

Appendix S: the relationship between large digital platforms and publishers

Introduction

1. Publishers of online content rely on Google and Facebook's user-facing services to host content or for referrals of traffic to their online properties, which they can then monetise by displaying advertising to these visitors. However, online publishers consider that they face an imbalance of bargaining power with Google and Facebook, which disadvantages their businesses in a number of ways.
2. This appendix draws on evidence we have received from a number of large publishers of online content aimed at a UK audience as well as from the large platforms themselves.¹ These publisher submissions come from a range of publishers including traditional news publishers – who have transitioned from print-based distribution of content towards either a mix of print and online distribution or solely online distribution – and 'digital-first' publishers.
3. Below we first set out the various relationships that publishers have with Google and Facebook's user-facing services, and our analysis of the potential imbalance in bargaining power between publishers and these platforms. We then identify some aspects of these relationships which might benefit from being covered by the proposed code of conduct.

Relationships between publishers and large digital platforms

4. Publishers typically identified Google and Facebook as being by far the most important digital platforms for their businesses. Apple, in its role as a supplier of a large mobile operating system and its Apple News service, was also mentioned as being important by a number of publishers, but its importance was generally rated as being significantly below that of Google and Facebook.

Publisher business models

5. Publishers describe employing three broad types of business models to monetise their digital properties. The type of digital business model affects how the publisher generates revenues and how it interacts with Google and Facebook. Broadly the three types of business model are:

¹ Guardian News and Media, Independent News and Media, Reach UK, Daily Mail, The Telegraph, Reach plc, Sky, Vice, JPI Media and eBay.

- Subscription based – where the prime focus of the publishers is to turn an engaged audience into paying subscribers;
 - Traffic/digital advertising based – where the main goal is to drive traffic towards the publisher’s webpages and monetise this in the form of targeted digital advertising displayed to visitors; and
 - Monetisation of content on posted third-party platforms – publishers post content on third-party platforms, usually social media, and use the monetisation tools made available by those platforms – such as the sharing of advertising revenues – to generate revenue.
6. In practice, most of the publishers we spoke to blend some aspects of all three of these business models and depend to a significant degree on Google and Facebook for the success of their business models.

Interaction with Google’s consumer-facing services

7. Table S.1 summarises the various ways in which online publishers interact with Google’s consumer-facing services.

Table S.1: Overview of publisher interactions with key Google user-facing services

<i>User-facing service</i>	<i>Flows of content</i>	<i>Flows of traffic</i>	<i>Flows of data</i>	<i>Flows of money</i>
Google Search	Publisher sites are listed in Google search ranking, including title and a 'snippet' of content	Source of significant proportion of traffic to publisher sites	Publishers can access aggregated data on users searching for their content	Publishers receive no revenues in relation to Google Search
YouTube	Publishers post content on YouTube	A source of referrals to publisher website	Publishers can access aggregated data on users' interactions with their content	Publishers above a certain size receive a share of Google's owned-and-operated ad revenues. For news publishers this is generally a very small percentage of overall digital ad revenue. For some digital-first publishers, the revenues can be significant
Google News	News publisher content appears in an aggregated and curated product on Google News site and app	A source of referrals to publisher website	Publishers can access aggregated data on users' interactions with their content on the Google news site/app	Publishers are eligible for a share of Google owned-and-operated ad revenue but currently Google does not display ads on Google News
AMP	Publication format for faster loading mobile web pages	Source of significant proportion of mobile web traffic. Publishers consider they must use AMP due to prominence given to AMP pages in Google mobile search rankings	Publishers can access data on individual users' interacting with their content As content is hosted in Google's ecosystem, Google can also observe user interactions with publishers' content	Publishers are able to sell open display advertising in a similar way to how they do this for other websites and apps. Some publishers have noted that there are some restrictions around, for example, header bidding and that per page revenues for AMP pages are generally lower than for similar pages. The importance of AMP pages as a source of revenue is likely to be roughly proportional to volumes of publisher AMP traffic compared to other types of traffic. For some publishers this can be very significant.

Source: CMA.

8. The most important interaction is with Google Search, which is a very significant referrer of traffic to online publisher websites both via organic and, to a lesser extent, paid search results.² Prominence in the organic results of relevant Google searches is considered extremely important by online publishers. As a result, significant resources are devoted to optimising the positioning of web pages in Google search results (an activity known as search engine optimisation – SEO). In addition, almost all publishers we spoke to told us that they engage in paid search activity to increase the prominence of their web pages in Google search results.
9. YouTube is also considered very important for publishers' businesses: all the publishers we received submissions from reported that they post content on YouTube. Their aims in doing so are threefold: firstly, to drive traffic back to their websites; secondly, to generate awareness to their content and brand;

² For search items considered to have 'news intent', Google Search may also present the consumer with a 'Top Stories' carousel sourced from news results that Google crawls and places in a separate index to its standard search index. Results delivered via the separate 'news' and 'video' tabs can also be important sources of traffic for online publishers.

and thirdly, to generate revenue via YouTube advertising and, to a lesser extent, subscriptions.

10. Publishers of content on YouTube can be eligible for a share of advertising revenue on YouTube if the content meets certain standards and a threshold number of views is passed. Where this is the case, YouTube sells advertising, displayed at the beginning of (pre-roll) and during (in-stream) publisher content. Publishers told us that they receive around 55% of any consequent advertising revenue. In some cases, publishers can arrange for their directly sold advertising to appear alongside their content. In addition, content publishers have the option to earn revenue from subscriptions.³
11. One other aspect of Google's consumer-facing services that is considered very important by publishers is Accelerated Mobile Pages (AMP). AMP is a publishing format for mobile devices that enables the fast loading of content in browsers. In order to enable fast page loading, AMP employs an optimised and restricted version of the code used to build web pages, and web pages are cached within the AMP ecosystem.⁴ As pages are cached, usually by the Google AMP Cache,⁵ consumers usually remain within the Google ecosystem whilst browsing an AMP page.⁶
12. Publishers told us that they need to put web pages in AMP format, because they think that gives them greater prominence in mobile search results. In particular, it was noted that only AMP pages will appear in the 'Top Stories' carousel that are shown in the results of searches with 'news intent'. Publishers considered that, for the most part, AMP pages operated in a similar way to their regular mobile web pages. However, as is discussed in more detail below, they have concerns around restrictions on their ability to monetise these pages and to access data generated from consumers' interaction with them.
13. Another Google consumer-facing service that was mentioned by several publishers was the Google News website and app. However, its importance was generally considered to be much lower than that of Google Search and YouTube. Google News collects information from other online publishers and presents it to consumers as a collated, curated product.⁷ In principle,

³ See [Google.com 'How to earn money on YouTube'](#).

⁴ This means that pages are effectively pre-loaded to the AMP system to enable faster upload to the consumer device. See ['how AMP pages are cached'](#).

⁵ See: [AMP on Google: Google AMP Cache](#).

⁶ There are two currently operators of AMP caches, Google and Bing. Content publishers do not choose which cache to use, as this is selected by the platform's themselves, such as Google.

⁷ Consumers are presented with news stories collated under a series of categories such as 'Top Stories', 'For You' and 'Your Local News'. Stories are presented as a headline, usually with a hyperlink back to the source website, although some content may be viewed within Google News itself.

publishers are eligible for a share of any advertising revenue for ads that appear alongside their content, but Google does not currently display advertising on Google News.

Interaction with Facebook’s consumer-facing services

14. Table S.2 summarises the various ways in which online publishers interact with Facebook’s consumer-facing services.

Table S.2: Overview of publisher interactions with key Facebook user facing services

<i>User-facing service</i>	<i>Flows of content</i>	<i>Flows of traffic</i>	<i>Flows of data</i>	<i>Flows of money</i>
Standard Facebook pages	Publishers post content on Facebook pages	Source of referrals to publisher sites and content discovery	Publishers can access aggregated data on users’ interactions with their content	Publishers receive no revenues in relation to standard Facebook pages
Facebook Watch	Publishers post content on Facebook Watch	A potential source of referrals to publisher website	Publishers can access aggregated data on users’ interactions with their content	Publishers above a certain size receive a share of Facebook’s owned-and-operated ad revenues. Generally, not regarded as a significant source of revenue for publishers
Instant Articles (IA)	Publication format for faster loading mobile web pages used in Facebook News Feed	Can be a source of referrals to publisher website Not all publishers use IA	Publishers can access aggregated data on users’ interactions with their content hosted in IA	Publishers are eligible for a share of Facebook owned-and-operated ad revenue for ads placed around IA. Publishers also have the option to place advertising they have sold directly to advertisers around IA, for which they receive 100% of revenue. Generally, not regarded as a significant source of revenue for publishers
News Feed	Personalised collection of Facebook items that a user might be interested in. A publisher’s content (page, Watch or IA) will surface in a user’s News Feed if the algorithm determines it would be of interest to the user	Method of discovering publisher Facebook content	Publishers can access aggregated data on users’ interactions with their content	Publishers receive no revenues in relation to standard Facebook News Fees per se

Source: CMA.

15. Facebook is another key source of consumer traffic for publishers. Publishers post content on their own Facebook pages with the aim of generating awareness of their content and brand and of referring traffic back to their websites and apps. Publishers report that they have little or no opportunity to directly monetise what might be termed standard content on their Facebook pages.

16. Facebook’s News Feed is another key source of traffic.⁸ To post content in Facebook’s News Feed, a number of publishers put their web pages into Facebook’s Instant Articles (IA) format. Similar to AMP, IA is a publication format that has been designed to allow mobile pages to load faster,⁹ but in the case of IA it is only in use on the Facebook mobile app. Publishers receive a [majority] share in advertising revenue generated by Facebook from adverts that appear alongside their IAs. They also have the option to insert their own directly sold advertising¹⁰ alongside their IA and, where they do this, they receive 100% of the advertising revenue.¹¹
17. Several publishers also post content on Facebook’s video hosting service Facebook Watch. Where they post content of Facebook Watch, publishers may be eligible for a share of advertising revenue where the content meets certain standards and a threshold of views is surpassed, and they receive approximately 55% of any revenue for advertising displayed alongside their content.¹²

Google and Facebook as a source of traffic for publishers

18. We have analysed website traffic data from a number of large publishers.¹³ This data shows that in 2018 and 2019 these publishers relied on Google and Facebook properties for around between 36% and 38% of total traffic to their websites, as shown in Table S.3 below.

Table S.3: Sources of Website traffic for online publishers

Year	Website traffic from Google, Facebook and Direct visits ¹⁴								
	All traffic			Desktop/Laptop			Mobile		
	Google*	Facebook	Direct	Google	Facebook	Direct	Google*	Facebook	Direct
2019	25%	13%	43%	26%	4%	55%	25%	17%	38%
2018	26%	10%	44%	29%	4%	52%	25%	14%	40%

* Some publishers included AMP in the ‘referred from Google properties’ data estimate but not all. If all AMP referrals were included, then this percentage would increase.

Source: CMA analysis of publisher data.

⁸ The News Feed is the constantly updating list of stories in the middle of a consumer’s home page, including status updates, photos, videos, links, app activity and likes from people, pages and groups the consumer follows on Facebook.

⁹ Facebook states that [Instant Articles](#) load 4 times faster than standard mobile web articles and that consumers read 44% more pages on Instant Articles average.

¹⁰ This is open display advertising sold directly to advertisers but not programmatically through real-time auctions.

¹¹ See [Facebook.com ‘Monetise your Instant Articles with Audience Network’](#).

¹² See [Facebook.com ‘About rules for monetisation’](#).

¹³ This analysis includes traffic data for the following websites: The Independent, The Sun, The Times, The Daily Mail, The Telegraph, Reach PLC websites, Sky websites, and all Vice websites.

¹⁴ Other visits come from what are termed ‘other third-party referrals’, for example referrals from Snapchat or Instagram. The publishers were not always consistent with how they reported the data. In some cases there are overlaps between categories so when third-party referrals are included in the analysis the source of data adds up to a value slightly greater than 100%.

19. Based on publisher submissions, in 2018 and 2019 (up until June) the average proportion of traffic to their websites that was referred via Google properties was 26% and 25% respectively (for 2019, the lowest proportion across publishers was 8% and highest was 57%). Referrals from Facebook properties were responsible on average for 10% of website visits in 2018 and 13% in 2019 (for 2019, the proportion ranged between 2% and 47% among publishers). Direct website visits were the most important source of traffic, with 44% of visits being direct in 2018 and 43% being direct in 2019 (ranging, in 2019, between 6% and 57%). Other visits come from what are termed ‘other third-party referrals’, for example referrals from Snapchat or Instagram.
20. In its response to our interim report, Google stated that ‘this data shows that news publishers are not dependent on Google for Traffic. Less than a third of their traffic comes from Google Search.’¹⁵ However, it is clear that, for the most part, the publishers from whom we received data are dependent on Google and Facebook for a significant part of their traffic.

The balance of bargaining power between online publishers and Google and Facebook

21. Publishers have told us that they view Google and Facebook as ‘must have’ partners. This is primarily due to a substantial proportion of the traffic referred to their websites coming from Google and Facebook properties and a degree of reliance on prominence on Google and Facebook properties for content discovery and brand awareness.
22. As a consequence of this reliance on Google and Facebook for traffic, publishers told us that they suffer from an imbalance of bargaining power when dealing with these platforms. This was an issue that was also raised as part of the Cairncross Review, which concluded that ‘Google and Facebook also increasingly control the distribution of publishers’ content online’ and that as a consequence ‘these platforms can impose terms on publishers without needing to consult or negotiate with them’.¹⁶
23. An example of the imbalance of bargaining power, cited by several publishers, was the approach taken by Google to updating its terms and conditions shortly before the introduction of GDPR. At the beginning of May 2018, just weeks before the GDPR came into effect, Google released its updated online T&Cs to cover changes to its advertising services. The terms describe Google as a co-controller of data for certain of their advertising products and require publishers to gain consumers’ consent on Google’s behalf to gather and

¹⁵ [Google’s response to our consultation on the Interim Report](#), paragraph 51.

¹⁶ [The Cairncross Review: A Sustainable Future for Journalism](#), February 2019, page 57.

utilise their data. Publishers consider that these changes were made in a non-negotiable way and that they had no choice but to accept this update to the T&Cs.

24. This potential imbalance of bargaining power has led to publishers expressing a number of concerns about how their relationships with Google and Facebook result in them being disadvantaged. The most significant of these are explained below.

Unexpected and unexplained changes to search and ranking algorithms

25. Most publishers expressed some concern about unexpected and unexplained changes to Google and Facebook algorithms, most notably in relation to Google Search and Facebook News Feed. Specific examples of several algorithm changes which significantly impacted website traffic were mentioned, including:
- a) Facebook's announcement on 18 January 2018 that from that day its News Feed ranking algorithm would prioritize 'meaningful interactions' from friends and family over content from brands.¹⁷ This change had the impact of deprioritising content from some online content publishers in the News Feed ranking of many Facebook users.
 - b) Google's announcement on 3 June 2019 that it was updating its core search algorithm the following day, which resulted in a step change in the daily traffic arriving at many news websites via Google Search. Some sites saw an increase in daily traffic arriving via Google Search, but others saw significant decreases.^{18,19}
26. Publishers have argued that a reduction in website traffic resulting from an algorithm change has a direct financial consequence for their business in the form of lost opportunities to monetise these visits through advertising. Furthermore, they told us that sudden, unexplained and significant algorithm changes make planning and financial decision-making more complicated and can lead to significant, potentially wasteful, expenditure on understanding these opaque algorithms and optimising content to appear high up in the rankings. Some publishers have also told us that they think that in some cases algorithm changes may be commercially motivated to favour the platforms or affiliated parties at the expense of other publishers.

¹⁷ See [Hootsuite blog, April 2018, 'How the Facebook algorithm works and how to make it work for you'](#).

¹⁸ See [Google SearchLiaison update 2 June 2019](#).

¹⁹ See [Searchengineland.com update 6 June 2019](#).

27. Publishers broadly are of the view that they do not get sufficient warning of algorithm changes or sufficient explanation of their impact or of what they might do to mitigate any loss of traffic. A number of publishers have suggested that there should be a role for a regulator to monitor and report on the main Google and Facebook search ranking algorithms.
28. Whilst some publishers feel very strongly about algorithmic transparency and considered it to be of critical importance to their businesses, for others it was less of an issue and was viewed more as a consequence of doing business with Google and Facebook which they have, to a degree, grown used to. However, in response to our interim report all publishers thought that provisions about the reasonable explanation of search algorithms and sufficient notification of changes should be covered by the code of conduct. For example, the Reach Plc stated that:

‘we believe it is vital that the Code requires Google and Facebook to provide us and other publishers with:

(a) within a short period of time after the Code is introduced, an explanation as to how their algorithms work; and

(b) going forward, sufficient notice prior to making any significant changes to their algorithms.’²⁰

The News Media Association (NMA) further suggested that ‘[c]omplaints about ranking practices should be referred to the digital markets unit, which must have the power to investigate and impose remedies. This complaints process should be available to publishers which have concerns about the impact of existing ranking algorithms on their traffic’.²¹

29. Whilst publishers have a number of issues with the approach taken by Google and Facebook to updating their algorithms, we note that there are legitimate reasons for regularly updating search algorithms as these are central to improving the consumer experience of Google and Facebook properties.
30. In its response to our interim report, Google set out a number of factors that place limits on the amount of information it can share on the operation of, and changes to, its search algorithms. Google argued that disclosing too much information might allow publishers to game the system, and that the information could be used by competitors to copy innovations and free-ride on its investments. Google also stated that the provision of information was

²⁰ [Reach PLC’s response to our consultation on the Interim Report](#), paragraph 2.24.

²¹ [NMA’s response to our consultation on the Interim Report](#), paragraph 3.7.1.

complicated by the fact that ranking may be governed by several different algorithms and added that it already provides ‘vast amounts’ of data on the criteria used to search rankings.²²

31. It is clear that many publishers rely on Google and Facebook for a significant proportion of their traffic and that changes to key search algorithms by either of these can have a significant impact on publisher businesses. We would, therefore, consider it reasonable that publishers have sufficient explanation of how these algorithms work and sufficient notification of changes to them where they might notably impact upon their businesses. We consider that provision to publishers of sufficient explanation about how the key search algorithms work as well as explanation and notification of changes to these are areas that would appropriately be covered by the proposed code of conduct (see Appendix U).

Potential de facto requirement to use the AMP and IA formats

32. As we discuss above, most publishers feel compelled to use the AMP publication format because only AMP pages appear in Google’s News Carousel. As one publisher puts it, ‘the Google News carousel on mobile is void of any non-AMP pages and is a hugely important part of mobile search distribution, without which audiences could in some cases be halved or worse.’
33. Publishers suggested to us that their ability to monetise content hosted on AMP is significantly reduced when compared to their standard mobile web pages. One of the reasons suggested for this was that AMP does not currently support client-side header bidding and that, whilst a form of server-side header bidding is supported, the number of partners they are able to integrate into this is limited to five or less.²³
34. In addition, if content is consumed within the AMP or IA publication of formats, then the consumer remains in the Google or Facebook ecosystem and, therefore, these platforms have access to data on the consumer interaction. Google and Facebook are able to use this data to develop their services, to deliver targeted digital advertising and improve their ability to undertake ad analytics, including ad attribution. This, publishers suggest, ‘hampers our ability to build our own unique datasets for targeting purposes and thus present a competitive threat to Google.’

²² [Google’s response to our consultation on the Interim Report](#), paragraphs 52 and 53.

²³ See [Github.com](#), ‘AMP Real time config’.

35. In addition to AMP, almost all publishers said that they also produce ‘mobile friendly’ versions of their webpages, which present the content in a mobile compatible manner and aim to load more quickly on mobile devices than standard web pages. Publishers told us that these pages give them more control over their data and that they typically receive higher levels of revenue from these pages than they do from AMP pages. However, whilst these ‘mobile friendly’ versions of publisher websites may have some advantages over AMP, AMP sites have the significant advantage of greater prominence in Google Search results.
36. Google submitted that ‘AMP is not a Google ecosystem, but an open-source technology that is the result of collaboration between developers, publishers, websites, distribution platforms and other companies. It is wrong to say that AMP does not currently support client-side header bidding. AMP currently supports 16 header bidding partners, including the top four in the UK’.²⁴ Our understanding is that this is referring to supporting server-side rather than client-side header bidding and that restrictions on the number of partners that can be used are in place. In addition, publishers have submitted to us that they receive lower levels of monetisation from AMP pages than from other types of mobile-friendly pages.
37. Google also asserted that ‘[t]he fact that [Google] collect[s] information about user interactions with certain online properties does not prevent the owners of those properties from collecting the same information. Nor does it reduce the value of that information’.²⁵ However, Google did not further explain why its access to the publisher user data does not undermine its uniqueness or its value, nor did it provide any evidence to support this. Nevertheless, it is not clear why the collection of this data by Google is necessary to facilitate the use by publishers of the AMP format. Publishers’ control over their data is also materially different when they use other mobile-friendly formats compared with when they use AMP – we discuss this further below.
38. We consider that there is an analogy between how Google is able to influence publishers to use AMP – potentially to its own advantage – through preferential rankings in its search results and its ability to leverage its search results to influence users to use its own specialised search services. A similar remedy, in this case preventing Google mobile search from favouring AMP pages over other mobile-friendly pages in its news carousel, may therefore be appropriate in this case. Greater prominence of other mobile-friendly pages in Google mobile searches may lead to an increase in visits to publisher non-

²⁴ [Google’s response to our consultation on the Interim Report](#), paragraph 56.

²⁵ [Google’s response to our consultation on the Interim Report](#), paragraph 57.

AMP pages, where they have greater control over their data and a greater ability to monetise the page. This is an area that would appropriately be considered within the proposed code of conduct (see Appendix U for more on our proposals for the code of conduct).

39. We are aware that publishers have expressed similar concerns about control of data and monetisation when using Facebooks IA. However, publishers did not feel that same need to use IA as they did AMP and a number stated that they did not use IA at all.

Use of publisher content for 'free' and keeping users within Google and Facebook ecosystems

40. Publishers provide content into the Google and Facebook ecosystems in three main ways: firstly, by posting content on their social media platforms, such as Facebook's main site and YouTube; secondly, through hyperlinks and short explanatory 'snippets'²⁶ of content that appear, for example, within Google's organic search ranking pages; and thirdly, when utilising the publication formats AMP and IA (although Google submit that AMP is not a Google ecosystem most AMP content is hosted on the Google servers).
41. Some publishers argue that Google and Facebook effectively 'free ride' on content produced by third-party publishers and that this includes professionally produced content – whether that be breaking news, analysis, features, entertainment or sport, produced by publishers under their editorial responsibility and legal liability. They argue that, without this, Google and Facebook would attract less traffic to their platforms and would consequently generate less advertising and have fewer opportunities to collect valuable user data.
42. Some publishers have also argued that there is an increasing tendency for content to be consumed within the Google and Facebook ecosystems without clicking through to the source websites. A European Commission report in 2016 reported that 47% of UK individuals surveyed said that when they access the news via news aggregators, online social media or search

²⁶ A snippet refers to the small amount of text, an image, or a short video that forms part of a link. When producing a list of hyperlinks in response to a search item (including news content), search engines often scrape websites that are the subject of the hyperlink for content and provide a snippet of content relevant to the website. The purpose of the snippet is to provide context to the hyperlink and an indication of the contents of the relevant website to the user, so that the user can evaluate the relevance of the website to their search query. While a snippet may be the first line or two of a news article, a snippet can also be extracted from the body of a news article if the search engine finds that information to be more relevant to answering the user's query.

engines, they most often browse and read the main news of the day without clicking on links to access the whole articles.²⁷

43. A couple of publishers referenced a recent study by the News Media Alliance (NMA). This estimated that Google receives \$4.7 billion in revenue from News Publishers' Content worldwide in 2018.²⁸ The methodology of this study is however limited²⁹ and, as Google pointed out in response, 'the overwhelming number of news queries do not show ads'³⁰ and no advertising is currently displayed on Google News.
44. Publishers have also mentioned that, while they consider that Google and Facebook benefit significantly from using their content, they cannot always easily monetise content that is hosted on Google and Facebook properties. As we note above, publishers do not benefit from advertising that is placed next to 'standard' Facebook content. Publishers can benefit from advertising revenue that is generated from their own content posted on YouTube, Facebook Watch and through IA. However, several publishers have suggested that the mechanism by which they receive the remuneration is opaque and that any revenues they receive are limited.
45. Google and Facebook have previously argued that they do not receive third-party content for free from online publishers, but that in fact the publishers receive a significant volume of web traffic in return for their content. In response to the NMA study, Google stated that 'Google News and Google Search drives over 10 billion clicks to publishers' websites, which drives subscriptions and significant ad revenue'.³¹ In a submission to the EU, as part of its development of the EU Copyright Directive, Google submitted research that it said showed that news publishers in the EU benefited significantly in financial terms from traffic referred to their websites by third parties (including Google Search).³²
46. One potential development in this area is the EU Copyright Directive,³³ which was approved in April 2019. The Directive provides media businesses with rights governing the online use of their content by information society service

²⁷ 'Internet users' preferences for accessing content online', European Commission, Flash Eurobarometer 437 (March 2016), page 33.

²⁸ See News Media Alliance, June 2019, 'New Study Finds Google Receives an Estimated \$4.7 Billion in Revenue from News Publishers' Content'.

²⁹ It takes an estimate stated by a Google executive for news related revenue in 2008 (\$100m) and simply extrapolates this to 2018 by assuming that Google revenue from news represents that same proportion of total revenue from Google properties in 2018 as it did in 2008.

³⁰ As referenced in [New York Times article, 9 June 2019](#).

³¹ As referenced in [Guardian article, 10 June 2019, 'Google made \\$4.7bn from news sites in 2018, study claims'](#).

³² [Deloitte \(2016\): The impact of web traffic on revenues of traditional newspaper publishers A study for France, Germany, Spain, and the UK \(commissioned by Google\)](#).

³³ See [EU Copyright Directive](#).

providers (which would include digital platforms). However, it explicitly states that this right shall not apply to the ‘acts of hyperlinking’ and ‘in respect of the use of individual words or very short extracts of a press publication’.

47. Publishers were sceptical that the Directive would have any material effect on their ability to negotiate with Google and Facebook over the use of their content if it were to be adopted into UK law. Although one noted that, in principle, it could enable publishers to negotiate licensing agreements for the distribution of journalism through search and social platforms with market power, the prevailing views were that either its impact was highly uncertain or that there would be very little impact. Particular points of concern were that implementation of the Directive would lead to less content appear on these platforms or that publisher would have little choice but to enter into licencing agreements with the platform for no remuneration in return. This has been the experience of publishers in jurisdictions which have tried to introduce measures to require platforms to compensate publisher for the use of their content.³⁴
48. In their responses to our interim report, some publishers proposed remedies around the use of their content that were in addition to those listed under the potential rules for the ‘fair trading’ principle of the code of conduct. The NMA articulated these as follows:
 - ‘Where the platforms extract value from news publishers’ content, whether directly or indirectly, they should negotiate fairly with publishers to determine how that value should be shared in order to ensure an appropriate level of compensation to publishers’.³⁵
 - As the producers of original content, news publishers should have the right to control the extent to which their content is scraped as well as the length of related snippets displayed by platforms’.³⁶
49. In response to publishers’ concerns about the use of their content, Google submitted that ‘our services generate a huge volume of free user traffic for publishers, which they can then use to grow their brands and earn ad and subscription revenues’.^{37,38}

³⁴ See [The Register, French monopoly watchdog orders Google to talk payment terms with French publishers](#).

³⁵ [NMA’s response to our consultation on the Interim Report](#), paragraph 3.5.3

³⁶ [NMA’s response to our consultation on the Interim Report](#), paragraph 3.5.4

³⁷ [Google’s response to our consultation on the Interim Report](#), paragraph 55.

³⁸ In a related development, Google recently announced that in Australia, Brazil and Germany (but not the UK) it intends to compensate some ‘high quality’ news publishers whose content is available only behind a paywall for the use of this content. See: [Google News Initiative: A new licensing program to support the news industry](#).

50. The current evidence from other jurisdictions,³⁹ where stronger requirements for platforms to compensate publishers for use of content have been put in place, suggests that remedies which seek to give publishers rights to compensation for use of their content by platforms make little difference or, worse, result in less traffic to publisher sites.⁴⁰ In addition, it is not clear from the evidence available to us at this time whether publishers do or do not receive adequate compensation for the use of their content through increased referral traffic. We note that there have been some recent moves in some other jurisdictions, notably France⁴¹ and Australia,⁴² to mandate Google and Facebook to negotiate with publishers fairly and in good faith. The impact of these measures should be closely monitored, and this issue may need to be revisited in the light of emerging evidence.
51. There may be a case for granting publishers more control over how their content is used by the platforms, particularly when they have little discretion over the use of content, such as when it is obtained by scraping their sites. The most significant example of this is when Google generates snippets to appear next to hyperlinks in search results. If a publisher opts out of allowing snippets, it may be ranked lower in organic search results, reducing the publisher's visibility to consumers and, accordingly, click-through rates, website traffic and monetisation opportunities. However, where a snippet is produced by Google the relevant snippet may reveal the substance of the media business' content. This could then have a direct impact on referral traffic by reducing click-through rates of organic search results. Currently publishers have little control over snippet length or content; if they had, they might be able to tailor them in a way that maximised referrals to their websites. This is an area which could be covered under the proposed code of conduct (see Appendix U for more on our proposals for the code of conduct).

Publishers cannot access user-level data for content hosted within the Google and Facebook ecosystems

52. Publishers do not have access to the same level of data on consumer interactions with their own content when this is hosted on Google and

³⁹ In 2013 and 2015, Germany and Spain enacted copyright laws that allowed charges to be levied against aggregators for republishing article snippets.

⁴⁰ In response to the introduction of new copyright laws in Germany Google stopped providing snippets from German publishers in search results. Some publishers saw a large drop in traffic as a result (see: [Reuters Germany's top publisher bows to Google in news licensing row](#)). In Spain Shortly before the introduction of new copyright laws Google withdrew Google News in the country (it still is not available in Spain) this resulted in losses of traffic to news publishers of around 10% with much of this impacting smaller publishers (see: [The Impact of News Aggregators on Internet News Consumption: The Case of Localization By Susan Athey, Mark Mobius, Stanford Business Working Paper No. 3353](#)).

⁴¹ See [Related rights: the Autorité imposes urgent interim measures on Google](#).

⁴² See [ACCC mandatory code of conduct to govern the commercial relationship between digital platforms and media companies](#).

Facebook properties. Publishers report that they generally receive data that is very aggregated and anonymised, which they cannot match to their own first-party data to create consumer profiles across domains. Publishers expressed specific concerns about this with regard to Facebook, Instant Articles displayed within Facebook and YouTube. Publishers told us that the reason why data is only provided in an aggregated and anonymised form is generally stated as being due to privacy. There appears to be less concern about AMP, where publishers report that they are able to track the interactions of individual consumers, although some have commented that they have difficulty matching this data to their own first-party data.

53. Publishers do not have the same level of control over their first party data as Google and Facebook do. One impact of this ability of Google and Facebook to observe consumer interaction on many publisher sites is to reinforce the advantages they have over most other online publishers in offering targeted advertising, due to their greater access to online data. In particular, this reinforces Google's and, to a lesser extent, Facebook's ability to track consumers across different web domains.⁴³ At the same time, publishers are limited in that they cannot build a complete user-level picture of all of the users who interact with their content.
54. In its response to our interim report, Google stated that 'although we understand that many publishers would like to match their proprietary data with our proprietary data in order to create more detailed user profiles, we are limited by privacy concerns from sharing anything that is too granular.'⁴⁴ On the other hand, publishers have put to us that 'SMS Platforms must share all data they collect from publisher content online including through platform services such as AMP and Instant Articles'.⁴⁵
55. The principle that publishers should, within the limits of data protection, be able to access data on user interactions with their own content and as a consequence be able to build a complete picture of their user base seems reasonable. We would, therefore, consider it appropriate that the code of conduct would, again within the limits of data protection and privacy laws, facilitate platforms sharing data with publishers about user interactions with their own content at a user level and in a format that would allow them to match it to user data generated from the publishers' own sites.

⁴³ A paper by [Engelhardt and Narayanan \(2016\)](#), 'Online Tracking: A 1-million-site Measurement and Analysis', suggests that Google is able to track 75%+ of web domains and Facebook around 25% of web domains.

⁴⁴ [Google's response to our consultation on the Interim Report](#), paragraph 57.

⁴⁵ [NMA's response to our consultation on the Interim Report](#), 3.5.5.

'Commoditisation' of publisher data

56. Google and Facebook are able to collect and use individual data from consumers who interact with content from third party publishers on their platforms. In addition, the use of Google and Facebook analytics services by a publisher leads to the placing of a cookie when the service is accessed or a pixel on the publisher's website, from which data on user interactions can be accessed by Google and Facebook.⁴⁶ Google is also able to observe user data on a publisher's webpage when it is viewed via its Chrome browser or when the page is hosted in its AMP cache. The sharing of data with Google and Facebook is not only a result of publishers' interactions with the platform user services but also of their use of adtech products such as Google Ad Manager or Facebook Audience Network.
57. There is a possibility that access to data by Google and Facebook on consumer interactions on many publisher sites may undermine the value of that data to the publishers themselves. Access to this data by Google and Facebook may lead to it being used by their adtech intermediaries for targeting of ads on sites other than the original publisher website. This 'data leakage', may mean that data on a publisher's unique audience may be 'commoditised' and used to target ads on cheaper sites and apps, which might undermine the value of advertising inventory on a publisher's own website. We discuss the issue of commoditisation more in Appendix M.

⁴⁶ Although not in all cases is this data used by the platforms. For example, in the case of Google Analytics publishers have the option to prevent Google using their data for its own purposes.