



# Joint Committee on the Draft Online Safety Bill

## Corrected oral evidence: Consideration of government's draft Online Safety Bill

Wednesday 27 October 2021

9 am

Watch the meeting: <https://parliamentlive.tv/event/index/47212f70-89bd-4c5e-8451-ab51fe3693c2>

Members present: Damian Collins MP (The Chair); Lord Black of Brentwood; Lord Clement-Jones; Baroness Kidron; John Nicolson MP; Dean Russell MP; Lord Stevenson of Balmacara; Suzanne Webb MP.

Evidence Session No. 12

Heard in Public

Questions 193 - 199

### Witness

I: Maria Ressa, CEO, Rappler.

### USE OF THE TRANSCRIPT

1. This is an uncorrected transcript of evidence taken in public and webcast on [www.parliamentlive.tv](http://www.parliamentlive.tv).
2. Any public use of, or reference to, the contents should make clear that neither Members nor witnesses have had the opportunity to correct the record. If in doubt as to the propriety of using the transcript, please contact the Clerk of the Committee.
3. Members and witnesses are asked to send corrections to the Clerk of the Committee within 14 days of receipt.

## Examination of witness

Maria Ressa.

Q193 **The Chair:** Welcome to this further evidence session of the Joint Committee on the draft Online Safety Bill. Our witness, Maria Ressa, joins us from Manila. Maria, thank you for joining us. On behalf of the committee, I want formally to congratulate you on your Nobel Peace Prize, which was a wonderful recognition of the very important work that you have done for people in the Philippines and as a beacon of light around the world. It is an honour to have you with us today.

**Maria Ressa:** Thank you for having me. I am looking forward to hearing your concerns.

**The Chair:** I would like to start with a question that is close to home for you—the nature of disinformation in a country such as the Philippines. One of the issues that we have to consider with the draft Online Safety Bill is its scope and to what extent it should include things such as harmful disinformation that can affect democracy and society.

One of the challenges that we constantly face when we discuss that question is, if you start to regulate the disinformation, will oppressive regimes not use that as a guide to suppressing speech in their own countries even more? As a journalist who operates in what is effectively for you an oppressive regime, what are your views on that? Can we afford to leave things as they are? Should we seek to create guidelines and regulations on harmful disinformation?

**Maria Ressa:** Regulation is our last pillar. We are having elections in the Philippines in May next year, and we will not have integrity of elections if we do not have integrity of facts. We have shown in our research—in the data that we have been able to gather—that the very platforms that distribute the news, and distribute facts at a large scale, are biased against facts.

Part of the reason that I wanted to speak with you today is that there will be a lot of pressure to delay and delay, but we have to live through this. I am at the front line on both ends: the weaponisation of the law—I am concerned about the nuances of the law—and at the same time the weaponisation of social media, and they are intertwined. I think that a law is necessary. Regulation is our last hope.

As you know, I am also a partner of Facebook. Rappler is one of two Filipino fact-checking partners. Since we discovered what was happening online—the insidious manipulation in 2016—we have constantly worked behind the scenes for a few years, and then finally I started calling it out because it was not moving. As you can see from the Facebook Files, almost everything that has come out are things that we have lived through, so we definitely need legislation.

The question, of course, is what is at stake for the UK. I love that you are pushing forward with it, and I thought that your earlier report that looked

at the global scope was unique. The problem is that you will be a model for everyone else around the world, so you must be a gold standard. That is tough. At the same time, if you do nothing, this also cannot stay, so please do something. I will try to help as much as I can from our experience.

**The Chair:** From what you say, it sounds as though, if we do nothing, effectively, it is still doing something. It could be sending a signal to the rest of the world that we do not think that disinformation is important.

**Maria Ressa:** Doing nothing pushes the world closer to fascism. The Oxford University Computational Propaganda Research Project earlier this year said that cheap armies on social media are rolling back democracy in 81 countries around the world. I know this myself. There was a big data study done by the International Center for Journalists, UNESCO, the University of Sheffield and our Rappler research team of almost 500,000 social media attacks against me. It showed that 60% of the attacks were meant to tear down my credibility and 40% were meant to tear down my spirit—dehumanising, sexist at best and misogynistic at worst.

What I have gone through is what every female journalist and every user of social media is going through. I am looking at the reaction in the United States to Frances Haugen, and we are talking about children and teenagers, but think about it as poison in the information ecosystem. That poison is going to attack kids as well as adults, as well as everyone else.

Doing nothing is not an option. I continue to appeal behind the scenes to all the tech platforms for enlightened self-interest because the road that they have set the world on is not sustainable.

**The Chair:** The problem that you experience in the Philippines seems to be not just that the Government are failing to regulate social media, but that the Government are an active participant in the spread of disinformation and the targeting and drowning out of dissident voices.

**Maria Ressa:** That is correct. They are a major state-sponsored force. I was in one of 12 research groups that were working under Camille François, who is now with Graphika. We saw that this is state-sponsored online abuse. It is meant to do two things. For the target, it pounds you into silence and it creates astroturfing—a fake reality. We in the Philippines even quantified that. How do you create a fake reality? You have sock puppet networks, meaning you create fake accounts, and those fake accounts follow each other and manufacture a reality.

This all sounds like words, but I can tell you what it means when you are the target. It means waking up to 90 hate messages per hour on your feed, having to deal with that and having to deal with your family and friends and anyone else you do not know asking you questions about outright lies. Government officials are targeted in similar ways, I suppose.

Here in the Philippines, it was one of the ways that a brutal drug war was accepted, because Facebook is essentially our internet. This is the sixth year in a row that Filipinos spent the most time online and on social media globally. The Cambridge Analytica whistleblower, Chris Wylie, called us the Petri dish for the parent company of Cambridge Analytica, SCL, which was operating here. It was trying these tactics of mass manipulation, and if they worked it used them on you guys. We were the guinea pigs. Beyond that, Facebook itself admitted that we were ground zero. Why is that important?

I will state the five main findings because we do our own research on this. What sets Rappler apart is that we have a database which we call the Sharktank. Some of this comes from our partnership with Facebook; we work with Facebook to try to make it safer. It is just not enough when it is exponential.

The first part was the discovery of information operations, let us call it, because I sometimes rail against what western media calls misinformation. Misinformation is, "I made a mistake". Disinformation is meant to manipulate you. It is a half-truth or a half-life; it is done with a purpose, it is done by an actor, and it is about power and money in many instances.

Our first discovery was that we found the Government were involved—these are content creators who are now employed by the Government—in a brutal drug war. There were times in July 2016 all the way to 2017 when there would be at least eight dead bodies a night. Later figures would say up to 33 people were killed every night in those early months. By January 2017, the Philippine police had said that 7,000 people had been killed, which they rolled back in plain sight to 2,000. As violence was happening in the real world, on Facebook in the virtual world, people were posting about it, but if you posted about it, you got clobbered. The first victims were ordinary citizens, so that stopped the questions. Then when the journalists started asking questions, we became the target.

Rappler was the third news organisation attacked by President Duterte. The methodology is bottom-up attacks on social media and then top-down attacks by Government, and they work hand in hand because online violence, as you know, does not stay online; it seeps into the real world.

The first was our discovery of it. The second was seeing the Government, after they took office and after the drug war, apply it in different ways instead of answering questions—the checks and balances of our American-style democracy. Our constitution is patterned after the United States. They used information operations to attack the journalists. I can send a link to that. There were information operations for something simple and minor, such as the Southeast Asian Games and allegations of corruption there.

The third was Marcos historical revisionism. As early as 2015, the son of our former dictator Ferdinand Marcos, Ferdinand Marcos Jr—his name is

Bongbong Marcos—was running for vice-president in 2016. The networks of disinformation began as early as 2015, and they have continued. We exposed them in 2019. Thirty-five years after a people-power revolt ousted the Marcos family and forced them into exile, Ferdinand Marcos Jr is running for President. He announced this a few weeks ago. Those networks of disinformation were the ones that also essentially attacked me and Rappler after the Nobel and after we kept exposing them. That is part of the problem. It has just become far more dangerous.

Finally, the last part, which is not just causing a chilling effect but creating an enabling environment, is something that is called red-tagging—kind of like McCarthyism. We, like Hong Kong, have a new anti-terror law. Right around the same time, the Government passed this law. Red-tagging means calling a human rights activist, a journalist or an opposition politician a “terrorist”. Under this new anti-terror law, you can be arrested without a warrant and jailed for up to 24 days.

We saw an upshot of violence at the same time as we saw an upshot of the same tactics from the drug war online, moved to target activists and human rights organisations, and then an increase, beginning in March this year, in the killings of human rights activists. Sorry, I have said so much. I will shut up and take your questions.

**The Chair:** That is all terrific, thank you. You have set this out. Rappler, your organisation, has set this out. As you say, you work with Facebook as a fact-checking partner in the Philippines. Facebook is well aware of these issues, and other people have also written about the Philippines. What is Facebook’s role? Do you see Facebook as being effectively complicit in the hijacking of its services to manipulate information and target people in the Philippines? What sort of relationship do you think that the company has with the Duterte Government?

**Maria Ressa:** Those are three different questions that are tough on their own. As I have said, Facebook and I will call ourselves frenemies. Part of it is the hierarchy that has been created inside Facebook, and the papers lay this out. When I brought the 90 hate messages per hour that I was receiving to Facebook in 2016, I was told, “You are a public figure”. I am actually not, and the Philippine constitution gives some protection to journalists. But I endured it for many years.

Just a few weeks ago, Facebook rolled out increased protection for journalists. “Move fast and break things” does not work in the real world when the product that you are making insidiously manipulates people’s minds. The question that I would ask is: would I be facing 10 arrest warrants from the Philippine Government if Facebook had not helped to enable an environment where that could happen? Would our democracy and our institutions have been weakened as quickly without Facebook?

I will not single out Facebook. YouTube is now No. 1 in the Philippines. Twitter has been a little bit better in terms of protection for me. That is part of what your legislation should look at. They are different in their

own ways. They have different definitions; it is little things. I will jump to the Bill itself.

The biggest problem so far in what we have seen globally is that people look at content moderation. That is when you run into these freedom of speech issues, but in the end that is not the problem. It is not a freedom of speech issue; it is actually the way the platforms are designed and the factors that are included in the algorithmic bias, because every platform chooses the algorithms it uses and A/B tested it to take advantage of the weaknesses of the human psyche. A/B testing chooses what will keep us on the site longer.

I know that you know this already, but part of what I would look for far more than I would look for content moderation is radical transparency of algorithmic bias, algorithmic distribution and algorithmic amplification, because those cause the most harms. Those assumptions that they build into the algorithms are the ones that help determine emergent behaviour, not just of one country but of humanity.

**The Chair:** Absolutely. In a country such as the Philippines, where Facebook and YouTube have very strong commercial positions, do they effectively go along with the Government of the country? Do they not cause too much trouble because they are making a lot of money and they are quite happy to keep it like that? Do they operate to a very different standard in the Philippines than would be acceptable in the UK or America?

**Maria Ressa:** Let me put it this way. There have been four Facebook takedowns in the Philippines. I would say that the period from 2016 to 2018 saw virtual impunity. As we called out the impunity of Rodrigo Duterte and the drug war, Mark Zuckerberg and Facebook had impunity. That was also when I saw my reputation, my company's reputation and human rights activists torn down with relative impunity.

We are always behind you, the west. At one point, someone told me, "Well, you don't bring in enough revenue", so I said, "Well, then maybe turn it off". We would be better off because things would move slower and then journalists would have a chance to catch up. What I know is that they are afraid of the same things. This is certainly not publicly stated, but when you have an office in a country such as the Philippines, we journalists face these threats. When the President threatens an organisation, it then has to take those risks.

The giant platforms are here, and their behaviour has been determined partly by the environment that they are operating in, like us, and the ability to kick things back up to the United States.

I do not really have much choice but to accept it, and that is part of the reason that I am looking globally for legislation that will help put these guardrails in place.

Q194 **John Nicolson:** Thank you so much for joining us, and many

congratulations on your wonderful award.

You talked about algorithmic amplifications. As a journalist, how do you think journalism can avoid just being the result of myriad clicks—clickbait? How do we preserve good-quality journalism?

**Maria Ressa:** It is increasingly difficult to do that, partly because of the incentive scheme of the entire internet. I am sure that you have had others speak to you about surveillance capitalism, but think about it like this. I grew up in an age when we had more time and money to do good-quality investigative journalism, but in the age of the internet everything has been reduced to a page view, and that is globally. Even the ad structures come down to that. The incentive scheme is not so much for quality journalism. We try. The news organisations continue to do that. It is the process of journalism, by the way, that creates good journalism; it is not one person writing on a Substack or in a newsletter. It is the fact that you have a reporter, an editor on top, another editor and legal review. That part is expensive. The incentive scheme reduces everything to the page view and it commodifies news. That is the beginning of the end. When it went there, it moved towards clickbait, as you rightly said.

I handled the largest news organisation in the Philippines, our primetime newscast on television. If I looked only at ratings, I would have had a primetime newscast of crime and entertainment because they brought the most ratings. Because I am a journalist, we tried to balance that with—Facebook used the words we used—vegetables, because it is important. The news agenda needs to be protected. Our citizens need to get the information. Even though it would cost us less money to bring you crime and entertainment, we would send a flyaway to a remote area where conflict was taking place.

The commoditisation of news has led to where we are. That is the first part. Now, you add the amplification and algorithms that amplify what keeps you on the site. Take entertainment and crime. As a television news group, you know that that is the case but you cannot do much beyond the square box; everyone will see the same thing. Put it in the multiple worlds that Facebook gives even all of you in the room, and your feed will be different from one another. When this was first explained to me in 2014, I asked, “What will happen to the public sphere?” Each of us is now creating our own worlds and is given our own worlds, so by necessity it tears the shared reality apart.

The algorithms also decide how you grow your network. How the platforms grow is based on friends of friends. That is straight across all the social media platforms—Twitter and Facebook. When you do that, you are also pulled apart.

In the Philippines in 2016, I watched this happen. At the beginning, we were not debating the facts; we were all in the centre. Then, when President Duterte came in, if you were pro-Duterte, “the friends of friends” algorithm pushed you further right, and if you were anti-Duterte

you were pushed further left. That is 2016, 2017 and 2018. This is the shared reality that is torn apart.

Beyond that, another algorithm comes in and keeps you there. If you watch a 9/11 conspiracy theory, you will be brought further down the funnel because it keeps you on the site.

All these things may be great for business—and they have been good for their business—but it has been horrendous for democracy, for the education of our people, for a thinking-slow process of all the problems in each of our countries that we each need to solve.

**John Nicolson:** I worked for the BBC. I grew up watching “Weekend World” with Brian Walden—45-minute, long-form interviews and only the best-briefed politicians could possibly get through them. I remember when CNN came along and we thought, “Oh my goodness. Repetitive, incredibly American and reports that are only one minute 50 seconds long. This is not proper journalism”. Now, we look at CNN and think, “My goodness, this is incredibly complex, detailed journalism compared to some of the stuff that we are watching online”. Is one of the problems that there is such a gulf between people who grew up at the time that we grew up and folk who are growing up now? Are we lamenting the loss of a golden era without necessarily the means to change the future?

**Maria Ressa:** I embraced this technology. When we created the idea for Rappler in 2011, my elevator pitch was, “We build communities of action and the food we feed our communities is journalism”. We started on a Facebook page, and if Facebook search had been better we probably would not have moved off.

There is a solution. I continue to work with all the platforms because I believe there is. I do not think that we can roll it back. Pandora’s box has been opened. The genie is out of the bottle. I do not know whether you saw this, but an anger emoji would be valued five times more than a like.

**John Nicolson:** We had a witness appear before the committee this week, the Facebook whistleblower, who said precisely that. She said that—I am paraphrasing—nothing grows the business like anger and rage, and Facebook has discovered that. The angrier they make people, the more people will spend time on Facebook. Joy does not sell.

**Maria Ressa:** I disagree with that because we have a mood meter and a mood navigator, which we rolled out in 2012. What we saw until 2016 was that “happy” was the top emotion because we did not put our fingers on the scale. That is what I mean by algorithmic amplification. They chose to put their fingers on the scale so that you would stay on longer and they would be able to make more money—surveillance capitalism.

When we did it, we had a mood meter, but we wanted to get a clear sense of how stories and emotions travelled through our society, and we did get that. The top emotion was “happy”. The second was “inspired”. The third was “anger”. Come 2016, “anger” was No. 1 by a huge margin.

What I keep saying is that this is insidious manipulation of our biology, and it is geared to that. It undercuts our thinking.

Reading parts of the OSB, I think it is fantastic that you guys are setting out to do this, but it cannot be a whack-a-mole approach. It is almost that you are going downstream to the product—the toxic sludge. The content is already tainted by the creator. The gatekeeping could have happened at the creator level or at the algorithms and the factors that go into them.

There is a solution, but it requires enlightened self-interest from the platforms, all of which say that they want legislation. I have had discussions. So, what do you want? What are you willing to do? Ultimately, it means making less money.

**John Nicolson:** The trending story here in the UK today, you might like to know, comes from talkRADIO. I declare an interest as having worked for talkRADIO. One of the presenters did an interview yesterday in which he said that he believed that concrete grew in the same way as trees grow. I think that it has 2.6 million views as things stand. It is radio being amplified online. It also shows you how, if a social media team really hate their employers, they can secretly post to cause maximum damage to their employers, which is an interesting thing to watch.

How do you cope with the horrible harassment that you get? I read how many times you had to post bail—10 times just in order to stay free. That must have had a devastating effect on your sense of well-being.

**Maria Ressa:** I knew that the world was upside down. We began to see it first on social media because we lived on social media. I do not know if you heard that I had brought the data to Facebook in 2016, and it was alarming to me what was happening in the Philippines. The people I had spoken with basically listened, and then I said, "Please come back to me immediately because I want to write this". I expected them to fix it. I did not realise how involved it was to fix it then. I said, "If you do not fix it, Donald Trump could win". This was in August 2016. At that point, it did not look like he would, and we all laughed.

After he won in November, they asked me for the data again. I saw it coming. I am a journalist. I wish I was not in the place that I am. It feels like the baton was passed to me and I became a news head at exactly the wrong time to be a news head.

My family has it worse. I am okay. I know what I went in for. I am fighting for justice in the real world and justice in the virtual world. That is what I feel has been taken away from me—justice. I hope you bring back some of it.

**John Nicolson:** You have a voice. We, at Westminster, and people watching around the world are hearing your voice now. We are considering the whole issue of online anonymity. A lot of witnesses want us to recommend that there should be an end to online anonymity and

that, even if you retain a nom de plume, for example, you should have to identify who you are. Some people think that that will reduce harassment and bullying. I often worry—and I have not heard an answer to this—about what we do for people who are in oppressive regimes and want to stretch out to parliamentarians like us, tweet us, send us direct messages and tell us what is happening to them. If we introduce verification, how are folk in Belarus going to get to speak to us?

**Maria Ressa:** It is one of the tougher questions that you, and we, are handling globally. I have several thoughts, and I will answer first with a lot of caveats. From my experience first, I will put it this way. Since 2016, the attacks that are exponential initially came from anonymous accounts because they are easy to make and easy to throw out. Even if they get blocked or are thrown out by the platform, it is just as easy to create them.

That is me. I have watched my credibility get whittled away. You cannot respond. If you are the journalist or if you are a government official, your hands are tied. You are responding to a no-name account. You just do not do things like that. The attacks are horrendous. That is my experience.

Having said that, I have also worked in countries where you have Governments that repress people. I worked under Suharto's Indonesia. Suharto was in power for almost 32 years. I came in at the tail end. I was president of the Foreign Correspondents' Club in his last year. I worked in China under Deng Xiaoping. We had put in place processes, when journalists were the gatekeepers, where people who needed help could come to us. I preface all of what I will say with that.

There must be some way whereby the platforms—and it is problematic because Governments can demand this data from the platforms—prevent the mass creation of new accounts. Each of the platforms says that every account is real. How do they verify that? What was that blue check mark for in the first place? Why was this not verified when I first formed my account? Why is it so easy to form an account? We now know it is because the more people you get, the more accounts you have, the more revenue you make. That is good for the platform. How about fraud—just basic fraud? Starting there. I am not directly answering your question—

**John Nicolson:** It sounds like you are supporting the idea of verification, but at the same time you are saying that the social media companies should not surrender data to oppressive regimes. I suppose they would say, "So, you are telling us we have to break the law in many of the countries in which we operate". If Putin says, "You have to surrender all the stuff", they simply say "No". I am guessing you would close them down.

**Maria Ressa:** That is correct. There are two points. One is that none of the platforms has signed up to be news organisations, which have dealt with things such as this. I opened the Jakarta bureau for CNN. It took

eight months to negotiate with the Indonesian Government so that we would be able to operate with the integrity we needed.

These companies are not setting out to protect the public sphere. They do not have a standards and ethics manual. We keep asking them, "What are the values that drive you, beyond an itemised list of content moderation that evolved over a period of time?"

Let us take fraud first before we go to anonymity. I do not want to answer you directly because it is so nuanced. How do you prevent fraud? You make sure, because it is your home, that people are not masquerading as different things; you make sure that fraud does not happen. If that first upstream thing happens, everything else follows.

On the issue of anonymity, if you are engaging with Apple, earlier on, or Google now you need to put in a credit card and a name. You need something that anchors you in the real world. My response to the Bill is that we want to re-emphasise that the laws in the real world apply in the virtual world. We do not want to create new ones, but we want to make sure that the laws are there. If you give that, nothing should stop you creating a persona, especially if you come from a repressive country.

The second point is that we have a list of repressive nations. The world press freedom index does this on a daily basis. It could just as easily roll out specific features for countries that are ruled by tyrants. That prevents the rest of us in other parts of the world who are fighting for our democracy from having to deal with exponential attacks from anonymous accounts.

Having said that, some of the most damaging attacks come from real people. This is the connection to that. How do real people get misled? They get misled by astroturfing and repetition. When you go on these platforms it is like the difference between an Excel sheet and big data. When you pound that lie a million times on these platforms it becomes a fact.

That goes right back to why I think that anonymity in general is bad for democracy. You can see what it has done to our politics all around the world and, at the same time, to the countries run by repressive systems. There is a lifeline there and we need to protect both, but not at the expense of each other.

**Q195 Lord Clement-Jones:** May I add to the Chair's congratulations on your Nobel Prize?

You very cogently described why disinformation and misinformation on social media platforms should be regulated in some form. You have described the operation of the algorithms that they use in amplification and the virality of the messaging in particular. Part of our draft legislation is looking at the whole question of what duties we need to impose on the platforms. At the moment, it does not include societal harms, but many of us think that it should include societal harms in the form of misinformation and disinformation.

The question is: what risk assessment should be imposed on the platforms as regards algorithms? As part of that requirement for risk assessment, is it practical to say that, for instance, social media platforms such as Facebook should limit the virality and amplification of misinformation and disinformation? In a sense, how do you identify what that misinformation and disinformation is?

There are some fairly tricky issues involved. You made a very strong distinction between misinformation and disinformation. You can fact-check misinformation, but you cannot really fact-check disinformation. It is basically a value judgment, is it not? How do you get to grips with the question of risk assessment that is required and would be, if you like, policed by a regulator?

**Maria Ressa:** I hear three questions. I think that the first one was about how to differentiate societal harms and how the Bill is focused on individual harms. The second was about how we are going to ask for radical transparency, because I think that is what you need when it comes to algorithms. I think the last one is whether we do things such as limit virality.

On the first one, it is impossible to talk only about individual harms because it is the pressure of the group and the emergent behaviour that is being created that is dangerous. With emergent behaviour, I always say that there is a devil and an angel on each of our shoulders and our conscience decides which way we go with difficult choices. All of us are on these platforms. Because of the algorithmic choices, what is amplified and fed to us in our individual feeds, we do not know what anyone else's feed looks like. What is fed to us is equal to gagging the angel and giving the devil a megaphone, because that is what is coded into the algorithms. Toxic sludge is what spreads.

I think it was a 2018 MIT study that showed that a lie is tweeted 70% more than a fact. Facts are really boring. I have spent my career trying to make them interesting, to make people care about the facts. The algorithms are biased against that.

The societal harm is that, if that is happening at an individual level and it is shaping the way we are going to go, is it any wonder that cheap armies on social media are rolling back democracy in 81 countries around the world? I use that statistic all the time, because I watched it grow from 10 to 27 to 48 and then to 81. Every year it gets worse.

The emergent behaviour that we are seeing kills democracy and brings fascism. You can see this in the type of digital authoritarians who are being elected and then take the reins of power and collapse institutions from within.

The answer to your question about exactly what that means is that the fact that we have to rely on whistleblowers to see what is happening is a major problem—in the same way that a news organisation will do a corrections page if it makes a mistake. It has a standards and ethics page and this is given to the public. These social media platforms are now the

operating systems of our societies. That is a given. They are the world's largest distributor of news; they are more powerful than any single news organisation.

What can we do? I suppose it is like product safety tests—in the same way they will say that there is A/B testing. That A/B testing may be part of the regulatory process. On blue sky media, I hope that the platforms will co-operate. I think they will. They can do this better, because they do it but, when they make the final choice, that choice is for greater revenue and greater engagement.

What the regulatory agency needs to do is make the public's fear part of that equation. Independent researchers should be given access to the data, which should include corporate data: product safety tests and the internal research that comes from a whistleblower. Then you can do an audit. This can be part of a thinking-slow debate, instead of being told, "We have already rolled that out, so this is the impact on you".

**Lord Clement-Jones:** That is fantastic. I know that Baroness Kidron will follow up on the issue of safety by design.

Q196 **Baroness Kidron:** Maria, thank you for the work that you do, and congratulations. It is incredibly clear.

I want to get down to the basics of what we should do. I am struck by not only your testimony but that of others that the eyes of the world are on us. They are saying, "Please do something sophisticated and effective and please do something now". That is the mantle that we have, and I want a little help from you on it.

Part of the Bill is focused on the idea of a market risk assessment by the regulator. The regulator says, "We are going to decide and ascertain what the risks are". Then it will create a risk profile for different kinds of companies. That picks up on the comment you made that not all companies are the same.

There is then another set of things. I think that the set of things we are looking at is that the company should do a risk assessment that matches it. That is where you get the "societal" bit that you referred to in answer to my colleague. Then it has to mitigate it.

I am interested in where you see the strength and weakness of that circle that goes all the way to whether they have failed if they do not mitigate against the harms that the regulator has identified. We have some work to do on enforcement but, were that to be an enforceable loop, do we have it?

**Maria Ressa:** I worry. Part of your problem in trying to regulate is that the best people to understand the technology are in these companies, so you need to get them on board.

I will throw out an idea. For us, whenever we need to get lots of news organisations working together, we get them together, and then they need to give you their assessment of where they are. Discussions I have

with them such as this are always about how everyone is getting it wrong. I just say, "If you guys had done this five years ago we would not be at this level". That is my frustration. Part of it is that if the regulator goes down the wrong path, it will continue because there are infinite possibilities to what can happen, how we can be manipulated and how the platforms can continue to protect their revenue streams.

Part of it is getting together a group that includes them and their inputs and giving short deadlines that you also set. I say this because I see the data from the Philippines. You will see it in the data, because it is not about words or content; it is about recidivist networks that are constantly pushing out lies. Why are they allowed to stay? Why are recidivist networks not taken down? They have the sophistication to go down to traits that define each of these accounts. Why is that not rolled out? It is because it does not necessarily make revenue.

The difference is between the content moderation of one platform and 180 pages of PageRank, for example. What I have learned from the work that we do is that everything on the internet can be gamed, so there needs to be that transparency. Even Wikipedia has been gamed. What we have seen so far is that what holds them to higher standards is transparency.

The first thing I would say in answer to your question—I know this is going out to everyone—is, "What are the biggest problems?" You already know what the biggest problems are. Why has no one been held accountable for genocide in Myanmar? Why is no one held accountable for these attacks? I think I lost 20 years of my life because of the attacks. These things are illegal in the real world. Why are they legal in the virtual world?

The first thing I would do is get them to say, "Give me your best because I will make it worse if you don't. Give me the problems you are facing. You know the societal problems you need to deal with", and give timelines. We have been talking about this for five years and in that period people have died in the Philippines; in other countries in the global south people have died. I do not have a quick answer.

**Baroness Kidron:** You did give a very clear answer. To push back, there are two problems in what you have just said. One is that we cannot rely only on their research because you can push it so that the incentive is not to do research and then you do not find out what is wrong. There is a little bit of a problem. You need independent research and to set the terms of what problems you are trying to solve.

I really take your point about timelines, mitigations and the real world. That is something I would like to bring up with you. A couple of times, the committee has heard from people that the best thing to do is not to start redefining harm but to import what we have already decided. To a degree, that supports your point about different countries doing slightly different things because we have slightly different rules. In the UK we would bring in our equality and anti-discrimination law, and if those

things were what we perceived as harmful we would be looking to see, first, that they are upheld and, secondly, the idea of radical transparency. Those two things should mean that the conversation is more realistic. I can see you nodding, but would you like to speak to that?

**Maria Ressa:** Yes. For example, let us take something that is videoed. Facebook has admitted that it overstated video by 90%. Yet news organisations fired people and shifts happened in the real world, but nothing happened. No one was held accountable. While you cannot see these things, they ripple through the real world, in the same way that five points for an anger emoji versus a like are rippling through the real world and encouraging behaviour. The further away you get from that algorithm the more gigantic the tidal wave becomes.

Those two things are great. I think that independent researchers play a large role in this. We have demanded data for a very long time—from the very beginning. The one time that data was given, there was a mistake in it and no one was held accountable for that.

People's lives are at risk in these things. That is what I keep coming back to. My country is on the verge of losing its democracy. We have watched it being eroded. I should not have 10 arrest warrants or face the rest of my life in jail, yet I do.

**Baroness Kidron:** For that, I am sorry. It is a travesty.

My final question picks up something around anonymity. If we start by saying what harm is and we have radical transparency—I hear your point that we make people and companies accountable—there is one obvious point around anonymity. Many of these companies have terms and conditions that prevent abuse, so why are they not taking down those accounts, whether or not we know who they are? Does it even matter that they are anonymous? Should they not be taken down on the basis of terms and conditions? Do you think that, although it cannot be the only thing because you have to set some basic terms, companies should be fundamentally responsible for their own terms and conditions?

**Maria Ressa:** A normal company in the old world would have been. The product impacts our minds, our version of reality and the facts we have to make decisions in our lives, and the shared reality we live in. I do not think that there is anything more important than that. I became a journalist because I believe that information is power. Now, the very delivery platforms for information are corrupted, so there needs to be some accountability for that.

The debate should go further upstream instead of going down to the much more problematic things. We have figured out systems of democracies; we have figured out that, while imperfect, the checks and balances are there. What has turned the entire world upside down is that those checks and balances are like Alice in Wonderland. We went down the rabbit hole and there was a new system in place where the old geopolitical power system had gone and the bad actors could use these

existing systems for money and to consolidate power. Only the people who have no conscience, do not care or have a scorched earth policy for the future would use this stuff. That is what you have.

**Baroness Kidron:** Thank you so much for being with us and for your fantastic testimony.

Q197 **Suzanne Webb:** I am very honoured to have met some amazing people, particularly women, whom I would not have met were it not for this Bill, so it is a delight to meet you, Maria.

You have talked about the accountability of the tech companies. To me, it is also very much about the immediacy of action and removing the harmful content. The human cost of harmful content concerns me greatly. Pornhub is a classic example, if I have got this right. It implemented verification very quickly. I do not believe that it was done out of the goodness of its heart but more that it saw it would hurt it financially if it did not.

This Bill will take time. What would convince the tech companies to act now and take down harmful content and resolve the algorithms spreading it? What I would like to see is immediacy to prevent what we know is the high personal cost of harmful content online. What are your thoughts on that? I have only one question.

**Maria Ressa:** For every platform there is a basic assumption that we do not really question, which is that if you do not like the information—we will go to the harmful part—and you are being attacked, mute it or block it. Yet it still exists in the public sphere. When was that okay? I will change my reality to make me feel better, but it is still out there. To me, that is harmful content if it is tearing down.

There are some basic assumptions in algorithmic bias and design that we should be asking questions about. Some of it is bad. There are multiple realities. Where are the shared realities? How do you bring that back versus tearing us apart? There is a human cost of harmful content.

Of course, the question is: what is harmful content? Every platform defines it slightly differently. They do not carry out everything they say they should. There are some great things in the Facebook policy, but they are not carried out. The attacks continue.

We should not be dealing with the after-effects of the neglect. This is a tough one. I guess that accountability will stop it. What stops journalists letting a lie go through? If someone lies to you and you are not able to check it, what stops you putting it out? You are accountable for it. Who is accountable for this? Hiding behind algorithms does not work because the algorithms are ultimately designed by a human being—the company.

**Suzanne Webb:** My real concern is that, we are doing work on this Bill with the help of you and brilliant people who have come along and given evidence, but with all this stuff it will take time to go through the process. At the moment it is just a draft Bill and it has to go through the House of Commons, the House of Lords and then come back. There will

probably be a bit of ping-pong. I guess that it will take a year or a couple of years before this gets through. I do not know. I am new to all of this, but I think that we are talking about a long time.

I would like to see the moral compass of those tech companies waking up and saying, "We do not need to wait for this Bill. We can do something about this now. I do not need a Bill to tell me my moral compass has been going in the wrong direction and I need to put measures in place to stop this", whatever we all decide. We should not even have to think about what harmful content is—it is pretty obvious—and the human cost of it. That is where I am coming from. To all those tech companies out there: get on with it now; get your moral compass in the right direction. Maria, am I right in thinking that?

**Maria Ressa:** Absolutely. I continue to ask for their enlightened self-interest. In the end, it is not in their interest to tear down democracies, to tear the facts down and manipulate people in this way. They looked at it and it happened by, I always say, death by a thousand cuts. A thousand cuts brought them to where they are—a machine.

The other interesting part is that they have compressed the human experience. In the real world we have six degrees of separation; on Facebook you have 3.5 degrees of separation, which means that the virality is faster. Everything moves faster. You can slow it down. In the old days a news organisation would put out a story and civil society, NGOs and Governments would take that story and have time to think it through. I am not saying it is for ever, but part of the disconnect in our society today is that the information comes out, you are pummelled by it, your perception is shifted, and yet civil society and Governments do not have the ability to turn it into action that makes a democracy stronger.

I watched this during my career. Technology has made things faster, so part of it is the pace of it. They can still have the platforms and some of these changes can happen gradually.

Please do not wait years. I understand that laws are different, but maybe in conjunction with the platforms these shifts can happen with the threat of the law. Here is the difference between Governments, news organisations and tech platforms in the way they work. Agile development in tech breaks down everything to the nuts and bolts—to the nails—and then it will build it up from scratch. News organisations and Governments start with principles and then have to think it through and never go down to the nuts and bolts.

I have always felt that old power and new power working together could be transformative for the world, but it is about the basic understanding of how they can fit together and make it better. Their process is also iterative. They will try it. If it works they keep moving down that path; if it does not work they drop it. It is a two-week or one-month build. Given that kind of more iterative experiment, more in tune with the technology that we are using and the way tech platforms work, why cannot a regulatory agency work that way with them and then report back to the

public about what it is like, "This is what we are trying to do this month"? It is an idea.

**Suzanne Webb:** I love the language that you are using, and "agile" is a very good way to describe perhaps how this could go along, but it is most definitely an iterative process. In basic language, we could have some sort of pilot to make sure that it is working, because it will not work from day one. I think that we are fully aware of that. There will have to be some tweaks to it. Maria, thank you very much.

Q198 **Lord Stevenson of Balmacara:** I was going to ask a bit about collective harms, but I think that you have already responded on that to my colleague Lord Clement-Jones.

What you have been talking about is the importance of having journalists and journalism of good quality to provide a context in which all these changes are happening and to keep an independent and moderating view on those who have power. What are journalists? Are they born? Are they made? How would you define them?

**Maria Ressa:** I think that my generation grew up in the golden age of journalism. I watched standards and ethics manuals get tested; I watched the mission of journalism, and I also watched the growth of legislation that brought along Fox. I have an American bent because I worked for so long with CNN. Then, there was the splintering of each of our worlds.

As for journalism, we have three pillars in Rappler. This is how I work now day to day. The three pillars are technology, journalism and community. Those three are interlocking circles for us. I look at information cascades in our society. At one point when we set up Rappler, I really thought that technology could help to build institutions bottom up. The rallying cry of Rappler at the beginning was social media for social good, so I did drink the Kool-Aid.

On the technology part, putting guardrails in place is the work that you are doing. To take CRISPR technology, for genetic research it was easier to put those guardrails in place for western nations, but I think that we have underestimated the impact of information and guardrails on tech. Journalists and I have been pushing to build tech. We are building tech that will roll out in November that will try to have fact and evidence-based discussions in time for our elections.

Journalists need to survive; independent media need to survive this period because, if you look at the dynamics and economics of it, the advertising model is dead for journalism. I have already talked a little bit about the commoditisation of news and the incentive structures for it, and yet the law in many democratic countries has journalists speaking truth to power. What makes journalists different? I think that it is courage. It is the standards, ethics and mission of journalism and then the courage to ask you questions that you do not want to answer, and to do it in a way that is professional. I love this part of my job.

The other part is that it was never about gotcha. The other thing the internet did was to make it all about trying to one-up everyone. The job of journalists was always to be a surrogate for the people. We were there to help you make the right choices moving forward because we believed in the collective good in the public sphere.

**Lord Stevenson of Balmacara:** That is very helpful.

Q199 **Lord Black of Brentwood:** Thank you for your inspiring and compelling evidence today. You said just now, as is absolutely right, "Please do not wait years for this to happen", although parliamentary processes are time consuming. You then said that shifts can perhaps happen at the same time as that legislative process is going on.

There is one other point to that. The flip-side of the coin of regulation is competition. You talked just now about how the business model is in effect dead for trusted independent journalism. Would your message to the UK Government be, "Don't just think about regulation. Get on and deal with the competition issues at the same time", as indeed they have in Australia?

**Maria Ressa:** On all fronts, the other important thing is data protection. That is another lever, but it is the surveillance capitalism model we all have to deal with. Using behavioural surplus data for monetisation leads to insidious manipulation. Antitrust in the US was a big thing.

The last part that we have not really talked about is data protection standards, because in the end the question I always have is: is it fair that the atomised post that I put on to Facebook can be pulled together by machine learning to create a model of me that knows me more than I know myself and is then sold to the highest bidder without my consent? Essentially, it is manipulating me. Those are some of the protections that, for us, are still missing. I do not know whether I answered your question. That is the exciting part.

**Lord Black of Brentwood:** You are tempting us down another road, but I think that you will have to come back another time. Thank you very much.

**The Chair:** Maria, that concludes the questions from the committee today. Thank you so much for joining us. We appreciate your testimony. It has been a real pleasure to hear from you.

**Maria Ressa:** Thank you for having me.