independent stations
- x1 small radio group

Within these organisations, we spoke to people with a range of different roles in order to try and get the broadest understanding of the issues and attitudes surrounding consumer data. It was important to get a data perspective, an editorial perspective, a strategy perspective and a policy perspective. The roles included but were not limited to:

[3] Throughout this report, we attribute the quotes we gathered during the interviews in a way that keeps the people we spoke to anonymous. We use the descriptions 'small', 'medium' and 'large' throughout. Small = 10 to 49 employees. Medium = 50 to 249 employees. Large = 250 employees or more.

- CEO
  Head of Data and Insight
- Head of Digital
- Editor Emeritus
- Data General Manager
- Director of Policy
- Chief Analytics Officer
- Head of Strategy

And to understand more about how media organisations go about using consumer data to generate revenue through advertising, we spoke to:

Media Agencies (x2)

- x2 large media agencies

Consent and anonymity

The organisations and individuals that we spoke to throughout this research have all been anonymised. Participants were informed at the beginning of the interviews that they and their organisation would be made anonymous. Organisations gave prior consent for the insight throughout this report to be shared on that basis.

Developing and refining the barriers to data use in the media sector

Throughout the research, we wanted to understand the different levers and barriers to data use. From the start of the research, we consistently consulted and held collaborative sessions with the team at DCMS, in order to explore how the different characteristics of organisations we spoke to impacted their data use and the barriers to data use. We mapped these drivers and barriers throughout these sessions, which helped to form the structure of this report and the key issues to cover as the project progressed.

4. The sectors

The organisations we spoke to all had differing relationships with consumer data. For some, data was central to their ongoing strategy. Many of these organisations were innovating and investing in leveraging consumer data to improve the quality of their content and increase revenue.

Others had a more simplistic relationship with consumer data and knew relatively little about their audiences. However, all were facing challenges in effectively collecting and utilising this kind of data.

As explained in section 2, investing in collecting and using data doesn't always pay off for organisations and some were struggling to compete in a landscape where big tech dominates. The organisations we spoke to within each sector were often facing similar challenges—these challenges centred around:

Access

Can media organisations access the type of data they need to make editorial and commercial decisions?

All organisations faced barriers to accessing data about their customers. Often this was due to the role that big tech organisations play as an intermediary between customers and the media companies—controlling access to content (for consumers) and access to customer data (for media companies).

Many challenges faced by media organisations stem from this initial lack of, or reduced, access.

Ethos

Can media organisations provide the type of content/service to the public that they want to or are expected to (by the public)?

Decisions about what data to collect and how to use it have a 'moral' element as well as a commercial one.

Data collection enables several things, including personalisation of content and an ability to generate revenue from advertising.

But there are some reasons a media organisation may be averse to pursuing a strategy of maximising volume of content and user engagement over 'quality' and curation of content; or on the other side, to putting exclusive content behind a paywall. These were alluded to in section 1, and will be expanded upon within this section.

Value

Does the data a media organisation collects provide suitable commercial value to offset the cost of collecting and using it?

Sheer volume of data is not always endlessly valuable to a media organisation, and because there is always a cost (in some form) to collecting data, it does not always make sense to simply collect as much as possible.

Resource

Does a media organisation have the technology, skills and resource to collect and use data in the ways they need to?

Some organisations were not able to invest in the technology, skills and time needed to collect and utilise data, which can be considerable. The value data can provide does not necessarily offset the costs of implementation.

As expected, different teams were likely to focus more on different barriers—for example, those more responsible for content/editorial decisions compared to those more responsible for commercial decisions. However, many of those we spoke to in this research were in senior positions who were able to comment on data use across the organisation.

Within this section we'll delve into each sector in detail, providing:

- An overview of the landscape for each sector
- How organisations are currently leveraging data
- Exploring each of the above challenge areas and the impact they have on press, radio and television within the UK

Some of the challenges mentioned during interviews were **current issues** whereas others were **future concerns,** based on trends in the data and media landscape. Where possible, we have differentiated between the two.

a. Press

Overview of the press landscape

While all media organisations have to navigate the concept of a value exchange with customers, some of the challenges in terms of editorial decisions, subscription models and access to third party data were particularly prominent in the press sector.

Within press there are traditional / legacy organisations and new digital / social media-first organisations, and both have slightly different relationships and approaches to data.

Traditional or legacy press organisations (e.g. national newspapers) are creating huge amounts of content that is increasingly consumed on or via third party platforms (i.e. predominantly social media and the big content distributors). They are often adapting the content they create to fit these new mediums, but were keen to maintain certain standards of editorial control, to ensure they were providing value to their audience. Some were using sign-in and subscription as a way of collecting more in-depth data about their audiences, whilst maintaining editorial control.

> *"A subscription model where you're putting out higher quality content allows a more considered approach to journalism. A reader who's got a propensity to subscribe is very different to someone getting their news free from somewhere"* – Medium sized media group

New, digital and social media-first press organisations were often more dependent on social media platforms to engage users. These platforms are, in some cases, the reason they exist in the first place. As a result, these organisations are likely to be entirely familiar with using built-in analytics on social media platforms to, for example, assess performance of certain posts and adapt what they are doing to maximise engagement levels.

> *"We get lots of data about what people read on Facebook and what gets shared. Then we get a deeper dive when it's a video. This side of things then informs editorial."* – Small, social media-first publisher

> *"All of the stuff that we've done [published] kind of comes from the audience, I think that is the cycle that our founders talk about, it's 'put something out there, see what the reaction is', and if it's good, we keep doing that sort of stuff. And that's how we've grown to the size that we have today."* – Large, social media-first publisher

Challenges to data use

Access

Social media and tech companies were perceived to have better access to user data than press organisations

Social media and tech platforms host and distribute a huge amount of the content that publishers produce. Increasing numbers of users are consuming this content through these platforms. When this happens, these host/distributor platforms have access to first party user data. The publishers, unless the consumer is asked for additional consent, do not.

> *"They will have what they call first party data because they can collect it across all their various services…We only have the data they give us"* – Large, national newspaper

Organisations with a large social media presence do benefit from extended reach and active promotion of their content to users by social media platforms. However, when it comes to data, social media-based views provide the organisation with very little information about the users who have engaged with their content. When users engage with content directly on publishers' websites and are asked to sign-in, organisations have access to a range of first party data, including name and certain demographics. Whereas:

> *"Users and readers on Facebook and Google are anonymised"* – Medium-sized, local newspaper

This rang true across the publishers and press organisations we interviewed. The data that social media companies are able to provide to publishers tends to be focussed on engagement metrics. This is a key piece of information for publishers as it enables certain types of editorial decisions to be made (i.e. how can content generate increased engagement). However, they are not able to access as much information about who the users are—which is ultimately much more valuable to advertisers.

"We monitor our social media analytics, so Facebook and Google give you some summary demographic information about the users who engage with that, but with users on Facebook and Google we don't know who each reader is" - Independent, local newspaper

The social media company, however, does know these things, and is able to add the fact the user read some content from that publisher to their bank of data on that user.

Some of the big tech companies who benefit from the free news content generated by UK media organisations have set up funds or programmes to provide some limited support or reimbursement, recognising the value they get from hosting content that generates engagement from users.

There are some existing partnerships between social media organisations and press

Facebook have set up the Facebook News tab, which funds some of the journalism from press organisations whose articles are shared on the platform.

However, one of the organisations we spoke to felt that smaller organisations had been excluded from the deal.

> *"It's fine for us, it's a deal that we're happy with, and is a three-year deal. We're getting significant funding out of Facebook. That said, there are a number of companies, smaller than us, who've been completely excluded from the project. And if I was in their shoes, I'd be mightily unhappy with it"* – Medium-sized media group

The smallest publisher they knew of that were involved in the project owned "two daily papers, and about five associated weeklies", and they guessed the "cut off point for digital audience size was around a minimum of around 15 million unique monthly browsers".

Other press organisations felt they didn't get a fair deal for their content being hosted on Facebook, but that it was better than nothing.

> *"We have a deal with Facebook. The terms are pretty much take it or leave it. There was very little room for negotiation. We've accepted it, but there are a lot of doubts about it. And there isn't really much clarity around how much traffic it's delivering. We also talked to Google, but the terms they offered were very poor and the contract was very restrictive."* – Large, national newspaper

Single / one-click sign-in options meant press organisations were receiving far less data

Single sign-on has the potential to represent a positive opportunity for online publications as it could provide the benefits of sign-in without as much friction for customers (i.e. customers can sign-in with one click using existing credentials which could theoretically enable the publisher to collect first party data, without the user having to create a separate account for that organisation). In reality, organisations who let their customers use single sign-on get very little data from this approach, while the single sign-on platforms that customers log in through benefit from capturing additional user behaviour data.

What is single sign-on?

Single sign-on (SSO) is an authentication scheme that allows a user to log in with a single ID and password to any of several related, yet independent, software systems.

An example of this is the option on an iPhone to sign in with Apple ID across different apps. Instead of having to create a new set of log-in details for each app, users can log in with their existing Apple ID credentials.

These options—signing into a website at one click with your Google, Apple or Facebook account, rather than registering directly with the platform—provide an easy option for users. *They reduce friction.* When this option is used, however, the user has not created an account with the media organisation itself, and so the organisations don't have access to the user's data in the same way they would if they registered an account and signed in normally.

This means they cannot link sign in data with other data collected elsewhere about that user.

> *"Apple introduced the new single sign-on solution, where you sign in with your Apple account. But actually when they do that, we only get a 'Mac' code number-type email address. We don't get a name or any information about that particular person who is subscribing through the app ecosystem. Or the alternative is to sign in with Google or sign in with Facebook, which are also the identity controllers from the other two major platforms, and they might share a name or an email address, but Apple doesn't even share names or email addresses. It maps all the details under a six or eight digit string of numbers and then @apple.com…"* – Medium-sized local paper

When asked if they then had the option to specify that users can't sign-in through Apple, Google or Facebook, many organisations felt that people were used to being able to sign-in to things in this way, and the increased effort would put users off and affect sign up rates.

> *"You can, but obviously this can impact your rate of sign-up. Because people are familiar with face ID or touch ID and then they sign in with their Apple or Facebook accounts. It's one thumb away. When you're asking someone 'tell me your name, tell me how old you are' etc. it's a slightly different proposition for people."* – Medium-sized local paper

It is also potentially unclear to the user that using this option prevents the service they are trying to access from benefiting from their data, while sharing it with the company they have signed in with. It is likely that many consumers are unaware that the big tech companies are able to collect and utilise their data while the organisations they know and trust are losing out.

Organisations were concerned about the end of third-party cookies reducing the amount of data they have access to

For all press organisations operating online or on a digital platform (e.g. an online newspaper), third party cookies offer a way to significantly increase the depth and quality of data they hold on their users for advertising purposes. This enables them to participate relatively effectively (as suppliers) in programmatic advertising.

Organisations are able to track the activities of visitors to their websites across different domains—this provides additional value to advertisers as it gives information about where customers have come from and where they go next after visiting their website and being exposed to an ad.

> *"We were at least able to [know more about individual users through cookies]. To be able to tell an advertiser, well, here's a group of people who are of a certain age, certain social backgrounds and have an interest in Off-Road vehicles, or whatever."* – Large, national newspaper

Many organisations are reliant on third party cookies to sell advertising space—it provides low-friction access to audience data without requiring consumers to sign-in or subscribe.

The cookies come from the browser the customer is using. However, Firefox and Safari have phased out third party cookies and Google are planning to do the same, meaning organisations will no longer receive any information about audience behaviour before or after they visit their site.

> *"The whole concept of tracking people around the internet is sort of coming to an end anyway. So you can't forge those private relationships with your readers."* – Medium-sized media group

Press organisations felt this was likely to lead to a loss in advertising revenue and a huge advantage in the advertising market for big tech companies, who would 'still be able to' track audience behaviour across their platforms, devices and browsers.

One large press organisation explained that they're finding it increasingly hard to fill all of their ad space in Apple environments because they know so little about those users, which significantly reduces the value an advertiser is willing to place on these spaces.

With access to third party cookies coming to an end, press organisations were having to find other ways of accessing and utilising user data. Some of these options have already been mentioned. One way is to move to a sign-in / subscription model. Another, to maintain a free *and* easy-access product, is to move to a more click-based, social media model.

For some of the media organisations we spoke to, neither of these were viable options.

Google Privacy Sandbox and FLOCs (ringfencing data)

Part of Google's push for greater privacy for individual users is to put an end to third party data tracking. Their proposed replacement is the Google Privacy Sandbox. Many of the media organisations we spoke to took issue with this.

"Under this proposal they'll make packages of data available to third parties like us, which they're calling FLOCs, which stands for federated learnings of cohorts. This isn't the same as having our own data and being able to go to other companies for data and being able to match data across a number of different sources. And it puts Google in complete control. We only have the data they give us. We can only get these packages. We can't chop it and change it and divide it up in the same way we can at the moment. And some of the key digital advertising services like 'attribution' which is the way advertisers can tell who's actually seen their ads, which may not work at all. Or at best, we'll be relying on Google to perform that for us. So as far as the user is concerned, their details are actually being tracked more closely than before, they're just being tracked by one company rather than many companies. But as far as we're concerned, Google reinforces its monopoly position." – Large, national newspaper

Ethos **Value**

Some organisations didn't want to follow a 'social media model' as it would compromise the quality of their content

Most publishers we spoke to were straddling the line between developing content they consider valuable to their consumers, and content that generates the highest engagement or 'clicks' (the model chosen by social media first organisations). Many organisations are trying to maximise the latter, without entirely losing the former.

> *"It's certainly the case that if mass market news publications can't fund themselves, they will be forced down a route of trying to use what you might call 'click bait news', which, you know, generates users, generates clicks and is cheap to produce. And we all know where that will take us: no shortage of Love Island stories."* – Large, national newspaper

While it is possible for organisations to make editorial decisions based exclusively on what would generate the most engagement—and therefore opportunities to place ads in front of customers—this is seen as detrimental to the quality of the content.

> "*For every 1000 page impressions you're probably generating a couple of quid. So you have to have a hell of a lot of free traffic to have a viable business—so this leads to lowest common denominator stuff"* – Medium-sized media organisation

Many of the press organisations we spoke with—of all sizes—echoed this sentiment. The more views they get, the more money they earn. While this provides a route to financial sustainability for some, there is a perennial

worry that this *"forces you to be click bait-y and social media-y"*—which many organisations feel they can't abide by.

This challenge is well documented, but there is so far no solution or support for media organisations who are producing content they see as having a public value.

Some, like a large national newspaper we spoke to, were determined to keep their news "open to everyone". However, they foresee that the only way to remain free and open to everyone may be by massively diluting the quality of their content. This was a huge concern to this particular newspaper, as they do not want a world where all quality journalism sits behind a paywall.

Some were using data to personalise content, but others felt there were limits to this

Part of the 'social media model' is the ability to host or produce massive amounts of content, and then to personalise that content based on user engagement.

There is a huge incentive for user-generated platforms like social media companies to increase levels of personalisation, as it means they can keep users scrolling and serve increased amounts of targeted advertising.

In contrast, traditional press organisations have conventionally based their model on offering comparatively limited, editorialised content, which provides value to their readers.

There are some who, based on "progressive profiling" and other forms of data capture, were pushing towards a more personalised experience in order to enhance the reader's experience and engagement—hence also creating greater advertising value:

> *"We're creating a kind of very tangible, measurable community. And so we've got progressive profiling of our users. They have quite a low barrier to sign-up initially just because you want to get mass reach. So we're just getting first name and last name, email address, post code and date of birth just to help segment. And what we want to do will be to profile a lot more views and political leanings…It will make sure that our editorial is matching the views of our audience. So it's sort of like a kind of community measurement data capture, but we're trying to provide personalized, tailored content based on information that they've given us already."* – Digital and TV news organisation

However, for other publishers, it was not necessarily seen as within their interest to personalise *all* their content. They felt that this could undermine their editorial control, and result in people ending up in 'filter bubbles', only seeing content based on algorithms or assumptions about what they would want to see.

> *"What's the point of a publisher these days, is it to just present a load of stuff that people are self-selecting and is self-determined, or is it to actually try to make sense of the chaos? Should we be saying we do think there is value in an editing process that does determine what the things are that we think you should know about—people don't know what they want to read until they read it"* – Group of local newspapers

And even if it was in their interest to personalise content, many struggled to have enough content to do so anyway, as described below by a publishing group of local newspapers.

> *"Personalised homepages comes up time and time again. You could. But bear in mind, we do not have a limitless bucket of content so there's not that much content for most publishers to actually draw on. If you're the Guardian and you've got millions of stories in that, then yes, that model could make sense, but if you've got 30 stories, and there's not massive variety of stuff to give them content of their choice on their page"* - Group of local newspapers

Sign-in and subscription were not suitable approaches for all press organisations

Sign-in and subscription models can be seen as a way to protect the quality of content while providing another option for collecting and utilising data. Sign-in allows the collection of first party data directly from users, which had value to publishers wishing to target advertising at specific users.

> *"We are in the process of implementing a user sign in system to read as well as comment on our articles, in a move to reduce our reliance on social media platforms for data. The goal would be to get as many people who*

> *read us to sign up for free and tell us, not necessarily their name, but their gender, age and where they live, and once we have collated this data we can leverage it for more targeted advertising, rather than just appealing to advertisers who want to target everyone living in [local area]"* – Small, local newspaper

However, this comes with its own set of problems. For some organisations, sign-in and subscription often meant a significant loss in audience numbers—as the friction placed on the consumer is too great—particularly in a market that is full of free-to-access content.

> *"[If there was a sign-in required] I think we'd still get lots of traffic but no one would actually read any content, and then we'd probably start to die off over a longer period of time as people start to realise that they weren't going to get anything without giving their details"* – Large, social media-first publisher

Successful utilisation of sign-ins and subscriptions requires a customer base who attach a certain value to a media organisation's products/service. Regardless of the exact product or brand, requiring sign-in or registration of some kind will reduce audience—there will always be some people who will only access a service when it is entirely free and frictionless.

For press organisations who provide a mass market product, the move to sign-in and/or subscription was seen as less feasible—the less targeted/specific an audience and content, the less likely people are to pay for it, directly or indirectly.

> *"We've run trials elsewhere. The subscription revenue has never matched the ad revenue that you lose. There are two big obstacles to subscription. One is the BBC. The second is: there will always be someone who's offering some sort of advertising funded model. And it is difficult to charge for English language news aimed at the mass market, because there is absolutely no shortage of people trying to supply it. It's very hard [to move to subscription] on a mass market product."* – Large, national newspaper

Press organisations felt they were left with a difficult choice if they want to better access and utilise data for advertising: lean more towards the 'click bait model' and lose the quality and editorial control that many of their current audience value, or go the other way and put in place sign-in and subscription barriers which often means they lose audience numbers and less people have access to their content.

Resource

Smaller organisations didn't always have the resource to invest in data collection and management software

As well as not benefiting from the ability to partner with big tech and social media companies, many smaller press organisations felt they couldn't afford to invest in better data collection and management infrastructure.

Software and digital tools such as CRMs or Data Management Systems can be expensive, raising important questions about the value of investing in this technology. While they may be beneficial, the value they produce may not outweigh the initial and ongoing investment required to implement and maintain them.

> *"It is massively expensive to get these subscription and customer relationship management platforms. Third party agencies might say it'll cost £100,000 up front and £20,000 per month thereafter or whatever—so we've decided to do it in house"* – Medium-sized media organisation

This same organisation acknowledged that many smaller or local organisations may not have the requisite resource and skills in-house to be able to do this. And to incur the kind of investment cost mentioned above is sometimes out of the question.

Cost aside, others spoke about a lack of 'off the shelf' solutions that they could use to collect and use reader data in the ways they wanted to.

> *"Whether it is £1,000 or £10,000 there is no one size fits all solution, there is no plug in. We'd need a metered paywall so after 10 visits you ask them to sign in with the same user ID across devices, then to sign up to emails, notifications etc., but all of the solutions you need are all quite disjointed. So we are looking at trying to integrate one or two solutions. And then once we gather all this data we need to leverage it into a new advertising system which will be the highest investment."* – Small, local newspaper

However, another small, local newspaper we spoke to did mention using an 'off the self' payment solution, to help them manage their subscription payments, which meant managing their subscribers was a lot "easier".

Summary

The biggest challenge the organisation spoken to were facing within the press sector was difficulty in accessing user data. Social media and tech organisations already have huge advantages in access to user data. And the ending of traditional methods of accessing user data is pushing traditional press organisations further down other routes, many of which aren't seen as viable.

Some traditional press organisations refuse to go down a 'social media, click-based model' because in doing so, they may sacrifice their editorial control and dilute the quality of their content; content which they see as having a specific value to their readers, or even society more broadly. They don't want to get to a point where access to quality journalism is no longer widely available.

Whilst some are choosing to go down a sign-in / subscription route, others can't reconcile this, both because they see their remit as providing a free and important service that they don't want to make more exclusive, and because of the potential drop-off in readership that would follow such a move.

Finally, for small organisations in particular, the investment associated with collecting and leveraging user data, and the lack of 'off the self' solutions to do this, were further barriers to unlocking the value of user data.

b. TV

Overview of the TV landscape

This research explored data use within both commercial TV channels, as well as publicly funded channels i.e. the BBC. Given there are fewer small, local broadcast TV channels, this research focused on larger, national channels.

The role of data for TV businesses is similar to press and radio—to make programming (scheduling), commissioning and marketing decisions, improve the user experience for customers through personalisation and drive advertising revenue. However, their experiences and challenges with accessing and using customer data are slightly different.

The value exchange with customers is different: generally the content is seen as having a higher value to the viewer, so the willingness to share data—endure friction in the user experience—or pay directly for services is greater (over half of households pay for a video subscription service[4]). Sign-in is also the norm across video streaming platforms, meaning it is less overtly obstructive than it appears to be on news and media websites.

While not in the same direct competition for audiences with social media companies as press organisations, TV companies have had to adapt quickly to the rise of SVOD businesses—large tech organisations such as Amazon and Netflix. All of these companies are providing competing video services, with a significant focus on personalisation.

Challenges to data use

Access

TV organisations felt they had limited access to data compared to big tech firms and SVOD platforms

Organisations felt it was extremely challenging to compete with SVODs when it came to collecting user data. SVOD organisations were perceived to have a greater amount of data, collected from a vast amount of content and users across the world, and a huge investment in technology to support data collection. In contrast the broadcast TV organisations spoken to relied on a mixture of BARB data and sign in data from their platforms. And many were critical how of much BARB could really offer.

> *"BARB has been way too slow in incorporating within what BARB does to measure all viewing. We've been advocating for some time for BARB to measure all viewing."* – PSB channel

> *"BARB is an incredibly historic tool*…very clunky." – PSB channel

What is BARB?

The Broadcasters Audience Research Board (BARB) has complied audience measurement and television ratings in the UK since 1981. BARB has two sources of data:

- **People-based data:** the BARB panel, a sample of households recruited to represent television viewing across the nation. These households have a meter installed on their TV set, PCs and tablets, and a handset for each household member to register that they are watching the show. This enables BARB to monitor what household members are watching on TV, using audio samples to identify what is being watched.
- **Device-based data for online TV viewing:** data is collected whenever anyone in the UK watches a BVOD service on PCs, tablets and smartphones. Content is tagged by broadcasters so

[4] Research news: 'Half of UK Households Subscribe to Streaming'

that when it is watched the embedded code generates data on what is being watched, for how long, and on what device. This data cannot say how many people are watching a certain show or who they are.

As of September 2021, BARB are going to start releasing, publicly, data coming from a new method of measurement called router meter. This will mean:

> *"We're going to be able to spot all sorts of things for the first time. Including YouTube; including SVOD, at content level, and actually report that independently and not rely on various companies to decide whether they're going to share it or not. It will give a much better picture and a much more complete picture in homes of how people's viewing is—of both broadcast and non-broadcaster content."* – PSB channel

Organisations also spoke about the number of 'touchpoints' that certain big tech firms have with their users, meaning viewing data can be combined with the multiple other data points that these firms hold on users. For example, viewing data from Amazon Prime Video being combined with data from Amazon Alexa and Amazon shopping data.

TV organisations felt there was unequal access to data—with SVOD platforms and big tech organisations 'controlling' data about their users and viewing behaviour. Broadcast TV organisations were provided with no, or very little, data from SVOD platforms who stream their content. Many were very welcoming of the upcoming BARB changes which will provide router data on SVOD watching.

> *"So Netflix famously don't share data… so we have to guess how well our programmes are doing, so on the one hand you don't know how good a deal you've done, and on the other hand you don't know how much [audience / revenue] you could have generated yourself. But now because of BARB we will be able to tell that"* – PSB channel

On top of this, some TV organisations were finding it increasingly difficult to be listed in the tech companies' environment in the same way they used to:

> *"Increasingly, if you think about the way in which we get our app in front of the consumer, there's now loads of intermediaries that are now playing a role in this. So an example might be, if we want to get into iOS devices, the way in which we get our app into the platform has to go through Apple's approval process. Now, historically, that was reasonably simple, but increasingly the things we need to do and the rules we need to abide by, that allow our app to be approved, are getting increasingly complex. And so Apple's a great example, again, where the most recent iOS 14 update in terms of the privacy options, it's creating a very complex ecosystem for us".* – PSB channel

Additionally, organisations with PSB status were struggling to stand out for users accessing TV through smart TVs. While international SVODs like Netflix are able to pay for 'top spots' on smart TVs, or to have a button on the remote, PSB channels felt they lacked control over how visible they were compared to having a clear ranking on a traditional TV guide—leading to lower audience figures and reduced data collection in comparison to the SVODs. PSB organisations felt that the 'prominence' they used to be granted in listings was under threat.

> *"As consumers are moving away from that linear environment into a kind of on-demand digital environment, how do you replicate that prominence? And this is a debate we're having right now with government. And it's really important to understand that these matters can't just be left to generic competition law or competition or the market to resolve, because it is essential in order to protect the kind of importance of public service broadcasting, the information consumers need in order to participate in their country and in a democracy, to give national broadcasters who have particular responsibilities that kind of prominence…*
>
> *So, with prominence: users should be able to find our content and should be able to find our channels when they turn the television on. Whether that's fresh out of the box or in the home page or in recommendations, PSBs should be prominent and available. Now, the flipside to that is obviously a [tech] platform could say, 'well, if we're forced to put you at the top spot, we're just not going to have you at all'."* – PSB channel

Organisations also spoke about the cost of listing their channel on linear TV, and the perceived inequality of having to follow regulations which other SVOD and big tech firms did not.

> *"Yes, they're [TV listings] incredibly expensive—millions of pounds a year, biggest expense outside staff"* – Digital and TV news organisation

TV operating systems were perceived to have increasing control over viewer data

Some organisations felt that data about their shows and viewers was being 'ringfenced' by the companies who control the operating systems on TVs—the TV manufacturers and large tech firms. The companies who control the built-in operating systems on TVs, such as Amazon, Google or Apple, were perceived to have a huge amount of control both over *what people see* (e.g. what content is prioritised on the home screen) and *what data is available* to the other media providers whose content is watched on them.

> *"Platforms [tv operating systems] are running personalisation and recommendation solutions, they run over the top so they see everything everyone is watching on this device, and they will recommend and surface certain shows and content"* – PSB channel

One company worried that in the future TV operating systems would exert more control, potentially inserting their own ads into content:

> *"All the new smart TVs ask for opt-ins, so when you open your TV you opt in, which gives the manufacturer or operating system—which tends to be say, Google Android operating system—permission to collect data, and they can [potentially, in the future] say we own the 'last mile' of delivery to the consumer, and overlay ads as they are in control of the viewing experience."* – PSB channel

As the operating system is 'upstream' from the individual platforms on which people watch content (e.g. BBC, C4, Netflix) the data that these companies have access to was perceived to often be far greater, and more powerful, than the individual platforms providing the content. While some organisations had deals with organisations, such as Sky and Virgin, the data they received was felt to be limited—such as total streams rather than any demographic information or granular information to understand consumer behaviour.

One of the traditional PSBs accessed third party data from companies like Virgin and Sky who host content from their channel. However, they felt they got limited value from it as it is very rudimentary and provides little insight into consumer behaviour:

- "The reporting comes through with a delay" – meaning data is out of date
- There are frequently errors in the data provided – meaning it is hard to rely on
- There is no demographic information about audiences – limiting the value of the data

This provides the station with very little to go on in terms of what types of content are popular with viewers, meaning the data has limited value when making decisions about what to invest in, and how to better personalise content. They noted that these challenges were unlikely to be down to active unwillingness to share, and more likely due to their systems and operating model.

However, an industry body for tech organisations, including TV manufacturers, felt that TV manufacturers "were not asking for anything more than them [broadcast TV channels]" from consumers, meaning they didn't possess any additional data that they could be sharing with the PSBs.

Companies controlling the devices users view content on were also perceived to increasingly restrict access to data

Similarly, organisations spoke specifically about Apple's restrictions on data collection as interfering with their relationships with consumers. For example, one organisation spoke about receiving reduced insight from their

consumers with the introduction of Apple's App Tracking Transparency feature[5], while Apple were perceived to continue to benefit from user's data. As Apple would not list their VOD app on the App store if they did not comply with these restrictions on data collection, the organisation felt they had no alternative options.

> *"When you have Google and Apple having a strangle hold on the data, it makes it harder for start-ups like us to reach a new audience in the UK"* – Digital and TV news organisation

Another PSB channel talked about the problems they faced with Apple's iOS updates: when their content is viewed on their app on an Apple device, customers are prompted to opt out of being tracked by the app, meaning they have access to the data of very few viewers.

More importantly for this station, they said the hoops they have had to jump through to even get their app listed on the App Store massively restricts what they can then do with the data they do have access to. They used to be able to merge their customer tracking data with third party datasets in order to target advertising (e.g. this segment of customers just moved house and therefore are prime for mattress adverts)—while this was considered legal under GDPR, they said that Apple would only list their app if they didn't do this.

They felt that the station has a really strong B2C relationship with their own customers based on trust and transparency and the intrusion of Apple is unfair.

> *"Why should Apple decide what we can and can't do if our own customers are happy with it? Especially when Apple will be doing the exact same thing themselves."*

They also felt that organisations like Apple and Google are using 'privacy' as a selling point to their customers but that it is actually a smokescreen for them to be able to ringfence data and reduce competition.

Organisations were concerned that advertisers are beginning to expect new metrics, determined by the big tech companies

Viewing data is the main metric on which TV advertising is sold. Specifically, BARB average viewing figures are what have been used in the TV industry to negotiate ad sales.

However, digital competitors, such as Google, Facebook and SVOD platforms, use their own metrics, relating to streaming figures. A number of organisations spoke about their concerns that without a standardised metric across the industry, large tech firms, with access to greater data, would be able to establish their metrics as the standard with advertisers, which could make their own platforms look more valuable compared to other TV channels.

> *"We're very nervous. This is a space into which the big platforms, and Google in particular, could step in exactly the same way that Apple has kind of owned the privacy conversation. Google could own the audience measurement conversation. And we are very, very wary of that because Google would do it in a way that suits Google. And the way that suits Google is to say, here you go, here's an ad right in the middle of Love Island, in an advertising break, on a connected TV, and we're going to weight that exactly the same as a six second ad in a piece of user-generated content on YouTube. Now you can see those two things are not equal but it completely elevates the value of the YouTube asset."* – PSB channel

Value

Mandatory sign-in for BVOD plays an important role in collecting valuable user data. However, this also has its limitations

Most of the TV organisations spoke about how important mandatory sign-in was to enable first party data collection from users. Most had already introduced mandatory sign-in and while they acknowledged the risk

[5] Apple's App Tracking Transparency feature was introduced in April 2021. It requires apps to ask the user's permission before being allowed to track them (i.e., link their data on their app to data collected from other companies' apps or websites), or access their device's advertising identifier. Pop-ups appear on apps asking users whether they want to allow tracking or ask apps not to track them, thus making it easier for users to opt out of tracking. All apps submitted to the App store must enable this feature.

that they may lose some users, this was outweighed by increased revenue from higher advertising fees. Indeed, being able to target specific demographics of users resulted in significantly higher advertising rates.

Organisations spoke about ensuring a fair value exchange for sign-in data and ensuring the content they would be able to access felt worth the registration data that users had to provide.

> "*they're happy to sign in if you get something in return, it's that world of people expect something for their data, which is fair enough. It's really important that everyone in the market does similar things, then it is just the norm of how broadcast works"* – PSB channel

Some of the challenges in using sign-in data mentioned were:

- Not knowing whether the person signed in was the person watching the TV—other platforms such as Netflix were seen to be ahead here with the introduction of user profiles to distinguish who is watching on an account
- User data not always being valid. This could be for a number of reasons, for example people providing false data (such as providing a fake age to allow them to access content that is for adults) or data going 'out of date' (such as their location or occupation).

One PSB channel had built mechanisms to collect a range of data about their audiences. They collect first party data at sign-in and use this to track and segment their users alongside BARB data. They were also using first party sign-in data to match with other third-party data sets, and provide more information about who their users were to advertisers. However, the quality of this third-party data was often seen as low.

They also described a limit to how much value they can get from first-party data, which meant the increase in the friction for the user incurred when collecting more detailed data isn't always worth it for the channel.

> "*We don't want to get in the way of people's primary use of our product, there is still resistance in asking for user data 'why are you asking for this', and we will only take what we actually need. Data is too abundant to have value intrinsically, it's about what you do with it—finding the things to do with the data you can turn into value which is where we have been concentrating our efforts"* – PSB channel

Ethos

Upholding PSB duties meant some channels weren't always able to, and didn't *want to,* use data to maximise ad revenue

Some PSB channels raised the limits of using data to maximise ad revenue – i.e. focusing on producing the types of content that will increase viewing among audiences who are more valuable to advertisers. They spoke about valuing diversity in content and audiences, and quality of content, as outlined in the PSB requirements[6], over ad revenue.

> *"Our responsibility across the breadth of our content is to deliver a valuable service to every household young, old, rich, poor, North, South, you know, so it's I think that there are certainly different challenges for us where there is the obligation of universality rather than maximisation of ad revenue. It's really a very different model"* – PSB channel

> *"our values must be persevered, and if it meant more ad revenue at the expense of impartiality or responsibility of the news, I think we would always side on the impartiality"* – PSB channel

Summary

One of the biggest challenges to the television sector is the rise of huge international SVOD organisations. With national channels struggling to compete in a highly competitive market, where streaming services are perceived to have access to much more data and to have struck lucrative deals with operating systems to get their services in front of viewers first. PSB stations who used to benefit from TV guide top billing to drive

[6] Source of PSB requirements: PSB regulatory framework

audience numbers, were having to adapt and diversify, and recognised the importance of effective data strategies to so do.

Similar challenges in accessing data were raised in relation to big tech firms like Apple and Google, who were seen to be reducing organisations' access to data about their audiences.

A lack of resource and skills were not raised as particularly large barriers in this sector, in comparison to the other issues that organisations mentioned.

c. Radio

Overview of the Radio landscape

Within the research we explored national and local stations, including media groups who owned several stations.

Some radio stations were struggling to innovate around data—partly due to limited access, particularly for those with a smaller online presence, meaning their main source of data comes through the audience measurement system RAJAR.

The move to more online listening has allowed many stations to understand much more about their audiences, with some giving the opportunity for listeners to access tailored radio content. Much like in press—there was a reluctance from some organisations to lose the shared experience and editorial control of un-tailored content, however this only seemed to be the case for smaller, local stations.

Radio stations are also facing similar challenges to TV and press around big tech players acting as intermediaries and gatekeepers to data with the rise of aggregator platforms and smart speakers. Even some cars now come integrated with Google technology, reducing even further opportunities for radio stations to collect audience data.

Challenges to data use

Access

Radio, particularly local stations, are relying on limited data sources, such as RAJAR

What is RAJAR?

RAJAR Ltd (Radio Joint Audience Research) operates a single audience measurement system for the UK radio industry[7].

The organisation produces quarterly figures on listener numbers for each station, broken down by demographics. It produces these through the continuous placement of diaries across 50 weeks of the year (100,000 p.a.), in which respondents record their live radio listening for one week.

Some radio stations had extremely limited access to data about their consumers. When listened to offline, FM and DAB radio stations receive no direct data from their audiences—the stations themselves can't directly collect data on how many people are listening, or any demographic information, unless they run their own surveys with their listeners.

> *"You have the broadcast device and the broadcast receiver. So you essentially have a big mast just feeding out this content feed to all these receivers. And we don't get any data back."* – Large radio group

Instead, RAJAR is used widely as a consistent source of data around listener numbers and audience types. However, many stations recognised the limitations of relying on RAJAR figures. The data produced isn't actual audience figures, but an estimate based on a sample of the population. For some of the smaller stations in particular, this can be incredibly problematic—within the sample size of 100,000, individual listeners to their station might be in the single figures—one of these individuals deciding not to listen to the station during their one-week diary can represent a huge drop in their audience estimates or a skew in their demographic data.

[7] Source: RAJAR website

> *"If one person that listens to us goes on holiday when the knock on the door comes, that can have a significant downwards effect on our business. If we get one extra respondent on the survey who listens, that can send us up."* – Small, independent station

As this station only received quarterly figures from RAJAR (based on the previous quarter) they said the effect of things like a new high-profile presenter will take months to show in their figures. The data produced is also limited to audience size, listener location and NRS social grade. However, this was the only data the station had access to about their offline listeners, and they felt *"it's better than nothing"*.

Some of the larger radio stations, with large audiences, were much more positive about the kind of data they are able to glean from RAJAR:

> *"We also would normally get some quite rich demographic insights, behavioural insights through RAJAR. It's a recall-based survey methodology. So it's not first party data on the whole audience, but it's a very large sample size and it's a very robust methodology. So, yes, it would give us a very reputable profile of the audience."* – Large radio group

For a station with a much larger audience and therefore more ability to effectively segment, RAJAR does an important and useful job in allowing them to target content at their audiences. This segmentation data also then serves as a launchpad for the targeting of advertising.

Stations with an online presence had other ways to access and use listeners' data

All of the stations we spoke to have some form of online element to their offer. Online, stations had more ability to collect data and react to it with their scheduling and programming. Stations were able to better understand consumption behaviours of users, provide tailored content and collect some first party data through sign-in on apps or websites.

For a large radio group we spoke to, data gathered from online listeners, particularly data from people signed into their website and app, was incredibly important. In terms of advertising, the data types in demand were demographics, income and user propensity around their particular product or service.

> *"We're able to understand, through online use, what people click on and what they listen to, how they listen—app, smart speaker, website—their location and when, including their historical behaviour. This is in addition to first party registered and third-party cookies."* – Large radio group

They use this data to run email campaigns, send push notifications and sell targeted audio advertising. They were also utilising data gathered through social channels to drive people to back to their websites.

For the most part, radio stations—particularly larger ones—sell their IP ad impressions separately to their conventional broadcast impressions. And one alternative method some larger (and smaller independent) stations use other than RAJAR to collect and utilise consumer data, is from their IP streams. Some of the big groups we spoke to had set up ventures where they combined the first party data they gather across their different affiliated publications in order to sell audio ad impressions at scale:

> *"It takes all of that inventory, overlays it with first party data from a number of sources: so log-in data, but also online browsing data, user ID data and device ID data from our own website and our products, third party data sources as well…it allows direct and programmatic advertisers to acquire audio ad impressions at scale across both of our inventories with that data overlaid so that the advertising campaigns aren't bought by brand, so you don't buy a [name of radio programme] package—you buy 'men', or you buy 'auto-intenders' [segment intending to buy a car] based on browsing history across our combined publishing sites that feature reviews of cars etc.*
>
> *So there's a model there which has some scale, some value and some strength, but it's only able to overlay data against a partial view of the audience. It's not able to overlay that data against the non-IP listening. But even of the IP listening, which is part of that sales proposition, it's only the IP listening that has that first party data that we can draw that data proposition against. And that generally excludes all the smart speaker listening and excludes all the listening on third party apps".* – Large radio group

Smart speakers and third-party listening platforms were creating a barrier to data access

Online listening, although giving access to more data, still came with barriers to effective data collection and use for stations. If a consumer listened to content via a smart speaker, such as an Amazon Echo or Google Home device—stations had little ability to access anything other than very simplistic data around audience figures.

> *"If they're on a smart speaker, we know next to nothing."* – Large radio group

And smart speakers are becoming an increasingly popular way to access radio[8]. One radio group believes a huge proportion of their audience listen through smart speakers, which means their access to most of their listeners' data is sparse.

> *"It's a tiny minority of the audience that we have any known data on. And it's a massive fundamental problem. If anything, it's our most significant strategic challenge…as we look over the next decade. Quite simply because in digital advertising and in digital media, there's a basic level of expectation around the quality of data, the quality of attribution and the quality of targeting that a media owner would have access to. But in radio, because the bulk of listening is 'off platform' and so much listening is not even on IP, we don't have an equivalent position even with what the traditional publishers have."* – Large radio group

Another large radio group had recently launched the ability for users to link their radio account to their Amazon Echo—enabling increased personalisation of content and limited incentives such as additional stations and premium content. However, this still came at a cost to data collection.

> *"We certainly don't get direct access to the data that Amazon or Google collects on our users on those platforms. None of that is shared. So, it's only really what we can garner from the connection to our streaming service primarily."*

Currently very few of their users are using this linking feature, however, huge numbers of their listeners do use smart speakers to access their stations, which they get even less data from.

> *"Between 50 and 60 percent of all of our IP listening is consumed on a smart speaker."*

There were concerns from this station that this percentage will inevitably increase.

> *"These platforms are increasingly an important route to access for radio and also for listeners to have access to radio. And they're in a gatekeeping position. Most of them. Not all of them, but the big ones are. And they're getting most of the data so we're having to make assumptions about what our listeners are doing. They're getting the precise data."*

According to one of these large radio groups, eventually internet connected devices will become the dominant way to listen to radio—with smart speakers already taking up a chunk of that market[9].

In this landscape, radio groups are trying to start a dialogue with some of the big tech companies in order to work out how best to go about getting users' consent to share their data with the radio station as well as the tech company themselves. But this is proving difficult:

> *"I think the bigger challenge […] is that there's a very significant imbalance, even between an organisation of our size and the tech platforms, which have very important intermediation roles between the listener and us as a broadcaster in actually accessing the data that we would like to be able to access on the listeners of our radio channels. Entering into a sensible dialogue with Amazon, about how we might be able to secure consent from listeners to one of our radio channels on Amazon Echo devices so we can start personalising our products and target advertising, is next to impossible.*
>
> *Theoretically there are ways to synchronize user accounts and data flows. But Amazon has no incentive to support us in doing that. They have much more incentive to hold the user relationship directly themselves. It's very easy for them to cite concerns around European data protection regulation and consumer privacy as justifications for not finding technical solutions that would allow them to give us an opportunity to capture consent for legitimate data practices and monetisation practices.*

---

[8] Publication: RAJAR MIDAS (Measurement of Internet Delivered Audio Survey) Spring 2020.

[9] Publication: RAJAR MIDAS (Measurement of Internet Delivered Audio Survey) Spring 2020.

Also, all these tech platforms are operating in an international sphere as well, so they don't tend to develop models on a localised basis. As broadcasters we tend to operate on a localised basis. So, there's a misalignment of incentives, there's a geographic imbalance and there's an imbalance in market power that simply denies us the opportunity to progress normal commercial discussions around access to data. We are therefore moving it into a regulatory discussion…about whether we can be provided with a framework to rebalance the terms of those discussions that we want to have with the platforms." – Large radio group

Ethos

Stations weren't always interested in being able to target content

Radio stations are often built around the idea of a collective listening experience; having developed their brands around certain genres and personalities.

> *"There's a real brilliant thing about radio, which is everyone else listening with you. And the presenter name-checks people who've called in or tweeted and obviously makes it that shared experience, something that TV's lost a lot of."* – Large radio group

Independent stations had far fewer opportunities to personalise content, particularly if they only had a small online presence, and their business models tended to be based on providing a range of localised content for a broad audience.

> *"It's for people who don't want to just listen to oldies but who are still interested in current music. And then we mix that with local news, travel information, weather—everything that you could want to know about or in the area is on the radio. We have one studio and a very small team."* – Small, local radio station

Even when stations were open to targeting content, this would require getting users to sign in, and there were worries around creating barriers to audience access, when—especially smaller stations—were mainly concerned with just increasing audience figures.

The small, independent station we spoke to did have an app, but this does not require sign in. The main barriers to introducing this were the perceived time and effort to make the change. But also, audience figures were the main focus for the station—targeting content wasn't really on their radar and anything that might reduce audience figures in the short term was seen as too risky.

> *"It's an interesting concept [getting more first party data through things like sign-in], but our real objective is to get the audience figures up. So if our audience was 60,000 rather than 30,000, we'd have double the audience to go in and to be able to say to advertisers, we've got 60,000 people now—rather than 1 in 10 people in the area we could say we have 1 in 5: 'advertise with us'."* – Small, local radio station

However, the bigger organisations were tailoring content online, streamlining across devices, serving targeted advertising and letting users have more control over their listening experience. Allowing sign in provides stations with more detailed information about their audiences and their listening behaviours.

> *"[We use sign in data for] creating a more personalised experience or, for example, allowing the user to be able to have a continuity of experience across devices. As we know, more people have lots of devices, connected devices these days. We want to kind of create a seamless experience across those."* – Large radio group

Value

Some stations felt they had limited need for more granular data

Stations did not always feel they needed more granular data about their listeners. Innovation was limited when advertisers were only interested in very simplistic data—audience numbers and listening hours were the main tools some of the stations we spoke to were using to sell advertising, and they felt there was a lack of interest in more detailed demographic or behavioural information.

> *"If we've got some restaurants placing ads with us, who goes to restaurants? Who goes to the pizza place in town? Well, the answer is: it's everybody. So as broadcasters, we don't need specific granular data. We just*

> *want broad-brush stuff that shows that we've got we've got 30,000 or 40,000 listeners or whatever it is."* – Small, local radio station

There is a perception that local advertisers in particular aren't interested in the more granular data, with the focus being on the largest amount of reach for the lowest price—as we've seen from the small, local station we spoke to, the focus from their advertisers was on the RAJAR listening hours data. In contrast, it was felt that national advertisers were much more willing to invest in more detailed data sets, using demographic data to better target advertising—as the joint venture between two large radio groups alluded to earlier.

The investment stations put into collecting and analysing data was understandably proportional to the demand or 'ask' from advertisers for that data. However, if local advertisers did begin to request more granular data, smaller stations may struggle to keep up with larger organisations who are able to invest more in the tools and expertise required.

> *"We've invested a lot of money in our infrastructure to really enable us to be able to capture more data on people [online] […] increasingly we are now capturing the behavioural data."* – Large radio group

Resource

Lack of resource and skills affected some smaller organisations

Cost and skills were an issue for some of the smaller stations. As mentioned above, developing the infrastructure to capture and utilise data can be expensive, and didn't always seem valuable enough commercially in the short term. However, one larger organisation felt that local radio stations should be investing in the resources and skills needed to utilise data, in order to be successful in the longer term, and that developing skills in house would be the most cost-effective way to do this:

> *"I think it's about having the knowledge within the business. So I think you need to make sure that knowledge already exists. And I think that's where some businesses will fall down. They don't invest in data experts. So I think that's absolutely key.*
>
> *I do think that businesses need to realise the importance of data and make sure they've got the right people to be able to deliver it. […] if you don't have loads of money in a business, you can't afford the consultants, we use consultancies. We are lucky in that respect. So, you need to be able to really recruit those experts. […] So it is making sure that you've got the resources, the expertise, the infrastructure. But you could do that on a much smaller scale. You don't have to use someone like Salesforce or Dalibor. There are loads of other companies out there that you can use on a much smaller scale. If you are a smaller organization."* – Large radio group

Summary

In terms of accessing data, broadcast radio was limited to RAJAR figures. However, many stations also had an online presence which required sign in, enabling the collection of first party data. The rise of smart speaker listening was concerning many of the stations we spoke to, as this significantly reduced the amount of data they could collect from listeners, compared to online or in-app.

There wasn't always felt to be proportional value for stations in increasing the amount and types of data they collected. One local station felt advertisers were interested primarily in the viewing figures provided by RAJAR data, and there wasn't enough demand for more detailed information.

Targeting and tailoring content was much less of a priority for smaller stations, as there was a focus on the shared experience of radio consumption. However, larger organisations and radio groups were investing in creating a far more personalised experience for online users.

5. Conclusions

The biggest challenges raised by media organisations throughout this research were linked to the explosion of competition in the media landscape—social media and online only content providers can often offer content to audiences in cheaper and easier to access formats, and big tech organisations are increasingly acting as gatekeepers to valuable data.

These issues are well documented, particularly in the Cairncross review, which highlights the challenges arising from an increasingly digital and social media dominated landscape, which is having a huge impact on the way people consume and engage with content.

Within this new landscape—where media organisations are having to compete with social media companies and big tech organisations for both audiences and advertising revenue—the Cairncross review states that "publishers will thus need to collect far more extensive information on their users" to "offer advertisers better targeting opportunities".[10]

However, many of the media organisations we engaged with felt this was easier said than done.

It was not always practical for organisations to collect more data as this can introduce increased friction for customers, potentially leading to reduced audiences. This is particularly true for press organisations with mass market appeal that are free to access, who worry customers would just go elsewhere if requirements to share more data were added.

Loss of audience was perceived to be less of an issue for big tech firms due to the relatively low effort of agreeing to share data compared to the value a user then gets. For example, if a user agrees to share data with a smart speaker, they will likely only have to do this once, which will then enable them to access a whole range of services on an ongoing basis. In comparison, an individual looking to read a single article on a news site would get far less in return for sharing their data, and may be able to find content on the same topic across multiple other news sites that have no data sharing requirements.

While the value outweighs the effort and cost associated with sharing data in the first example, the opposite is likely true in the latter.

Media organisations are currently operating on an unequal playing field

Media organisations were often struggling to find the right balance in this value exchange, and at the same time were facing additional challenges, which made it more difficult for them to collect and utilise data in comparison to big tech organisations:

- **Unequal access to data** – tech firms often own the interfaces between the user and the content, meaning their access to user data is often greater (e.g. social media platforms, online browsers, smart devices). This is exacerbated by a perceived lack of negotiating power for small and often even large media organisations, who feel they are not in a position to demand data from tech firms, or even have a conversation about it.
- **Unequal control over how data is used** – tech firms often set the agenda for how data can be used by media organisations (e.g. Apple privacy tracking, Google third party cookies).
- **Unequal regulation to abide by** – tech firms are often less heavily regulated, meaning they can focus on providing the most value to users, without any restrictions on what their content should include. In contrast UK based media organisations are often overseen by regulatory and industry standards bodies, for example, local commercial radio stations can be required by Ofcom to provide a certain level of local content (news, travel etc.). This means that media organisation can't always be completely data led in their decision making.
- **Unequal importance placed on providing 'public value'** – The values of many media organisations don't always align with the implications of pursuing data driven model. Many of the

[10] Source: The Cairncross Review: a sustainable future for journalism, 2019.

organisations we spoke to were keen to continue to provide audiences with free-to-access, high quality and varied content, which they felt might be compromised by some of the available routes for increased or more lucrative data collection and use. Big tech firms in comparison, particularly social media organisations, have always operated via a data driven model.

A note on how these findings might relate to other industries

This research only engaged individuals from within the media sector, so cannot provide definitive insight into the barriers around data use experienced by other industries. Press, TV and radio are also somewhat unique in that the content they provide is often free, resulting in a significant reliance on data for advertising for many organisations.

However, many, if not all industries need information about their customers, and the value exchange to gather that data—as outlined in the executive summary—will apply to any organisation trading data for content online. There will always be consequences for businesses to weigh up when collecting user data, which will affect their ability to utilise data to its maximum potential.

Big tech organisations are also likely to be having an impact across sectors—with the major players dominating the consumer data market. An obvious comparison is the retail sector—where an international organisation like Amazon, with the ability to use consumer data gathered across its retail, technology and web service products, mean they easily outperform many retail organisations in the UK.

The ability to access and use data is clearly important across industries, and many businesses will likely require more support to survive and thrive in the current data landscape.

6. Annex: Organisation profiles

Deep dives into six organisations

Press

Small, independent, local news organisation

"If I was more hardnosed, I'd ask for more [data], but the attraction to my site is that I don't ask for any data, you don't even have to give your real name".

Head office location: Southwest

Organisation size: 1-5 employees

Format: Online only

Content: Local news

Audience size: Approx. 800 subscribers

Business model: Subscription

This was a small, online, local newspaper. They focussed on providing high quality local journalism and utilised a subscription model, which meant they didn't have to serve adverts to their readers**.** Something they were determined not to do.

They collected very little data about their subscribers, just asking for an email address and username—they wanted to keep it that way for 'ethical reasons'. Whilst they wanted their audience to grow, they wanted to keep it very local, and to keep their loyal customer base.

Current data collection methods and uses

- *Subscription data* which only included email addresses, so tells them very little about their audience
- *Google analytics* to understand what type of stories their readers liked and what the best times to post were
- *Social media analytics* to work out which platforms were generating the most traffic to their website
- *Surveys* with their subscribers to understand more about their audience demographics, attitudes towards adverts and opinions on the price of subscriptions.

Barriers to data collection and use

Access: Access was less of an issue for this organisation, as they did not feel they needed any additional data.

Ethos: They held strong principles about their content and monetisation model. Editorial control was seen to be very important—they went as far as to continue writing articles they knew very few people would read, as they felt a duty to report on all local news.

Value: They felt including advertising on their website would put off the sort of customer who accessed their content, so had instead gone down the subscription route to fund the organisation, and even when subscribing, users were asked to part with minimal data.

This organisation was happy with their current size and model, and therefore felt they had limited use for additional data.

Given that most of their subscribers renewed their subscriptions and seemed to like the content, they felt that to improve the business' finances the focus should be on getting new subscribers—which they didn't see data playing as much of a role in.

Resource: As a micro-business, there was limited time and capacity to think about the other potential uses for data. However, they were very happy with their current size and audience so may have chosen not to invest in additional data skills and tools even if the option became available.

Large, national newspaper with a strong international online presence

"You have to give your advertisers segments of the population that they want to sell advertising to. We don't know the personal identities of our readers. We don't know their names or anything. But we at least have been able, through cookies, to tell an advertiser, well, here's a group of people who are of a certain age, certain social backgrounds and have an interest in Off-Road vehicles".

Head office location: London

Organisation size: 2500-3000 employees

Format: Print and online

Content: National news

Audience size: Average daily readership of 2.180 million

Business model: Free to access content, reliance on advertising revenue

This is a traditional and long-standing press organisation that has fostered a strong online presence. Their online content is free to access with sign-in only required to leave comments. They are keen to keep it that way, both to ensure access for all and because they were sure their customers would go elsewhere if required to sign in or subscribe. Despite producing some 'click-bait' content they were also committed to maintaining their ability to produce high-quality journalism.

They relied heavily on programmatic advertising, with a large number of adverts per article. Without sign-in and subscription they were using third party cookie data to sell advertising, and support content decision making. They felt this data provided them with rich profiles of their customers and were very concerned about Google's plans to put an end to third party cookies.

Current data collection methods and uses

- *Third party cookie data* was relied upon by the organisation to sell advertising and understand how users are interacting with the site, helping them to make editorial decisions
- *First party IDs* from users logging in to leave comments was used by the organisation to track engagement and further profile customers

Barriers to data collection and use

Access: They were struggling to access data that was being collected and held onto by big tech organisations. Google putting a stop to third party cookie tracking will make it much harder for this organisation to continue with their current free to access business model and still sell programmatic advertising in the same way. They currently rely on this data to segment their users, and once this ends, they will know very little about what their users are doing off site.

Apple were also causing this organisation issues. Apple now automatically turn off data collection on apps meaning they are getting very little information about their in-app users, similarly the option for single sign-in through an Apple ID means far fewer users are creating a separate account with this organisation, meaning they lose the benefits of first party sign-in data on Apple products—Facebook and Google offer similar single-sign-in options.

They feel that privacy-led regulation has allowed Google, Amazon, Apple etc. to restrict access to data in the name of privacy, while enabling them to collect more for themselves.

Ethos: The difficulty of accessing data through third party cookies is that free to access, mass market news is becoming less lucrative. However, this organisation were adamant they did not want to change their business model. They felt the two alternatives available to them: a subscription / sign-in model or relying more on click based content goes against their principles as an organisation. Sign-in and subscription would reduce the accessibility of their news, and a reliance on click-bait would mean lowering their standards of quality and editorial control.

Value: Not turning to a subscription or sign-in model wasn't just about ethos for this organisation—they felt that because the content they offer is so mainstream and people can access similar content elsewhere, including from the BBC, their readers would jump ship if they increased the hoops they were required to jump through to access content.

This organisation believed that subscription models only work well for smaller titles with a niche audience.

Resource: As a large newspaper they didn't lack the skills or money to invest in data collection and use. They felt their options were becoming limited for collecting data and they were seeing advertisers increasingly turn to Google, Apple and Amazon to buy data.

TV

Established PSB channel

"So what we do is we collect personal information, so age, gender etc. And we also use the viewing data, so the first party data we collect about how they use our platform. And we use that data to build data products to segment our users into specific groups. So, this might be based on models of demographic groups. It could be based on their interests. It could be on any kind of audience defined by an advertiser. We then build machine learning models based on a sample of survey data we collect to predict which segment the user is in, and that forms the basis of targeted advertising products"

Head office location: London

Organisation size: 800-1000 employees

Format: Linear TV channel and VOD

Content: General entertainment

Audience size: Roughly 51.1 million quarterly viewers

Business model: Advertising revenue, recent introduction of ad-free subscription

This channel felt they were ahead of the curve with regards to their data strategy and already weren't relying on third party data alone to sell advertising space, so weren't too worried about the upcoming cookie changes.

They had built their own data suite to securely sell segmented data, and profile their audiences to provide a more personalised experience for customers within their streaming service.

However, they still faced barriers to making the most of data, one of which was advertisers' low expectations for the quality of the data they are able to collect.

Current data collection methods and uses

- *BARB* data was something they relied *on* a lot for viewing data. It is valued by advertisers because it is recognisable and trusted. But couldn't provide data on digital viewers. Now it is branching out into digital devices which they are really excited about.
- *Advanced data suite* – they have their own digital ad products that help advertisers target ads on VOD, this means they don't have to disclose personal data to clients.
- *Log in data* tells them that a profile/device is being used and means they can tailor content. However, they don't necessarily know who is watching.
- *Third party data* is something they use a lot less now. They have found the quality of third party data to be very poor, although advertisers still like it because it's so abundant.

Barriers to data collection and use

Access: Operating systems are becoming gatekeepers for data (battle between manufactures and tech giants). You often now get a built-in operating system on a TV—Google, Apple, Amazon etc. This means they have a huge amount of control both over what people see and what data is available to the media organisation. There is also the potential for telecoms companies to control data. This organisation was also worried about the levels of control telecoms companies might exert with the use of IP data.

Ethos: The amount of data audiences are willing to disclose proved a barrier. Even with sign-in there's a limit to how much data audiences are willing to share—they don't want it to become a barrier to entry—age and gender is about as far as you can go. They have a permission-based philosophy—only take what they really need and if they do need more they will specifically ask for it.

Value: This organisation have trialled third party data measurement systems and found them to be flawed. They feel that advertisers don't tend to care about the quality of data—they're more interested in abundance and measurability. This means moving away from third party data can be a hard sell and only really benefits the big players with large audiences that can provide lots of first party data. They also talked about there being a 'shelf life' for data. You could try and get loads of information about someone, but often that information is only correct for a certain amount of time—job, location etc.

Resource: Being a large organisation, resource wasn't much of an issue. Their concerns were more surrounding the kind of the data they were able to collect, and the utility of that data both for their own editorial purpose, and in terms of value to prospective advertisers.

Digital news organisation

"We're creating a very tangible, measurable community. And so we've got progressive profiling of our users. They have quite a low barrier to sign-up initially just because you want to get mass reach. So we're just getting first name and last name, email address, post code and date of birth just to help segment. And what we want to do will be to profile a lot more views and political leanings…It will make sure that our editorial is matching the views of our audience. So it's sort of like a kind of community measurement data capture, but we're trying to provide personalized, tailored content based on information that they've given us already"

Head office location: London

Organisation size: 100-200 employees

Format: Linear TV channel and VOD

Content: News and entertainment

Audience size: Unknown

Business model: Advertising revenue with a plan to introduce subscription

This is a TV news channel and digital platform based in London. They felt they had recognised a gap in news market and wanted to make their content appeal to the country as a whole, rather than having a London-centric perspective and appeal.

They found that not being one of the big, established channels presented a big challenge: "TV is naturally geared towards to incumbents".

They see themselves as offering a digital product to a national audience in the future, one which they can personalise based on things like political leaning.

Current data collection methods and uses

- *BARB data* provides information on demographics and viewing figures, both on Linear and VOD. Soon they will also receive router data on SVOD watching.
- *Digital services, subscription* was something this station will be investing more of in the future. Ultimately, they will offer full digital package with sign-in, app etc. to provide a personalised experience
- *First party website sign-in data* is collected through their website, they plan to use this more in the future to tailor content.
- *Newsletter sign ups* is something they want to offer later down the line, possibly segmented by region, even by things like political leaning.
- *Qualitative and quantitative user research* with different audiences, polls, focus groups etc. is helping them to design their programming

Barriers to data collection and use

Access: They feel that Google and Apple have "a stranglehold on data which makes it hard for start-ups to reach new UK audiences". Individuals' access to content is often within Google, Apple and Amazon environments, and these platforms don't share all the information they have, so "there's much more information about our audience available to Apple than there is to us. […] They keep back quite a lot of insight for their own gain, and that's very frustrating because it's our data, our audience and our content."

Ethos: This organisation wants to create a "very tangible, measurable community" where the digital audience are able to participate in the show. For this, they need a lot of first party data on their users/viewers. They're depending at the moment on sign-in for this data, which they acknowledge can lead to drops in potential audience numbers.

Value: This organisation felt it was incredibly important for them to collect audience data in order to personalise content. However**,** the amount of content they have to produce is vast, both in order to rival their competitors and to be able to start properly personalising. They're starting by segmenting by location, political leanings etc. While they are focused on investing in data collection—they acknowledged this will take a while to pay off.

Resource: This organisation acknowledged how hard it is for a new TV channel to join the market. When asked about licensing, they said: "They're expensive—millions of pounds a year, the biggest expense outside staff'. There's also contestability of funding—new players in the environment are competing with some very established brands for funding. As a new company, the level of liability for GDPR is huge and can cripple you if

you're not fully compliant. This presents a challenge for a young, growing company—they don't have special compliance department or anything, so they're relying on people upskilling themselves."

Radio

Independent, local radio station

"The real objective is to get the audience figures up—we want to be able to say to advertisers that we have 1 in 5 people in the area rather than 1 in 10"

Head office location: Southwest

Organisation size: 5-10 employees

Format: FM, DAB, online and app

Content: News and entertainment

Audience size: Approx. 32,000 weekly listeners

Business model: Advertising revenue

This was a successful independent local radio station which included a traditional radio station and online listening through a website and app. Their access to and use of data was fairly limited, with an emphasis on RAJAR audience figures to sell advertising. They felt their local advertisers wouldn't be interested in more granular data about their audiences, so were not planning to invest in other data gathering methods such as sign-in online. When asked about improving their ability to collect data, they felt RAJAR should collect more robust data on factors like age of users, rather than something they should be investing in themselves.

Current data collection methods and uses

- *RAJAR figures* provide them with the majority of data about their audiences and its size which they use to measure success and sell advertising space.
- *Consumer insight work* such as street surveys has provided them with more information about the appeal of their station.
- *Digital stream data* lets them know how many devices are streaming the station—from website, app, and smart speakers.
- *Competition data* includes information on about 1000 listeners who have entered competitions—e.g. where they live and their gender.

Barriers to data collection and use

Access: As a station where a significant amount of their listeners were still listening through DAB and FM radio, they relied on RAJAR to provide the majority of data about their audiences. However, as the station was relatively small, they had struggled with the fluctuations in their estimated audience numbers, which they felt was a consequence of the RAJAR methodology—one listener being on holiday during the RAJAR reporting period can have a big impact on the figures.

Value: The station felt there would be limited value in collecting more granular data on their audiences, as this was not something their local advertisers wanted**.** The majority (65% post-pandemic but 85% pre-pandemic) of their ad sales are to local advertisers who were described as being 'unsophisticated' and mainly caring about total listener figures. There was a perception that having more rich data on listeners (e.g. demographics, interests) would not be valued by local advertisers.

The station also felt it would not be worth investing in gathering data via their app. The station have an app which does not require sign-in. The main barriers to introducing this were 'time' and the perception that the most lucrative objective for their organisation would be to increase audience figures, rather than add any friction to access for their users.

Resource: The lack of perceived value to increasing data collection for this organisation was linked to resource concerns—they felt the time and investment required to develop improved data collection and analysis infrastructure would not provide them with equivalent benefits.

National radio and digital broadcasting group

"I think the bigger challenge for the whole industry is that there's a very significant imbalance, even between an organisation of our size and the tech platforms, which have very important intermediation roles between the listener and us as a broadcaster, in actually accessing the data that we would like to be able to access on the listeners of our radio channels."

Head office location: Northern Ireland

Organisation size: 1000+ employees

Format: FM, DAB, online and app

Content: Music, sport, entertainment

Audience size: Approx. 5-6 million annual listeners

Business model: Advertising revenue

This is a large radio group owned by a global media group. They run several radio shows with large, national audiences. They have exciting plans for the near future with ideas around increased personalisation of their broadcast streams, and even podcasts. However, they are also facing serious strategic challenges in the realm of user data. The most significant of which is the rise of smart radio developed by companies like Amazon and Google, which pose serious concerns for this organisation when it comes getting fair access to their listeners' first party data.

Current data collection methods and uses

- *RAJAR figures* had been accessed by the station on a quarterly basis. But this was suspended as of March as they RAJAR reviewing alternative methodologies due to COVID. '*We know almost nothing at the moment, would usually get some quite rich data from Rajar'*
- *Web streaming data* provided real time data from unique devices on number of streams, and where they are from in the UK; sessions, time spent etc. (e.g. a log file provides connection made to stream, IP location, unique device, listening session length)
- *Smart speakers* provided them with very little data as they felt it was impossible to get consent on smart devices, tech companies will not yield on discussing how to share this data
- *Web & app* provided first party data from sign in to provide a personalised experience
- *Third Party Cookies* were being used by the station to understand more about their online listeners
- *Customer insight* in the form of independent surveys and qualitative research provided more information about the demographics and preferences of their listeners
- *Direct attribution* through signed in listening, mostly through smartphones allows them to target advertising much more effectively

Barriers to data collection and use

Access: The main barrier for this organisation was around access. Increasingly, their listeners are listening through smart speakers. They as an organisation are not able to get consent for data collection through smart/connected devices, so there's very limited ability to learn about this growing audience. They are unsure if people are being directed accurately to their content by the smart speakers—there's no transparency. Tech companies (e.g. Amazon, Google) are not incentivised to share or help.

Ethos: Unlike some of the smaller radio stations with smaller audiences, this organisation was actively pursuing a strategy of personalisation, knowing that it could provide more value to their users. Their main issue was being increasingly unable to collect first party data from listeners who are more and more listening to content through smart speakers.

Value: Again, unlike some of the smaller stations who found it hard to see the value in collecting more data, this station is constantly trying to gain more first party data from their listeners, knowing that it provides value through things like increased personalisation. Again, their principal barrier is actual access to this data, which is increasingly becoming ringfenced.

Resource: Resource wasn't an issue for this organisation as they were a large group owned by a global media company. They acknowledged, though, that resource can be an issue for smaller, independent stations who may not be able to afford CRM tools or to invest in their own research methodologies.

7. Annex: Glossary of terms

SVOD	Subscription Video On Demand
BARB	Broadcasters' Audience Research Board
CPM	Cost Per Mile
SSO	Single Sign-On
FLOC	Federated Learning Of Cohorts
IP	Internet Protocol
CRM	Customer Relationship Management
PSB	Public Service Broadcaster
RAJAR	Radio Joint Audience Research
VOD	Video On Demand
iOS	Apple's Operating System
GDPR	General Data Protection Regulation
BVOD	Broadcast Video On Demand
DAB	Digital Audio Broadcasting
FM	Frequency Modulation
NRS	National Readership Survey
IPSO	Independent Press Standards Organisation