

COMPETITION COMMISSION OF SOUTH AFRICA
In the matter of
MEDIA AND DIGITAL PLATFORMS MARKET INQUIRY ("MDPMI")
held in hybrid format at
Dtic Campus, Sunnyside Pretoria and virtually via MS TEAMS
on 18 March 2024

Chairperson: Chief Economist and Acting Deputy Commissioner:
Competition Commission: Mr. James Hodge

PANEL MEMBER:

Ms. Paula Fray

Day 11: School of International and Public Affairs: Columbia University

START OF PROCEEDINGS ON 18 MARCH 2024

CHAIRPERSON: Good afternoon and welcome to week three of the public hearings of the Media and Digital Platforms Market Inquiry. Today we just have one stakeholder, very fortunate, privileged to have Dr Anya Schiffrin, who is director of technology, media and communications at the School of International and Public Affairs at Columbia University in New York. And she will be joined by Haaris Mateen who is a fellow researcher on their paper, Paying For News.

10 So, welcome Anya, I know it's morning there in New York and thank you for making the time to come and speak to this inquiry.

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Thank you all for having me. Thank you to the panel and thank you to the competition commission as well as to James Hodge and other friends who I see on stage.

CHAIRPERSON: Great, Anya. I believe you've got a presentation so we're going to let you run with that and then we'll have questions afterwards. We won't interrupt you, but yes, our modus operandi here
20 has been to just use James and Paula and we'll call you Anya if that's okay?

DR. ANYA SCHIFFRIN SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Please do. Yeah, I'm just looking to see how to switch to presentation. I seem to have a lot of things up here. Here we go. Okay, let me get into presentation mode.

Yeah. So I'm Anya Schiffrin and I've been working at Columbia University for more than 20 years. And I was asked to come up with a study as to what publishers would be owed by Google and Facebook, were we to have an Australian style law in the United States. I did not receive any payment for this funding and my lead author, Haaris Mateen, also worked pro bono. And our co-authors on the study are Patrick Holder and Haris Tabakovic and they work for a firm called Brattle which does witness prep for very, very large lawsuits including some that deal with compensation and technology.

10 They work in the technology practice, from what I understand. So they were able to assist with checking some of the calculations and contributing otherwise to the paper. Haaris finished his PhD at Columbia last May. He is now a doctor and he is as assistant professor of economics at the University of Houston. So, we are two academics who took this project on out of interest. This is a summary of my presentation. We're going to lay out the main problem which is that platforms have benefitted from news content but we don't know how much they've benefitted or the dollar value. They have no provided exact data that many have asked for so that the exact

20 amounts can be measured. And so scholars around the world have studied this question and so have news organisations and they've made different estimates about how much they're probably owed, but all of these estimates suggest that the funds paid out by the platforms are far lower than the amount owed. And this underpayment is due to the market power of the platforms and there's an imbalance of power

and negotiating power and information power and so the platforms tend to pay whatever they feel they can get away with. This is a global problem and different countries have addressed the problem in different ways which we'll be discussing. And platform remuneration policies, which are policy decision, have certainly raised the amounts paid in different countries. However, policy choices and how to deal with this overarching problem, belong to each government. And then I'll finish by talking a little bit about the rise of generative AI which is already causing further disruption to the news media industry. I'd like

10 to start by saying that Google has created an incredible search engine and they did a very good job in their testimony here last week describing how it works and how they do links and how they create search and how they advertise from it. But the point, I think, for all of us, is not so much the ins and outs of what Google's technology is but the point that because the engine's so powerful and useful, the firm now has huge amounts of market power. It's really an oligopoly. It's pretty much unrivalled in many areas. And the October revenue of Alphabet is \$307 billion with approximately about 70 billion of prophet in 2023 and their Search revenue is based on their advertising. So

20 these are very large amounts that will give you a sense of how big and powerful this firm is. And you can see that in 2020 they pledged about a billion dollars to support journalism. And this was not actual cash, this included all kinds of access to their tools and training on how to help their businesses grow. And the actual amounts are actually unknown so we don't know how much they really paid. I won't

get here into taxes and the tax avoidance by the large platforms, but that should also be taken into account when designing policy measures. So, part of our argument is that as Google was designing its very powerful search, they knew that they were making click through to news unnecessary. In December there was a lawsuit filed by Helena World Chronicle, LLC v Google and Alphabet and then the complaint is quite long, but they also point out that Google has been extracting information from publishers and using it to build knowledge panels on the front of the Search page and those end up hurting referral traffic. And this is also true of their AI patents, which they alleged in 2016 patent, and when they said that they knew it would damage referral traffic and their patent for featured snippets which use the lead from news articles to make the search engines results page look like an encyclopaedia entry. And this reduces the need for people to click on links to news. So Google's own design, which they have stated in their patents, is to make sure that people don't click on news. And that's relevant because they talk about the importance of subscription revenue generated by referral traffic while knowing that they were going to minimise that. What's happened recently is that Spanish publishers are now suing Meta and Chilean publishers are suing Google and it's partly over the self-preferencing of the ad practices and partly because of this problem of revenue not going to them. So, we view the main problem is really one of platform remuneration. And our perspective is on intellectual property is actually different than the platforms which are using it and benefiting

from it. Google and Meta do not think that they owe money to publishers. They say that they're just using fair use and fair use under copyright means no payment is required. But we believe that they do owe money, very clearly because they've benefitted from news and they've profited off news. Google and Meta get quite emotional when you talk to them and they say, well we don't owe anything and why are you asking us this and you should be glad you're getting all this traffic and Search is helping. And they often say well they're critics, they're just a bunch of luddites who don't trust technology and they

10 say, well the horse and buggy drivers didn't accept the rival of the automobile. It's very common in tech circles to say, well the monks who made manuscripts didn't appreciate the printing press. So they try to say that the rest of us are outdated, legacy journalism is outdated, audience habits have changed. But their argument is a little contradictory because it's basically we don't owe journalists any money and anyway, we gave out lots of grants to newsrooms. So our perspective is different. We believe that Google and Meta, quite possibly other platforms and certainly large language models soon, do owe money to content producers including news publishers

20 because the platform profits would not be what they were had it not been for their taking content without really paying for it. Many economists, who we have talked to, believe that the amounts that have been paid out are far too small. We believe that news is valuable to Google and Meta and also to the large language models. Basically, news content in large part, benefits from audience impressions on

platforms. Platforms can monetise impression's data when selling adverts. Impressions are also valuable to platforms because they capture user preference when users look at content on the platform. Credible news content builds habit and trust for users and this has spillovers into other types of profitable searches. People, as I mentioned, are consuming news already by a search and social media without clicking through to the source and that has become a direct substitute for consuming news via a newspaper. Their systems were designed by Google deliberately to hurt this click through and

10 because of their market traffic, the platforms pay peanuts for their use of news content in search and social media. I like to say, when I'm teaching to my students, I like to compare it. If you're a newspaper, imagine that you're a muffin baker and you have beautiful crusty muffins and you sell them every day and then every morning somebody comes through and takes off the top of the crusty old muffin, the delicious part with all the blueberries and all the nuts and they leave you with a little dry bit underneath. And then it's hard to sell that dry bit and then the platforms say well you need a new business model. And I would say no, there's been cream skimming because

20 the platforms have taken off all the best parts, they've taken off the headlines, they've taken off the main gist of it. So here we are in a situation where they've taken the best part, that's resulted in lower sales and subscriptions to news outlets and it also means they're not getting the ad revenue that they were getting. So, in the situation where the platforms have refused to share their profits with the news

outlets, what we've seen is that the threat of legislation has actually pushed the platforms to pay for content in some countries, although it has not been enough. And this slide was adapted from the testimony from economist, Hal Singer, who testified in hearings in California last fall. So, the question I was asked by the commission was how to measure the value of journalism to the platforms. So, the simplest would be if we got data from Google and Meta, right? It would be helpful to know which part of their revenues come from the South African market, which part of the revenues comes from attention to news and then we could throw in accounting of direct and indirect value added together. We could then deduce the referral traffic, which is the number of people that click through the [bios description? 00:11:29] and have a pretty good idea of what is owed. With Facebook, if we could attribute the advertising revenue attributable to each user and measure the share of time that each of these users spent consuming and interacting with news content, then we could allocate the advertising revenue associated with each user to news content versus other types of content on the platforms. Or, if researchers had access to the share of news content related informational searches done on Google, they could calculate the share of advertising revenues that might reasonably have attributed to news content. But Google and Meta do not reveal those numbers. What Google tells everyone, and I even saw this in Bahasa, in a document given to the Indonesian government, is that only 2% of direct searches are news related, but they don't define what means

that 2% is. Does that mean that just people that go to Search and type in the New York Times, what is that? What is that 2%? It appears everywhere. We heard Rasmus Kleis Nielsen last week say 3%, nobody has defined it. I think 2% is a bit silly. It's obviously very low and I've noticed that they have actually stopped saying that in the US and I think that's probably a testimony to the work that researchers have done, including ourselves, suggesting that's silly. And you can look up the Washington Post op ed by Jims Albrecht that got published a few weeks ago where he doesn't even mention the 2%.

10 So, my take is I adore Googling, I do it all the time. Sometimes I'm looking for a restaurant or a hairdresser or what time the supermarket opens, but a lot of the time I'm looking for something news related. It could be an article I remember reading or it could just be information that's reported on the news. So clearly, I think, the 2% figure is so narrowly defined, whatever that definition is, it's utterly meaningless but we don't even have the definition. It's also interesting that they say 2% all over the world. Surely there might be some different habits in different parts of the world, but they just keep saying 2%. Anyway, bottom line is that since we have no data from Google or Meta, the
20 team of economists I worked with had to make their own calculations as to what Google and Meta owe US publishers. And there've been other estimates as well. The parliamentary budget office of the Canadian government released report in 2022 estimating that their Bill C18, which was like the platform remuneration law of Australia, would release 392 million Canadian dollars per annum. Australian

competition commission used their own methodology and their numbers were lower than ours. The news media alliance in the US, which is the Lobbying Association of News Publishers, hired a well known economist called Hal Singer to do a confidential study for them. All of these different methods came to a similar conclusion about the amounts owed to publishers. When we peer reviewed our working paper, my Columbia journalism school colleague, Emily Bell, put “It’s billions, not millions of dollars that are owed in many large markets like the US.” Canadian scholar, [Sean Hugh Roy? 00:14:46] who I just
10 found his work online and contacted him, uses CrowdTangle data, which by the way Meta is about to shut down CrowdTangle, and then he used GDP figures from the world bank regional figures from both companies for the rest of the world for the Meta and MEA. And the Fair Mateen Brattle methodology, which I’ll talk about. And Roy came to a fair payment of \$145.8 million from Meta and Google to South African publishers per year. So that’s his figure. When I was looking all around the world last spring, trying to find methodologies that we could use in our paper, one of the papers I came across was a Spring
20 2023 study done by the very well-known and respected behavioural economist Ernst Faure, which looked at media content and I’ll explain more of his study but they, his hypotheses that when searching for information, people are concerned with quality completeness, trust, independence, up to dateness and orientation, that a large number of Swiss people in their survey are informed about news by Google and that people stay on Google when they get the information they need.

And that journalism content makes Google attractive to users and that media content is significant, contributes significantly for 70% of information searches. In his study, they chose random individuals, divided into two groups. One was given Google with news and the second was given a Google without news. So they had two arms of their study. And what they found was that 50% of searches are related to information, that's 55%. 70% of searches are related to news in some way, that's of the information. 53% of users stay on Google and do not click on links. And without news, subjects were less likely to
10 look at Google. So, if 35% of Google revenue in Switzerland is attributable to news searches, the question then becomes how do you split it? And the fair study decided 60% to go to the platforms, 40% to publishers and that amount would be \$174 million a year. They have then gone on to replicate the study in other European countries where it has worked and these numbers are being used in negotiations between platforms and publishers. What we did was use different methodologies, we used the fair methodology for Google and then we used a different methodology, which we'll explain, for Facebook. And what we found, Haaris and the economists on the team, is that Google
20 owes US publishers \$13.6 billion a year and that Facebook owes \$1.9 billion a year. And if we scale the pay shares for US publishers by South African GDP, South Africa's GDP is approximately 2%, we would find that Google owes South African publishers \$244 million a year and Meta and Facebook would owe \$34 million a year. So that's about 270 million a year more. South Africa, I'm looking to see if that's

consistent, that's more than the [roy? 00:18:15] amount. So you can see, Haaris can explain better some of these slides, but you can see here how news appears on the sort of basic Google Search page, right? So there's an about but there's all kinds of news about Ukraine. Here as well you've got a happy birthday message, but underneath you also have a news. So the key findings from our working paper were that first of all value is jointly created by both the platforms and the publishers working together. And another way I'm going to explain that is I'm going to tell a story that in December, a friend of mine had
10 a fall and decided it was time to go and live in assisted living. And her daughter called me and said, I don't know what to do because my mom's apartment needs to be sold and it's the biggest asset our family has. And I think that the apartment is probably worth about \$2 million and we have a neighbour and her apartment is worth about \$4 million. She wants to buy my mom's apartment and combine it and I think that together, in a combined apartment with a river view and three bathrooms, could be worth about \$15 million. So that is a \$9 million surplus from working together. How should we divide that? What should we do? How much should I ask for? She called me and I said,
20 well that's so funny because I'm working on this exact same question right now of platform and publishers. And here's what the economists say. Economists say they've been working on this since the 1950's and if you want a more recent book then read, Split the Pie by the game theorist, Barry Nalebuff at Yale. And what economists find is that actually the fair split would be 50%. So 50% of news related

content revenue earned by Google and Facebook should be split with publishers and that's how we got the number. And then we also looked at okay, not just economic theory, but what's really been happening in the world over the last few decades. And the answer is traditionally licensing agreements were about 50%. And then along came Google and then to Meta and with all their enormous market power and we found that they were making payments that were really, really below what we would think of as a fair payment. So, as I mentioned, news is valuable to Google. It provides better, more current and more reliable search [indistinct 00:20:51] than anybody else and that means everybody wants to use it in the future. There are spillover effects from more profitable searches such as commercial searches. And in other words, news content contributes to the attractiveness of Google overall and that increases its value, credibility and completeness. And this is amplified by the network effect, by attracting more users, the platform attracts more paying advertisers. The more people use the search, the more data Google collects and that allows it to improve both its Search and its results and the targeting of its ads and the reach of its ads. So, all of this better Search results enhance user experience and improve outcomes for advertisers, therefore Google increases the overall value of its Search income on multiple sides of the platform. So, what we did was we calculated relevant advertising revenue for the platform. The share of advertising revenue attributed to the presence of news content and then the appropriate split share between platform and publishers. And Haaris and the economists on

the team used data from company filings, state of the art studies and then historical benchmarks as well as economic theory. So, what we found, to go into detail, is how many people use Google because it provides them with information and it would stop if they didn't get information. How much of the information content comes from news? Using the fair 35%, we arrived at the 10 billion a year, that's accounting for the split. For Facebook, we used different data. Currently, Haaris kind of inferred his figures of time spent consuming news based on an Allcott at All Paper, a 2020 paper in the American

10 Economic Review which is a top economics journal in the United States, probably in the world. He wrote a paper called Welfare Effects of Social Media. Haaris took some of those numbers and then imputed, which is what scientists do all the time. Haaris also looked at testimony that Facebook gave to the American House about social, Facebook use and so that's how he came to the 2 billion a year. And what we've got here, this methodology can be replicated anywhere in the world. People have already started replicating it, including some of the earlier people, Rod Simms tried it. And so you can see, it's very simple to do this. So you can do it yourself, you can get your

20 economics department to do it at your local university and here's the Google methodology as well. So it's not a complicated calculation and it would be better if we had better data. And then also Haaris can talk about these slides later but revenue of newspapers as a percentage of US GDP has fallen. And news consumption on digital platform Search is stable. This is 20 to 22. We explain here exactly why

impressions are important and, as I've said before, platforms monetise impressions data when they sell adverts. But they're also valuable because they capture user's preferences when user's look at content on the platform. So they're fully aware, when you're looking at something, what you're looking at and they can use that to see how to adjust the algorithm, the design of the platform in order to maximise engagement from both users and advertisers increases the change that users will view and notice adverts. The impressions itself is a metric use to quantify the number of times a user views a website
10 where an advertisement is displayed. It measures user engagement with the content and it increases the chance that users will view and notice the adverts. There are feedback effects and network effects. Click through is downstream from this. Our methodology is totally transparent, anybody is welcome to copy it and to use it in their own countries and they are and they can use their own data. And it really doesn't matter if you adjust some parts of our methodology. In almost any scenario, the numbers dwarf current payments made. Publishers are getting far less than our findings. So you could say, okay I think publishers should only get a 30% split. Or you could say well, maybe
20 only 20% are news searches. Either way, you will still come up with hundreds of millions or billions of dollars that are owed to news publishers. And if there were a law that would require disclosure of data to pin down the numbers then we would be able to do something even more accurate. So, again, what we're talking about is there are surplus value generated by the platforms and the publishers and this

needs to be share fairly. Historically, many licensing agreements with news outlets and distributors accounted for a 50% split. For the digital platforms, news media provides a source of high quality content to keep users engaged and coming back to the platforms. For news publishers, the platforms provide popular, easy to use channels through which they can distribute news content to a larger audience. So, there is some benefit to publishers of course, but the market power of the platforms means they under pay. And I can talk more about that later today or on camera. Here's some examples. How

10 Singer calculated that Australia's \$134 million was 4% of Australian Google Search revenues. 4% in the US would be 3.2 billion, 4% for California where they've been discussing a similar law to Australia would be 0.5. Canada settled and in \$74 million, Singer found that's 1.7% of Canadian Google Search revenues which would be 1.4 billion for the US, 1.7% for California. So we can really see through this that publishers in Australia got a larger share of advert revenues from Google than Canada did. Here are some of the questions that we get. One, how about the subscription revenue that's already been received by publishers, where does that fit in? And the answer is if you know

20 that number you can absolutely offset it from the amount of revenues earned by the platforms. We can also talk more later about what happens if platforms reduce news content as they've already been doing and that depends on a lot of things, you know, substitution patterns, the importance of the market so they're more likely a reduced content, in news content to a market that's not important for

them. They may do it for a short time or a long time, Google cut off news in Spain for seven years. In Australia, Facebook reduced news just for a few days. I think it was August 2020. And the really important thing to remember, even if you don't like these laws, is that deals are already being signed with publishers. Google has gone around the world and I'll talk about this in more detail, signing agreements with large publishers and they're just doing it completely opaquely without any transparency and people who sign those deals have to sign nondisclosures. Where we can also talk, I'm going to finish up talking
10 a little bit about generative AI and how we calculate what news is worth to large language models. We would say that there's even a stronger case for payments to be made for large language model results related to current affairs, but also that the payments should be given for legacy information generated by news. I think there's a case with all of this for reparations or rather for back pay, but I don't know of any government that's trying to collect that. So, what are other governments trying to do? I think I've made clearly the point that there's a huge market imbalance, that governments and publishers are powerless and the only way to extract payments from Google and
20 maybe Meta, is to talk about legislation. This has concentrate the mind and has led to further agreements with publishers all over the world. In the US we have a couple of bills right now. 886 is similar to the Australian Bill and that's just had hearings. In the US we've had something called the Journalism Competition and Preservation Act which has been proposed a couple of times by bipartisan senators.

Most recently, in 2023 by Amy Klobuchar and Republican, John Kennedy, from Louisiana. Both Australia and Canada passed laws which pushed Meta and Google to negotiate. We've seen the amounts of money they got. In Indonesia, in February 2024, they issued a decree and details are still being worked out. I'll be speaking with the press council probably this week. And in Europe, James Hodge will remember that after Australia passed their law, we gathered a Columbia on Zoom a number of competition authorities from all over the world including South Africa. The Europeans said they were less likely to do what Australia had done because they were busy joining the transposition of the EU copyright directive. And that has led to Google settling one legal case, at least, in France. Axel Springer, I think, has been trying for more than 12 years to get money from Google. That's been very difficult. I think they've finally started to get some payouts. And as we've seen in some countries, publishers are finally suing Google. But most of all, Google is going and making settlements with publishers around the world and these are covered by NDA's. A couple of things that are really important about the Australian law, first of all, they were the first government that said you know, this is a competition problem. We're not looking at copyright and we're not going to look at licensing. The problem is that social media are so powerful that any negotiations for payment would yield unfair outcomes. So the first thing that the law did was require Google and Facebook to negotiate with publishers. And then, if they can't come to an agreement, they go to arbitration with a government

appointed arbitrator. And this is very important, each side has to bring their final offer. And this is to avoid a situation where you have Google with millions of lawyers saying to the New York Times, we owe you \$10 million. New York Times says, you owe us 100 million and then they spend the next ten years in court and Google wins because they have more lawyers. So, you come in with your last best offer and then the mediator decides which one is the one, that's it. And because of this, nobody has actually gone to public arbitration in Australia, they've all negotiated privately. And again, because of these laws,

10 Google has started increasing payments to publishers. Even last summer in 2023 and then Google settled copyright disputes in France. The Australian treasury did an evaluation of the codes first year of operation and they found that the code was a success to date and that more than 30 commercial agreements between digital platforms, Google and Meta, and Australian news businesses have been struck. I think the number may be even higher now because some of the smaller outlets got together and negotiated as a unit and they got money as well. My understanding is the smaller outlets got more money per journalist I think than the larger ones, but these were all

20 covered by NDAs. Facebook and Google opposed to the news media bargaining code. As I mentioned, Facebook temporarily blocked Australian news in 2021, but however, they have paid out, Facebook, more than 200 million Australian per year and this has created jobs to fund local journalism. There are a bunch of scholars in Australia who follow the [won? 00:33:09] and have found that job advertisements

have been going. Anecdotally, when I was there, I think it was two years ago, a year and a half ago, the journalism professors I met told me that their students were all getting jobs which was unusual. But definitely there's concern in Australia and the world, Bill Bruskin has written about this, that there's a lack of transparency in how the producers are being compensated and concern over information and financial symmetries. Canada passed C18 in 2023 and Facebook cut off news. Indonesia has just passed their regulation and we expect that the US JCPA has stalled. So platform response. Number one, 10 they have been lobbying governments vociferously. I sat next to a European official who told me that Facebook had been there December the year earlier and said we're going to pull Italian news. I think this has happened in many, many countries. Meta is now threatening to pull out of news in Illinois if the Journalism Preservation Act is enacted. I assume this means that they won't distribute Illinois news to Illinois audiences, but I don't know the details. They cut Canadian news in 2023 and in Spain in 2014, Spanish publishers got together and said let's put on a levy on Google and then we're going to divide it amongst Spanish outlets and include the smaller ones as 20 well as the bigger ones. And then Google actually dropped Spanish news for seven years until EU copyright directive. There's lots of examples of the kinds of things that they're doing. You may have seen that in Brazil last May, when Brazil was looking at their platform remuneration law and discussing it, Google put something onto search, so if you went to Google search, even if you were just looking

for, I don't know, what time a supermarket was opening, Google put up a little message saying this new law is going to destroy the internet. And when I wrote about that, they called me up and very high handedly insisted on a correction because they didn't like the word I used to describe what they had done. They said it's legitimate campaigning. And then the other thing that's going on is they're making private deals with publishers all over the world and I'm getting a sense of the funding. What's basically happening is if you're a large US outlet and you go to Google and you say you know, really we think
10 you owe us about \$100 million, Google says we'll give you \$10 million and then they walk away. If you go to Google and say you know, we really think you owe us \$10 million, Google will say we're giving you \$1 million, take it or leave it. And if you're a small outlet in a small, low income country, Google is saying, okay we'll give you \$20 thousand a year. So the publishers are feeling very, very frustrated and they don't feel like it's a real negotiation. So, you know, do you want to legislate? Do you want to impose a platform remuneration policy? This is a question for South Africa to decide. The evidence suggests that
20 planning a law will lead to increased payouts for publishers, but these are powerful companies that don't believe that they actually owe money. Intellectual property rules are social constructions designed to provide incentives for production of knowledge. The breath of fair use and what exceptions you're willing to make is a social decision. I would say personally, 20 years ago, I was a huge believer in a very broad definition of fair use and intellectual property. It seemed to me

that it was essential that, for example, African scientists and African students should be able to read western textbooks without having to pay a lot. In my day, fair use meant somebody sending an email and saying, could I photocopy a page of your textbook or your book to use in my class? That is what we thought of as fair use. What's happened now is completely different which is that these enormous outlets are sucking up all kinds of information and profiting off it. So I'm not sure, I don't believe any more that this is fair use. I think they're hiding behind fair use. And I think that unfortunately, this huge imbalance of market power, suggests that some form of government intervention is required to make sure that social media pay appropriate amounts for their use of information and knowledge which they have not created. So I think that our methodology provides a persuasive framework for determining the relevant amounts. Now remember that the platforms have a very strong financial interest in believing what they believe and they have an ideological interest in believing what they believe. They get angry and emotional when we suggest otherwise. I don't have a lot of contact with some of these folks in Silicon Valley, but when I do meet them you know, sometimes I get shouted at. They feel very, very strongly. They're not working in the spirit of collaboration. Their view is you're barking up the wrong tree, I don't know why you think we owe money and they, you know they say this in a very strong way. I got an email recently from someone apologising for shouting at me. But the other problem is okay, they don't want to pay for intellectual property that they're profiting from but they don't really want to pay

taxes either. This isn't about taxes, this hearing, but if you look at you know, the history of Google and Facebook and Amazon and all of these companies, you know, they've been quite famous for all kinds of things like moving headquarters to Ireland or Luxembourg or places or Singapore where they don't have to pay taxes. So there's you know, that makes them less credible. I think they just don't feel like paying money out. Having said that, I'm sure you're well aware that we drafted some principles for these platform remuneration codes at Columbia University and then we improved upon them and ratified

10 them at a meeting held last July 2023 at GIBS. And Michael Markowitz, I'm sure, has talked about it. I think about 200 people signed them, but you know, we certainly think that the next set of laws can improve upon Australia's law. We'd like to see more transparency about who is getting what. We would like guarantees that the money will go to journalism. We're not interested in laws that just give CEO's or hedge funds which own outlets, more money. And we'd also like to see equity. We'd like to see smaller outlets get funding as well as the big ones. Again, for all of you who don't like these laws, remember the reality, Google is paying publishers around the world and requiring

20 nondisclosure agreements and interviews that we've done in the summer of 2023, suggest that the payments have increased because of the threat of laws. Google and Meta are spending money on lobbies to fight this and they're saying that news isn't important and they don't care about news and nobody wants news. Well then why are they fighting so hard to avoid paying? I asked somebody from Google

about this, off the record, recently and he said, well it's a precedent. We don't want to pay now because then we'll have to pay all over the place. So, one of the questions I was asked is are there alternatives to codes? And so I did a bit of research and I found dozens of ideas that have come up really in the last five years, probably older, for taxes to pay for journalism. So you tax the big tech companies and you pay for journalism. My renowned colleague, Victor Pickard, at Annenberg in 2018, said let's tax 1% on the revenue of Google and Facebook, put the money into a fund and seed an endowment for independent
10 journalism. 2019, the non-profit media advocacy group, Free Press, suggested 2% tax on digital advertising. Donor funded non-profit digital advocacy group, Public Knowledge, proposed the tech companies that conduct content moderation, pay a federal user fee and put that into a journalism superfund. In Australia, Andrew Jaspen, put forward a proposal for a levy on big tech that would be administered by an independent agency. Nobel laureate, Paul Rohmer, in 2019 proposed a progressive tax on digital advertising. Reports Without Borders has called for European tax on the tech
20 giants that would be levied on the turnover they generate in the EU regardless of where their headquarters are physically located. The tax revenues would fund journalism in Europe and redistribute part of the value being captured by the largest digital platforms to the detriment of European media. So there's been many ideas for taxes to fund journalism. In my understanding, not one has happened anywhere in the world. So that suggests that either the governments aren't making

them happen or that they're being fought by the platforms. I won't go too much into the US case, although we're doing a study right now on tech taxes for Steve Waldman and the coalition to Rebuild Local News. I think Maryland passed a law and they're now being sued by something like 20 different companies. There's other ideas as well. Tax credits to support the hiring of journalists, giving vouchers to citizens to pay for subscriptions, creating journalism funds. New Jersey has one, many European countries have them. Scholarships for young journalists so they can go work in news outlets.

10 Governments and philanthropies have successfully backed many approaches and it isn't just Canada and Scandinavia, but the economist research unit did a study for UNESCO a couple of years ago and found examples in Indonesia, South Africa, Tunisia. We've got states that are doing programs to support journalism. We've got a federal bill for tax credit in the US or an employment credit, just like in Canada. And some places get VAT reductions as well. So you've got a lot of different ideas out there. However, as I mentioned, so far it's been, you haven't had successful examples of collecting taxes from big tech platforms in order to pay for news. Another question I was

20 asked is, what will happen if the platforms exit South Africa? Now, companies always do that right? They always say if you tax us we're going to leave. We saw in the pandemic, plenty of businesses said we don't like being in New York City. We don't want to pay the taxes here, it's ridiculous. Individuals said we're moving to Florida where there's no personal income tax. So you have to ask yourself, you know, will

Meta and Google leave? My understanding is that Google needs news and is not going to exit news. This is through informal conversations, but they've also said, very clearly, they don't want to pay for it. Now, one thing you have to assess as policy makers is you know, how likely is it that they will leave? And remember that even if they had to do platform remuneration, they would still be left with a very large part of the surplus that is created. So there isn't really an incentive for them to leave except as a lesson for other countries. And then the other thing is will higher payments for news cause reductions in news contents on platforms or you know, something like TikTok right, which focuses on entertainment. It seems that with [indistinct 00:44:39] it would, but we do not know. And then you have to think about the consequences. Let's say Facebook left South Africa [indistinct 00:44:47] a vacuum, would this be a chance for African technology? Would there become an Africa Facebook could fill the void? Developing domestic technology firms would have long term benefits. Again, these are policy questions and questions your government needs to decide. Yes, some more thoughts. So again, Google has made it clear through their words and their deeds that they need news, will profit from news and they will not pay for news. Meta, yes, has left. Brazil and Indonesia have talked about a must carry provision so that almost turns these platforms into public utilities right? You're saying you cannot retaliate. Some laws have said you cannot retaliate and others have said you have to carry news. You cannot drop news just because you have to pay for it. On the panel on the right, I just

have again some clarifying questions. Business has always threatened to leave and then you have to decide, as the government, do you want to succumb to these threats? And then you have to think about these threats. Do they lack credibility? What are the consequences? Would the consequences even be positive? Now, another thing, I'm going to bring us into the open AI question. Open AI is a huge, existential threat to Google. Generative AI and the large language models will become even better search engines than Google. And the piece that Jim Albrecht wrote for the Washington
10 Post, which I referred to recently, is fascinating because what he's saying is that you know, Google was great, they gave loads of traffic to publishers, everybody loves traffic, but what they want is for open AI to give money to Google and to give them eyeballs. So they want Open AI to return the favour by sending them traffic. That's what I mean by giving money. So, it's quite interesting and inconsistent that they're now asking for money. Google made most of the money off the traffic they got, a lot of it that came from news so that journalists, audiences would not have to go to the publishers and now Google is saying okay, well we want traffic too. So here we are with Google
20 facing all kinds of trouble and that's making them less reluctant to pay out. I know there was some discussion last week about do we need quality journalism anyway? I think that Rasmus Kleis Nielsen talked a little bit about elite news and not designed for the people. People aren't willing to pay for it and I just want to remind everybody of the Windhoek +30 Declaration of Information as a public good which

reinforced the importance of information as the public good. And since that time, social scientists have used new, rigorous research methods including natural and randomised experiments to confirm the beneficial effects of journalism in the global south as well as in the north. But as a public good, quality journalism is subject to classical free rider problems as the economists call it. And so it's unlikely to be sufficiently provided in the absence of support from government and donors. And here's a little bit about the scholarship. So, we have all kinds of studies, Julia Cage, the leading media economist showed the
10 press coverage reduces rent extraction in foreign aid. In India we have studies that show voters revised their vote choice for local government in India in response to newspaper reports about spending and corruption, that the prospect of newspaper report cards on Canada performance makes Canada, it's more likely to focus on pro poor spending. They also respond to media coverage by supporting high performing. Sierra Leone, Katherine Casey found local radio coverages associated with voters being less susceptible to voting along ethnic lines and links reductions in ethnic voting to more equitable allocation of campaign resources. 2019, they used, some
20 scholars used a randomised control trial to show the public debates between political candidates, build voter knowledge and interest and link this increase in citizen engagement to increased investment in development constituency engagement campaign effort. I think I know what country this is in but I'm not going to say because I'm not sure. Recent research in Italy and the US and Rwanda and Russia suggests

that the spread of low information, highly polarised and propagandic news sources can undermine many of the benefits just listed. So I think we know that news is incredibly important and what is the counterfactual? If we had no quality news, where would we be? So, I'm going to move now to the last part of my presentation before we take questions and I wanted to say that large companies, like Open AI are going to need to pay for content. And this creates an opportunity for collective negotiation with publishers. So you know, we wish, many of us, that publishers had begun to negotiate collectively

10 with Google and Meta. They would have developed the muscle to do that. To develop a really good, large language model, I was told in August, so I'm maybe out of date, but in August I was told it cost \$100 million and that there were only seven good ones. And so the AI companies were in an arms race for reliable content. They're in more of a hurry for once than the journalism organisations and, from what I was told, is that if you're producing content in Spanish or Portuguese, you're in a good position to negotiate, and I'm going to talk more about other things that journalists can sell. Again, we have a vast lack of transparency. We don't know what's being hoovered up by all of these

20 large language models. I'll talk a little bit more about the New York Times lawsuit against Open AI. The French Conseil national du numérique is already proposing a third party platform where data of that large language models can be stored to understand and evaluate the relative importance of information, it's quality and variety, develop tools, establish metrics. One thing that's important for publishers to

understand is a few things. One, open AI is going to pay for content, they're already doing it. And what they say, the bigger [indistinct 00:51:07] companies is that the small open source companies have absolutely no compunction and they will steal content and it will be hard to do anything about it. Right now you cannot figure out easily whether your content has been used and that's a huge problem. Paywall does not protect news content all because they can still crawl and what you need to do is put up terms of use that need, say that there's a prohibition on using content for training, but in many places

10 the damage has been done. The content has been taken to train the large language models. How are people dealing with this? In the US, publishers are trying to negotiate and then they're suing. So the New York Times has a lawsuit against Open AI. You can read the lawsuit, it's online. They found examples where their articles were copied or copied and barely changed in the large language models. And they found examples where the large language models said the New York Times have printed things that the New York Times had not printed. Authors are also finding their writing is used in large language models. So one of the outstanding questions is how do you value this content

20 when the LLM companies say, oh we're losing money, we're in startup mode and also liability and defamation. Who's responsible if there's defamation? Who's responsible if there's defamation? Or, for example, the New York Times can say you're destroying our brand by saying things, by the large language model returning results attributed to the New York Times when we hadn't said them. Can you sue? Barry

Media and Digital Platforms Market Inquiry

Diller, who's the chairman of a massive media holding company in the ISC, has not filed his lawsuit yet, but he was trying to lead publishers in legal action to protect their work from being stolen by AI companies and used in LLM's. He wants a complete change to fair use and changes to the US copyright law. He believes this litigation is the only solution and it may encourage legislation that will help publishers deal with generative AI. He supports collective action and he thinks it would be an opportunity for smaller outlets to join a larger, more powerful group and gain protections. The reality, however, is that separate
10 deals are being made. If you're interested in his views, he had an interview in August with Kara Swisher on her podcast, which I recommend, and in preparation for the hearings, I got in touch with his general counsel's office through an intermediary and I'm trying to arrange a meeting, but I've learned that so far they have not filed any cases. I also just wrote an article that got published on Friday where I interviewed a lot of people involved in Open AI negotiations. And what I'm hearing is that it's turning into very much like what happened with Google and Meta. I'm hearing that the Open AI companies, the large ones are going to give low millions to the large outlets, hundreds
20 of thousands of dollars US to the medium outlets and small outlets to low and middle income countries, probably about \$20 thousand a year. As you can imagine, the US companies Open AI, they're calling for legislation are being accused that they to pull up the ladder behind them, that legislation and guard rails will create oligopolies, big US companies. It will squeeze out the open source language models and

the smaller competitors. So the European regulators and smaller firms are worried that regulation in the US will just freeze out the competition. What Open AI is doing now is nailing down non-exclusive agreements with large, reputable outlets in the US and in Europe. So you've probably read About Axel Springer, Associated Press and they're postponing talks with small outlets and outlets in the rest of the world. I contacted Open AI, they declined to comment. What I'm hearing is that Open AI is not returning phone calls of news publishers in South America, for example. If anybody in South Africa has made
10 a deal, I'd love to hear about it and one of the people I know who's been talking with many of these, involved in these discussions, David Klinch, has said that agreements being made now are setting the table for the mechanisms and metrics for later deals to be made with local outlets. So, what's valuable? What do publishers have that they can actually sell to the large language models? Original images and video are valuable but they will be made more valuable if their work can be authenticated and it's provenance being known. As you know, there's several alliances and standards right now, it might be better if there were one. True Pic is one of them. Many of the newspapers are
20 signing up for this and there's definitely a belief amongst some of these organisations that governments are going to eventually pass some sort of watermarking or content provenance mandate. And even if they don't, one big company like TikTok or Meta, which I think Meta's already done for political advertising, then others will follow. So one company goes ahead, that will become the industry standard. It's

very, it's harder to make watermarking for video because video has more data than images, but when I spoke to Klinch, he said that it can be done and he pointed to You Tube has a content ID that matches rights owned videos with newly loaded videos. The hope is that once the content is clearly marked, more money will flow from advertising and licensing deals. Historical archives. These can be used for training models and so access to them can be sold by the publishers who have them. And these will inevitably be the larger and older publishers so that may not help some of the South African outlets as
10 much. And also valuables current content that comes with journalists who can answer questions about the reporting process and how they got the news. Another key thing to remember is that, for these large language models to be effective, they also have to be able to search current information right? So they have to be a combination of like an encyclopaedia plus a Google and this is why it's so threatening to Google. So someone said to me, they might not care normally about Indiana or Indonesia, but if something happens you know, there's an earthquake or a fire or some political news, all of a sudden they're going to need a lot of news from those places. So they're going to
20 need current information as well as a complete archive. I view this as like a stop option. You don't know that you're going to need to buy news from Indonesia, but if you do you want to have the right. Another question is writing quality. I asked does writing quality still matter? Like do we care about the New Yorker being beautifully written or The Economist being beautifully written? They said, right now it matters,

it's something you can sell. It might not matter as much later. Just like news outlets optimise for search, so they may start optimising for AI prompts and that could change completely how they present information. I already talked about Japan, Spain and Portugal. I was just in Iceland a couple of weeks ago presenting our research to the government and while I was there I stopped by and spent a couple of hours with the large language model people in Iceland. And they just said, you know what, we're a small language and we need to be in Open AI's large language model. So we're just going to give them our

10 information and hope that that will make Open AI large language model in general much more accurate. Another thing that publishers are talking about is they don't want it to be just a licensing but almost like a syndication. And I don't know if they're going to get that, but you know, if Reuters gets its news or the BBC gets its news in a large language model, they don't want to be just buried in this big, big model where no one can tell what's their content. They want guard rails, they want attribution, they want prominence and they want context. They don't want some extreme quote with some horrible person they interviewed, being taken out of context and thrown in to the results

20 from the LLM. So, that's really important. They're trying to think about it more like a syndication. When you see a syndicated piece of content online, I mean on a newspaper, you know this is AP, they reported it from Duluth and this is the name of the author. Again, reliable content is critical because if the large language models don't have reliable content, then they start training on itself and then it gets worse and

worse and worse, right? So training on unreliable content means that it's going to be a terrible model. As I mentioned, unique content is useful. And this is going to be a problem for small outlets that don't have a lot of original content. Many of the local newspapers in the US have been starved for funds and rely on wire service stories. If you take that out, there's just not a lot left to sell to the general AI companies. And then finally, contextualized information. So, it's not just about the, you know, finding the bus schedule and giving it to a user when you do a search. It's also about context like, hey, was there
10 a snowstorm? Did the mayor warn people in the town that the 3pm bus service might get cancelled? So, there's a lot of things there that publishers are going to have to think about as they negotiate with the large language models. I'd like to thank you for your time. I've come in just exactly in an hour. I co-authored this paper with three other people. The lead economist was Haaris Mateen, and he's going to be on the call for taking questions. Just want to say again that Harris and I did not get any money for our research and the Brattle Company received a fifty thousand fee from Omidyar Network to pay for some of their costs. So that is my presentation. You're welcome to keep it.
20 And I'll be glad to submit written testimony as soon as there is a deadline. Thank you very much. I will mute right now. I don't know if we're taking a break or going straight to questions.

MS. PAULA FRAY: Thank you very much, Anya. And that's just an incredible body of work. And we're really thankful that you've put it together for us. I wonder, though, if we could step back just for a little

bit. I mean, you argue that 2% figure on the news queries raises the question on how to define news. So, I'm wondering how you think we should define it for our purposes. How should news be defined?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: I would let Haaris take that one.

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Hi. So, the way to define it would technically be how many of news queries lead to search results where the primary source of impressions for users is news content, right? So, what you want to see is that if I'm searching for, as the example states in our paper, Ukraine, and I spend my time, my engagement, or as we like to measure it, impressions on news snippets and news articles related to Ukraine rather than, say, a Wikipedia article or something else, that would be the fraction of search queries that are related to news. That would be a much more accurate method. And it goes to what Anya already presented and said about having data from Google and Facebook or Meta exactly about the amount of engagement that users have with news content while they're using the platform.

MS. PAULA FRAY: Welcome, Haaris. Good to see you. I wonder if you could perhaps just give us your definition then of what a news story would be. So, I mean, I know that how you define it when you're doing, when you're looking at in terms of searches, but what would you define as news, is weather news, for example?

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Yes. So, just to be

very precise, a news story is any story that is being published by publishers that involves some kind of journalistic input. So, this includes politics, but it also includes entertainment, sports, weather news, practically anything that, you know, good health and gardening news, anything that is being produced by publishers that involves journalistic input is defined as news content. That does involve production costs, journalist costs, and is typically timely, credible and
10 constantly updated. So, yeah, it's a very broad definition. It doesn't just stop with, let's say, politics or yeah.

CHAIRPERSON: And you propose search results that are primarily news. So, how would you define primarily?

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Primarily with respect

to when I was talking about the measurement of the of the news content, is that what you're referring to?

CHAIRPERSON: Yeah, correct. So, I mean, if we took your measure of news, which is anything published by publishers, so we could use
20 their URLs for that, their main domain and in the search results, are you saying where maybe more than 50 percent return had content from publishers, then that's a news related search, even if it is on entertainment or health or, and that would be the basis, you would say that we would measure. So primarily, is that sort of more than 50 percent to the impressions?

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND

PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: So, from the data that

we have, 50 percent of searches that take place on, say, Google Search are information related searches, and out of those 70 percent of searches are for information that comes from news content. Right.

So, I would be a little more conservative rather than saying 50 percent, I'd say about 35 percent of search demand is from users that are looking for news content.

CHAIRPERSON: Yeah, I think the question is more because I think

10 as Paula said, because I mean, we have the powers to ask for information. So, we're hopefully in a better position than you may be.

But just looking for what information you think we should ask for. And you said that search results that primarily show news results would be what you would say is how you define the news queries. So, I assume it's not just showing one news outlets result and then a whole lot of books or other things. You mean it's a search result that sends back impressions where maybe more than 50 percent are on news content.

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND

PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: I see. I see. OK. So, I

20 better understand your question. Yes. So, there are two parts to it, right? If I end up in a position to ask Google the first question that I would like, you know, would like to see as a researcher would be exactly when out of all the search results that are generated, how many of the search results are primarily news related in the sense of displaying news related search results. Let's say on the top half of our

screen. Right. So, if it's 50 percent or more is a nice benchmark. That's one. That's just the supply of news on the search result page. That's one. So, that's an important thing to measure. The other thing is that, out of all the search results that are there, what is the fraction of impressions or user engagement with news content? And that can be measured. So, if user impression is basically the idea that how much time is a user using or spending consuming content that is displayed. So, the other way, so, the first way that we discussed primarily news content is kind of what we like to call an extensive margin, which is just the number of results that are displayed that are 10 related to news. And let's say it's more than 50 percent that would be classified as news related. The other is that whatever results are being displayed, it might be a mix of news content and not news content, how much of user engagement is on the news content as a fraction of the total? So that's another thing that all platforms measure. And it should be available if they're willing to share.

CHAIRPERSON: Yeah, no, that helps. And just to wrap it up, I mean, I understand from the methodology that for Facebook, what's being used is the data that was used in the 2020 American Economic 20 Review article. And that the Columbia University team then interrogated that to see how much was news or engagement on news. It's not clear from your paper how you interrogated Facebook, how you define news or engagement with news. I don't know if you can just unpack that for us from a Facebook perspective as it's different to Google.

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND

PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Yes, so the

methodology for the Facebook study is based on, as you stated, on a journal peer-reviewed article called The Welfare Effects of Social Media published in the AER. The nice thing about the paper is that it gives you raw data for the study that they conducted. The study that they conducted was essentially looking at the consumption that people have in terms of time usage on various kinds of content. News content is again defined by content that is being consumed from news publishers. The way we get to that information is we get the raw data which is freely available, publicly available. The raw data has the time that people, each individual anonymized obviously, but each individual is spending on Facebook which is about 48 minutes. And then for each individual you can also see the amount of time that they spend consuming news content right, and that fraction of time that you spend consuming news content when averaged across all users is 13 percent. So, it's very transparent and well defined in the data that's available. Other kinds of inferences can also be made. So, yeah, it is available on the website of the journal.

10
20 CHAIRPERSON: I mean so that's helpful. It can sort of push us in the right direction. Sorry Anya.

DR. ANYA SCHIFFRIN- SCHOOL OF INTERNATIONAL AND

PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Not at all. I would add that we can send you a list of questions that you can ask for, I mean there's a lot of very broad questions you can ask including what are

their own studies that they've done and have they replicated other studies such as the one done by the fair group and what were their findings. Further you can look at for example, as we've done, the number of news related queries in Alexa and that is 20 percent and includes finance and sports. so, there are other questions that you can certainly ask and other places where you can try to get comparable data.

CHAIRPERSON: I mean as I understand your google, there's an argument about direct value and indirect value. The indirect value the
10 credible news or credible content also obviously the driver towards Google as an overall platform. For Facebook it seems it's just share of almost the feed and attention. Do you think there's additional value that comes to Facebook apart from just the attention on news?

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: I can take this. So, yes there is. Arguably there are benefits of having news content as a source of engagement but relative to Google's positioning as a platform versus Facebook, I think the relationship for Google is much more clear because Google is primarily a search for, is a platform
20 where you search for things. It can be commercial; it can be information related and therefore the value that you have from having a ready source of rich and credible information coming from news publishers is extremely valuable which can be measured in many ways. I think with Facebook the question then becomes that Facebook is used as a platform for social interaction and social interaction

combines many forms of content not only news but also related to what people in your friend's circle and community are doing and therefore it becomes less clear, at least as researchers to be able to measure if there is an additional benefit of news content that comes from this indirect route as compared to the others. So, I think as researchers, we took the conservative position that it's much more difficult to attribute the indirect benefits that Facebook has from news content. Although I would say that the same set of arguments do follow in the sense that if I know that Facebook, my social media
10 platform, Facebook, has news content coming in from very credible sources, it makes me more likely to be there because the quality of engagement, not only with news, but with my friends becomes much higher. So, there is definitely a sense in which the way we measure it, the value that news generates for Facebook is a lower bound, a conservative lower bound. But I think from a very rigorous research standpoint, it makes sense because the differentiation is less with respect to Google in this case.

CHAIRPERSON: I just want to ask a question about Google before coming back to Facebook. I think Google's criticism has been that,
20 well, if everything is worth more than its share of impressions, then the sum exceeds the total substantially because you could say the same about other forms of content that's placed on Google's platform, that its value is not just its contribution to the overall advertising revenue of Google, but if it has some additional credibility and added value, then all the parts would sum to more than the whole. I don't

know if you've got a response to that criticism.

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Oh, that's actually

amusing because, so the way it works is that the value add that comes from news content is important because arguably the value, the information that you get from news content is objective and not driven by commercial interests. It is credible in the sense that there are independent ways of assessing the credibility of these information

sources. The easiest being that this comes from a big publisher, but
10 others as well. It's timely, it's driven by its own set of production costs.

And therefore, it sort of builds a sense of habit and trust that people have on the product, the Google Search platform itself, coming from all of these very much more objective criteria of quality, right? This does not in any way, and I think I can say that with a lot of confidence,

extend to all kinds of search results. If you were searching, if I was searching for shoes or, you know, or any other commercial product, I will get a set of results. There is no sense in which the advertised products that I receive versus other products that I receive are

enhancing the trust that I have for other kinds of products that are
20 there on the platform. So, this externality thing is not available for everything in the same way. Yes, if any search engine provides good

results in general, it enhances the value that I have for the entire platform as a whole. There's no illogical inconsistency in this, right?

There are many products which have individual value, which when summed up, enhance the value of the total product. So that

externality, the positive externality thing can be there without any kind of inconsistency. But to compare the value that something like news content has, and its positive externality on the trust and habit and the value that it gives to Google with other kinds of search is, I think stretching the logic too much. I think that there is a fair case to be made that a big differentiator for example between Google Search and Amazon is that Google Search gives me a lot more information. Amazon also allows me to search for products that are sold through Amazon, sure, but Amazon is a huge platform. I think the primary
10 differentiator between Google and Amazon is that I know that Google also provides me information that comes from these outside sources that do not necessarily have a direct or as stark an objective of selling things as say other commercial products have. So, I don't think there's good logic in that. First answer is that there can exist externalities operating from each part of the business that is not counted in direct value that adds to the indirect value for the entire system. That's just, there's no inconsistency in that. The second point is that there are clear differences in the quality, credibility, and objectivity of news content versus other kinds of content on the Search platform that
20 makes the economics of the case very clear that one side is contributing a lot more to the overall value.

CHAIRPERSON: Anya, I just wanted to pivot back to you on an important question we're tackling and maybe some of this context has given it but should we not be thinking about Search and social media differently because this lumping them together continually when Meta

is clearly behaving differently to google where the value may be more difficult to determine at least on an indirect effect, should we be starting to think about these two very separately and just distinguish them in our analysis and even if we were to get to a remedy at any point, even the remedies may be different.

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY:

Well, I think that the advantage to coming later to this compared to Canada and Australia is, you can also think about whether you want to include other
10 platforms and what you want to think about generative AI. So, I think when you're designing policy, there's kind of two questions right? One is what in an ideal world absence, you know the counterfactual, absence their position, their opacity, their power would you be able to
um calculate as the worth and then the second one is what can you realistically collect or ask for and how to do that. So, I think you've got a lot to consider as you think about sharing revenue from these firms. I assume that they've made some proposals and so another question is what they have expressed you know, a willingness to do. So, I think it's a whole lot of things that you're going to have to take into account
20 as you design remedies. Yeah, that's what I think. Haaris, I don't know if you have a thought specifically on google versus Meta.

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY:

No, I think I broadly agree with you. I think the use of content by all these platforms is necessary to be accounted for because and having a broad

framework is important because people's habits change. For example, we know that a lot more news is now consumed on Instagram and even more on TikTok. Now the quality and how it's being consumed changes but having a broader framework including all of these sources is important.

CHAIRPERSON: Yeah, and that we've certainly done. And we've also included the broadcasters as well in AI, but just fundamentally I suppose Meta will say, look google is there to offer you the correct answer. It crawls the web, it fetches your content and displays it, but
10 on social media you post. So, the news media will make posts. It's not as though Facebook has gone and fetched that. So, does this difference not warrant thinking about the two differently?

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: It certainly does, and we account for that difference in our study. There is there is a clear difference between providing news content directly as a search result which enhances your value versus people posting and people also consuming from content that's already been posted but I think the important thing here to understand is that that's the design of the
20 platform. The platform is so designed that people exchange information. People exchange content and people link to content that comes from news sources that enhances value not only for the people, but it enhances value for Facebook. So, if Facebook was not benefiting at all from the presence of news content on their platform this argument would make perfect sense to me. It would be that you

know, we can't do anything. These people are putting up these new stories and that's about it, but what's really happening is that on the platform people are sharing news stories. These news stories are typically linked to publishers or their discussions of these published articles. There is a discussion that happens on Facebook because of that, like that leads to higher engagement and then Facebook benefits from that engagement, right? So, the fact that produced content again, journalistic produced content again is producing value for Facebook means that that value should be shared because the counterfactual is, if there were no such news publishers present then Facebook would not be benefiting from this kind of engagement. And therefore, there is one side that is bearing the cost of producing this content and there's another side that's benefiting from the engagement that's being produced by it with very little that's going to the other side through click-throughs. So, I think even if people sharing and we take note of this in our study that this is a different method from google search, there is value being generated that the platform is benefiting from which [indistinct 01:25:47 – 01:25:54]

CHAIRPERSON: I think we lost you for a moment there Haaris but it's okay. So, Anya, I just want to come back to you and just, I mean Facebook's behaviour of cutting off news, I mean does it not challenge the whole notion that they do in fact benefit and they would I think they released at least a statement in the last sort of few days saying yeah in Canada they've in fact increased numbers. The engagement hasn't deteriorated, so where is this evidence that they value this so much

and that they benefit so much? So, does that not fundamentally challenge that notion, or how do you see it?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: A couple of things.

One is, remember they've been getting out of news for a long time, so yes, I think, and I can't speculate as to why that is, whether they don't want to pay for it, if that's the main reason, or if there's other reasons. The second thing is when they say increased numbers, you always have to ask for a whole lot of other information that you don't usually
10 get. It's like when they talk about, say for example, in the olden days with mis- and disinformation, all the platforms would say, well, we've taken down 20 million pieces of content, but they never said, compared to you know, the rise in consumers and the rise in use. So, they tend to just throw a lot of numbers out without giving actual context. So, when they say the number of users have gone up, how does that compare to other countries with news? Do you see what I mean? Like that number in and of itself doesn't really tell us that much. So that's the first thing. I think their numbers were always a bit confusing and misleading and sort of lacked context. Second of all,
20 I'm absolutely willing to believe that they're getting out of news. And then you also have to think to yourself, and these become sort of policy questions. It's like everything. Do we say, yikes, they're really huge, Facebook's actually pretty useful, therefore we're just going to give up? And those are policy questions that you have to sort of decide when you're dealing with really huge companies. I mean, the

other thing I want to say is that none of this is fun or good or nice. It would be much better if these weren't huge companies. It would be much better if you could negotiate with them fairly. It would be much better if they weren't playing tough. So those are all problems. And then you also have to think to yourself, to what extent will South Africa be a model for the rest of the region? Will other countries in Africa follow the footsteps of what the South African Competition Commission is doing? Are you a large block? Would they be able to just exit South Africa and then stay everywhere else? Or in fact, would
10 all of the other regulators say, wait a minute, we don't want to be pushed around and we want to collect? So yeah, these are tough choices. I can't give you an answer to tell you what to do. I'll say that the data that we have, Taylor Owens has been collecting data on what's happened in Canada. And it seems that the smaller outlets are being harder hit by the declines in traffic. I have a PhD student who's done a chapter, which he unfortunately hasn't published, where he looked at when Facebook suddenly downranked news in a whole lot of countries overnight without any warning. I think it was Bolivia, Cambodia, somewhere in Eastern Europe. And then he went back a
20 year later and he found that the publishers actually developed their own websites and search, and they were able to kind of overcome and ended up stronger a year later. So, yeah, they could pull, yes, that would be painful, and these how to deal with that are all policy questions, I would say. And there are some studies on what happens. We have some studies that have been done on what happened when

Google dropped Spanish news for seven years. So, there's definitely things you can look to get an idea. But as I talked about earlier, they're entrenched ideologically. And yeah, they're very entrenched for sure.

CHAIRPERSON: Thanks, Anya. And I think it would help us to get some of those studies because there are questions, I think that arise, which is, and a question one has naturally is, is no news on Facebook good news or bad news? Do people work around? Do people find alternative channels? And so, in the end, is it better or worse? But I think also the issue of missing disinformation that may rise and fill that
10 vacuum, what effect does that have? Because then is credible news still an important counter?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: I'll be happy to find you some of these studies. And then I think, as I mentioned, in Brazil and Indonesia, the laws include a must carry provision. So that's another option for South Africa to explore. Can you require these outlets to carry news? The First Amendment, that would be difficult, but that may be something along with the do not retaliate that you can insist on or include in your legislation.

20 CHAIRPERSON: I mean, just in terms of the remedy, and I think we've been careful to indicate that whilst there are these things happening elsewhere, as you rightly point out, one can learn and we also need to make findings in the South African context. And we have a different country, a different history and different world of news concentration in the media. There are many challenges. But Erasmus Nielsen, I

think you said you'd seen some of his inputs. But the one input he made was that are we not creating the wrong incentive? We're saying, in a sense, either make news more expensive, we will incentivize to pay. But then that incentivizes to show less. Maybe that is more difficult in the Google context, where people may do search queries. But it's clearly been possible in the case of Meta, and you mentioned the deprioritization of news. So, I don't know what your response to that comment is and this is why we were asking about is search in social media different? And should we look at it differently? But does
10 it create the wrong incentive, at least in social media, one that may be detrimental or not, depending on the view you take?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Yeah, I don't have enough insight into what Facebook is doing to know what incentives they respond to. But a must carry provision or a law that requires them to carry news may be a way of insisting, of requiring. Why they're pulling out of news, whether it's because they don't want to pay for it, whether it's because of something else, whether South Africa actually makes a difference, it's very hard to know. Is South Africa a big
20 enough market that they won't drop news? Are South African preferences different? These are all things that I don't have insight to because we simply don't have the data. I'm sure that they're having conversations with you all, and I would imagine that they're saying they're going to exit news, but I don't have particular insights into Facebook's thinking. I think all I can say is that threatening to leave is

exactly what you would expect from these big companies. So that's a threat that they're always going to make.

CHAIRPERSON: Yeah, we're very used to threats from companies about not investing or this and that.

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: In Canada, in Australia, they did come back, but overall, they're definitely lowering. And it'd be better to regulate now and force them now before they've already dropped news, right?

10 CHAIRPERSON: And that, I suppose, is getting back to the question of whether having no news on that platform is a loss or a win. I mean we had, on the very first day in fact, we had Khadijah Patel from the International Fund for Public Interests and Media and she raised an interesting point about – so they're looking at the fund as more of a short term stop gap, prevent di compli exodus a certain media, but they're also looking as part of their mandate as a longer term solution. And she raised the issue of is there not a – I'm going to call it a win, win situation or an alignment where incentives are aligned and maybe the precise share has not been exactly that correct, but where one
20 finds a solution that is not an intent to then to destroy the distribution of news which has this public good value as well. And I think a number of the media came and said, look we recognise that the platforms have also helped us reach a broader audience and that has also had some benefits so we should recognise that as well. And so we don't want to literally throw the baby out with the bath water type situation.

I mean have you thought about whether there are just these sort of win-win scenarios that could play out or are the only options in your head regulate, enforce?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: So I have been saying

for ages that I think these platform remuneration policies will not be a long term source of income for a news publisher. To me it's a very, very short window before they're destroyed by open AI or everything else. But what I have seen around the world is that they don't pay
10 until there's the threat regulation and then they start paying more.

CHAIRPERSON: So if in your head there are not long term, I mean I suppose we're grappling with that long term market structure, competitive dynamic and this is a competition inquiry, do you think there are other measures that provide that long term solution? And long term solution is more likely to be win, win because it some respects regulation is more tricky in that context?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: So a couple of things,

20 one is short term is actually really important right now, right. South Africa has outlets that are closing, many places have outlets that are closing, so I would not pooh, pooh anything that can help get some serious amounts of money in a short period of time. I think that's really important. And then long term as I mentioned, sure it would be great to have some kind of user levy that is not cast onto consumers or some kind of tax, but the problem is that so far they've been unwilling

to pay any of these things. So that is why I don't know what that thing would be that would actually work. Everybody has tried everything, people have gone to [indistinct 01:38:19] Facebook over the years with all sorts of suggestions and ideas and they have not wanted to pay for any of it. So I'm not sure what – I'd love to create incentives for everybody but I'm just not sure what those would be.

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Can I also add that in

the very fast changing market landscape this short term, even if it's
10 short term regulatory intervention would ensure that the terms of the debate are modified in the sense that the idea that platforms do owe something to news publishers based on rigorous foundations, is something that by itself would have a lot of long term value. Even if the exact contours of the kind of incentives or the kinds of regulations or the deals that are struck between news publisher's platforms and regulators changes. So there's a lot of value in the long term just by setting the terms of the debate based on sound rationales.

CHAIRPERSON: That certainly does make sense, thank you both for that input. I mean going back and your answer sort of slipped out tax
20 and I suppose tax seems like it may be potentially investment or investment neutral if it's determined regardless of how much news you've showed. I mean where are your ideas going on the tax front?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND

PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Ja, I think my problem with the tax and like I said I'm working on a paper about this right now,

is I don't see evidence that Google and Facebook are willing to pay taxes and I think that in many cases politically it adds an extra element which is public support. So in Australia all you needed to do was persuade the Competition Commission right, but if you all of a sudden like in the US, have to get State Legislators on board and then you face a million lawsuits. So what is the quickest and most efficient to get some of what is owed. For Google and Facebook the quickest and most efficient way has to be private arrangements with large publishers paying. And the bargaining codes have helped make that

10 happen even if the amounts have been small. So painful as it is, time consuming as it is, it may be easier than trying to put a tax on tech and then also creating a fund, and then finding a fair way to allocate the funds. So I always say, Dean Baker when I spoke to him years ago about different policies and had a lot of conversations with different – Andrea Crad, and Julia Cage and all these media economists and what's the best thing that we should be putting our eggs in that basket. Is it subscriptions vouchers, is it tax credits, is it employment tax credits, like what should we be doing. And the answer from the economists is always do whatever actually works,

20 what you can get done. So this relates to the protocol context. The bargaining codes have pushed Google and Facebook or at least Google to negotiate with the publishers. So that's really my practical approach, it's an emergency, journalism is having all kinds of problems, money is owed, what is the quickest and most efficient way to get some of that money to the publishers so that they can use it for

news. But I am certainly agnostic as to the method.

CHAIRPERSON: And I mean related to that you mention in your presentation that Jeanne Hughes Roy has used Crowd Tangle to do some estimates and we had a South African estimate, we'd appreciate getting that methodology as well.

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Absolutely and I think also think just to add to that, so one of the things Haaris and I did was we signed non-disclosure agreements with some of the researchers
10 who calculated funding for the US. I think as the Competition of South Africa, you may be able to talk to some of your counterparts in other parts of the world and find out how they did their calculations as well, if that would help you. But I also think that the methodology that Haaris came up with as we've got the methodology right there for you and you can absolutely use it to do your own calculations. And then also adapt it. So if your publishers say hey we've got this amount of subscription traffic, or we really think that or Google tells you well actually advertising revenue attributable South Africa is X, you can tinker with it. Harris even had a friend make a little widget so to get a
20 ballpark figure you can look at GDP so there's a lot of ways that you'll be able to those calculations and a lot of people you can talk to, to compare notes.

CHAIRPERSON: Ja, I suppose the one thing that we have picked up and you mentioned that this is a fast changing world, but Google Discover which is part of the I suppose the Search app on the mobile

seems to be generating a huge amount of traffic, at least in South Africa and many publishers are now seeing – I think one had up to eight times more impressions and clicks on Discover than on news, Google News. Now this is a product which seems to have been developed to create engagement because it is interest driven, it is about a feed and the inevitable scrolling addiction, I mean this certainly seems to be a little different but it does seem that crawled content and a lot of news for that matter, because it's about fresh content is driving the product and hence the user engagement. Now, 10 I've thrown that in just more because I was wondering whether this is something you have picked up, it may be that in South Africa the engagement is different to the US, but this is not passive search anymore, this is developing a feed and populating that. And on that certainly Google does monetise because there are ads in the feed, much like your Facebook feed or your TikTok feed. But should we think about that slightly differently to the search and if so how would you suggest we do?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY:

20 Google Discover but [indistinct 01:45:37] my co-author Helena Herrera Beaumont and I did a short piece on what happened in Spain after Google dropped news. And they launched Google Discover I think in 2017, the personalised news feed that you allude to, and what we were told by publishers is that it actually really promotes a lot of click bait and it doesn't help revenue. So their experiences with

Discover have not been very good. I'm pretty sure it's also been launched in Brazil. I don't think it exists in the US as far as I know, so I've heard bad things about it but again there's definitely people with more experience who we could connect you to.

CHAIRPERSON: Yeah, it seems to be on any android phone if you swipe to the side you get Google Discover, at least that is in South Africa, I don't know if it's elsewhere. The other aspect is just the copyright and fair use. I mean from what we can see at least this approach, Google seems to base their remuneration, this is on
10 publicly available information on the ads that may appear in search pages where news appears and then some sort of share of the news in that page. But I think we are told that not many appear where there is news because most of the ads appear where you're doing a product search, it's intent based advertising. So is this likely, given your belief around sort of value being indirect as well? I mean is this likely to capture that sort of value, going a copyright approach if that is how it's implemented?

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Any copyright
20 approach should be taking into account the full value of the intellectual property that is being accrued to the other party. It just so happens that the strategy and it's absolutely true, and we mentioned this in our paper that the direct advertising is not this real source of value that comes from this intellectual property, it comes from the indirect value art. So I don't think there's any conflict with respect to the principles

of the value generated with intellectual about being the value that news generates. It's kind of like a cross-subsidisation arrangement where news is providing positive spill overs to other parts or more monetizable parts of Google search, while news itself is not being directly monetised, so it's perfectly consistent with the principles of intellectual property and copy right.

MS. PAULA FRAY: Anya, I wonder if we could just take a look at what happened in Australia, I mean it's been argued that the solutions in Australia benefited the large media houses most. And I think the
10 same concerns have also been raised for California, is that justified and if it's justified, how would South Africa attempt to avoid that?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: So a few things, one is that is absolutely a Google talking point. And my husband always jokes that when he was in the Whitehouse, the only companies that ever said they cared about the little guy were always the big ones. So that is a talking point for sure. You can look at it in a few ways, one is you could say it's market distorting for small outlets with no traffic to get as much money as the big outlets. But the entire success of
20 what happened and part of how Australia was even able to do this, was of course Rupert Murdoch was absolutely behind the law, but the Green Party came in and said there's going to be so much money paid out that even the small ones are going to benefit. And that is exactly what happened in Australia. So sure, the big outlets with the most market share, got the most money but the small outlets were able to

collectively negotiate and it was Emma McDonald at the Minderoo Foundation who worked with them as – she's a lawyer. They banded together in a class and they did get payment. So now most of the Australian outlets, large and small, have gotten payments. According to Rod Simms the amounts paid to the smaller outlets were larger per I think employment, per journalist than the big outlets. And you can absolutely check everything I just said with the Australians because I would hate to give you misinformation. And this is part of the problem by the way with the lack of transparency is that we're all sort of

10 guessing and curating it and Rod Simms said call the outlets and ask them questions. This is why we would like to see more transparency. The same fight is absolutely happening in California right now, while the small outlets are saying this isn't fair, this is going to be given to Legacy Media Outlets it's just not fair, and that is part of why I was asked by the Rebuild Local News Coalition to do a paper on taxes. Would a tax be suitable and would it be more fair. So two things, one first you have to pass the – as I mentioned, first you have to get it tax approved, you have to get politicians on board with the tax and then you have to create some sort of fund that allocates the money fairly.

20 And this is exactly what Sabine tried to do in 2014. He said let's have some kind of levy on the platforms and then we're going to create a fund. And then we're going to divide the money equally and the small outlets are going to benefit as well as the big ones. And what happened, Google pulled out for seven years until the EU copy right directive forced them back in. So what can I say, it's just – yeah, and

I think and I'm also, yes, it's important for small outlets to get money, absolutely and that is why we passed this code of principles in Johannesburg last summer is to say that whatever new laws come down the pipe, they have to include the small outlets and they have to include transparency. But that is going to be a question regardless of what measures you implement and it doesn't matter whether it's a tax, whether it's a fund, whether it's a platform remuneration bargaining code type deal.

MS. PAULA FRAY: You know Erasmus Neilson described it as the
10 current media landscape as a winner takes most scenario and it's been really interesting here to be listening to the array of different types of media who've come to the Commission from really the top big media houses to the small community media who are living hand to mouth and publishing when they get sufficient. And I'm wondering, I hear what you're saying in terms of how this kind of support but what are the kinds of things that you would be thinking of as a Commission thinking of in order to protect the plurality of voices when we are looking at solutions?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND
20 PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Ja, and by the way, as I mentioned this is what's already happening. So the private arrangements that Google is making around the world are already giving a lot more to the large outlets than the small ones. And that is what's going to happen to open AI as well. So it's not like right now it's this fantastic fair system and under a new system the local small

outlets are going to get screwed, right. So that is the first thing to remember. So then you have to start thinking, we believe from what we've seen that collective bargaining is the way to go. Haaris and I have been going around the world begging the publishers, hold hands and bargain together. And when you talk to open AI include the small ones. This is not what we've seen. We've seen everybody under cover of darkness cut their own deals and benefit themselves. Everyone is so desperate that they just will take whatever they can get and even the small outlets in California they're nervous too. It's not like they're stupid, I mean they understand the implications but some of them are already getting a little bit of money from Google. They say, well gee bird in the hand. So I think if you're thinking about what the Government can do then the question is how can you design some sort of redistribution into the process. Now you know of course that in the Scandinavian countries they've got all kinds of terrific criteria. So I think it's Denmark and I've written extensively about this in the saving journalism reports that the second largest news outlet in a given town will get funding but not the first largest, because they want to make sure that there's plurality. Or think of all the countries that have funded local language publications or local language music or cinema. So there's lots of ways the Government can design subsidies for small outlets. I think in Paris the newspaper kiosks only get their licences if they agree to sell about 40 different kinds of magazines. And you can argue that this will lead to all sorts of unprofitable outlets being kept alive or you can say, hey, this is a

public good, we need to have a diversity of voices and we need to design policies accordingly. I mean there is absolutely no reason that if you were designing a code in South Africa, you couldn't also say 10% surplus on top of what's being given to the largest publishers needs to be set aside for smaller publishers. I mean there's all kinds of ways you can do that. And you have publishers' associations, you have [indistinct 01:56:00] I think you're already in fact having some of these debates internally about how would the pie get divided.

MS. PAULA FRAY: Thank you very much, Anya.

10 CHAIRPERSON: Anya, just in the few minutes that remain I just wanted to cut across to AI. In the first week we had Professor Mariate from the University of Pretoria and I suppose he is also researching AI and he came more to educate us, but also to give a perspective. And I suppose his concern was that, I mean already at least from an African perspective and a South African perspective, most of these AI models are built on Europe and US data sets, they include inherent biases. Some of the things – well they don't include context, that might be South African context, but they also may include certain racist or stereotypical perspectives of Africa. And so I think his view
20 was yes, one needs to potentially pay for some of this but it's worse to be out of it from the perspective of these AI models, than continuing with that bias and being used in South Africa with that bias. I don't know if you've got a perspective on that?

DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Absolutely. So as I

mentioned I just spent a couple of hours with the Icelandic Large Language Model people two weeks ago, and they were very much of the view that it's important to be in the model and to just give everything away. Others I've spoken to have said that that was a mistake and they should have held out for some kind of payment. Again, happy to put you in touch with anybody you want who has a perspective on this. But again what's clear to me is that the large language models are going to give small amount of money to the smaller outlets, and it's going to be repeat of what happened with Google already and social media. Paris is of the view that it's really important for local publishers to band together and collectively negotiate. As I explained earlier it's not just the archives and the content that has existed before which may already have been hoovered up. But it's also in the future when big news happens in South Africa and these models need current information as well. So there may be a chance there to do some kind of negotiating. But that is going to be a [indistinct 01:59:02] situation where you throw up your hands and say, okay just take it versus, hey wait a minute, you're going to need it to have a decent model and therefore you should be paying for it. Haaris may have something to add there.

MR. HAARIS MATEEN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Just a very small point. That point about South Africa and African news or African news in general not being used as an assumption, I think there are good grounds to assume that these large language models are already

using news that is available archival news that is available from African sources and South African sources. I mean there's very little that has been – that stops them from using it. The question really is that when new content is being produced which again costs money and has many welfare values, do we think that that should be paid for, or that should be compensated for. And I think the answer then is yes, and that would still allow for more – the biases to not be present. And just to wrap this up by saying that the importance of Competition Commission or regulatory authorities cannot be underestimated

10 because from an economics perspective the whole problem is really of coordination. I think it's clear that news is producing a lot of value, but when there are individual publishers big and small, but many of them that can be individually negotiated with, who do not have any kind of bargaining power in the process, it becomes very important for one body to act as we like to call it, a coordinating device. That by just that act itself it ensures a lot more bargaining power on the others.

CHAIRPERSON: Well thank you both for making the time, I know you got up early. Now you've found employment elsewhere, Haaris, I'm sure it's even earlier than Anya, but I do appreciate making the effort,

20 it means a lot to us to have the kind of quality inputs into this inquiry from yourselves and your institutions. And certainly we will be in touch about following up on some of that information and studies. From the inquiry's perspective we're only running tomorrow and Wednesday this week and then we pick it up next week as there is a long weekend. So tomorrow we have Media Monitoring Africa and SOS Save our

Media and Digital Platforms Market Inquiry

SABC. We were scheduled to have Google AdTech tomorrow night but that's been reallocated to the 26th where we can get a little more time with them. And then Wednesday we have iMedia and YFM in the morning and then Omnicom RTB House and Prime Media in the afternoon and that would wrap up this week. Next week we will have Open AI, YouTube, Google AdTech and Meta. And most of those will happen in the evenings just to accommodate the different time zones. But thank you again, Anya and Haaris and that closes our proceeding for today.

- 10 DR. ANYA SCHIFFRIN - SCHOOL OF INTERNATIONAL AND PUBLIC AFFAIRS: COLUMBIA UNIVERSITY: Thank you for inviting us. And let us know the deadline for written testimony submissions please?

CHAIRPERSON: Thank you.

END OF PROCEEDINGS ON 18 MARCH 2024

Media and Digital Platforms Market Inquiry