

DMG Media response to the CMA consultation on how algorithms can reduce competition and harm consumers

1. This response is made on behalf of DMG Media, publishers of the MailOnline, metro.co.uk and inews news websites, the Daily Mail, Mail on Sunday, Metro and i newspapers, and New Scientist magazine.
2. Algorithms are central to the operations of our digital publications, insofar as they govern both distribution, through search and social media, and advertising, through a range of categorisation and blocking measures.
3. The algorithms with which we have to engage are all operated by digital platforms, each of which functions as a near-monopoly in the markets it dominates. The platforms run their algorithms as a 'black box', releasing very little information about how they work, the likely effects of any changes made to them, or even giving warning when changes are to be made. Very rarely is it possible to obtain explanation or remedy when algorithm changes damage publishers' businesses.
4. We described in previous evidence to the CMA's Digital Platforms and Online Advertising Market Study how a core algorithm change made by Google in June 2019 cut MailOnline search visibility by 50pc, while that of rivals improved. No warning or useful explanation was given, though after we complained to the CEO of Google's owner, Alphabet, our search visibility was restored in September that year. However MailOnline is the fourth-largest English-language news website in the world¹, with a strong political presence in Washington as well as Westminster. That remedy would not have been open to a local or regional publisher – indeed they would have difficulty in finding anyone at Google even prepared to speak to them.
5. This submission focuses on more recent developments concerning two different sets of algorithms – those which run Google Discover, which has supplemented Google Search as a source of traffic and therefore revenue, and those which govern content categorisation and inventory blocking in Google DV360 and Google Ad Manager.

Google Discover

6. The public are familiar with Google Discover as the recommended content, generally news items, which appears below the search bar when they access the Google Search homepage. Google Discover automatically shows personalised

¹ <https://www.pressgazette.co.uk/top-50-largest-news-websites-in-the-world-right-wing-outlets-see-biggest-growth/>

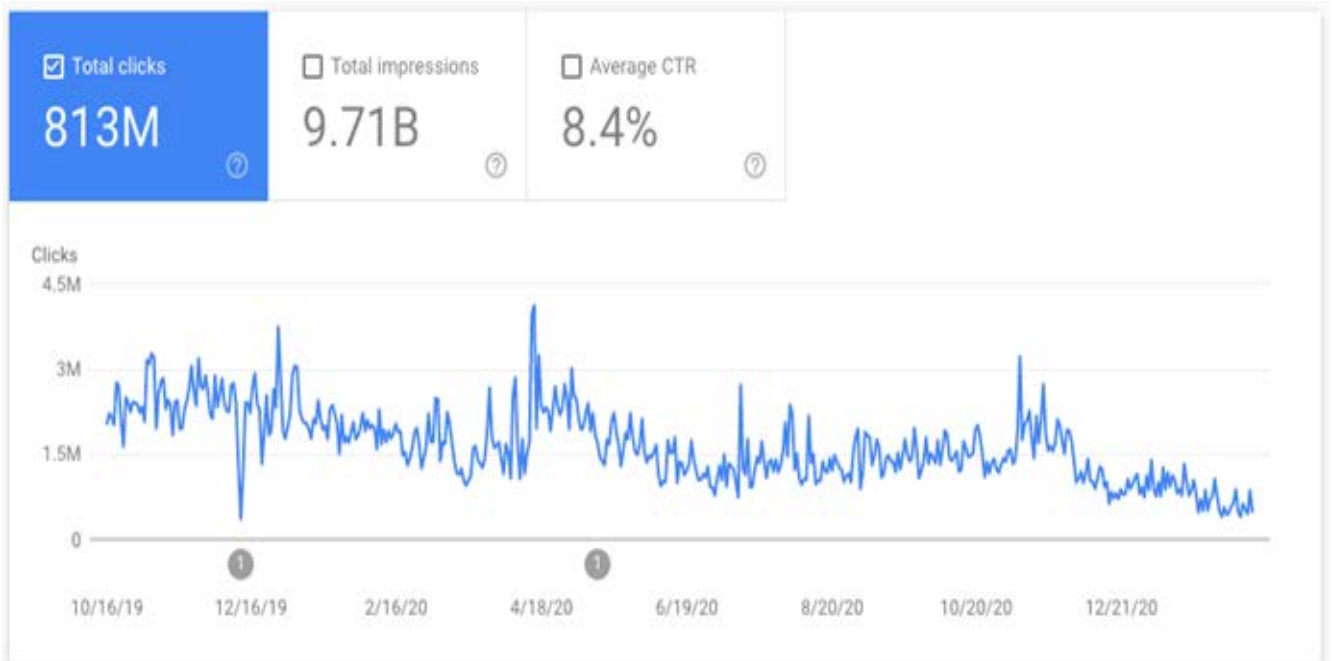
content based on a user's search history, browser history, app activity and location (dependent on Google account set up and privacy settings). It uses users' information to identify and display topics Google believes will be relevant to the user, known as "interests".²

7. Discover was launched in September 2018 and by the middle of 2019 it was starting to have a meaningful impact on MailOnline's traffic – not least because the personalisation algorithm that powered it enabled us to recover more quickly than we had feared from the down-ranking we received from Google's general search algorithm in June 2019.
8. More recently, however, Discover traffic has become a significant cause for concern. MailOnline's Discover traffic has reduced from 14m+ page views per week, to a current figure of 3.5m. As always, when these concerns are taken to Google they say that, as the problem is related to their ranking system, no support can be offered
9. The most serious decline began with the release of Google's December Core Algorithm Update on December 3, 2020, which reduced MailOnline's global Discover traffic by 48% week-on-week. As December progressed Discover traffic continued to deteriorate, with a further 40% from December 14 to 16 – from 1.02m global visits per day to an unprecedented low of 641,000.
10. The most significant losses were to US traffic. In November 2020, DailyMail.com, our US website, averaged 424,000 visits per day from US Discover users. By December 19 this had reduced to 45,000, a 90% reduction in traffic. DailyMail.com had been in discussions with the Google's US Web Stories team to develop the quality and diversity of Discover results. At no point was it brought to their attention that our US traffic was at risk of decimation.
11. A traffic drop of this severity suggests a fundamental change in how the Discover platform operates, yet no details were offered to key partners, nor have we been able to find any made publicly available. One possible clue is an article on 9to5Google, which promotes the addition of Google's own 'What to Watch' units and Shopping technology to Discover³. The date of this article, December 17, coincided with MailOnline's worst losses, which leads us to believe Google is seeking once again to promote its own products over the interests of long-standing media partners.
12. The following charts show the performance of Google Discover over the last 16 months. Chart 1 shows how MailOnline's global Discover traffic fell from around 2m visits per day in October 2019 to around 600,000 a day by December 2020.

² <https://www.searchenginejournal.com/google-discover/361142/#close>

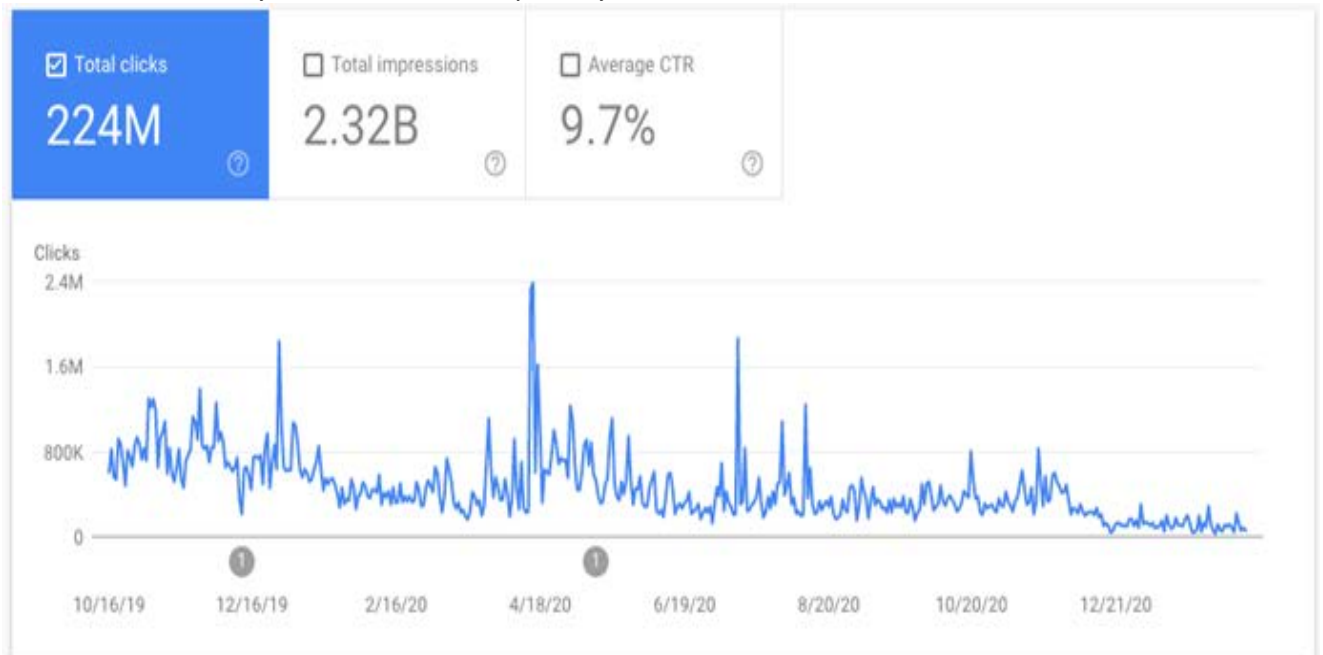
³ <https://9to5google.com/2020/12/17/google-discover-feed-what-to-watch-beauty-apparel-recommendations/>

Chart 1: MailOnline global visits per day via Discover



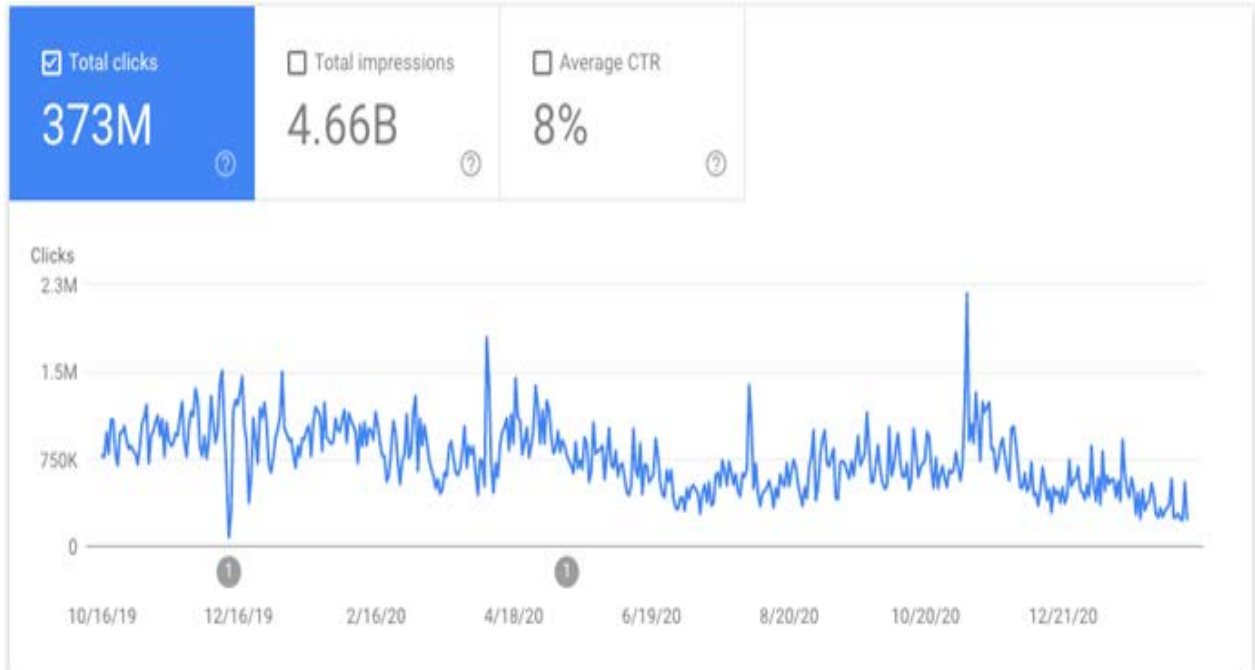
13. Chart 2 shows how the damage has been particularly marked in the US market, where traffic from Discover dropped from around 800,000 visits per day in October 2019 to around 130,000 a day by December 2020.

Chart 2: DailyMail.com US visits per day via Discover



14. The position in the UK, which remains MailOnline’s largest market, is little better. Chart 3 shows how traffic dropped from around 950,000 visits per day in October 2019 to 375,000 in December 2020.

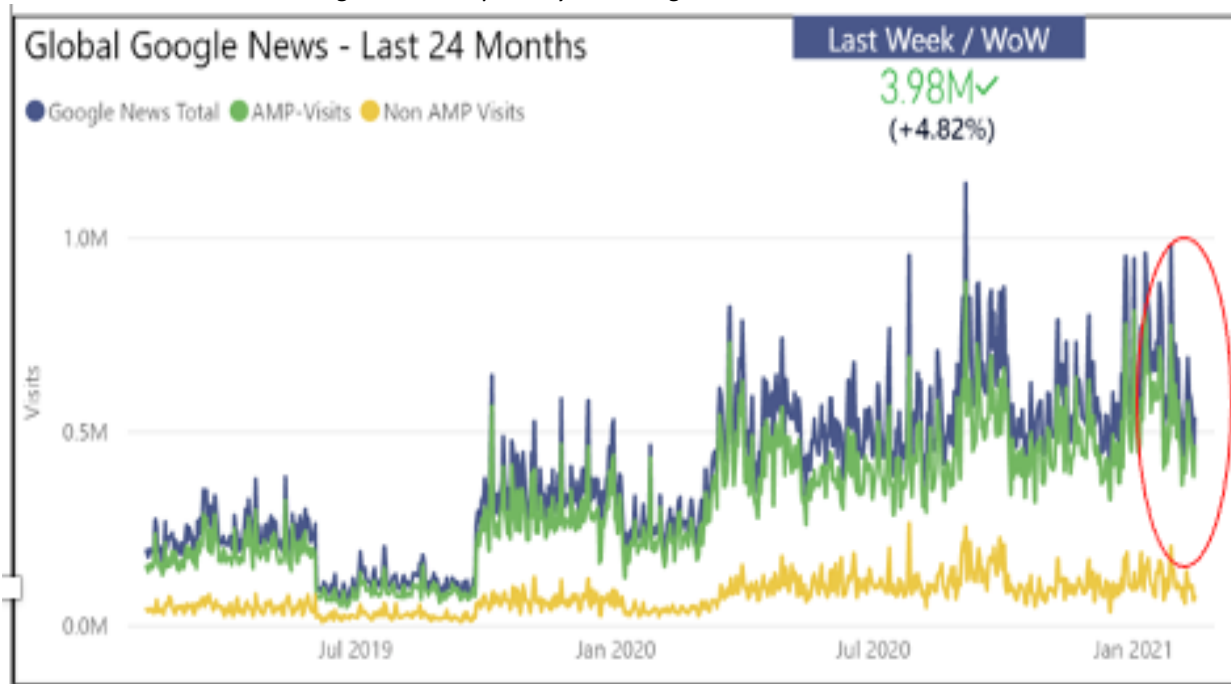
Chart 3: MailOnline UK visits per day via Discover



15. This trend has continued in the first two months of 2021, with traffic falling even further. A source of traffic that was once worth 2.8m page views a day globally, on February 26 reached a new low – 153,000. Our average page views per day in February were 492,000, and from the middle of the month onwards, we were unable even to reach this threshold. No explanation - public or private - has been offered to help us understand this.

16. Another source of Google traffic, also driven by algorithms, is its news aggregator, Google News. During the course of 2020 Mail Online was optimistic that growth in Google News traffic would mitigate some of the losses in Discover and Search. During the early part of the year global traffic grew from an average of 250,000 visits per day to around 625,000, briefly peaking at 1m. But Google News traffic was also hit in December, for reasons which have not been explained, falling back to 500,000 (see Chart 4).

Chart 4: MailOnline global visits per day via Google News



15. During the course of 2020, changes to algorithms have resulted in MailOnline losing 1.25–1.5m page views per day from Discover and gaining around 250,000 from Google News - an overall loss of 1m page views per day. A loss of 1m page views a day is not a trivial matter for any web business. Page views drive advertising volumes, advertising volume drives revenue, and without adequate revenue publishers cannot fund their journalism, on which democracy depends. We are fortunate in that our Google Search traffic remains steady at between 5–7m page views per day, but we understand that for some publishers Discover is now 50% of their web traffic.

16. We strongly recommend that an investigation into how Google Discover's algorithms function, and why user traffic has dropped so dramatically, is a priority for the Digital Markets Unit when it is launched.

Content Categorisation and Inventory blocking

17. The platforms do not only use algorithms to direct user traffic to the news websites of their choice. Algorithms also play a major role in the distribution of advertising demand. Amongst other functions, they are used by Google to apply various blocks to our advertising inventory. Some of these operate across Google's buy side within

DV360 and Google Ads, others on the sell side within Google Ad Manager. There are five different blocks/classifications that we understand are applied using a combination of algorithms and human review:

- Digital Content Labels in DV360;
- Sensitive Categories in DV360;
- Demand Restriction in GAM;
- Policy Violation in GAM;
- Confirmed Click penalty in GAM.

18. Digital Content Labels (DCL)⁴ and Sensitive Categories (SC)⁵ are classifications which, when applied within DV360, will exclude demand from running against a publisher's inventory. Digital Content Labels are based on the American video classification system and divide content into five broad categories, ranging from 'content suitable for general audiences' through to content only suitable for mature audiences'.

19. Sensitive Categories is a list of seventeen types of content which Google will block if advertisers don't want their ads to appear against it. Some are unexceptional if correctly applied – such as 'adult', 'weapons', 'drugs' or 'profanity'. Others cause serious problems even when applied correctly. For instance 'Politics: Political news and media, including discussions of social, governmental, and public policy' potentially defunds virtually all serious journalism. This problem had particular impact at the start of the Covid pandemic, when the government had to intervene to discourage advertisers from using 'Tragedy: Content related to death, disasters, accidents, war etc' to prevent ads appearing against Covid content, which at the time made up the majority of all news output.⁶

20. [REDACTED]

21. Policy Violation and Demand Restriction are applied on the sell side within Google Ad Manager. Demand Restriction⁷ covers 11 categories of content, ranging from "sexual content" to 'unapproved pharmaceuticals and supplements' against which ads may not be served. Policy Violations covers 22 categories, some legal, some technical, and others deemed harmful to users⁸, again against which ads may not be served

⁴ (<https://support.google.com/displayvideo/answer/2735881?hl=en>)

⁵ (<https://support.google.com/displayvideo/answer/6327207?hl=en>)

⁶ <https://www.theguardian.com/media/2020/apr/08/british-news-outlets-could-fail-due-to-coronavirus-ad-blocking>

⁷ https://support.google.com/admanager/answer/9335567?hl=en&ref_topic=7316904

⁸ https://support.google.com/admanager/answer/9335564?hl=en&ref_topic=7316904

22. Demand Restrictions block Google ADX running against the URL. These restrictions can apply to any URL we publish, and in order to maintain advertising supply we have to a) establish when a block has been applied (there is no Google alert or warning), and b) manually request Google to re-review and unblock the URL.
[REDACTED]
23. [REDACTED]
24. [REDACTED]
25. [REDACTED]
26. [REDACTED]
27. [REDACTED]
28. The fifth type of block we encounter is a Confirmed Click⁹ block. When Google deems that the inventory generates too many unintended clicks on ads they will require the user to click twice in order to reach the ad's landing page. This second click required from the user reduces the performance of our inventory, in turn causing Google DV360 and Google Ads to stop buying. These blocks can be applied account wide, and so impact all sites are running within GAM. We do not want to generate unintended clicks, and therefore conscientiously make best efforts to comply with Google's best practice advice. [REDACTED]
29. [REDACTED]
30. [REDACTED]
31. This incident is a perfect illustration of the deficiencies of algorithms, and the one-sided nature of the business relationship between platforms and publishers. It is wholly unreasonable that DMG Media – and probably many other publishers – should suffer significant revenue losses for ten months, while Google failed to find a remedy and changed its story to avoid taking responsibility for the damage inflicted.
32. The most worrying aspect of this problem – and of those we have encountered with Google Search and Discover - is that the imbalance of power between the platforms and publishers means there is no transparency. Changes to algorithms are made without warning, explanation or redress. We have no idea why Google switches traffic on or off – it could be for commercial or political reasons, or it could be that it does not properly understand and control the operation of its own algorithms. The platforms have never been held to account on this issue and it is time they are.

⁹ <https://support.google.com/admanager/answer/10025624?hl=en>

33. In the digital age democracy depends on vibrant, independent and pluralistic journalism just as much as it ever has in the past. That journalism cannot be sustained unless algorithms are operated in a way which is fair, consistent and transparent. That will only happen when the monopoly power of the platforms is balanced by regulation which ensures the platforms are required to give warning and explanation of all significant algorithm changes, and there is redress for publishers when their business is damaged.

34. At the risk of repetition, we believe it is vital the Digital Markets Unit starts to examine these problems as soon as it is open for business, both to establish whether any remedies can be effected under existing competition law, and to prepare codes of conduct for the future.

Peter Wright
Editor Emeritus
DMG Media
March 2021