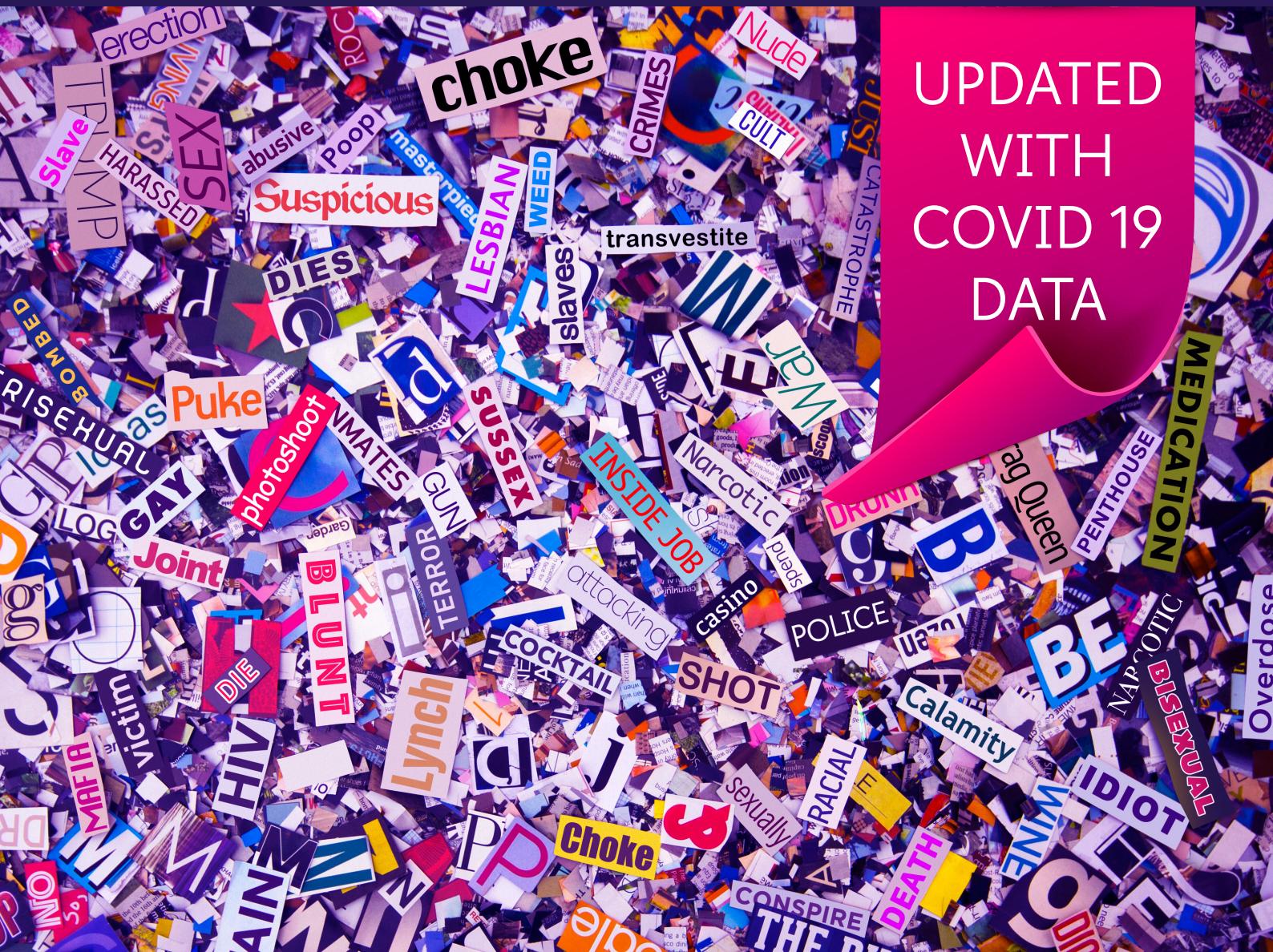


THE ECONOMIC COST OF KEYWORD BLACKLISTS FOR ONLINE PUBLISHERS



UNIVERSITY OF **BALTIMORE**

CHEQ

REVISED EDITION MAY 2020

Since the first version of this study was published, the world of online news publishing has faced an unprecedented challenge. The economic impact of keyword blacklisting set out in our first edition of this study has been exacerbated, in a way no-one could have envisaged, through a news cycle dominated by COVID 19. Despite record traffic for news stories during this pandemic, brand safety is preventing ads from appearing next to stories mentioning keywords such as "virus" "Coronavirus" or "COVID19". This revised edition includes new figures on the economic cost of this specific and systemic keyword blacklisting.

Following this general analysis, we delve into the wider problems caused by the continuing dominance of blunt brand safety efforts and its impact on the online economy. This sets out that online publishers face a hotbed of challenges in funding news. Rising pressures include economic instability; competition from giant tech companies, and [thousands of layoffs](#) across digital media. This comes at the same time as the industry has witnessed only a small rise in the number of people paying for news¹. What is clear is that the challenge of COVID 19 has hastened the need to adopt clearer brand safety measures for the long-term survival of the online news ecosystem.

¹ The Drum: 4 surprising media trends from the Reuters Institute Digital News Report

THE ECONOMIC COSTS OF BLACKLISTS DURING COVID 19

CHEQ

This study commissioned by CHEQ and undertaken with the Merrick School of Business at the University of Baltimore analyzes the revenue lost from safe, blocked, inventory through keyword blacklists. The following section is updated with specific information on the estimated costs of blocking online advertising next to Coronavirus stories on brand safety grounds. The effect of this blocking has damaged the revenues of online news sites across the world. It has united advertising leading advertising bodies and online publishers in the call for smarter brand safety as COVID continues to dominate news coverage. Reuters CEO Michael Friedenberg spoke for many when he said: "Factual news coverage on the global pandemic is essential and must be supported. It is incumbent upon the ad tech and verification industry to ensure that trusted news from reputable sources about Covid-19 is not blocked by safety filters." In the UK, culture secretary Oliver Dowden pleaded with marketers to rethink coronavirus keyword blocking, He said: "I am asking companies and the advertising industry to act and do all they can to resolve this issue," urging them to review their policies and check they are not "inappropriately" blocking ads from appearing next to the news providing a vital public service."

METHODOLOGY

COVID19 will generate at least one-third of page views for news sites between February 2020 and June 2020. ¹In 80% of cases, keyword blocking is used, by brand safety is preventing ads from appearing next to stories mentioning keywords such as "virus" "Coronavirus" or "COVID19". These figures are generated by applying this loss as a percentage of the online ad revenue for news sites. The analysis is based on 2019 online news revenue for these countries and assuming that COVID 19 continues to make up at least one-third of news page views until June.

¹ Based on Chartbeat data showing that COVID19 news stories make up one-third of page views, based on an analysis of 26 million articles

FINDINGS: COST OF COVID BLACKLISTING

CHEQ

In the period from February to June 2020, the revenue loss breaks down to \$1.3 billion for US news sites. In the UK, online news sites are set to lose \$105 million due to Coronavirus blacklists (£86 million); Japanese news sites will lose \$47.4 million (5.1 billion Yen) and Australian online news publishers will see a \$42 million revenue loss (\$66.7million AUS).

USE OF AI IN CORONAVIRUS BRAND SAFETY

The use of AI-based brand safety for brands is likely to allow vastly more ads on COVID inventory. CHEQ, which uses AI rather than keyword blacklists for brand safety advertising verification has found that in fact, around 60% of Coronavirus blocked stories are considered brand-safe for most advertisers. This would include stories such as USA Today ("Let's Give thanks to Tom Hanks: Watch his 10 Essential Feel-good roles"), because of mention of Hanks contracting the virus). The vast majority of top inventory about COVID 19 stories involve topics such as economic strategy, movie and book recommendations, tips for home working, and the global search for a vaccination, which are blocked by blanket keyword blocking. CHEQ's Health and Diseases category has been trained on more than 20,000 sources of data including news, and medical research and journals like such as Science, Nature, and, as well discussions forums for diseases to ensure accurate filtration of disease content. This is distinguished from other classifications such as "death", or "movies and entertainment".

ECONOMIC IMPACT OF COVID19 KEYWORD BLOCKING ON ONLINE NEWS PUBLISHERS, FEBRUARY TO JUNE 2020

COUNTRY	ANNUAL REVENUE FOR ONLINE ADS	MONTHLY REVENUE	MONTHLY LOSS	5-MONTH BLACKLIST LOSS
USA	\$12 billion	\$1 billion	\$264 million	\$1.3 billion
UK	\$955 million	\$79.5 million	\$21 million	\$105 million
Japan	\$431 million	\$35.9 million	\$9.4 million	\$47.4 million
Australia	\$382 million	\$31.8 million	\$8.4 million	\$42 million

THE ECONOMIC LOSSES OF KEYWORD BLACKLISTS: THE ANNUAL COSTS FOR PUBLISHERS

Brand Safety describes the controls that advertisers employ to protect their brands against negative impacts to their reputation associated with specific types of content and related loss of return on investment⁴. The principal protection offered by first-generation brand safety solutions to police brand safety incidents have been "keyword blacklists". These lists according to publishers have grown to up to [3000 words](#). One ad verification company alone has blocked [9.8 billion ad impressions](#) in a single year using this technology.

Publishers themselves permit the use of keywords. The UK's *FT.com*, in its brand safety guidelines, offers that "buyers can provide FT with their own keyword blocklists to run against their campaigns." News Corp Australia alerts advertisers to "the ability to suppress advertising against an agreed list of keywords that may appear within the content on the page or that the video is tagged for."

However, publishers have complained that this blunt approach denies ad dollars to news sites—with multiple campaigns blocked often with no explanation. The complaints about overblocking occur as such keywords fail to understand context or nuance. Standard keywords used across industry standard blacklists list words such as "kill" "death" "injury" or "sex" end up pulling over safe stories that mention that LeBron James 'killed It', or the "death" of a character in a Game of Thrones review. The mere mention of certain words instantly blocks ad dollars to publishers. LGBT publishers are hit hardest by the technology with *Pink News* citing the inclusion of certain keywords in blacklists, such as 'lesbian' a "death sentence" for their advertising revenue.

⁴ Brand safety reached fever pitch almost three years ago when The Times in the UK exposed how well-known brands were funding terrorism through adverts that appeared on YouTube videos posted by Islamic extremists. Such brand safety incidents have continued to appear periodically such as brands appearing next to stories such as the [Christchurch](#) massacre in New Zealand.

METHODOLOGY

CHEQ

In our analysis to determine the annual cost of the problem, researchers determined that 80% of ads served on premium news sites are subject to keyword blacklists. Through extensive analysis, researchers discovered that around 40% of global premium media inventory is brand safe — that is neutral or uncontroversial content, not in contravention of leading advertisers' brand safety guidelines. Nevertheless, the study found that, of this safe content, 57% of neutral or positive articles are being incorrectly blocked by keywords⁵.

The economic loss calculation involves applying the 57% rate of blocking against the 40% of brand safe inventory across online news media. This, in turn is set against the 80% of online new media inventory governed by blacklists. The final loss is calculated as a percentage of the online ad revenue for news sites.

This level of over blocking, and subsequent loss in revenues, is also corroborated by data and estimates supplied by publishers who consistently discover incorrect blocking affecting up to 60% of their safe inventory. While unable to quantify the exact amount of loss, most report the figure to be "very high" and a "highly significant." On an individual campaign level, executives claim they can lose between 10% and 30% of ad revenue in a single campaign. Some are much higher — for LGBT sites, the number of safe articles blocked rose to 73% due to the widespread presence of words such as "same sex" or "lesbian" on blocklists.

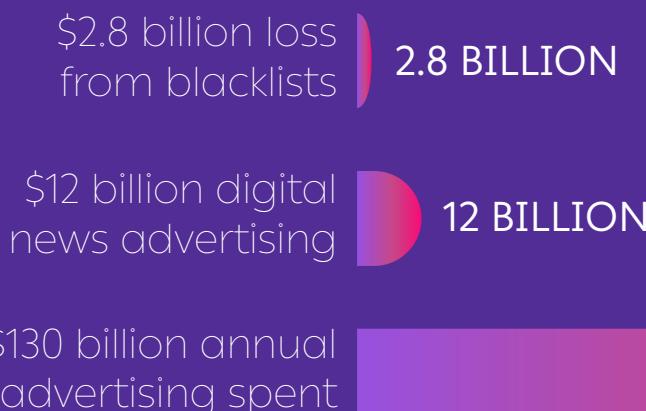
In this report, CHEQ also reviews the top trending stories of 2019 on Google in each of our countries, generating billions of impressions to news sites analyzed against industry standard keywords to reflect the size of the problem. Some 55% of consumers access news through Google, social media, or news aggregators⁶ but keyword blacklists blocked the top safe stories of the year, including the Rugby World Cup, and Avengers. This made such topics off limits for advertisers bleeding publishers of essential revenue on their most trafficked stories.

⁵ CHEQ: How Keyword Blacklists are Killing Reach and Monetization <https://www.cheq.ai/Keyword-Study>



THE US

**\$2.8 BILLION EACH YEAR
LOST DUE TO INCORRECT
BLOCKING OF SAFE CONTENT
ON PREMIUM NEWS SITES.**



US news publishers are losing \$2.8 billion each year due to the incorrect keyword flagging of their online content by brand safety tools. This amounts to publishers losing out on nearly one in every four dollars of potential programmatic ad revenue.

This economic loss is based on current online spending for US online news of \$12 billion annually⁷. This comprises two main forms of advertising: digital consumer magazine advertising (\$6.6 billion) and \$5.4 billion for digital newspapers including titles such as the New York Times and Washington Post (\$5.4 billion).

The importance of programmatic ad revenue in the US is rising in comparison to print ad revenue. For instance, *The New York Times* recently saw digital advertising revenues ([\\$103 million](#)), surpass print (\$88 million), as it achieved \$709 million in overall digital revenues. Given the importance of these programmatic ad revenues, all news sites have suffered from the blacklisting challenge. Guthrie C. Collin, chief analytics officer at Dow Jones says a standard keyword blacklist blocks advertising on up to 20% of the site's content. Collin [says](#): "[The journalists] could write something like, 'death to the earnings of...' or 'it's killing their margin' and that would be blocked."

⁷ PWC Global Entertainment and Media Outlook 2019-2023

CASE STUDY: DISNEY PLUS STORIES BLOCKED

CHEQ



The most-searched phrase on Google in 2019 in the United States was "Disney Plus", however the most-read news sites stories about the new streaming site, were unmonetizable. The most read article on the service, [Disney+ Complete List In Alphabetical Order!](#), shared 459,000 times on social media⁸, from publisher Just Disney, a media site powered by Disney fans, was flagged for a brief mention of *Star Wars Attack of the Clones* (keyword: "attack"). One of the other most-read posts, CNET's "[Everything to know about Disney's new streaming service](#)" was shuttered because of the word "adult" (the offending sentence: "Hulu is where Disney's more adult content is housed").

⁸ Data on social shares from buzzsumo.com

ONLY 9 OF 100 OF THE NEW YORK TIMES' 2019 MOST-READ ARTICLES WERE BRAND SAFE

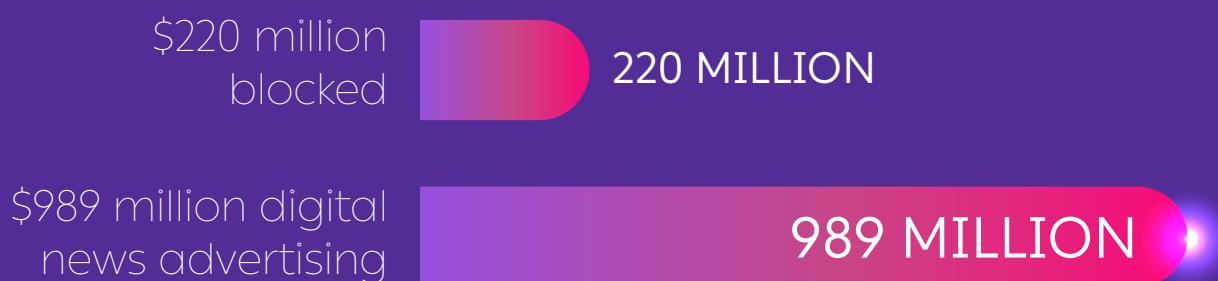
CHEQ

The issue is highlighted by the fact that of the 100 most read News articles from *The New York Times* in 2019, only 9 were considered brand safe by keywords blacklists effectively killing monetization. While some of their most-read articles most read are rightly blocked — several shootings, and the scandal of Jeffrey Epstein, other pieces of journalism starved of monetization are baffling choices. For instance journalist Kevin Roose's story about his detox from his mobile phone is blocked for the keyword of "alcohol" (Unlike alcohol or opioids, phones aren't an addictive substance so much as a species-level environmental shock"). No less than four entertainment stories featuring in their vaunted top 100, were blocked. This includes these blacklist offenders: "The 50 Best Movies on Netflix Right now" (blocked for mention of "sexuality": 1967's Bonnie and Clyde "mixes sexuality, danger, restlessness and ennui"); "10 Recent Netflix Originals worth your time" (blocked for "sex"); a review of Game of Thrones (blocked for "death") and a review of The

Science stories are also victimized — including the first ever image of a black hole (blocked for "violent" ("There, a black hole several billion times more massive than the sun is unleashing a violent jet of energy some 5,000 light-years into space"). The keyword of "slavery" also blocked ad revenue: this included a travel story mentioning Williamsburg, known for the arrival of the "first African slaves to North America"; a story about the Duchess of Sussex, Meghan Markle's giving birth to a baby boy ("The entry of Meghan Markle, the descendant of plantation slaves, into the royal family resonated deeply with many people of African descent") and an in-depth investigation into 400 years of American slavery. The blocking of LGBTQ news is also evident: the story about Jonathan Van Ness of TV show, Queer Eye, was deemed unmonetizable through the deployment of any number of common industry-standard keywords including "HIV", "same-sex" and "sexual".

UK

UK NEWS SITES LOSE \$220 MILLION (£167 MILLION) IN REVENUE FROM INCORRECT KEYWORD BLOCKING



In the UK, the cost to news sites from overblocking was \$220 million (£167 million) in 2019. This is based on annual brand spend on online digital news in 2019 of \$989 million (£758 million). The Advertising Association found that total internet advertising grew to £11.5bn but a small proportion went to online news. National news brands picking up £275m in digital advertising, regional news brands £212m and magazine brands, £271m⁹. Overall, more than one in every five dollars of ad revenue is lost by UK news sites because of incorrectly blocked keywords.

Discussing the daily struggle against keywords, one national UK newspaper site publisher explained: "We have had times in the past when block rates are high and the clients have turned around and said they are not going to rebook, or where we have had to refund them.

"We recently had a campaign running that flagged 40% of impressions (on all inventory) that day as unsafe. We published an article on how the Chernobyl disaster zone is becoming a sanctuary for wildlife, which was being blocked because it had the word 'disaster' in it."

James Wildman, CEO of Hearst UK publisher of *Cosmopolitan*, adds: "We are getting blocked for keywords like 'shoot'. We write about photo shoots constantly. The word 'Manchester' still gets blocked, as they [advertisers] haven't updated their blocklists [since the Manchester Arena bombing in 2017]. We are doing quite a bit of work to shine a light on how crazy it is."

⁹ <https://www.pressgazette.co.uk/uk-advertising-spend-grew-to-record-22bn-in-2017-as-duopoly-prosper-and-newsbrands-are-squeezed>

CASE STUDY UK: RUGBY WORLD CUP BLOCKED

CHEQ



In the UK, the most searched phrase on Google in 2019 was the Rugby World Cup (it was the second most popular Google search in Australia, and third in Japan). However, two of the top five most read UK news articles about the tournament were blocked on brand safety grounds. This bled ad revenue for stories which together generated 273,000 social media shares among rugby fans. The first article, from *Lad Bible* ("[England Are In The Final Of The Rugby World Cup](#)")

protected brands from the word "attack". Coach [Eddie] Jones was quoted as saying: " I thought when we had opportunities to attack, we attacked well." The second most popular UK news article on the tournament, (*The Daily Mail's "[England players refuse to put on runner-up medals after heavy Rugby World Cup final defeat](#)"*) was blocked because of the presence of the word "injury" ("Prop Sinckler - replaced after two minutes with a head injury, ripped his [runners-up medal] straight away").

JAPAN

JAPANESE PUBLISHERS LOSING \$98 MILLION IN REVENUE FROM BLACKLISTS

In Japan, the world's third largest advertising market (behind the US and China) losses from blacklist blocking stands at \$98 million (¥10.6 billion yen). This is based on online digital news spend of \$431.5 million (¥ 46.9 billion yen), estimated by leading Japanese agency, Dentsu.

\$98 million
blocked

98 MILLION

\$431.5 million digital
news advertising

431.5 MILLION

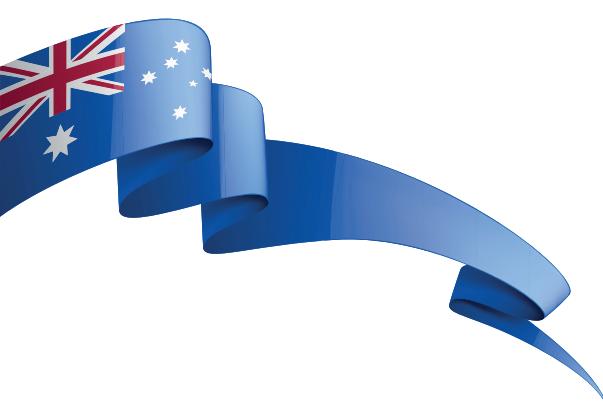
BRAND SAFETY BLOCKING SUMO AND BASEBALL

CHEQ



Brand safety blacklists naturally include words in all languages, including Japanese. However, to give the example of English content blocked, *Japan Times*, the country's largest and oldest English-language daily newspaper, two of the country's most popular sports — sumo and baseball — are regularly blocked programmatically by blunt brand safety. For instance, legendary sumo veteran Aminishiki is deemed unmonetizable because of the [keyword, "injury"](#).

("[Aminishiki] called it a day early in the tournament after an injury-forced withdrawal meant that he would have dropped out of the salaried ranks"). On the same news site, another sports article was blocked for "shots". In this case, the Nishinomiya Storks baseball team's new coach, is considered unsafe by first-generation brand safety for letting it be known that the new appointment, Christos Marmarinos would be "calling the shots".



AUSTRALIA

In Australia the revenue cost of blacklists amounted to \$84 million annually (\$122 million AUS). This is based on digital news advertising spend of \$382 million (\$559 AUS)¹⁰. Overall, one in every five dollars of ad revenue is lost by news sites because of incorrectly blocked keywords.

\$84 million
blocked

84 MILLION

\$382 million digital
news advertising

382 MILLION

¹⁰ DPwC Entertainment and Media Trends: Australia

CASE STUDY: AVENGERS ENDGAME

CHEQ



The blacklists denied advertising dollars to many popular entertainment stories read by engaged Australians. The country's fourth most-searched trend searched trend on Google in 2019 was "Thanos", a character from the Avengers movies, while the fifth-most searched term was "Avengers: Endgame". Both of these armies were ultimately defeated by keyword blacklists. For instance, one of ABCNet's popular articles ("[The Timeline leading up to Avengers Endgame](#)")

was shared 3000 times on social media, but ultimately was blocked for mentioning a New York "attack" against the character Tony Stark in Iron Man 3. In another primer article, by News.com.au, an article providing an [explanation of time-travelling mechanics in the superhero movie](#) was caught in the brand safety net for using the word "kill" ("Thanos wouldn't be on the Asgardian refugee ship to kill Heimdall and Loki, or sacrifice Gamora on Vormir").

KEYWORD BLOCKING: A CLASSIC EXTERNALITY

In economic terms, brand safety blocklists are an example of unintended economic consequences known as "negative externalities". In 1920, Arthur Pigou introduced this concept in *The Economics of Welfare*, using the example of alcohol which produces economic impact, including adding costs for police and health care systems.

MARKET FAILURES

Negative externalities costs affecting a party that did not choose to incur it¹¹. This theory also points out that negative externalities suggest market failure. Like in our case of keyword blocking, market failure is defined as an inefficient distribution of goods and services in the free market, where individual incentives for rational behavior do not lead to rational outcomes for the group.

INDIRECT COSTS OF BLACKLISTS

So far, we have discussed direct economic costs of the blacklist problem. This addresses the direct revenue lost by publishers. In contrast, indirect costs are the opportunity costs and less tangible effects of harm. These include executive time spent dealing with a situation, unintended consequences, or chilling of innovation. The economic consequences of blacklists satisfy many of these definitions.

¹¹ Externality: James M. Buchanan and Wm. Craig Stubblebine

EXECUTIVES LOSE VALUABLE TIME DEALING WITH THE CHALLENGE

The amount of executive time spent cannot be underestimated in the battle against blacklists representing a large indirect cost. Kelly Facer, Vice President of Revenue Operations, Sightline Media, the world's largest publisher of military, defense news and history sites, says that her fight against blacklists is a time-consuming "daily

She says: "We continuously have to seek exceptions to the rule. It is just a guessing game about what could be going wrong and why we cannot figure out delivery until someone chooses to disclose."

PUBLISHERS FORCED TO FIND ALTERNATIVE AND LESS PREDICTABLE REVENUE

The keyword-first approach is also causing a chilling of innovation, forcing publishers to find other sources of revenue. In particular, LGBT publications have had to turn to non-advertising channels after being frozen out by blacklists. The challenge issue is growing—the [Association for Online Publishing](#) found 88% of news publishers have escalated non-advertising growth as a top priority.

CHANGING JOURNALISTIC QUALITY

It is also hurting the ultimate product of news, changing the focus on independence and quality. Bedir Aydemir Head of Audience and Data, Commercial at News UK, publisher of *The Times* and *The Sun*, says: "We would never let advertising impact how we write about something. But for some smaller publishers I am hearing that they are completely changing the way

in their articles to try and circumnavigate these very blunt blocking tools. "This frightening scenario of undermining independent journalism represents an unknown and disturbing economic development. This trend is likely to have a significant impact on the reputation, quality and premium nature of the online news ecosystem.

VERIFICATION SLOWING SITES

Other unknown indirect costs include complaints that keyword blocking brand safety tags are slowing sites. Used as part of keyword technology, such wrappers [add between two and four seconds load time](#), at a time when Google reports that [53% of people will leave a mobile page](#) if it takes longer than 3 seconds to load.

LOSS OF REACH FOR BRANDS

CHEQ

Indirect costs are also borne by brands themselves. In particular rigid and unfair keyword blocking has seen brands edged away from engaged news audiences. One analysis [found that](#) ads in quality news environments are 42% more effective for advertisers based on engagement, viewability and dwell time. The race to brand safety could see brands appear on less credible news sites with poor, or irrelevant, content.

In the words of *Washington Post*, CRO Joy Robins: "Ultimately, all these brand safety or keyword blocklists are going to do is end up putting advertisers who want to reach a B2B audience or the sophisticated audiences that are news consumers, and put them in places that are sports, or entertainment which isn't necessarily the type of alignment they are looking for."

In addition, blacklists can harm their openly stated strategic or economic advertising strategies. For instance, more than 100 brands paid to collaborate with the hit HBO show, Game of Thrones, yet at the same time, 66% of online news articles about the hit show were blocked from serving on reviews and articles about the hit show (due to the presence of words such as "death", "sex" and "bastards"). In another instance [publisher, Hearst](#) say that Meghan Markle, is regularly flagged by blacklists. This is because her title, (Duchess of Sussex) contains the word "sex". Once again, this occurs despite association with the princess, in many cases, causing a [brand's turnover to double](#).

SOCIETAL DAMAGE: LGBTQ+ NEWS PUBLISHERS FACING STRUGGLE FOR SURVIVAL

Equally less easy to gauge economically is the societal damage caused by defunding news, particularly hurting minority groups. LGBTQ+ publishers see nearly three-quarters (73%) of neutral or positive LGBT online news incorrectly flagged by brand safety blacklists denying them vital ad revenue. Orlando Reece, CEO of Pride Media, which owns LGBTQ+ site, *The Advocate*, says: "In the end, blacklists hurt us all. We create content from an LGBTQ+ lens

and use words that our community, as well as the general community, do not find offensive, such as 'lesbian,' and 'queer'. These words are part of our lives and how we communicate and identify with one another. Brand safety needs to be a conversation involving people, not technology that is ill-equipped for today's digital publishing landscape." Michael Weaver, Senior VP of Business Development and Growth at Al Jazeera Media, adds that the economic impact of blacklists is "large". He warns: "The preventing of ads next to hard news could

CONCLUSION

CHEQ

Despite our findings, the use of blunt keywords is increasing. One ad verification provider has said that last year alone they saw a 250% increase in blocking. They added that keyword use has seen a [10% annual increase](#) in the number of clients that use the feature. In addition, 43 percent of media buyers say they now explicitly [avoid advertising next to the news](#), despite publishers investing heavily in brand safety policies.

Furthermore, the lack of transparency evident in decision-making remains a challenge with parties having little access to why large swathes of inventory are being blocked. Information asymmetry (when parties do not have access to the same information) has been shown to underpin the most inefficient markets (including the challenge [of ad fraud](#)) and acts as a barrier to reform.

More positively, better technology, particularly [the use of AI](#) which understands the context of content on news sites, has

It will be fascinating to see the effect that this is having with deep contextual understanding replacing keywords, promising the return of brands to news. It comes as some brands seeking better advertising outcomes, and more engagement in the causes their consumers care about, are slowly turning the taps back on news site advertising. In one example, Bank of America committed [30% of its entire digital budget to news advertising](#) at the same time as gearing its brand safety towards next-generation protection.

Regardless of the approach taken, any strategy must understand the deep economic importance of well-functioning publishers, and the impact of denying them revenue, particularly in an untransparent fashion. The modern brand safety approach must evolve to ensure stringent protection for advertisers, while seeking to maintain and restore the role of online news publishers as a democratic and economic force.